

Quantifying the use of large language models in published surgical abstracts

Kyle Lam PhD¹, Andrew Yiu MRCS¹, Catherine Simister MBBS¹, Nadir Omar MBBS¹, Jianing Qiu, PhD²
¹Imperial College London, UK ²Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE

Correspondence to Kyle Lam
at k.lam@imperial.ac.uk



Background

Large Language Models (LLMs) have revolutionized scientific writing with **human-like** text generation capabilities.

The use of LLMs threatens the **integrity** of surgical research if not properly disclosed.

It is becoming increasingly evident that LLMs are being used to publish scientific papers with **disclosure** potentially lagging behind actual use.

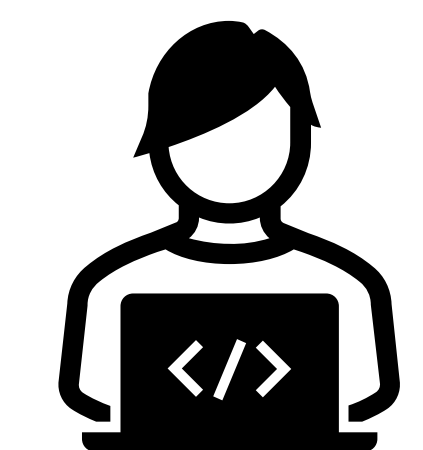
We aimed to quantify the use of LLMs among published abstracts across leading international surgical conferences and assess changes over time and across submission categories

Methods

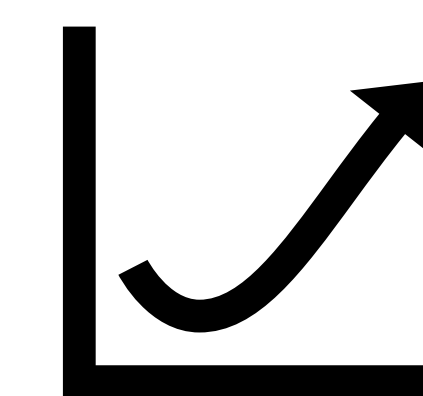
Study Design: Retrospective analysis of consecutively published abstracts in British Journal of Surgery between March 2022 and May 2025.

LLM Detection: A transformer based binary text classifier distinguishing human vs LLM-generated text. Model validated on 100 abstracts from 2015 paired with LLM-generated synthetic abstracts with identical titles achieving 100% accuracy.

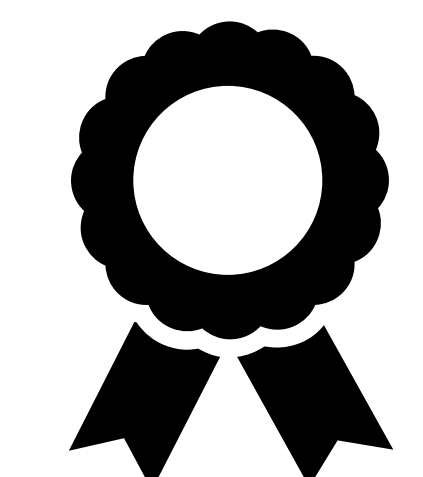
Results



8797 abstracts across **34 conferences** analyzed.



Pooled rates of estimated LLM use rising from 1.98% in 2023 to 38.3% in 2025.



Prize abstracts had highest rate of LLM use (28.8%) compared to oral abstracts (10.8%) and poster abstracts (9.2%) p<0.0001.



Higher rates of LLM use in European conferences vs UK in 2023 (OR 2.49, p = 0.032) and 2024 (OR 1.58, p = 0.001).

Key Messages

Over 1 in 3 abstracts from 2025 show probable LLM use with an over **threefold increase** year on year.

LLMs are powerful tools that can break down significant barriers for the communication of science, particularly where English is a second language but transparent disclosure of use is essential.

There is a need for **journal policies** on LLM disclosure, refined **detection methods** to improve sensitivity and specificity, and ongoing **discussions** around the role of AI in academic medicine.

