

# Assessing the Potential Role of ChatGPT in Spine Surgery Research

Ari Ruben Berg<sup>1</sup>, Dhruv Mendiratta, Isabel Herzog<sup>1</sup>, Ashok Para, Neil Kaushal, Michael Vives

<sup>1</sup>Rutgers New Jersey Medical School

**INTRODUCTION:** In less than two months, Chat Generative Pre-trained Transformer (ChatGPT), a complex machine learning model, has garnered more than 100 million users worldwide. ChatGPT's recent role as an author on scientific literature has academicians calling its validity into question. While it is generally able to maintain factual basis in its writing, the scholarly writing capabilities still have ways to go. Nevertheless, ChatGPT's ability to identify gaps of knowledge from the existing literature may allow it to function in novel research generation. Systematic reviews have gained popularity in surgical literature due to their ability to consolidate available data into one comprehensive analytical study. Therefore, ChatGPT may have a niche in identifying areas that need to be addressed via a systematic review. The aim of this study is to assess ChatGPT's ability to generate novel systematic ideas on topics within spine surgery.

**METHODS:** ChatGPT was instructed to give ten novel systematic review ideas for five popular topics in spine surgery literature: microdiscectomy, laminectomy, spinal fusion, kyphoplasty, and disc replacement. Therefore, a total of 50 systematic review ideas were generated. ChatGPT's ability to generate novel systematic review ideas was evaluated by conducting a literature search in PubMed, CINAHL, EMBASE, and Cochrane. The number of non-systematic review articles and number of systematic review papers that had been published on each ChatGPT-generated idea were recorded.

## RESULTS:

Overall, ChatGPT had a 66% accuracy rate in creating novel systematic review ideas. These 50 research ideas and their associated novelty assessments are shown in **Table 1**. More specifically, the accuracy rates were 80%, 80%, 30%, 70%, and 70% for microdiscectomy, laminectomy, spinal fusion, kyphoplasty, and disc replacement, respectively. However, there was a 32% rate of ChatGPT generating ideas for which there were 0 non-systematic review articles published. There was a 71.4%, 50%, 22.2%, 50%, 62.5%, and 51.2% success rate of generating novel systematic review ideas, for which there were also non-systematic reviews published, for microdiscectomy, laminectomy, spinal fusion, kyphoplasty, and disc replacement, and overall, respectively.

**DISCUSSION AND CONCLUSION:** Despite being limited in information past the end of 2021, ChatGPT still successfully generated novel ideas at an overall rate of 66%. As this technology improves through reinforcement learning and database expansion, there is a role for ChatGPT as a research tool. Its utility in generating new ideas can save spine surgeons time in having to parse through hundreds to thousands of articles by themselves. With more specific instructions and user-driven modulation, the accuracy of ChatGPT can be increased; future studies should investigate this.

Systematic Review	Number of Systematic Reviews	Number of Non-Systematic Reviews	Novelty Score	Success Rate
<b>Microdiscectomy</b>				
Comparison of minimally-invasive microdiscectomy (MID) and open discectomy	10	0	100	100%
Evaluation of the efficacy of minimally-invasive microdiscectomy (MID) compared to open discectomy	10	0	100	100%
The impact of patient factors such as obesity on microdiscectomy outcomes	10	0	100	100%
Review of the effectiveness of different approaches to microdiscectomy	10	0	100	100%
The role of patient comorbidities, including obesity, on microdiscectomy outcomes	10	0	100	100%
The impact of patient comorbidities and overall patient health on microdiscectomy outcomes	10	0	100	100%
The effectiveness of minimally-invasive microdiscectomy (MID) compared to open discectomy	10	0	100	100%
A comparison of MID to the effectiveness of microdiscectomy compared to alternative minimally-invasive techniques	10	0	100	100%
The role of patient comorbidities in determining outcomes after microdiscectomy	10	0	100	100%
A review of patient comorbidities associated with microdiscectomy including obesity and overall health	10	0	100	100%
<b>Laminectomy</b>				
Comparison of minimally-invasive laminectomy (MIL) and open laminectomy for spinal stenosis	10	0	100	100%
Review of the efficacy of minimally-invasive laminectomy (MIL) compared to open laminectomy	10	0	100	100%
Review of the impact of patient comorbidities on laminectomy outcomes	10	0	100	100%
Review of the role of patient comorbidities, including obesity, on laminectomy outcomes	10	0	100	100%
Review of the role of patient comorbidities and overall patient health on laminectomy outcomes	10	0	100	100%
Comparison of minimally-invasive laminectomy (MIL) and open laminectomy for spinal stenosis	10	0	100	100%
Review of the efficacy of minimally-invasive laminectomy (MIL) compared to open laminectomy	10	0	100	100%
Review of the impact of patient comorbidities on laminectomy outcomes	10	0	100	100%
Review of the role of patient comorbidities, including obesity, on laminectomy outcomes	10	0	100	100%
Review of the role of patient comorbidities and overall patient health on laminectomy outcomes	10	0	100	100%
Comparison of minimally-invasive laminectomy (MIL) and open laminectomy for spinal stenosis	10	0	100	100%
Review of the efficacy of minimally-invasive laminectomy (MIL) compared to open laminectomy	10	0	100	100%
Review of the impact of patient comorbidities on laminectomy outcomes	10	0	100	100%
Review of the role of patient comorbidities, including obesity, on laminectomy outcomes	10	0	100	100%
Review of the role of patient comorbidities and overall patient health on laminectomy outcomes	10	0	100	100%
<b>Spinal Fusion</b>				
The impact of patient comorbidities on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities, including obesity, on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities and overall patient health on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities, including obesity, on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities and overall patient health on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities, including obesity, on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities and overall patient health on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities, including obesity, on spinal fusion outcomes	100	0	100	100%
The impact of patient comorbidities and overall patient health on spinal fusion outcomes	100	0	100	100%
<b>Kyphoplasty</b>				
Comparison of minimally-invasive kyphoplasty (MIL) and open kyphoplasty	10	0	100	100%
Evaluation of the efficacy of minimally-invasive kyphoplasty (MIL) compared to open kyphoplasty	10	0	100	100%
Review of the impact of patient comorbidities on kyphoplasty outcomes	10	0	100	100%
Review of the role of patient comorbidities, including obesity, on kyphoplasty outcomes	10	0	100	100%
Review of the role of patient comorbidities and overall patient health on kyphoplasty outcomes	10	0	100	100%
Comparison of minimally-invasive kyphoplasty (MIL) and open kyphoplasty	10	0	100	100%
Evaluation of the efficacy of minimally-invasive kyphoplasty (MIL) compared to open kyphoplasty	10	0	100	100%
Review of the impact of patient comorbidities on kyphoplasty outcomes	10	0	100	100%
Review of the role of patient comorbidities, including obesity, on kyphoplasty outcomes	10	0	100	100%
Review of the role of patient comorbidities and overall patient health on kyphoplasty outcomes	10	0	100	100%
Comparison of minimally-invasive kyphoplasty (MIL) and open kyphoplasty	10	0	100	100%
Evaluation of the efficacy of minimally-invasive kyphoplasty (MIL) compared to open kyphoplasty	10	0	100	100%
Review of the impact of patient comorbidities on kyphoplasty outcomes	10	0	100	100%
Review of the role of patient comorbidities, including obesity, on kyphoplasty outcomes	10	0	100	100%
Review of the role of patient comorbidities and overall patient health on kyphoplasty outcomes	10	0	100	100%
<b>Disc Replacement</b>				
Comparison of minimally-invasive disc replacement (MIL) and open disc replacement	100	0	100	100%
Evaluation of the efficacy of minimally-invasive disc replacement (MIL) compared to open disc replacement	100	0	100	100%
Review of the impact of patient comorbidities on disc replacement outcomes	100	0	100	100%
Review of the role of patient comorbidities, including obesity, on disc replacement outcomes	100	0	100	100%
Review of the role of patient comorbidities and overall patient health on disc replacement outcomes	100	0	100	100%
Comparison of minimally-invasive disc replacement (MIL) and open disc replacement	100	0	100	100%
Evaluation of the efficacy of minimally-invasive disc replacement (MIL) compared to open disc replacement	100	0	100	100%
Review of the impact of patient comorbidities on disc replacement outcomes	100	0	100	100%
Review of the role of patient comorbidities, including obesity, on disc replacement outcomes	100	0	100	100%
Review of the role of patient comorbidities and overall patient health on disc replacement outcomes	100	0	100	100%
Comparison of minimally-invasive disc replacement (MIL) and open disc replacement	100	0	100	100%
Evaluation of the efficacy of minimally-invasive disc replacement (MIL) compared to open disc replacement	100	0	100	100%
Review of the impact of patient comorbidities on disc replacement outcomes	100	0	100	100%
Review of the role of patient comorbidities, including obesity, on disc replacement outcomes	100	0	100	100%
Review of the role of patient comorbidities and overall patient health on disc replacement outcomes	100	0	100	100%

Table 1. ChatGPT-generated systematic review ideas and respective accuracy rates for microdiscectomy, laminectomy, spinal fusion, kyphoplasty, and disc replacement. Number of non-systematic reviews (non-SRs), number of systematic reviews (SRs), and novelty score shown for each ChatGPT-generated idea.