Sixth International Joint Meeting on THORACIC SURGERY

Barcelona - 20th, 21st and 22nd November 2024
Auditorio Foment del Treball Nacional. Barcelona (Spain)

11th International Meeting on General Thoracic Surgery
Clínic Barcelona | Universitat de Barcelona | Hospital Universitari Sagrat Cor

5th Meeting of the Thoracic Oncology, Thoracic Surgery, Techniques & Transplant, Respiratory Nursing and Respiratory Physiotherapy Areas of the Spanish Society of Pneumology and Thoracic Surgery (SEPAR)
Sociedad Española de Neumología y Cirugía Torácica (SEPAR)

10th International Workshop on Surgical Exploration of the Mediastinum and Systematic Nodal Dissection
Hospital Universitari Mútua Terrassa | Universitat de Barcelona

10th International Workshop on Surgical Exploration of the Mediastinum and Systematic Nodal Dissection
Hospital Universitari Mútua Terrassa | Universitat de Barcelona

3rd Joint Meeting of the Spanish Society of Thoracic Surgery (SECT)
SECT Sociedad Española de Cirugía Torácica

30th Congress of the "Asociación Iberoamericana de Cirugía Torácica" AIACT

Sant Pau Campus Salut Barcelona | UAB Universitat Autònoma de Barcelona | MAR | upf Universitat Pompeu Fabra Barcelona

CAN AN ARTIFICIAL INTELLIGENCE TOOL LIKE SCHOLARGPT ANSWER THE OUTPATIENT CONSULTATIONS OF A THORACIC SURGEON?

Judith Marcè i Igual; Carlos Javier Déniz Armengol; Ivan Macia Vidueira; Francisco Rivas Doyague; Anna Muñoz Fos; Marina Paradela De La Morena; Camilo Andrés Moreno Mayorga; Inés Seratosa de Queralt; Marta Garcia Miró; Tania Rodriquez-Martos Reppetto et Amaia Ojanguren Arranz.

Thoracic Surgery Department, Hospital Universitari de Bellvitge, Barcelona (Spain)

Background: Artificial Intelligence tools like ScholarGPT (a Large Language model by OpenAI) are becoming more accessible to patients and prevalent in everyday clinical practice. However, questions are raised about their accuracy and usefulness in healthcare settings.
Objective: This study aimed to compare the management of thoracic surgery outpatient consultations between thoracic surgeons and ScholarGPT.
Materials and Methods: We selected a week of thoracic surgery outpatient consultations (n=81), introduced the reason for consulting to the LLM, and assessed the responses on a scale of 0 (minimum) to 5 (maximum) using a predefined rubric that measured the likeness with the surgeons' management. We studied the type of visit (physical or telephone), type of consult (1stvisit or control), diagnosis, and attending surgeon.
Results: The mean score obtained was 3.67. The ANOVA and t-score didn't find significant differences between the type of visit (p=0.35) and type of consult (p=0.67). We also didn't see significant statistical differences in the scores obtained depending on the diagnosis (p=0.14). The main reasons ScholarGPT didn't score top qualifications were that it asked for too many complimentary tests or extra treatments the surgeons did not consider necessary.
Limitations: The assigned score was decided by one individual and their subjective evaluation of the concordance. Also, ScholarGPT couldn't access the patients' full clinical records.
Conclusions: This study shows that LLMs couldn't answer thoracic surgery consultations by themselves but might be a useful tool in assisting daily clinical practice. However, further research with a bigger sample or a prospective study is needed.