Large language models for neurosurgery

- Clinical feasibility of AI Doctors: Evaluating the replacement potential of large language models in outpatient settings for central nervous system tumors
- Role for Left Dorsomedial Prefrontal Cortex in Self-Generated, but not Externally Cued, Language Production
- Do LLMs Have 'the Eye' for MRI? Evaluating GPT-4o, Grok, and Gemini on Brain MRI Performance: First Evaluation of Grok in Medical Imaging and a Comparative Analysis
- Knowledge Graph-Enhanced Deep Learning Model (H-SYSTEM) for Hypertensive Intracerebral Hemorrhage: Model Development and Validation
- Performance analysis of large language models in multi-disease detection from chest computed tomography reports: a comparative study: Experimental Research
- Transforming neurosurgical practice with large language models: comparative performance of ChatGPT-omni and Gemini in complex case management
- Retrieval-Augmented Generation with Large Language Models in Radiology: From Theory to Practice
- Comparison of quality, empathy and readability of physician responses versus chatbot responses to common cerebrovascular neurosurgical questions on a social media platform

---

A specialized type of machine learning model, specifically designed to understand, generate, and manipulate natural language text.

Large language models (LLMs), such as ChatGPT, have transformative potential in neurosurgery by enhancing clinical decision-making, patient care, and research.

Clinical Decision Support

LLMs can assist neurosurgeons by processing complex clinical data and providing recommendations based on vast medical literature:

- **Preoperative Planning**: Analyzing patient-specific data to suggest surgical approaches, risks, and alternatives.

- **Intraoperative Guidance**: Supporting neurosurgeons with real-time queries during procedures, such as anatomical details or technique clarifications.

- **Postoperative Care**: Offering evidence-based guidelines for managing complications or optimizing recovery protocols.

—

2. Patient Communication and Education Neurosurgical procedures are often complex and daunting for patients. LLMs can: - **Simplify Medical Information**: Translate technical jargon into accessible language. - **Personalized Patient Interactions**: Generate tailored explanations of

conditions like brain tumors, spinal disorders, or vascular malformations. - **Virtual Assistance**: Provide 24/7 access to patient queries, improving satisfaction and understanding.

—

3. Research and Knowledge Management LLMs streamline access to and analysis of vast scientific resources: - **Literature Summarization**: Condensing thousands of research articles into key takeaways. - **Data Synthesis**: Identifying trends, gaps, and future directions in neurosurgical research. - **Hypothesis Generation**: Suggesting innovative research questions by linking interdisciplinary knowledge.

—

4. Training and Education LLMs can enhance neurosurgical education through: - **Simulated Learning**: Creating realistic scenarios for decision-making training. - **Question and Answer**: Serving as a resource for residents to clarify doubts. - **Surgical Protocol Guidance**: Detailing step-by-step instructions for procedures.

—

5. Administrative Efficiency Neurosurgery often involves substantial administrative work. LLMs can automate tasks such as: - **Documentation**: Generating operation notes, discharge summaries, and referral letters. - **Scheduling and Coordination**: Assisting in the logistics of complex surgical cases. - **Regulatory Compliance**: Ensuring documentation aligns with institutional and legal standards.

—

6. Predictive Analytics LLMs integrated with machine learning models can: - **Outcome Prediction**: Anticipating surgical outcomes based on patient data. - **Risk Stratification**: Identifying high-risk patients for tailored interventions.

—

7. Ethical and Practical Considerations While LLMs have immense potential, they must be used responsibly: - **Bias and Accuracy**: Ensuring outputs are evidence-based and free from bias. - **Confidentiality**: Protecting sensitive patient information. - **Reliability**: Using LLMs as adjuncts, not replacements, for professional judgment.

—

Future Directions Integration of LLMs with neurosurgical tools like neuronavigation systems, robotic platforms, and real-time imaging could revolutionize intraoperative and perioperative care. Collaborative AI-driven platforms can also foster global neurosurgical research and training.

Literature

1: Ozenbas C, Engin D, Altinok T, Akcay E, Aktas U, Tabanli A. ChatGPT-4o's Performance in Brain Tumor Diagnosis and MRI Findings: A Comparative Analysis with Radiologists. Acad Radiol. 2025 Feb

8:S1076-6332(25)00089-3. doi: 10.1016/j.acra.2025.01.033. Epub ahead of print. PMID: 39924377.

2: Aydin S, Karabacak M, Vlachos V, Margetis K. Navigating the potential and pitfalls of large language models in patient-centered medication guidance and self-decision support. Front Med (Lausanne). 2025 Jan 23;12:1527864. doi: 10.3389/fmed.2025.1527864. PMID: 39917061; PMCID: PMC11798948.

3: Gautam D, Kellmeyer P. Exploring the Credibility of Large Language Models for Mental Health Support: Protocol for a Scoping Review. JMIR Res Protoc. 2025 Jan 29;14:e62865. doi: 10.2196/62865. PMID: 39879615.

4: Roy JM, Self DM, Isch E, Musmar B, Lan M, Keppetipola K, Koduri S, Pontarelli MK, Tjoumakaris SI, Gooch MR, Rosenwasser RH, Jabbour PM. Evaluating Large Language Models for Automated CPT Code Prediction in Endovascular Neurosurgery. J Med Syst. 2025 Jan 24;49(1):15. doi: 10.1007/s10916-025-02149-4. PMID: 39853605.

5: Genel O, Kalyal N, Perera A, Lee KS, Chowdhury YA, Lavrador JP. Letter to the Editor Regarding “Bridging the Gap: Can Large Language Models Match Human Expertise in Writing Neurosurgical Operative Notes?”. World Neurosurg. 2025 Jan 29;195:123668. doi: 10.1016/j.wneu.2025.123668. Epub ahead of print. PMID: 39814177.

6: Alber DA, Yang Z, Alyakin A, Yang E, Rai S, Valliani AA, Zhang J, Rosenbaum GR, Amend-Thomas AK, Kurland DB, Kremer CM, Eremiev A, Negash B, Wiggan DD, Nakatsuka MA, Sangwon KL, Neifert SN, Khan HA, Save AV, Palla A, Grin EA, Hedman M, Nasir-Moin M, Liu XC, Jiang LY, Mankowski MA, Segev DL, Aphinyanaphongs Y, Riina HA, Golfinos JG, Orringer DA, Kondziolka D, Oermann EK. Medical large language models are vulnerable to data-poisoning attacks. Nat Med. 2025 Jan 8. doi: 10.1038/s41591-024-03445-1. Epub ahead of print. PMID: 39779928.

7: Heisinger S, Salzmann SN, Senker W, Aspalter S, Oberndorfer J, Matzner MP, Stienen MN, Motov S, Huber D, Grohs JG. ChatGPT's Performance in Spinal Metastasis Cases-Can We Discuss Our Complex Cases with ChatGPT? J Clin Med. 2024 Dec 23;13(24):7864. doi: 10.3390/jcm13247864. PMID: 39768787; PMCID: PMC11727723.

8: Zong H, Wu R, Cha J, Wang J, Wu E, Li J, Zhou Y, Zhang C, Feng W, Shen B. Large Language Models in Worldwide Medical Exams: Platform Development and Comprehensive Analysis. J Med Internet Res. 2024 Dec 27;26:e66114. doi: 10.2196/66114. PMID: 39729356; PMCID: PMC11724220.

9: Nasra M, Jaffri R, Pavlin-Premrl D, Kok HK, Khabaza A, Barras C, Slater LA, Yazdabadi A, Moore J, Russell J, Smith P, Chandra RV, Brooks M, Jhamb A, Chong W, Maingard J, Asadi H. Can artificial intelligence improve patient educational material readability? A systematic review and narrative synthesis. Intern Med J. 2025 Jan;55(1):20-34. doi: 10.1111/imj.16607. Epub 2024 Dec 25. PMID: 39720869.

10: Fennig U, Yom-Tov E, Savitsky L, Nissan J, Altman K, Loebenstein R, Boxer M, Weinberg N, Gofrit SG, Maggio N. Bridging the conversational gap in epilepsy: Using large language models to reveal insights into patient behavior and concerns from online discussions. Epilepsia. 2024 Dec 10. doi: 10.1111/epi.18226. Epub ahead of print. PMID: 39655574.

11: Bongco EDA, Cua SKN, Hernandez MALU, Pascual JSG, Khu KJO. The performance of ChatGPT versus neurosurgery residents in neurosurgical board examination-like questions: a systematic review and meta-analysis. Neurosurg Rev. 2024 Dec 7;47(1):892. doi: 10.1007/s10143-024-03144-y. PMID: 39643792.

12: Alyakin A, Kurland D, Alber DA, Sangwon KL, Li D, Tsirigos A, Leuthardt E, Kondziolka D, Oermann

EK. CNS-CLIP: Transforming a Neurosurgical Journal Into a Multimodal Medical Model. Neurosurgery. 2024 Dec 5. doi: 10.1227/neu.0000000000003297. Epub ahead of print. PMID: 39636129.

13: Almekkawi AK, Caruso JP, Anand S, Hawkins AM, Rauf R, Al-Shaikhli M, Aoun SG, Bagley CA. Comparative Analysis of Large Language Models and Spine Surgeons in Surgical Decision-Making and Radiological Assessment for Spine Pathologies. World Neurosurg. 2024 Dec 23;194:123531. doi: 10.1016/j.wneu.2024.11.114. Epub ahead of print. PMID: 39622288.

14: Brügge E, Ricchizzi S, Arenbeck M, Keller MN, Schur L, Stummer W, Holling M, Lu MH, Darici D. Large language models improve clinical decision making of medical students through patient simulation and structured feedback: a randomized controlled trial. BMC Med Educ. 2024 Nov 28;24(1):1391. doi: 10.1186/s12909-024-06399-7. PMID: 39609823; PMCID: PMC11605890.

15: Luo X, Rechardt A, Sun G, Nejad KK, Yáñez F, Yilmaz B, Lee K, Cohen AO, Borghesani V, Pashkov A, Marinazzo D, Nicholas J, Salatiello A, Sucholutsky I, Minervini P, Razavi S, Rocca R, Yusifov E, Okalova T, Gu N, Ferianc M, Khona M, Patil KR, Lee PS, Mata R, Myers NE, Bizley JK, Musslick S, Bilgin IP, Niso G, Ales JM, Gaebler M, Ratan Murty NA, Loued-Khenissi L, Behler A, Hall CM, Dafflon J, Bao SD, Love BC. Large language models surpass human experts in predicting neuroscience results. Nat Hum Behav. 2024 Nov 27. doi: 10.1038/s41562-024-02046-9. Epub ahead of print. PMID: 39604572.

16: Patil A, Serrato P, Chisvo N, Arnaout O, See PA, Huang KT. Large language models in neurosurgery: a systematic review and meta-analysis. Acta Neurochir (Wien). 2024 Nov 23;166(1):475. doi: 10.1007/s00701-024-06372-9. PMID: 39579215.

17: Azkia H, Harritshøj LH, Nielsen C, Agerlin N, Jensen MG, Hillingsø JG, Andersen PC, Krogsgaard MR. A local, non-commercial tissue bank connected to an organ donor program can produce musculoskeletal allografts of uniform quality at very low costs - ten years' experience. Cell Tissue Bank. 2024 Nov 23;26(1):1. doi: 10.1007/s10561-024-10151-2. PMID: 39578258; PMCID: PMC11584506.

18: Aydin S, Karabacak M, Vlachos V, Margetis K. Large language models in patient education: a scoping review of applications in medicine. Front Med (Lausanne). 2024 Oct 29;11:1477898. doi: 10.3389/fmed.2024.1477898. PMID: 39534227; PMCID: PMC11554522.

19: Staartjes VE, Zanier O, da Mutten R, Serra C, Regli L. Machine Intelligence in Cerebrovascular and Endovascular Neurosurgery. Adv Exp Med Biol. 2024;1462:383-395. doi: 10.1007/978-3-031-64892-2_23. PMID: 39523278.

20: Di Ieva A, Stewart C, Suero Molina E. Large Language Models in Neurosurgery. Adv Exp Med Biol. 2024;1462:177-198. doi: 10.1007/978-3-031-64892-2_11. PMID: 39523266.

21: Lan Z, Chen Y, Rushmore J, Zekelman L, Makris N, Rathi Y, Golby AJ, Zhang F, O'Donnell LJ. Fiber Microstructure Quantile (FMQ) Regression: A Novel Statistical Approach for Analyzing White Matter Bundles from Periphery to Core. bioRxiv [Preprint]. 2024 Oct 22:2024.10.19.619237. doi: 10.1101/2024.10.19.619237. PMID: 39484397; PMCID: PMC11526951.

22: Hu J, Fu J, Zhao W, Lou P, Feng M, Ren H, Feng S, Li Y, Fang A. Characterizing pituitary adenomas in clinical notes: Corpus construction and its application in LLMs. Health Informatics J. 2024 Oct-Dec;30(4):14604582241291442. doi: 10.1177/14604582241291442. PMID: 39379071.

23: Su Z, Tang G, Huang R, Qiao Y, Zhang Z, Dai X. Based on Medicine, The Now and Future of Large

Language Models. Cell Mol Bioeng. 2024 Sep 16;17(4):263-277. doi: 10.1007/s12195-024-00820-3. PMID: 39372551; PMCID: PMC11450117.

24: Sarikonda A, Isch E, Self M, Sambangi A, Carreras A, Sivaganesan A, Harrop J, Jallo J. Evaluating the Adherence of Large Language Models to Surgical Guidelines: A Comparative Analysis of Chatbot Recommendations and North American Spine Society (NASS) Coverage Criteria. Cureus. 2024 Sep 3;16(9):e68521. doi: 10.7759/cureus.68521. PMID: 39364514; PMCID: PMC11448007.

25: Tong L, Zhang C, Liu R, Yang J, Sun Z. Comparative performance analysis of large language models: ChatGPT-3.5, ChatGPT-4 and Google Gemini in glucocorticoid-induced osteoporosis. J Orthop Surg Res. 2024 Sep 18;19(1):574. doi: 10.1186/s13018-024-04996-2. PMID: 39289734; PMCID: PMC11409482.

26: Arfaie S, Sadegh Mashayekhi M, Mofatteh M, Ma C, Ruan R, MacLean MA, Far R, Saini J, Harmsen IE, Duda T, Gomez A, Rebchuk AD, Pingbei Wang A, Rasiah N, Guo E, Fazlollahi AM, Rose Swan E, Amin P, Mohammed S, Atkinson JD, Del Maestro RF, Girgis F, Kumar A, Das S. ChatGPT and neurosurgical education: A crossroads of innovation and opportunity. J Clin Neurosci. 2024 Nov;129:110815. doi: 10.1016/j.jocn.2024.110815. Epub 2024 Sep 4. PMID: 39236407.

27: Isch EL, Sarikonda A, Sambangi A, Carreras A, Sircar A, Self DM, Habarth- Morales TE, Caterson EJ, Aycart M. Evaluating the Efficacy of Large Language Models in CPT Coding for Craniofacial Surgery: A Comparative Analysis. J Craniofac Surg. 2024 Sep 2. doi: 10.1097/SCS.0000000000010575. Epub ahead of print. PMID: 39221924.

28: Brown EDL, Maity A, Ward M, Toscano D, Baum GR, Mittler MA, Lo SL, D'Amico RS. ChatGPT as a Decision Support Tool in the Management of Chiari I Malformation: A Comparison to 2023 CNS Guidelines. World Neurosurg. 2024 Nov;191:e304-e332. doi: 10.1016/j.wneu.2024.08.122. Epub 2024 Aug 28. PMID: 39214295.

29: Niset A, El Hadwe S, Englebert A, Barrit S. AI in emergency medicine: Building literacy or castles in the air. Am J Emerg Med. 2025 Jan;87:145-146. doi: 10.1016/j.ajem.2024.08.020. Epub 2024 Aug 21. PMID: 39198071.

30: Li S, Wei W, Li X, Ma L, Li Q, Sun X, Chen X. Impacts of blended learning with BOPPPS model on Chinese medical undergraduate students: a comprehensive systematic review and meta-analysis of 44 studies. BMC Med Educ. 2024 Aug 23;24(1):914. doi: 10.1186/s12909-024-05917-x. PMID: 39180016; PMCID: PMC11344447.

31: Ali A, Kumar RP, Polavarapu H, Lavadi RS, Mahavadi A, Legarreta AD, Hudson JS, Shah M, Paul D, Mooney J, Dietz N, Fields DP, Hamilton DK, Agarwal N. Bridging the Gap: Can Large Language Models Match Human Expertise in Writing Neurosurgical Operative Notes? World Neurosurg. 2024 Dec;192:e34-e41. doi: 10.1016/j.wneu.2024.08.062. Epub 2024 Aug 15. PMID: 39153569.

32: Sommer J, Dierksen F, Zeevi T, Tran AT, Avery EW, Mak A, Malhotra A, Matouk CC, Falcone GJ, Torres-Lopez V, Aneja S, Duncan J, Sansing LH, Sheth KN, Payabvash S. Deep learning for prediction of post-thrombectomy outcomes based on admission CT angiography in large vessel occlusion stroke. Front Artif Intell. 2024 Aug 1;7:1369702. doi: 10.3389/frai.2024.1369702. PMID: 39149161; PMCID: PMC11324606.

33: Nietsch KS, Shrestha N, Mazudie Ndjonko LC, Ahmed W, Mejia MR, Zaidat B, Ren R, Duey AH, Li SQ, Kim JS, Hidden KA, Cho SK. Can Large Language Models (LLMs) Predict the Appropriate Treatment of Acute Hip Fractures in Older Adults? Comparing Appropriate Use Criteria With Recommendations From ChatGPT. J Am Acad Orthop Surg Glob Res Rev. 2024 Aug 9;8(8):e24.00206. doi:

10.5435/JAAOSGlobal-D-24-00206. PMID: 39137403; PMCID: PMC11319315.

34: Mohamed AA, Lucke-Wold B. Apple Intelligence in neurosurgery. Neurosurg Rev. 2024 Jul 15;47(1):327. doi: 10.1007/s10143-024-02568-w. PMID: 39004685.

35: Li J, Zong H, Wu E, Wu R, Peng Z, Zhao J, Yang L, Xie H, Shen B. Exploring the potential of artificial intelligence to enhance the writing of english academic papers by non-native english-speaking medical students - the educational application of ChatGPT. BMC Med Educ. 2024 Jul 9;24(1):736. doi: 10.1186/s12909-024-05738-y. PMID: 38982429; PMCID: PMC11232216.

36: Lang SP, Yoseph ET, Gonzalez-Suarez AD, Kim R, Fatemi P, Wagner K, Maldaner N, Stienen MN, Zygourakis CC. Analyzing Large Language Models' Responses to Common Lumbar Spine Fusion Surgery Questions: A Comparison Between ChatGPT and Bard. Neurospine. 2024 Jun;21(2):633-641. doi: 10.14245/ns.2448098.049. Epub 2024 Jun 30. PMID: 38955533; PMCID: PMC11224745.

37: Mohamed AA, Lucke-Wold B. Text-to-video generative artificial intelligence: sora in neurosurgery. Neurosurg Rev. 2024 Jun 13;47(1):272. doi: 10.1007/s10143-024-02514-w. PMID: 38867134.

38: Dagli MM, Oettl FC, Gujral J, Malhotra K, Ghenbot Y, Yoon JW, Ozturk AK, Welch WC. Clinical Accuracy, Relevance, Clarity, and Emotional Sensitivity of Large Language Models to Surgical Patient Questions: Cross-Sectional Study. JMIR Form Res. 2024 Jun 7;8:e56165. doi: 10.2196/56165. PMID: 38848553; PMCID: PMC11193072.

39: Rao SJ, Isath A, Krishnan P, Tangsrivimol JA, Virk HUH, Wang Z, Glicksberg BS, Krittanawong C. ChatGPT: A Conceptual Review of Applications and Utility in the Field of Medicine. J Med Syst. 2024 Jun 5;48(1):59. doi: 10.1007/s10916-024-02075-x. PMID: 38836893.

40: Brown EDL, Ward M, Maity A, Mittler MA, Larry Lo SF, D'Amico RS. Enhancing Diagnostic Support for Chiari Malformation and Syringomyelia: A Comparative Study of Contextualized ChatGPT Models. World Neurosurg. 2024 Sep;189:e86-e107. doi: 10.1016/j.wneu.2024.05.172. Epub 2024 Jun 1. PMID: 38830507.

41: Lawson McLean A. Constructing knowledge: the role of AI in medical learning. J Am Med Inform Assoc. 2024 Aug 1;31(8):1797-1798. doi: 10.1093/jamia/ocae124. PMID: 38812088; PMCID: PMC11258398.

42: Daungsupawong H, Wiwanitkit V. Comment on “Incorporating large language models into academic neurosurgery: embracing the new era”. Neurosurg Rev. 2024 May 24;47(1):230. doi: 10.1007/s10143-024-02470-5. PMID: 38787480.

43: Kumar RP, Sivan V, Bachir H, Sarwar SA, Ruzicka F, O'Malley GR, Lobo P, Morales IC, Cassimatis ND, Hundal JS, Patel NV. Can Artificial Intelligence Mitigate Missed Diagnoses by Generating Differential Diagnoses for Neurosurgeons? World Neurosurg. 2024 Jul;187:e1083-e1088. doi: 10.1016/j.wneu.2024.05.052. Epub 2024 May 16. PMID: 38759788.

44: Barrit S, Hadwe SE, Carron R, Madsen JR. Letter to the Editor. Rise of large language models in neurosurgery. J Neurosurg. 2024 May 17;141(3):878-880. doi: 10.3171/2024.3.JNS24610. PMID: 38759241.

45: Fan C, Hahn N, Kamdar F, Avansino D, Wilson GH, Hochberg L, Shenoy KV, Henderson JM, Willett FR. Plug-and-Play Stability for Intracortical Brain- Computer Interfaces: A One-Year Demonstration of

Seamless Brain-to-Text Communication. Adv Neural Inf Process Syst. 2023 Dec;36:42258-42270. PMID: 38738213; PMCID: PMC11086983.

46: Aamir A, Hafsa H. “Incorporating large language models into academic neurosurgery: embracing the new era”. Neurosurg Rev. 2024 May 10;47(1):211. doi: 10.1007/s10143-024-02452-7. PMID: 38724772.

47: Lawson McLean A, Senft C. Assessing the efficacy of large language models in spinal health information dissemination. Brain Spine. 2024 Apr 12;4:102815. doi: 10.1016/j.bas.2024.102815. PMID: 38690090; PMCID: PMC11059440.

48: Pressman SM, Borna S, Gomez-Cabello CA, Haider SA, Haider C, Forte AJ. AI and Ethics: A Systematic Review of the Ethical Considerations of Large Language Model Use in Surgery Research. Healthcare (Basel). 2024 Apr 13;12(8):825. doi: 10.3390/healthcare12080825. PMID: 38667587; PMCID: PMC11050155.

49: Niset A, El Hadwe S, Barrit S. Did GPT-4 really perform electrocardiography assessment? Am J Emerg Med. 2024 Jun;80:217-218. doi: 10.1016/j.ajem.2024.04.008. Epub 2024 Apr 9. PMID: 38604875.

50: Siow I, Narasimhalu K, Lee KS, Tan HK, Ting SKS, Hameed S, Chang HM, De Silva DA, Chen CLH, Tan EK. Predictors of post stroke cognitive impairment: VITATOPS cognition substudy. J Stroke Cerebrovasc Dis. 2024 Jun;33(6):107718. doi: 10.1016/j.jstrokecerebrovasdis.2024.107718. Epub 2024 Apr 10. PMID: 38604352.

51: Ali YH, Bodkin K, Rigotti-Thompson M, Patel K, Card NS, Bhaduri B, Nason- Tomaszewski SR, Mifsud DM, Hou X, Nicolas C, Allcroft S, Hochberg LR, Au Yong N, Stavisky SD, Miller LE, Brandman DM, Pandarinath C. BRAND: a platform for closed-loop experiments with deep network models. J Neural Eng. 2024 Apr 17;21(2):026046. doi: 10.1088/1741-2552/ad3b3a. PMID: 38579696; PMCID: PMC11021878.

52: Huang KT, Mehta NH, Gupta S, See AP, Arnaout O. Evaluation of the safety, accuracy, and helpfulness of the GPT-4.0 Large Language Model in neurosurgery. J Clin Neurosci. 2024 May;123:151-156. doi: 10.1016/j.jocn.2024.03.021. Epub 2024 Apr 4. PMID: 38574687.

53: Mehandru N, Miao BY, Almaraz ER, Sushil M, Butte AJ, Alaa A. Evaluating large language models as agents in the clinic. NPJ Digit Med. 2024 Apr 3;7(1):84. doi: 10.1038/s41746-024-01083-y. PMID: 38570554; PMCID: PMC10991271.

54: Waisberg E, Ong J, Masalkhi M, Lee AG. OpenAI's Sora in medicine: revolutionary advances in generative artificial intelligence for healthcare. Ir J Med Sci. 2024 Aug;193(4):2105-2107. doi: 10.1007/s11845-024-03680-y. Epub 2024 Apr 4. PMID: 38570404.

55: Guo E, Gupta M, Sinha S, Rössler K, Tatagiba M, Akagami R, Al-Mefty O, Sugiyama T, Stieg PE, Pickett GE, de Lotbiniere-Bassett M, Singh R, Lama S, Sutherland GR. neuroGPT-X: toward a clinic-ready large language model. J Neurosurg. 2023 Oct 6;140(4):1041-1053. doi: 10.3171/2023.7.JNS23573. PMID: 38564804.

56: Waisberg E, Ong J, Masalkhi M, Lee AG. Concerns with OpenAI's Sora in Medicine. Ann Biomed Eng. 2024 Aug;52(8):1932-1934. doi: 10.1007/s10439-024-03505-0. Epub 2024 Apr 1. PMID: 38558354.

57: Stengel FC, Stienen MN, Ivanov M, Gandía-González ML, Raffa G, Ganau M, Whitfield P, Motov S.

Can AI pass the written European Board Examination in Neurological Surgery? - Ethical and practical issues. Brain Spine. 2024 Feb 13;4:102765. doi: 10.1016/j.bas.2024.102765. PMID: 38510593; PMCID: PMC10951784.

58: Lawson McLean A, Wu Y, Lawson McLean AC, Hristidis V. Large language models as decision aids in neuro-oncology: a review of shared decision-making applications. J Cancer Res Clin Oncol. 2024 Mar 19;150(3):139. doi: 10.1007/s00432-024-05673-x. PMID: 38503921; PMCID: PMC10951032.

59: Van Veen D, Van Uden C, Blankemeier L, Delbrouck JB, Aali A, Bluethgen C, Pareek A, Polacin M, Reis EP, Seehofnerová A, Rohatgi N, Hosamani P, Collins W, Ahuja N, Langlotz CP, Hom J, Gatidis S, Pauly J, Chaudhari AS. Adapted large language models can outperform medical experts in clinical text summarization. Nat Med. 2024 Apr;30(4):1134-1142. doi: 10.1038/s41591-024-02855-5. Epub 2024 Feb 27. PMID: 38413730; PMCID: PMC11479659.

60: Maillard AM, Romascano D, Villalón-Reina JE, Moreau CA, Almeida Osório JM, Richetin S, Junod V, Yu P, Misic B, Thompson PM, Fornari E, Gygax MJ, Jacquemont S, Chabane N, Rodríguez-Herreros B. Pervasive alterations of intra-axonal volume and network organization in young children with a 16p11.2 deletion. Transl Psychiatry. 2024 Feb 14;14(1):95. doi: 10.1038/s41398-024-02810-5. PMID: 38355713; PMCID: PMC10866898.

61: Wang Y, Zuo J, Duan C, Peng H, Huang J, Zhao L, Zhang L, Dong Z. Large language models assisted multi-effect variants mining on cerebral cavernous malformation familial whole genome sequencing. Comput Struct Biotechnol J. 2024 Feb 1;23:843-858. doi: 10.1016/j.csbj.2024.01.014. PMID: 38352937; PMCID: PMC10861960.

62: Schonfeld E, Pant A, Shah A, Sadeghzadeh S, Pangal D, Rodrigues A, Yoo K, Marianayagam N, Haider G, Veeravagu A. Evaluating Computer Vision, Large Language, and Genome-Wide Association Models in a Limited Sized Patient Cohort for Pre-Operative Risk Stratification in Adult Spinal Deformity Surgery. J Clin Med. 2024 Jan 23;13(3):656. doi: 10.3390/jcm13030656. PMID: 38337352; PMCID: PMC10856542.

63: Kasthuri VS, Glueck J, Pham H, Daher M, Balmaceno-Criss M, McDonald CL, Diebo BG, Daniels AH. Assessing the Accuracy and Reliability of AI-Generated Responses to Patient Questions Regarding Spine Surgery. J Bone Joint Surg Am. 2024 Jun 19;106(12):1136-1142. doi: 10.2106/JBJS.23.00914. Epub 2024 Feb 9. PMID: 38335266.

64: van Diessen E, van Amerongen RA, Zijlmans M, Otte WM. Potential merits and flaws of large language models in epilepsy care: A critical review. Epilepsia. 2024 Apr;65(4):873-886. doi: 10.1111/epi.17907. Epub 2024 Feb 2. PMID: 38305763.

65: Dubinski D, Won SY, Trnovec S, Behmanesh B, Baumgarten P, Dinc N, Konczalla J, Chan A, Bernstock JD, Freiman TM, Gessler F. Leveraging artificial intelligence in neurosurgery-unveiling ChatGPT for neurosurgical discharge summaries and operative reports. Acta Neurochir (Wien). 2024 Jan 26;166(1):38. doi: 10.1007/s00701-024-05908-3. PMID: 38277081; PMCID: PMC10817836.

66: Zangrossi P, Martini M, Guerrini F, DE Bonis P, Spena G. Large language model, AI and scientific research: why ChatGPT is only the beginning. J Neurosurg Sci. 2024 Apr;68(2):216-224. doi: 10.23736/S0390-5616.23.06171-4. Epub 2024 Jan 23. PMID: 38261307.

67: Wu Y, Zhang Z, Dong X, Hong S, Hu Y, Liang P, Li L, Zou B, Wu X, Wang D, Chen H, Qiu H, Tang H, Kang K, Li Q, Zhai X. Evaluating the performance of the language model ChatGPT in responding to

common questions of people with epilepsy. Epilepsy Behav. 2024 Feb;151:109645. doi: 10.1016/j.yebeh.2024.109645. Epub 2024 Jan 19. PMID: 38244419.

68: Murphy Lonergan R, Curry J, Dhas K, Simmons BI. Stratified Evaluation of GPT's Question Answering in Surgery Reveals Artificial Intelligence (AI) Knowledge Gaps. Cureus. 2023 Nov 14;15(11):e48788. doi: 10.7759/cureus.48788. PMID: 38098921; PMCID: PMC10720372.

69: Truhn D, Loeffler CM, Müller-Franzes G, Nebelung S, Hewitt KJ, Brandner S, Bressem KK, Foersch S, Kather JN. Extracting structured information from unstructured histopathology reports using generative pre-trained transformer 4 (GPT-4). J Pathol. 2024 Mar;262(3):310-319. doi: 10.1002/path.6232. Epub 2023 Dec 14. PMID: 38098169.

70: Waisberg E, Ong J, Masalkhi M, Zaman N, Sarker P, Lee AG, Tavakkoli A. Meta smart glasses-large language models and the future for assistive glasses for individuals with vision impairments. Eye (Lond). 2024 Apr;38(6):1036-1038. doi: 10.1038/s41433-023-02842-z. Epub 2023 Dec 4. PMID: 38049627; PMCID: PMC11009354.

71: Mole J, Nelson A, Chan E, Cipolotti L, Nachev P. Characterizing phonemic fluency by transfer learning with deep language models. Brain Commun. 2023 Nov 28;5(6):fcad318. doi: 10.1093/braincomms/fcad318. PMID: 38046096; PMCID: PMC10691875.

72: Karabacak M, Schupper AJ, Carr MT, Hickman ZL, Margetis K. From Text to Insight: A Natural Language Processing-Based Analysis of Topics and Trends in Neurosurgery. Neurosurgery. 2024 Apr 1;94(4):679-689. doi: 10.1227/neu.0000000000002763. Epub 2023 Nov 21. PMID: 37988054.

73: Stroop R. Answer to the letter to the editor of H. Daungsupawong et al. concerning “Large language models: Are artificial intelligence-based chatbots a reliable source of patient information for spinal surgery?” by R. Stroop et al. (Eur Spine J [2023]; doi:10.1007/s00586-023-07975-z). Eur Spine J. 2024 Jan;33(1):362. doi: 10.1007/s00586-023-08038-z. Epub 2023 Nov 18. PMID: 37980277.

74: Peng C, Yang X, Chen A, Smith KE, PourNejatian N, Costa AB, Martin C, Flores MG, Zhang Y, Magoc T, Lipori G, Mitchell DA, Ospina NS, Ahmed MM, Hogan WR, Shenkman EA, Guo Y, Bian J, Wu Y. A study of generative large language model for medical research and healthcare. NPJ Digit Med. 2023 Nov 16;6(1):210. doi: 10.1038/s41746-023-00958-w. PMID: 37973919; PMCID: PMC10654385.

75: Daungsupawong H, Wiwanitkit V. Letter to the editor concerning “Large language models: Are artificial intelligence-based chatbots a reliable source of patient information for spinal surgery?” by R. Stroop et al. (Eur Spine J [2023]; doi:10.1007/s00586-023-07975-z). Eur Spine J. 2024 Jan;33(1):361. doi: 10.1007/s00586-023-08035-2. Epub 2023 Nov 16. PMID: 37973670.

76: Van Veen D, Van Uden C, Blankemeier L, Delbrouck JB, Aali A, Bluethgen C, Pareek A, Polacin M, Reis EP, Seehofnerová A, Rohatgi N, Hosamani P, Collins W, Ahuja N, Langlotz CP, Hom J, Gatidis S, Pauly J, Chaudhari AS. Clinical Text Summarization: Adapting Large Language Models Can Outperform Human Experts. Res Sq [Preprint]. 2023 Oct 30:rs.3.rs-3483777. doi: 10.21203/rs.3.rs-3483777/v1. Update in: Nat Med. 2024 Apr;30(4):1134-1142. doi: 10.1038/s41591-024-02855-5. PMID: 37961377; PMCID: PMC10635391.

77: de Kort FAS, Coenen M, Weaver NA, Kuijf HJ, Aben HP, Bae HJ, Bordet R, Cammà G, Chen CPLH, Dewenter A, Duering M, Fang R, van der Giessen RS, Hamilton OKL, Hilal S, Huenges Wajer IMC, Kan CN, Kim J, Kim BJ, Köhler S, de Kort PLM, Koudstaal PJ, Lim JS, Lopes R, Mok VCT, Staals J, Venketasubramanian N, Verhagen CM, Verhey FRJ, Wardlaw JM, Xu X, Yu KH, Biesbroek JM, Biessels GJ. White Matter Hyperintensity Volume and Poststroke Cognition: An Individual Patient Data Pooled Analysis of 9 Ischemic Stroke Cohort Studies. Stroke. 2023 Dec;54(12):3021-3029. doi:

10.1161/STROKEAHA.123.044297. Epub 2023 Oct 30. PMID: 37901947; PMCID: PMC10664782.

78: Stroop A, Stroop T, Zawy Alsofy S, Nakamura M, Möllmann F, Greiner C, Stroop R. Large language models: Are artificial intelligence-based chatbots a reliable source of patient information for spinal surgery? Eur Spine J. 2024 Nov;33(11):4135-4143. doi: 10.1007/s00586-023-07975-z. Epub 2023 Oct 11. PMID: 37821602.

79: Lawson McLean A. Navigating the ethical and practical challenges of large language models in telehealth. J Telemed Telecare. 2023 Oct 8:1357633×231205060. doi: 10.1177/1357633×231205060. Epub ahead of print. PMID: 37807675.

80: Diaz Milian R, Moreno Franco P, Freeman WD, Halamka JD. Revolution or Peril? The Controversial Role of Large Language Models in Medical Manuscript Writing. Mayo Clin Proc. 2023 Oct;98(10):1444-1448. doi: 10.1016/j.mayocp.2023.07.009. PMID: 37793723.

81: Barrington NM, Gupta N, Musmar B, Doyle D, Panico N, Godbole N, Reardon T, D'Amico RS. A Bibliometric Analysis of the Rise of ChatGPT in Medical Research. Med Sci (Basel). 2023 Sep 17;11(3):61. doi: 10.3390/medsci11030061. PMID: 37755165; PMCID: PMC10535733.

82: Wang L, Ambite JL, Appaji A, Bijsterbosch J, Dockes J, Herrick R, Kogan A, Lander H, Marcus D, Moore SM, Poline JB, Rajasekar A, Sahoo SS, Turner MD, Wang X, Wang Y, Turner JA. NeuroBridge: a prototype platform for discovery of the long-tail neuroimaging data. Front Neuroinform. 2023 Aug 31;17:1215261. doi: 10.3389/fninf.2023.1215261. PMID: 37720825; PMCID: PMC10500076.

83: Roman A, Al-Sharif L, Al Gharyani M. The Expanding Role of ChatGPT (Chat- Generative Pre-Trained Transformer) in Neurosurgery: A Systematic Review of Literature and Conceptual Framework. Cureus. 2023 Aug 15;15(8):e43502. doi: 10.7759/cureus.43502. PMID: 37719492; PMCID: PMC10500385.

84: Majoie CB, Cavalcante F, Gralla J, Yang P, Kaesmacher J, Treurniet KM, Kappelhof M, Yan B, Suzuki K, Zhang Y, Li F, Morimoto M, Zhang L, Miao Z, Rinkel LA, Huang J, Otsuka T, Wang S, Davis S, Cognard C, Hong B, Coutinho JM, Song J, Chen W, Emmer BJ, Eker O, Zhang L, Dobrocky T, Nguyen HT, Bush S, Peng Y, LeCouffe NE, Takeuchi M, Han H, Matsumaru Y, Strbian D, Lingsma HF, Nieboer D, Yang Q, Meinel T, Mitchell P, Kimura K, Zi W, Nogueira RG, Liu J, Roos YB, Fischer U; IRIS collaborators. Value of intravenous thrombolysis in endovascular treatment for large-vessel anterior circulation stroke: individual participant data meta-analysis of six randomised trials. Lancet. 2023 Sep 16;402(10406):965-974. doi: 10.1016/S0140-6736(23)01142-X. Epub 2023 Aug 25. PMID: 37640037.

85: Ali R, Tang OY, Connolly ID, Zadnik Sullivan PL, Shin JH, Fridley JS, Asaad WF, Cielo D, Oyelese AA, Doberstein CE, Gokaslan ZL, Telfeian AE. Performance of ChatGPT and GPT-4 on Neurosurgery Written Board Examinations. Neurosurgery. 2023 Dec 1;93(6):1353-1365. doi: 10.1227/neu.0000000000002632. Epub 2023 Aug 15. PMID: 37581444.

86: Rau A, Rau S, Zoeller D, Fink A, Tran H, Wilpert C, Nattenmueller J, Neubauer J, Bamberg F, Reisert M, Russe MF. A Context-based Chatbot Surpasses Trained Radiologists and Generic ChatGPT in Following the ACR Appropriateness Guidelines. Radiology. 2023 Jul;308(1):e230970. doi: 10.1148/radiol.230970. PMID: 37489981.

87: Lawson McLean A. Artificial Intelligence in Surgical Documentation: A Critical Review of the Role of Large Language Models. Ann Biomed Eng. 2023 Dec;51(12):2641-2642. doi: 10.1007/s10439-023-03282-2. Epub 2023 Jun 20. PMID: 37338711; PMCID: PMC10632277.

88: Ali R, Tang OY, Connolly ID, Fridley JS, Shin JH, Zadnik Sullivan PL, Cielo D, Oyelese AA, Doberstein CE, Telfeian AE, Gokaslan ZL, Asaad WF. Performance of ChatGPT, GPT-4, and Google Bard on a Neurosurgery Oral Boards Preparation Question Bank. Neurosurgery. 2023 Nov 1;93(5):1090-1098. doi: 10.1227/neu.0000000000002551. Epub 2023 Jun 12. PMID: 37306460.

89: Jiang LY, Liu XC, Nejatian NP, Nasir-Moin M, Wang D, Abidin A, Eaton K, Riina HA, Laufer I, Punjabi P, Miceli M, Kim NC, Orillac C, Schnurman Z, Livia C, Weiss H, Kurland D, Neifert S, Dastagirzada Y, Kondziolka D, Cheung ATM, Yang G, Cao M, Flores M, Costa AB, Aphinyanaphongs Y, Cho K, Oermann EK. Health system-scale language models are all-purpose prediction engines. Nature. 2023 Jul;619(7969):357-362. doi: 10.1038/s41586-023-06160-y. Epub 2023 Jun 7. PMID: 37286606; PMCID: PMC10338337.

90: Waisberg E, Ong J, Masalkhi M, Zaman N, Kamran SA, Sarker P, Lee AG, Tavakkoli A. Generative Pre-Trained Transformers (GPT) and Space Health: A Potential Frontier in Astronaut Health During Exploration Missions. Prehosp Disaster Med. 2023 Aug;38(4):532-536. doi: 10.1017/S1049023X23005848. Epub 2023 Jun 2. PMID: 37264946; PMCID: PMC10445113.

91: Májovský M, Černý 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 May 31;25:e46924. doi: 10.2196/46924. PMID: 37256685; PMCID: PMC10267787.

92: Mariager T, Bjarkam C, Nielsen H, Bodilsen J. Experimental animal models for brain abscess: a systematic review. Br J Neurosurg. 2024 Dec;38(6):1294-1301. doi: 10.1080/02688697.2022.2160865. Epub 2022 Dec 29. PMID: 36579498.

93: Yang X, Chen A, PourNejatian N, Shin HC, Smith KE, Parisien C, Compas C, Martin C, Costa AB, Flores MG, Zhang Y, Magoc T, Harle CA, Lipori G, Mitchell DA, Hogan WR, Shenkman EA, Bian J, Wu Y. A large language model for electronic health records. NPJ Digit Med. 2022 Dec 26;5(1):194. doi: 10.1038/s41746-022-00742-2. PMID: 36572766; PMCID: PMC9792464.

94: Woolnough O, Donos C, Curtis A, Rollo PS, Roccaforte ZJ, Dehaene S, Fischer- Baum S, Tandon N. A Spatiotemporal Map of Reading Aloud. J Neurosci. 2022 Jul 6;42(27):5438-5450. doi: 10.1523/JNEUROSCI.2324-21.2022. Epub 2022 May 31. PMID: 35641189; PMCID: PMC9270918.

95: Masud JHB, Shun C, Kuo CC, Islam MM, Yeh CY, Yang HC, Lin MC. Deep-ADCA: Development and Validation of Deep Learning Model for Automated Diagnosis Code Assignment Using Clinical Notes in Electronic Medical Records. J Pers Med. 2022 Apr 28;12(5):707. doi: 10.3390/jpm12050707. PMID: 35629129; PMCID: PMC9146030.

96: Cheng Y, Tang B, Zhang G, An P, Sun Y, Gao M, Zhang Y, Shan Y, Zhang J, Liu Q, Lai CSW, de Villers-Sidani É, Wang Y, Zhou X. Degraded cortical temporal processing in the valproic acid-induced rat model of autism. Neuropharmacology. 2022 May 15;209:109000. doi: 10.1016/j.neuropharm.2022.109000. Epub 2022 Feb 17. PMID: 35182575.

97: Sveikata L, Vasung L, El Rahal A, Bartoli A, Bretzner M, Schaller K, Schnider A, Leemann B. Syndrome of the trephined: clinical spectrum, risk factors, and impact of cranioplasty on neurologic recovery in a prospective cohort. Neurosurg Rev. 2022 Apr;45(2):1431-1443. doi: 10.1007/s10143-021-01655-6. Epub 2021 Oct 7. PMID: 34618250; PMCID: PMC8976790.

98: Zhang W, Tang Y, Liu H, Yuan LP, Wang CC, Chen SF, Huang J, Xiao XY. Risk prediction models for intensive care unit-acquired weakness in intensive care unit patients: A systematic review. PLoS One. 2021 Sep 24;16(9):e0257768. doi: 10.1371/journal.pone.0257768. PMID: 34559850; PMCID: PMC8462700.

99: Zayed M, Allers K, Hoffmann F, Bantel C. Transversus abdominis plane block in urological procedures: A systematic review and meta-analysis. Eur J Anaesthesiol. 2021 Jul 1;38(7):758-767. doi: 10.1097/EJA.0000000000001453. PMID: 34101639.

100: Cao H, Erson-Omay EZ, Günel M, Moliterno J, Fulbright RK. A Quantitative Assessment of Pre-Operative MRI Reports in Glioma Patients: Report Metrics and IDH Prediction Ability. Front Oncol. 2021 Jan 29;10:600327. doi: 10.3389/fonc.2020.600327. PMID: 33585216; PMCID: PMC7879978.

101: Muhlestein WE, Monsour MA, Friedman GN, Zinzuwadia A, Zachariah MA, Coumans JV, Carter BS, Chambless LB. Predicting Discharge Disposition Following Meningioma Resection Using a Multi-Institutional Natural Language Processing Model. Neurosurgery. 2021 Mar 15;88(4):838-845. doi: 10.1093/neuros/nyaa585. PMID: 33483747.

102: Maragkos GA, Filippidis AS, Chilamkurthy S, Salem MM, Tanamala S, Gomez-Paz S, Rao P, Moore JM, Papavassiliou E, Hackney D, Thomas AJ. Automated Lateral Ventricular and Cranial Vault Volume Measurements in 13,851 Patients Using Deep Learning Algorithms. World Neurosurg. 2021 Apr;148:e363-e373. doi: 10.1016/j.wneu.2020.12.148. Epub 2021 Jan 6. PMID: 33421645.

103: Lakretz Y, Ossmy O, Friedmann N, Mukamel R, Fried I. Single-cell activity in human STG during perception of phonemes is organized according to manner of articulation. Neuroimage. 2021 Feb 1;226:117499. doi: 10.1016/j.neuroimage.2020.117499. Epub 2020 Oct 24. PMID: 33186717.

104: Agnoletto GJ, Meyers PM, Coon A, Kan PTM, Wakhloo AK, Hanel RA. A Contemporary Review of Endovascular Treatment of Wide-Neck Large and Giant Aneurysms. World Neurosurg. 2019 Oct;130:523-529.e2. doi: 10.1016/j.wneu.2019.06.201. Epub 2019 Jul 4. PMID: 31279111.

105: Ramkumar V, Nagarajan R, Shankarnarayan VC, Kumaravelu S, Hall JW. Implementation and evaluation of a rural community-based pediatric hearing screening program integrating in-person and tele-diagnostic auditory brainstem response (ABR). BMC Health Serv Res. 2019 Jan 3;19(1):1. doi: 10.1186/s12913-018-3827-x. PMID: 30606168; PMCID: PMC6318860.

106: Magnotti JF, Beauchamp MS. Published estimates of group differences in multisensory integration are inflated. PLoS One. 2018 Sep 19;13(9):e0202908. doi: 10.1371/journal.pone.0202908. PMID: 30231054; PMCID: PMC6145544.

107: Uhl C, Markel M, Broggini T, Nieminen M, Kremenetskaia I, Vajkoczy P, Czabanka M. EphB4 mediates resistance to antiangiogenic therapy in experimental glioma. Angiogenesis. 2018 Nov;21(4):873-881. doi: 10.1007/s10456-018-9633-6. Epub 2018 Jul 10. PMID: 29987450; PMCID: PMC6208883.

108: Zech J, Pain M, Titano J, Badgeley M, Schefflein J, Su A, Costa A, Bederson J, Lehar J, Oermann EK. Natural Language-based Machine Learning Models for the Annotation of Clinical Radiology Reports. Radiology. 2018 May;287(2):570-580. doi: 10.1148/radiol.2018171093. Epub 2018 Jan 30. PMID: 29381109.