The Challenges and Benefits of Generative AI in Kidney Care

Emel Hamilton, MD, MSN/INF, CNN
Zuwen Kuang, MS
Luca Neri, PhD
Hanjie Zhang, PhD

While generative Artificial Intelligence (AI) has the potential to improve kidney care, it also poses substantial challenges. Applications under discussion touch on every aspect of treatment, including the creation of new prognostic tools and methodologies, personalized medical education for professionals and patients, and protocols that alleviate the burden of administrative tasks. FME is developing an AI framework for clinical workflows that considers both the benefits and risks that AI poses for the future of patient care.

The development of generative Artificial Intelligence (AI) has created excitement and prompted vigorous debate across various industries, including healthcare.[1] Dubbed "Gen AI," this novel technology transcends conventional rule-based systems, data analysis, and predictions. Departing from the familiar confines of traditional AI, generative AI ventures into uncharted territory where machines wield the power of creativity sans human intervention.[2]

Understanding Generative AI's Unconventional Pathways

Generative AI stands apart because of its ability to create human-like content—images, text, melodies, or even entire narratives—using complex computer algorithms. These differ from conventional machine learning algorithms that can generate simple outputs. Instead, generative models create new content based on an assortment of data on which they have been trained by weaving together semantic patterns and knowledge structures.[2,3]

Potential Benefits of Generative AI in Kidney Care

Generative AI systems, particularly large language models (LLMs), hold numerous potential applications and may revolutionize several aspects of healthcare.[4,5,6]

- **Clinical Insights and Powerful Prognostic Tools:** A recent systematic review highlighted that most published studies focus on the use of LLMs as medical chatbots and to generate patient information and clinical documentation as well as for patient education and to simplify imaging reports.[7] Generative AI and multimodal LLMs may have direct clinical applications, such as generating diagnostic [8,9,10,11] and prognostic [12,13] predictions, given their ability to encode medical knowledge and/or interpret medical signs and symptoms similar to semantic elements.[12,13] For instance, Kanda and colleagues utilized an early natural language processing (NLP) architecture, word2vec, to analyze chronic kidney disease (CKD)

literature, accurately predicting death and end-stage kidney disease (ESKD) onset. With the advent of more advanced LLMs, coupled with fine-tuning in the medical domain, highly accurate outcome predictions can be generated directly from medical notes, referral letters, and patients' narratives without the need to document medical encounters in structured electronic health record systems, thus reducing documentation burden and limitations due to incomplete ontologies.[12,15]

- **Personalized Care:** New LLM architectures like pre-trained transformers offer broader possibilities for analyzing multimodal data and detecting nuanced associations. These advancements enable language-understanding technologies to learn patterns across various data types, such as comorbidity codes, lab tests, images, clinical narratives, and patient-reported outcomes. For example, Savcisens et al. demonstrated the effectiveness of this approach in predictive modeling using life-events data, showing that these models could accurately predict diverse outcomes such as early mortality and personality nuances by learning patterns from detailed event sequences.[16]

Departing from the familiar confines of traditional AI, generative AI ventures into uncharted territory where machines wield the power of creativity sans human intervention.[2]

• **Efficiency and Cost Savings:** Generative AI can alleviate the administrative burden on healthcare staff, including time-consuming non-medical tasks.[17,18,19,20,21,22,23] Streamlining these tasks can save time, minimize disruptions, and potentially enhance patient-clinician interactions. Studies show that LLMs can summarize medical notes and dialogues with high accuracy.[24,25] For instance, FME Global Medical Office and Santa Barbara Smart Health developed software leveraging GPT-4 to transcribe patient-physician interactions, achieving reliable abstraction of 33 medical elements, including pre-existing medical conditions, drug prescriptions, biochemical parameters, active problems, and treatment plans. The system produced a reliable and accurate summary of medical concepts in a small proof-of-concept study. FME is exploring how generative AI might streamline the process of collecting patient referral information, with the potential to expedite referrals and admissions and enhance data entry accuracy. We are also investigating the development of a ChatGPT-like tool to assist staff in offering targeted guidance for handling non-clinical tasks, with the goal of reducing staff burden and supporting new clinical leaders. This includes

FME is exploring how generative AI might streamline the process of collecting patient referral information, with the potential to expedite referrals and admissions and enhance data entry accuracy.

examining how the tool could navigate intricate requirements related to Worker's Compensation and the Conditions for Coverage for ESKD Facilities. Additionally, FME aims to reduce patient attrition and improve their experience.[26,27,28] By considering the implementation of an AI-guided referral pathway and AI-powered case management, we hope to assist FME's Continuity of Care team in identifying patients at high risk of attrition, conducting root cause analyses, and providing data-driven insights to case managers (Figure 1).
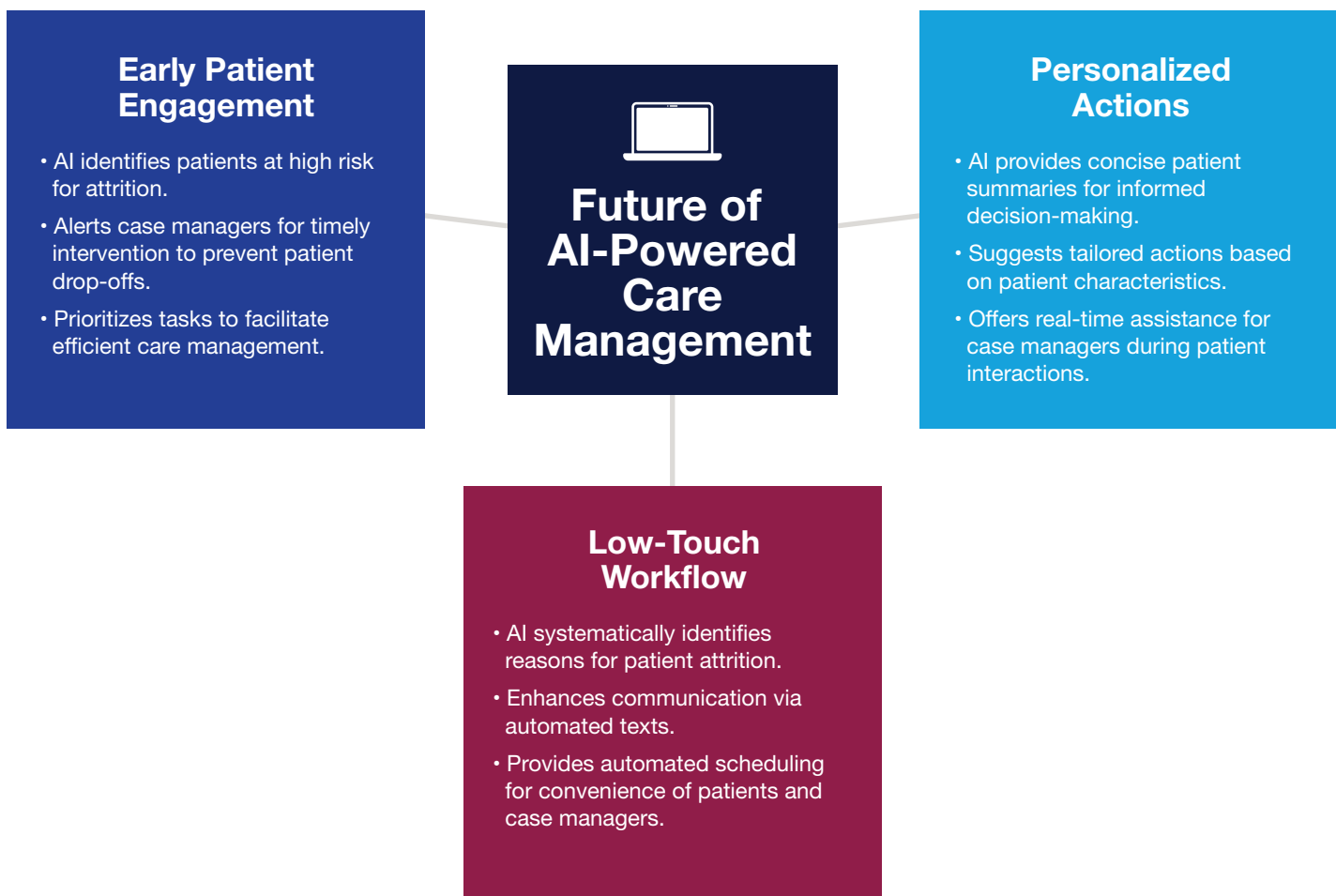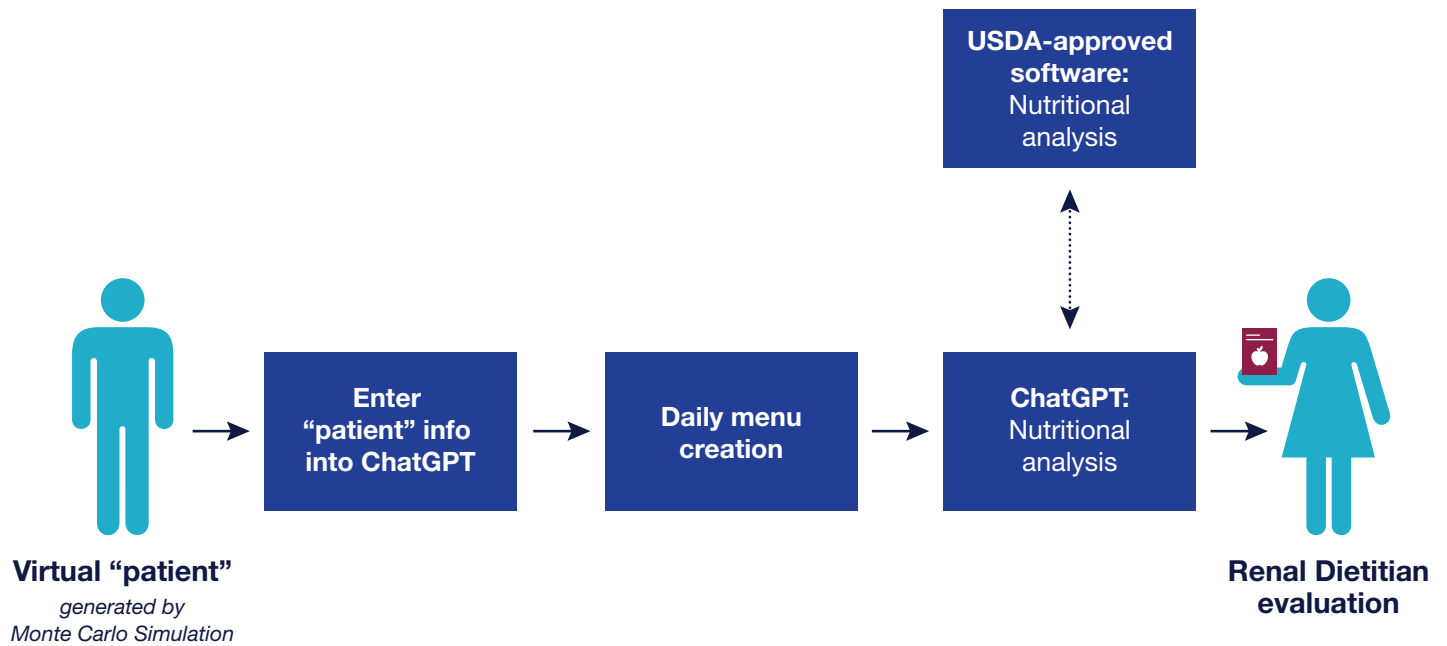
FIGURE 1 | AI-POWERED CARE MANAGEMENT

Early Patient Engagement
• AI identifies patients at high risk for attrition.
• Alerts case managers for timely intervention to prevent patient drop-offs.
• Prioritizes tasks to facilitate efficient care management.

Future of AI-Powered Care Management

Personalized Actions
• AI provides concise patient summaries for informed decision-making.
• Suggests tailored actions based on patient characteristics.
• Offers real-time assistance for case managers during patient interactions.

Low-Touch Workflow
• AI systematically identifies reasons for patient attrition.
• Enhances communication via automated texts.
• Provides automated scheduling for convenience of patients and case managers.

Virtual "patient"
*generated by
Monte Carlo Simulation*

Enter "patient" info into ChatGPT → Daily menu creation → ChatGPT: Nutritional analysis

USDA-approved software: Nutritional analysis
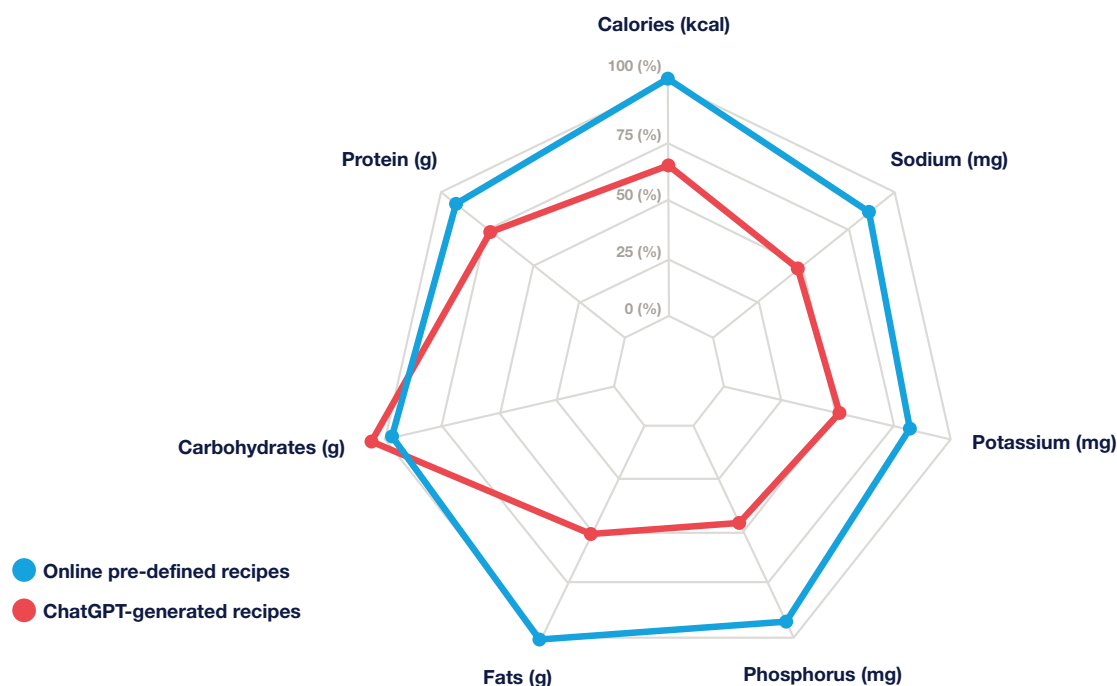
Renal Dietitian evaluation

- **Tailored Medical Education:** Personalizing medical education for healthcare professionals and patients is another promising area of application for generative AI.[26,27,28] We utilized retrieval-augmented generation (RAG), a novel AI-driven approach, to efficiently process and extract meaningful information from published literature on uremic toxins. The process involved preparing a curated literature database, creating a vector database from curated literature, retrieving relevant information based on queries, and generating responses using LLMs incorporating retrieved information. Although RAG has significantly improved content generation, the potential for "hallucinations" persists, and the enhanced LLM outputs still require human verification. For more information on the hallucination topic, refer to "Potential Risks" below.

- **Comprehensive Use of Data and Knowledge:** Dietary management is crucial for patients with kidney failure undergoing dialysis, but personalized advice is challenging due to varying food preferences and other factors. By leveraging LLMs, there is potential to integrate patient demographics, clinical data, and food preferences to create tailored recipe recommendations.[29,30] Renal Research Institute (RRI) tested the emergent ability of LLM to generate sound nutritional advice for people with CKD (Figure 2). While this approach has limitations in precise nutritional analysis for people with CKD, it's important to note that this evaluation of LLM sheds light on the current knowledge base. For instance, in RRI's study, ChatGPT underestimated calories, protein, fat, phosphorus, potassium, and sodium content on ChatGPT-generated recipes when compared with U.S. Department of Agriculture (USDA)-approved software. These discrepancies are much smaller with online pre-defined recipes (Figure 3). While the underlying knowledge basis of GPT-4 falls short in supporting nutritional analysis for people with kidney disease, incorporating LLMs in more complex architectures may enhance the accuracy of nutritional estimation.[31,32,33]

Personalizing medical education for healthcare professionals and patients is another promising area of application for generative AI.[26,27,28]

Potential Risks of Generative AI in Kidney Care

Generative AI offers unprecedented potential to revolutionize patient care, diagnosis, and treatment methodologies. However, substantial risks remain.

- **Biased Outputs from Training Data:** Generative models learn from the data on which they are trained. If their training samples and datasets include biases, then those models can generate outputs that are ethically questionable.[6] In the realm of kidney care, such biases could propagate treatment disparities or inequalities.

- **Privacy and Security Concerns:** Generative AI's ability to generate synthetic data, which resembles real data, is tremendously useful in research and model training, but this capability comes with privacy implications. If the original datasets used to train the generative AI are not adequately secured, there is a risk that the synthetic data could inadvertently reveal sensitive personal information. Furthermore, machine learning systems in sensitive domains such as healthcare are particularly vulnerable to adversarial AI attacks where malicious actors can manipulate or exploit the models by introducing carefully crafted inputs to the system.[34,35]

- **Hallucinations in AI Responses:** In the context of generative AI, "hallucinations" refer to the generation of responses that are not logically or semantically coherent or are not relevant to the input prompt.

These hallucinations can occur when generative AI formulates responses based on patterns or associations it has learned from its training data without fully understanding the meaning or context of the input prompt. This could pose serious risks to patient safety and well-being if implemented without proper verification or oversight.[36]

- **Transparency and Explainability Challenges:** Unlike traditional rule-based AI systems where decision-making logic is explicit and interpretable, generative AI models often operate as "black boxes," making it difficult for clinicians and patients to comprehend how generative AI arrived at a particular decision.[36] Addressing this risk requires meaningful human-AI collaboration, which involves integrating AI systems seamlessly into clinical workflows to enhance efficiency, accuracy, and patient outcomes while preserving the critical role of human expertise, empathy, and judgment in delivering high-quality care.[37]

Generative AI offers unprecedented potential to revolutionize patient care, diagnosis, and treatment methodologies. However, substantial risks remain.

Reflecting on Possibility

In our relentless pursuit of innovation, FME recognizes the immense potential of generative AI in revolutionizing clinical workflows. However, this potential must be harnessed responsibly. At FME, we are developing a trustworthy AI framework—one that prioritizes safety, security, and ethics. Our commitment extends beyond compliance to encompass the thoughtful integration of organizational values and change management principles. In this new era of healthcare, we remain steadfast in our mission to elevate patient care while upholding the highest standards of integrity and excellence.

At FME, we are developing a trustworthy AI framework—one that prioritizes safety, security, and ethics.

FME's use of generative AI tools such as ChatGPT is focused on research or quality assessment purposes and not used for patient care. Renal Research Institute is a wholly owned subsidiary of Fresenius Medical Care.

Dr. Emel Hamilton
Vice President and Global Leader of Clinical Systems
Global Medical Office

Dr. Hamilton is a multifaceted healthcare professional with a unique blend of medical expertise, nursing proficiency, and a deep understanding of clinical informatics. She earned her medical degree (MD) from Dokuz Eylül Üniversitesi in Turkey, laying a solid foundation for her comprehensive perspective on healthcare challenges. Further enhancing her skill set, she also holds a master's degree in nursing and clinical informatics (MSN/INF). Dr. Hamilton is a recognized expert in Health Information Exchange (HIE) and Electronic Health Records (EHR) nurse practice standards and has contributed several significant publications in the field. Her expertise extends beyond traditional healthcare boundaries into the realm of AI and her balanced perspective allows her to adeptly navigate the AI landscape, always with a keen eye on harnessing its potential while ensuring paramount importance to patient safety and data security. Currently, Dr. Hamilton holds the position of Vice President and Global Leader of Clinical Systems at Fresenius Medical Care. Her approach to healthcare challenges is inherently collaborative, recognizing the crucial roles played by both humans and machines.

Zuwen Kuang
Senior Vice President, Global Head of Data and Analytics
Fresenius Medical Care

Zuwen Kuang is a skilled and experienced digital healthcare executive with a proven track record of success in the field of data and analytics. As the Global Head of Data and Analytics in Digital Technology and Innovation at Fresenius Medical Care, she leads a team responsible for building enterprise big data platforms, delivering advanced analytics, AI/GenAI solutions and digital interoperability capabilities. This enables the transformation of healthcare data into intelligence that predicts trends and reveals actionable insights that influence long-term business growth.

Dr. Luca Neri
Senior Director, Clinical Advanced Analytics, EMEA, AP, LATAM
Global Medical Office

Dr. Neri leads the Data Science division for the EMEA, APAC, and LATAM regions of the Global Medical Office - Clinical Advanced Analytics department at Fresenius Medical Care. He joined Fresenius Medical Care in 2016. With almost 20 years of experience in epidemiology, outcomes research, and data science, Dr. Neri has acquired a broad range of analytical and methodological skills in the field of advanced analytics. The GMO-CAA Data Science team blends profound expertise in state-of-the-art machine learning techniques with dedication to the clinical integration of AI solutions in medical care.

Before joining Fresenius Medical Care, he held a postdoctoral position at the St. Louis University Center for Outcomes Research and was later appointed adjunct instructor of Health Management and Policy at the same institution. He also served as a scientific consultant and advisory board member for several commercial, scientific, and industrial clients and academic institutions. He earned his medical degree at the University of Milan School of Medicine, where he also earned a specialty degree and a PhD in environmental and occupational medicine.

Dr. Hanjie Zhang
Senior Director of Computational Statistics & Artificial Intelligence
Renal Research Institute

Hanjie Zhang, PhD joined Renal Research Institute in 2014. She received a master's degree in statistics from Columbia University, New York, and a PhD in medical science from the University of Maastricht, The Netherlands. Hanjie has been involved in the design of several large cluster-randomized clinical trials and complex statistical analysis. She is also involved in designing, developing, and deploying enterprise solutions across the artificial intelligence spectrum, such as machine learning, and deep learning. During her tenure with Renal Research Institute, Hanjie has authored over 30 research articles in leading kidney journals.

References

1. McKinsey, "The State of AI in 2023: Generative AI's Breakout Year," https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-AIs-breakout-year. Accessed April 13, 2024.

2. A. Zewe, "Explained: Generative AI," *MIT News*, Massachusetts Institute of Technology, November 9, 2023, https://news.mit.edu/2023/explained-generative-ai-1109. Accessed April 13, 2024.

3. E. Heaslip, "Traditional AI vs. Generative AI: A Breakdown," CO, US Chamber of Commerce, October 16, 2023, https://www.uschamber.com/co/run/technology/traditional-ai-vs-generative-ai. Accessed April 13, 2024.

4. B. Marr, "Revolutionizing Healthcare: The Top 14 Uses Of ChatGPT in Medicine And Wellness," *Forbes*, March 2, 2023, https://www.forbes.com/sites/bernardmarr/2023/03/02/revolutionizing-healthcare-the-top-14-uses-of-chatgpt-in-medicine-and-wellness/?sh=fc47ebb6e547. Accessed April 13, 2024.

5. E. Berger and M. Dries, "Getting the Most Out of Generative AI in Healthcare Today," Bain & Company, August 7, 2023, https://www.bain.com/insights/getting-the-most-out-of-generative-ai-in-healthcare/. Accessed April 13, 2024.

6. J. Neiser, K. Srikumar, M. Lee, et al., "Biopharma's Path to Value with Generative AI," Boston Consulting Group, October 9, 2023, https://www.bcg.com/publications/2023/biopharma-path-to-value-with-generative-ai. Accessed April 13, 2024.

7. F. Busch, L. Hoffmann, C. Rueger, et al., "Systematic Review of Large Language Models for Patient Care: Current Applications and Challenges," Preprint, submitted on March 5, 2024. doi.org/10.1101/2024.03.04.24303733.

8. T. Savage, A. Nayak, R. Gallo, E. Rangan, and J.H. Chen, "Diagnostic Reasoning Prompts Reveal the Potential for Large Language Model Interpretability in Medicine. *Nature Partner Journals Digital Medicine* 7, no. 1 (2024). doi.org/10.1038/s41746-024-01010-1.

9. S.Y. Lin, C.C. Jiang, K.M. Law, et al., "Comparative Analysis of Generative AI in Clinical Nephrology: Assessing ChatGPT-4, Gemini Pro, and Bard in Patient Interaction and Renal Biopsy Interpretation," *Preprints with The Lancet*, posted on February 1, 2024. doi.org/10.2139/SSRN.4711596

10. Z. Kanjee, B. Crowe, and A. Rodman, "Accuracy of a Generative Artificial Intelligence Model in a Complex Diagnostic Challenge," *Journal of the American Medical Association* 330, no. 1 (2023): 78–80. doi.org/10.1001/jama.2023.8288.

11. R.J. Chen, T. Ding, M.Y. Lu, et al., "Towards a General-Purpose Foundation Model for Computational Pathology," *Nature Medicine* 30, no. 3 (2024): 850–62. doi.org/10.1038/s41591-024-02857-3.

12. L.Y. Jiang, X.C. Liu, N.P. Nejatian, et al., "Health System-Scale Language Models Are All-Purpose Prediction Engines," *Nature* 619 (2023): 357–62. doi.org/10.1038/s41586-023-06160-y.

13. F. Liu, T. Zhu, X. Wu, et al., "A Medical Multimodal Large Language Model for Future Pandemics," *Nature Partner Journals Digital Medicine* 6, no. 1 (2023): 1–15. doi.org/10.1038/s41746-023-00952-2.

14. E. Kanda, B.I. Epureanu, T. Adachi, T. Sasaki, and N. Kashihara, "New Marker for Chronic Kidney Disease Progression and Mortality in Medical-Word Virtual Space," *Scientific Reports* 14, no. 1: 1–11. doi.org/10.1038/s41598-024-52235-9.

15. X. Yang, A. Chen, N. PourNejatian, et al., "A Large Language Model for Electronic Health Records," *Nature Partner Journals Digital Medicine* 5, no. 1 (2022): 1–9. doi.org/10.1038/s41746-022-00742-2.

16. G. Savcisens, T. Eliassi-Rad, L.K. Hansen, et al., "Using Sequences of Life-Events To Predict Human Lives," *Nature Computational Science* 4, no. 1 (2023): 43–56. doi.org/10.1038/s43588-023-00573-5.

17. N. Khamisa, K. Peltzer, and B. Oldenburg, "Burnout in Relation to Specific Contributing Factors and Health Outcomes Among Nurses: A Systematic Review," *International Journal of Environmental Research and Public Health* 10, no. 6 (2013): 2214–40. doi.org/10.3390/ijerph10062214.

18. R.M. Ratwani, E. Savage, A. Will, et al., "A Usability and Safety Analysis of Electronic Health Records: A Multi-Center Study," *Journal of the American Medical Informatics Association* 25, no. 9 (2018): 1197–1201. doi.org/10.1093/jamia/ocy088.

19. J.M. Ehrenfeld and J.P. Wanderer, "Technology as Friend or Foe? Do Electronic Health Records Increase Burnout?" *Current Opinion in Anesthesiology* 31, no. 3 (2018): 357–60. doi.org/10.1097/aco.0000000000000588.

20. C. Sinsky, L. Colligan, L. Li, et al., "Allocation of Physician Time in Ambulatory Practice: A Time And Motion Study In 4 Specialties," *Annals of Internal Medicine* 165, no. 11 (2016): 753–60. doi.org/10.7326/m16-0961.

21. T.R. Yackel and P.J. Embi, "Unintended Errors with EHR-Based Result Management: A Case Series," *Journal of the American Medical Informatics Association* 17, no. 1 (2010): 104–107. doi.org/10.1197/jamia.m3294.

22. B.G. Arndt, J.W. Beasley, M.D. Watkinson, et al., "Tethered to the EHR: Primary Care Physician Workload Assessment Using EHR Event Log Data and Time–Motion Observations," *Annals of Family Medicine* 15, no. 5 (2017): 419–26. doi.org/10.1370/afm.2121.

23. J.F. Golob, J.J. Como, and J.A. Claridge, "The Painful Truth: The Documentation Burden of a Trauma Surgeon," *Journal of Trauma and Acute Care Surgery* 80, no. 5 (2016): 742–7. doi.org/10.1097/ta.0000000000000986.

24. B. Chintagunta, N. Katariya, X. Amatriain, and A. Kannan, "Medically Aware GPT-3 as a Data Generator for Medical Dialogue Summarization," *Proceedings of Machine Learning Research* 149 (2021): 66–76. doi.org/10.18653/V1/2021.NLPMC-1.9.

25. D. Van Veen, C. Van Uden, L. Blankemeier, et al., "Adapted Large Language Models Can Outperform Medical Experts in Clinical Text Summarization," *Nature Medicine* 30 (2024): 1134–42. doi.org/10.1038/s41591-024-02855-5.

26. R.G. Hughes, K.L. Bobay, N.A. Jolly, and C. Suby, "Comparison of Nurse Staffing Based on Changes in Unit-Level Workload Associated with Patient Churn," *Journal of Nursing Management* 23, no. 3 (2015): 390–400. doi.org/10.1111/JONM.12147.

27. E.T. Roberts and C.E. Pollack, "Does Churning in Medicaid Affect Health Care Use?" *Medical Care* 54, no. 5 (2016): 483–9. doi.org/10.1097/MLR.0000000000000509.

28. A. Lemmens and S. Gupta, "Managing Churn to Maximize Profits," *Marketing Science* 39, no. 5 (2020): 956–73. doi.org/10.1287/MKSC.2020.1229.

29. M. Haman, M. Školník, and M. Lošťák, "AI Dietician: Unveiling the Accuracy of ChatGPT's Nutritional Estimations," *Nutrition* 119 (2024). doi.org/10.1016/j.nut.2023.112325.

30. H. Sun, K. Zhang, W. Lan, et al., "An AI Dietitian for Type 2 Diabetes Mellitus Management Based on Large Language and Image Recognition Models: Preclinical Concept Validation Study," *Journal of Medical Internet Research* 25 (2023): e51300. doi.org/10.2196/51300.

31. G. Cenikj, L. Strojnik, R. Angelski, N. Ogrinc, B. Koroušić Seljak, and T. Eftimov, "From Language Models to Large-Scale Food and Biomedical Knowledge Graphs," *Scientific Reports* 13, no. 1 (2023): 1–14. doi.org/10.1038/s41598-023-34981-4.

32. A. Rostami, R. Jain, and A.M. Rahmani, "Food Recommendation as Language Processing (F-RLP): A Personalized and Contextual Paradigm," Preprint, submitted on February 12, 2024. https://arxiv.org/abs/2402.07477.

33. Z. Yang, E. Khatibi, N. Nagesh, et al., "ChatDiet: Empowering Personalized Nutrition-Oriented Food Recommender Chatbots Rhrough an LLM-Augmented Framework," Preprint, submitted on February 18, 2024. http://arxiv.org/abs/2403.00781.

34. M. Shipman, "AI Networks Are More Vulnerable to Malicious Attacks than Previously Thought," ScienceDaily, December 4, 2023, https://www.sciencedaily.com/releases/2023/12/231204135128.htm. Accessed April 13, 2024.

35. M. Sallam, "ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns," *Healthcare* 11, no. 6 (2023). doi.org/10.3390/HEALTHCARE11060887.

36. H. Alkaissi and S.I. McFarlane, "Artificial Hallucinations in ChatGPT: Implications in Scientific Writing," *Cureus* 15, no. 2 (2023): e35179. doi.org/10.7759/cureus.35179.

37. W. Brinster, "The Human-AI Partnership: Unlocking The Power of Healthcare Innovation," *Forbes*, September 19, 2023, https://www.forbes.com/sites/forbestechcouncil/2023/09/19/embracing-a-human-ai-partnership-to-harness-the-power-of-healthcare-innovation/?sh=51aa18b96073. Accessed April 13, 2024.