Integrating Large Language Models into Personalized Diabetes Care: A Systematic Review of Clinical Applications, Model Adaptation Strategies, and Ethical Implications

J.A.U.S. Jayakody[*], S.M.M.S. Jayawardhana, Dr. P.R.H.N.G. Thilakarathne

Department of Information and Communication Technology, Faculty of Technology, Rajarata University of Sri Lanka.

usjayako@tec.rjt.ac.lk*

ABSTRACT

This systematic review examines the integration of Large Language Models (LLMs) into personalized diabetes care, focusing on clinical applications, adaptation strategies, and ethical considerations. As diabetes management demands increasingly personalized approaches, LLMs including GPT-3 and GPT-4 show promise for patient education, clinical decision support, and diagnostic assistance. This review synthesizes findings from studies published between 2018 and 2025 to evaluate LLM clinical applications and assess adaptation techniques. Key applications include conversational agents for patient education, personalized decision-making systems, and predictive modeling for diabetes-related complications. Model adaptation through domain-specific training and multimodal integration demonstrates enhanced performance in clinical settings. However, significant challenges persist, including data privacy concerns, model fairness issues, and limited real-world validation. Ethical considerations encompass training bias and data security, highlighting the need for privacy-preserving approaches. The review identifies critical gaps in current research and proposes future directions emphasizing explainable AI models to build trust among healthcare professionals and patients. While LLMs offer transformative potential for personalized diabetes care, their responsible integration requires addressing technical, ethical, and regulatory challenges. This synthesis provides a foundation for advancing LLM applications in diabetes management while ensuring patient safety and equitable care delivery.

Keywords: *Large Language Models (LLMs), Personalized Diabetes Care, Clinical Decision Support Systems, Patient Education, Model Adaptation, Ethical Considerations, AI in Healthcare.*

INTRODUCTION

Diabetes affects over 400 million people worldwide, with projections reaching 628.6 million by 2045 (Standl et al., 2019). This chronic condition requires personalized management approaches that consider clinical, personal, and psychosocioeconomic characteristics to reduce complications and improve outcomes (Khalifa & Albadawy, 2024; Vuohijoki et al., 2020). Large Language Models (LLMs), including GPT-3 and GPT-4, represent transformer-based deep learning systems increasingly adapted for healthcare applications through fine-tuning on domain-specific data. These models show promise in clinical decision-making, self-management support, and patient engagement. However, despite growing interest in healthcare LLM applications, limited consolidated evidence exists specifically addressing diabetes care integration. While emerging studies explore LLM potential in diabetes management, a comprehensive synthesis of clinical applications, adaptation strategies, and ethical considerations remains lacking. This gap is particularly significant given diabetes care's complexity and the need for personalized, culturally sensitive approaches that LLMs could potentially address. This systematic review addresses these needs by examining current evidence on LLM integration in diabetes care. Our objectives are to: (1) **evaluate clinical applications** of LLMs in patient education, diagnostic support, and treatment management; (2) **examine adaptation strategies** employed in LLMs for diabetes-related tasks; (3) **discuss ethical considerations** related to LLM integration in diabetes care, focusing on data privacy, fairness, and model transparency; and (4) **identify research gaps** and provide recommendations for responsible LLM implementation in personalized diabetes care.

METHODOLOGY (PER PRISMA 2020)

A comprehensive literature search was conducted across four academic databases: PubMed, Scopus, IEEE Xplore, and Web of Science to identify relevant studies for this systematic review. The search focused on the integration of LLMs into personalized diabetes care and covered publications from 2018 to 2025. Search terms encompassed four key categories: diabetes-related terms (including "diabetes care," "Type 1/2 diabetes," "diabetes management," and "diabetes prevention"), personalization terms ("personalized medicine," "patient-centered care," "precision medicine," and "chronic disease management"), AI/LLM-related terms ("Large Language Models," "GPT-3/4," "BERT," "Med-PaLM," and "clinical decision support systems"), and ethical/regulatory terms ("trust in AI," "model transparency," "FDA/EMA positions," and "AI tools regulation").
Studies were included if they met the following criteria: (1) peer-reviewed publications in English; (2) published between 2018 and 2025; and (3) involved the application of LLMs in diabetes care. Exclusion criteria comprised non-human studies, non-peer-reviewed sources (e.g., editorials, opinion pieces), and studies unrelated to either LLMs or diabetes care. The study selection process is summarized in the PRISMA flow diagram (See Figure 1). The findings were thematically organized into five categories: (1) clinical applications of LLMs, (2) model adaptation strategies, (3) ethical considerations, (4) limitations and (5) future directions.

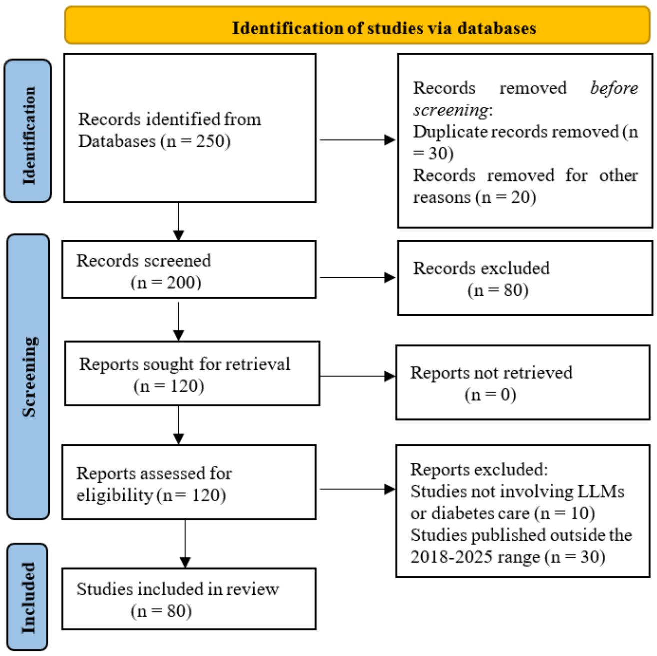

Figure 9:PRISMA flow diagram

CLINICAL APPLICATIONS OF LLMS IN DIABETES CARE

LLM-powered conversational agents show strong promise for personalized patient education. Dao et al. (2024) developed a multimodal system using 466 Q&A pairs for educational interventions, while Kelly et al. (2024) achieved 75% alignment with clinical guidance through a ChatGPT-based educational chatbot. These applications effectively support patient self-management through accessible, personalized interactions. Clinical decision support applications yield mixed results that depend heavily on implementation. Liu et al. (2023) found AI-generated recommendations were rated as relevant and safe, though performance varied significantly with prompt design. Williams et al. (2024) demonstrated through analysis of over 250,000 Emergency Department notes that GPT-4 matched or exceeded physician performance in specific areas like antibiotic prescribing, but showed clear limitations in complex clinical judgment tasks. Predictive modeling emerges as the most promising application area. Choi et al. (2025) used ChatGPT-4 in a code-free pipeline to develop risk calculators for diabetic retinopathy and diabetic macular edema, achieving AUCs of 0.786 and 0.835 respectively. These results are comparable to traditional machine learning methods, such as random forest models achieving AUC of 0.90 for end-stage renal disease prediction (Zou et al., 2022), while offering enhanced interpretability and natural language interfaces. LLM integration with electronic health records shows substantial potential for improving clinical workflow efficiency. Verma et al. (2025) demonstrated GPT-4o's capability to generate verifiable summaries of longitudinal patient records while maintaining clinical accuracy. Van Veen et al. (2024) showed GPT-4's superior performance over medical experts in clinical summarization with fewer hallucinations. These capabilities are being operationalized through real-world deployments, including Epic Systems' partnership with Microsoft for patient response drafting and visit summarization.

MODEL ADAPTATION AND TECHNICAL STRATEGIES

Successful deployment of Large Language Models in diabetes care requires sophisticated adaptation strategies encompassing domain-specific training, prompt optimization, multimodal integration, and clinical evaluation. Fine-tuning approaches, exemplified by Med-PaLM's use of Low-Rank Adaptation (LoRA) techniques on medical datasets, enhance accuracy for clinical tasks (Bui et al., 2025; Yang et al., 2023), while prompt engineering optimizes performance with limited data, as demonstrated by Pachetti and Colantonio (2024) through structured clinical inputs like blood glucose levels for personalized treatment decisions. Multimodal integration represents a transformative advancement, with AlSaad et al. (2024) showing how combining Continuous Glucose Monitoring data with LLMs enables real-time treatment plan adjustments. Clinical evaluation frameworks ensure reliability and safety, with Singhal et al. (2023) introducing the MultiMedQA benchmark to assess accuracy and reasoning while revealing gaps like hallucinations and biases, complemented by Ho et al. (2024)'s emphasis on patient-centeredness metrics and Park et al. (2021)'s observation that high AUROC values require real-world validation to confirm clinical utility.

ETHICAL, LEGAL, AND SOCIAL IMPLICATIONS

The integration of Large Language Models (LLMs) into personalized diabetes care presents significant Ethical, Legal, and Social Implications (ELSI). The literature reveals that biases within training datasets can lead to disparities in model performance across different demographic groups. For example, dermatological AI systems trained primarily on fair-skinned individuals have demonstrated reduced diagnostic accuracy for conditions such as melanoma in darker-skinned patients (Marko et al., 2025). Such biases are not limited to dermatology but extend across healthcare AI applications, where models trained on unrepresentative datasets fail to accurately diagnose conditions in marginalized populations. Chinta et al. (2025) underscore that these biases could exacerbate existing healthcare disparities, leading to less accurate recommendations or even missed diagnoses in communities that already face significant healthcare challenges. Addressing these issues requires the diversification of training datasets to ensure LLMs provide equitable care. Additionally, the ethical and legal frameworks governing AI in healthcare must evolve to promote inclusivity and mitigate systemic biases in AI tools. However, the successful deployment of LLMs must also address another pressing concern: data privacy and security. Unauthorized access or data breaches, particularly through advanced techniques such as model inversion or membership inference, could compromise patient confidentiality (Dodevski et al., 2024).

CHALLENGES AND LIMITATIONS

The integration of LLMs into personalized diabetes care faces multifaceted challenges that significantly hinder widespread adoption. Technical barriers pose the most prominent obstacles, particularly the substantial computational resources required to handle large, unstructured healthcare data (Nagarajan et al., 2024) and the complexities of integrating LLMs with existing hospital systems that rely on outdated technology unsuited for seamless AI integration (Mackenzie et al., 2024; Nagarajan et al., 2024). Real-world validation remains critically limited, as AI models demonstrating potential in controlled environments lack documentation of performance in diverse clinical settings, especially within underrepresented populations where healthcare systems are less equipped for advanced AI deployment (Campanella et al., 2024; Guan et al., 2023). Furthermore, gaps in interdisciplinary collaboration between clinicians and AI developers create a disconnect between practical healthcare needs and technical expertise, resulting in AI models that fail to align with clinical realities and limiting their practical usefulness.

Overcoming these technical barriers, evidence limitations, and collaboration gaps is essential for successful LLM integration into clinical diabetes care.

FUTURE DIRECTIONS

The future of LLM applications in diabetes care encompasses four interconnected trajectories requiring coordinated advancement across technological, privacy, cultural, and clinical integration dimensions. Personalized AI agents represent the most transformative frontier, with systems continuously evolving through dynamic learning from health data and behavioral patterns, as demonstrated by early prototypes providing tailored feedback based on real-time glucose levels (Ellahham, 2020) and advanced AI-driven insulin pumps showing proven effectiveness in autonomous treatment delivery (Campanella et al., 2024). Privacy-preserving federated learning frameworks offer promising solutions for collaborative model development while maintaining patient confidentiality through decentralized training (Hasan et al., 2024; Takahashi et al., 2021), though challenges persist regarding computational limitations and the need for high-quality, representative datasets across diverse healthcare systems (Schachner et al., 2020; Yang & Li, 2025). Ultimately, successful clinical integration necessitates explainable AI systems that build trust through interpretable recommendations aligned with clinical practices (Saraswat et al., 2022), coupled with patient-centered co-design processes ensuring AI tools meet individual health needs and real-world diabetes management demands (Ayobi et al., 2021; Turchi et al., 2024), collectively indicating that the practical impact and clinical adoption of LLM-based diabetes care solutions will depend on interdisciplinary collaboration across these convergent elements.

DISCUSSION & CONCLUSION

This systematic review examined the integration of LLMs into personalized diabetes care, focusing on their clinical applications, model adaptation strategies, and ethical implications. Table 1 presents an overview of the various LLM applications in diabetes care, summarizing different use cases, target conditions, AI/technologies employed, and their respective purposes.

Table 1. Overview of LLM Applications in Diabetes Care

Application	Models	Purpose	Supporting Literature
Patient Education & Conversational Agents	Custom Multimodal LLM, GPT-3.5	Enhance adherence, simulate diagnostic dialogues, evaluate diabetes education	(Dao et al., 2024; Kelly et al., 2024)
Clinical Decision Support	GPT-3.5/4, AI-SCE	Provide clinical recommendations, support treatment planning	(Liu et al., 2023; Williams et al., 2024)
Predictive Modeling	GPT-4, DeepDR-LLM, Time-LLM	Risk prediction for complications, glucose forecasting	(Choi et al., 2025; Lara-Abelenda et al., 2025)
EHR Integration	GPT-4o, Med-PaLM 2, AMIE	Automate record summarization, streamline documentation	(Van Veen et al., 2024; Verma et al., 2025).

Building on these findings, the implications for various stakeholders, clinicians, patients, policymakers, and developers are multifaceted. Table 2 explores the model adaptation strategies employed to optimize LLMs for the specific needs of diabetes care,

Table 2. Model Adaptation Strategies for Healthcare LLMs Grouped by Task Type

Task Type	Key Models	Adaptation Methods	Main Benefits	Supporting Literature
Clinical Decision Support	GPT-4, Claude, BERT, T5	Prompt engineering, fine-tuning, reinforcement learning	High accuracy, clinical relevance, privacy preservation	(Barra et al., 2025; Wang et al., 2021)
Medical QA & Decision	Med-PaLM, BioBERT, ChatGLM	Domain pretraining, instruction tuning, LoRA fine-tuning	Superior biomedical reasoning, patient-centered responses	(AlSaad et al., 2024; Yang et al., 2023)
Patient Engagement	ChatGPT, ClinicalBERT, Gemini	Prompt chaining, QLoRA fine-tuning, medical dialogue training	Empathetic communication, enhanced provider interactions	(Alavi et al., 2024; Song et al., 2025)
Diagnostic Support	GPT-4, LLaVA-Med, SciBERT	Domain pretraining, visual integration, prompt-based tuning	Improved diagnostic precision, automated reporting	(Alavi et al., 2024; Bautista et al., 2023)
Multimodal Integration	PandaGPT, LLaMA2, Time-LLM	Multimodal encoders, time-series modeling, knowledge injection	Robust reasoning, complex domain communication	(Bui et al., 2025; Song et al., 2025)

These findings collectively suggest that while LLMs offer substantial promises for transforming diabetes care, they should be viewed as augmentative tools, not replacements for healthcare professionals. Table 3 highlights the ethical, legal, and social implications (ELSI) related to the deployment of LLMs in diabetes care. The table underscores key issues such as bias, fairness, data privacy, and the need for explainable AI models.

Table 3. Ethical, Legal, and Social Implications of LLMs in Personalized Diabetes Care

LLM Models	Key Applications	Main Challenges	Ethical Considerations	Supporting Literature
GPT-3/4, Med-PaLM, LLaMA	Clinical decision support, diagnostics	Privacy, fairness, transparency	Patient data protection, bias prevention	(He et al., 2025)
GPT-3/4, BioBERT	Patient care optimization, diagnosis	Data bias, privacy issues	Medical data ethics, decision transparency	(Nassiri & Akhloufi, 2024)
GPT-4, ClinicalBERT	Medical analysis, virtual assistants	Inaccurate diagnosis, interpretability	Healthcare fairness and accountability	(Wang & Zhang, 2024)
GPT-3.5/4, PaLM	Documentation, predictive modeling	Inconsistent performance, AI errors	Bias mitigation, outcome transparency	(Mesko & Topol, 2023)
GPT-3.5, Clinical-T5	Healthcare robotics, diagnostics	Physical task accuracy, data bias	Autonomous medical robot ethics	(Pashangpour & Nejat, 2024)

This systematic review addressed three key objectives regarding LLM integration into personalized diabetes care. The evaluation of clinical applications revealed significant potential in patient education, diagnostic support, and clinical decision-making, contributing to more personalized and data-driven diabetes management. The examination of adaptation strategies, including fine-tuning, domain-specific training, and multimodal integration, demonstrated their importance in optimizing model performance for diabetes-specific needs. The analysis of ethical considerations emphasized critical concerns regarding data privacy, fairness, and transparency that must be addressed for responsible clinical implementation.

By synthesizing current knowledge, this review provides recommendations for future research to refine LLMs for optimal use in personalized diabetes care while addressing identified literature gaps.

REFERENCES

1. Alavi, A., Cha, K., Esfarjani, D. P., Patel, B., Than, J. L. P., Lee, A. Y., Nebeker, C., Snyder, M., & Bahmani, A. (2024). *Harnessing Large Language Models to Uncover Insights in Diabetes Wearable Data*.
2. AlSaad, R., Abd-Alrazaq, A., Boughorbel, S., Ahmed, A., Renault, M. A., Damseh, R., & Sheikh, J. (2024). *Multimodal Large Language Models in Health Care: Applications, Challenges, and Future Outlook*. https://www.ncbi.nlm.nih.gov/pubmed/39321458
3. Ayobi, A., Stawarz, K., Katz, D., Marshall, P., Yamagata, T., Santos-Rodriguez, R., Flach, P., & O'Kane, A. A. (2021). Co-Designing Personal Health? Multidisciplinary Benefits and Challenges in Informing Diabetes Self-Care Technologies. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1-26. https://doi.org/10.1145/3479601
4. Barra, F. L., Rodella, G., Costa, A., Scalogna, A., Carenzo, L., Monzani, A., & Corte, F. D. (2025). From prompt to platform: an agentic AI workflow for healthcare simulation scenario design. *Adv Simul (Lond)*, *10*(1), 29. https://doi.org/10.1186/s41077-025-00357-z
5. Bautista, Y. J. P., Lima, V., Theran, C., & Aló, R. (2023). *Health disparities through generative ai models: comparison study using A domain specific large language model*. https://arxiv.org/abs/2310.18355
6. Bui, N., Nguyen, G., Nguyen, N., Vo, B., Vo, L., Huynh, T., Tang, A., Tran, V. N., Huynh, T., Nguyen, H. Q., & Dinh, M. (2025). Fine-tuning large language models for improved health communication in low-resource languages. *Comput Methods Programs Biomed*, *263*, 108655. https://doi.org/10.1016/j.cmpb.2025.108655
7. Campanella, S., Paragliola, G., Cherubini, V., Pierleoni, P., & Palma, L. (2024). Towards Personalized AI-Based Diabetes Therapy: A Review. *IEEE J Biomed Health Inform*, *28*(11), 6944-6957. https://doi.org/10.1109/JBHI.2024.3443137
8. Chinta, S. V., Wang, Z., Palikhe, A., Zhang, X., Kashif, A., Smith, M. A., Liu, J., & Zhang, W. (2025). AI-driven healthcare: Fairness in AI healthcare: A survey. *PLOS Digital Health*, *4*(5), e0000864. https://doi.org/10.1371/journal.pdig.0000864
9. Choi, E. Y., Choi, J. Y., & Yoo, T. K. (2025). Automated and code-free development of a risk calculator using ChatGPT-4 for predicting diabetic retinopathy and macular edema without retinal imaging. *Int J Retina Vitreous*, *11*(1), 11. https://doi.org/10.1186/s40942-025-00638-9
10. Dao, D., Teo, J. Y. C., Wang, W., & Nguyen, H. D. (2024). Llm-powered multimodal ai conversations for diabetes prevention. Proceedings of the 1st ACM Workshop on AI-Powered Q&A Systems for Multimedia,
11. Dodevski, Z., Pavleska, T., & Trajkovik, V. (2024). *Leveraging Federated Learning for Secure Transfer and Deployment of ML Models in Healthcare* Proceedings of Chat Gpt in Medicine 2024,
12. Ellahham, S. (2020). Artificial Intelligence: The Future for Diabetes Care. *Am J Med*, *133*(8), 895-900. https://doi.org/10.1016/j.amjmed.2020.03.033
13. Guan, Z., Li, H., Liu, R., Cai, C., Liu, Y., Li, J., Wang, X., Huang, S., Wu, L., Liu, D., Yu, S., Wang, Z., Shu, J., Hou, X., Yang, X., Jia, W., & Sheng, B. (2023). Artificial intelligence in diabetes management: Advancements, opportunities, and challenges. *Cell Rep Med*, *4*(10), 101213. https://doi.org/10.1016/j.xcrm.2023.101213
14. Hasan, M. R., Li, Q., Saha, U., & Li, J. (2024). Decentralized and Secure Collaborative Framework for Personalized Diabetes Prediction. *Biomedicines*, *12*(8). https://doi.org/10.3390/biomedicines12081916
15. He, K., Mao, R., Lin, Q., Ruan, Y., Lan, X., Feng, M., & Cambria, E. (2025). A survey of large language models for healthcare: from data, technology, and applications to accountability and ethics. *Information Fusion*, *118*. https://doi.org/10.1016/j.inffus.2025.102963
16. Ho, C. N., Tian, T., Ayers, A. T., Aaron, R. E., Phillips, V., Wolf, R. M., Mathioudakis, N., Dai, T., & Klonoff, D. C. (2024). Qualitative metrics from the biomedical literature for evaluating large language models in clinical decision-making: a narrative review. *BMC Med Inform Decis Mak*, *24*(1), 357. https://doi.org/10.1186/s12911-024-02757-z

17. Kelly, A., Noctor, E., & Van de Ven, P. (2024). Design, architecture and safety evaluation of an AI chatbot for an educational approach to health promotion in chronic medical conditions. *Procedia Computer Science*, *248*, 52-59.
18. Khalifa, M., & Albadawy, M. (2024). Artificial intelligence for diabetes: Enhancing prevention, diagnosis, and effective management. *Computer Methods and Programs in Biomedicine Update*, *5*. https://doi.org/10.1016/j.cmpbup.2024.100141
19. Lara-Abelenda, F. J., Chushig-Muzo, D., Peiro-Corbacho, P., Wägner, A. M., Granja, C., & Soguero-Ruiz, C. (2025). Personalized glucose forecasting for people with type 1 diabetes using large language models. *Computer Methods and Programs in Biomedicine*, *265*, 108737. https://doi.org/https://doi.org/10.1016/j.cmpb.2025.108737
20. Liu, S., Wright, A. P., Patterson, B. L., Wanderer, J. P., Turer, R. W., Nelson, S. D., McCoy, A. B., Sittig, D. F., & Wright, A. (2023). Using AI-generated suggestions from ChatGPT to optimize clinical decision support. *J Am Med Inform Assoc*, *30*(7), 1237-1245. https://doi.org/10.1093/jamia/ocad072
21. Mackenzie, S. C., Sainsbury, C. A. R., & Wake, D. J. (2024). Diabetes and artificial intelligence beyond the closed loop: a review of the landscape, promise and challenges. *Diabetologia*, *67*(2), 223-235. https://doi.org/10.1007/s00125-023-06038-8
22. Marko, J. G. O., Neagu, C. D., & Anand, P. B. (2025). Examining inclusivity: the use of AI and diverse populations in health and social care: a systematic review. *BMC Medical Informatics and Decision Making*, *25*(1), 57. https://doi.org/10.1186/s12911-025-02884-1
23. Mesko, B., & Topol, E. J. (2023). The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *NPJ Digit Med*, *6*(1), 120. https://doi.org/10.1038/s41746-023-00873-0
24. Nagarajan, R., Kondo, M., Salas, F., Sezgin, E., Yao, Y., Klotzman, V., Godambe, S. A., Khan, N., Limon, A., Stephenson, G., Taraman, S., Walton, N., Ehwerhemuepha, L., Pandit, J., Pandita, D., Weiss, M., Golden, C., Gold, A., Henderson, J., Shippy, A., Celi, L. A., Hogan, W. R., Oermann, E. K., Sanger, T., & Martel, S. (2024). Economics and Equity of Large Language Models: Health Care Perspective. *J Med Internet Res*, *26*, e64226. https://doi.org/10.2196/64226
25. Nassiri, K., & Akhloufi, M. A. (2024). Recent Advances in Large Language Models for Healthcare. *BioMedInformatics*, *4*(2), 1097-1143. https://doi.org/10.3390/biomedinformatics4020062
26. Pachetti, E., & Colantonio, S. (2024). A systematic review of few-shot learning in medical imaging. *Artif Intell Med*, *156*, 102949. https://doi.org/10.1016/j.artmed.2024.102949
27. Park, S. H., Choi, J., & Byeon, J. S. (2021). Key Principles of Clinical Validation, Device Approval, and Insurance Coverage Decisions of Artificial Intelligence. *Korean J Radiol*, *22*(3), 442-453. https://doi.org/10.3348/kjr.2021.0048
28. Pashangpour, S., & Nejat, G. (2024). The Future of Intelligent Healthcare: A Systematic Analysis and Discussion on the Integration and Impact of Robots Using Large Language Models for Healthcare. *Robotics*, *13*(8). https://doi.org/10.3390/robotics13080112
29. Saraswat, D., Bhattacharya, P., Verma, A., Prasad, V. K., Tanwar, S., Sharma, G., Bokoro, P. N., & Sharma, R. (2022). Explainable AI for Healthcare 5.0: Opportunities and Challenges. *IEEE Access*, *10*, 84486-84517. https://doi.org/10.1109/access.2022.3197671
30. Schachner, T., Keller, R., & F, V. W. (2020). Artificial Intelligence-Based Conversational Agents for Chronic Conditions: Systematic Literature Review. *J Med Internet Res*, *22*(9), e20701. https://doi.org/10.2196/20701
31. Singhal, K., Azizi, S., Tu, T., Mahdavi, S. S., Wei, J., Chung, H. W., Scales, N., Tanwani, A., Cole-Lewis, H., Pfohl, S., Payne, P., Seneviratne, M., Gamble, P., Kelly, C., Babiker, A., Scharli, N., Chowdhery, A., Mansfield, P., Demner-Fushman, D., Aguera, Y. A. B., Webster, D., Corrado, G. S., Matias, Y., Chou, K., Gottweis, J., Tomasev, N., Liu, Y., Rajkomar, A., Barral, J., Semturs, C., Karthikesalingam, A., & Natarajan, V. (2023). Large language models encode clinical knowledge. *Nature*, *620*(7972), 172-180. https://doi.org/10.1038/s41586-023-06291-2
32. Song, Z., Yan, B., Liu, Y., Fang, M., Li, M., Yan, R., & Chen, X. (2025). *Injecting Domain-Specific Knowledge into Large Language Models: A Comprehensive Survey*. https://arxiv.org/abs/2502.10708

33. Standl, E., Khunti, K., Hansen, T. B., & Schnell, O. (2019). The global epidemics of diabetes in the 21st century: Current situation and perspectives. *Eur J Prev Cardiol*, *26*(2_suppl), 7-14. https://doi.org/10.1177/2047487319881021
34. Takahashi, Y., Chang, H.-t., Nakai, A., Kagawa, R., Ando, H., Imakura, A., Okada, Y., Tsurushima, H., Suzuki, K., & Sakurai, T. (2021). Decentralized Learning with Virtual Patients for Medical Diagnosis of Diabetes. *SN Computer Science*, *2*(4). https://doi.org/10.1007/s42979-021-00564-1
35. Turchi, T., Prencipe, G., Malizia, A., Filogna, S., Latrofa, F., & Sgandurra, G. (2024). Pathways to democratized healthcare: Envisioning human-centered AI-as-a-service for customized diagnosis and rehabilitation. *Artif Intell Med*, *151*, 102850. https://doi.org/10.1016/j.artmed.2024.102850
36. Van Veen, D., Van Uden, C., Blankemeier, L., Delbrouck, J. B., Aali, A., Bluethgen, C., Pareek, A., Polacin, M., Reis, E. P., Seehofnerová, A., Rohatgi, N., Hosamani, P., Collins, W., Ahuja, N., Langlotz, C. P., Hom, J., Gatidis, S., Pauly, J., & Chaudhari, A. S. (2024). Adapted large language models can outperform medical experts in clinical text summarization. *Nat Med*, *30*(4), 1134-1142. https://doi.org/10.1038/s41591-024-02855-5
37. Verma, R., Alsentzer, E., Strasser, Z., Chang, L., Roman, K., Gershanik, E., Hernandez, C., Linares, M., Rodriguez, J., Thakral, D., Unlu, O., You, J., Zhou, L., & Bates, D. (2025). Verifiable Summarization of Electronic Health Records Using Large Language Models to Support Chart Review. *medRxiv*. https://doi.org/10.1101/2025.06.02.25328807
38. Vuohijoki, A., Mikkola, I., Jokelainen, J., Keinanen-Kiukaanniemi, S., Winell, K., Frittitta, L., Timonen, M., & Hagnas, M. (2020). Implementation of a Personalized Care Plan for Patients With Type 2 Diabetes Is Associated With Improvements in Clinical Outcomes: An Observational Real-World Study. *J Prim Care Community Health*, *11*, 2150132720921700. https://doi.org/10.1177/2150132720921700
39. Wang, D., & Zhang, S. (2024). Large language models in medical and healthcare fields: applications, advances, and challenges. *Artificial Intelligence Review*, *57*(11). https://doi.org/10.1007/s10462-024-10921-0
40. Wang, J., Shi, E., Yu, S., Zihao Wu, C. M., Haixing Dai, Qiushi Yang, Yanqing Kang, Jinru Wu,, Huawen Hu, C. Y., Haiyang Zhang, Yiheng Liu, Yi Pan, Zhengliang Liu, Lichao Sun,, & Xiang Li, B. G., Xi Jiang, Dajiang Zhu, Yixuan Yuan, Dinggang Shen, Tianming Liu, and Shu Zhang. (2021). Prompt Engineering for Healthcare:Methodologies and Applications. *JOURNAL OF LATEX CLASS FILES*, *14*(8). https://doi.org/http://dx.doi.org/10.48550/arXiv.2304.14670
41. Williams, C. Y. K., Miao, B. Y., Kornblith, A. E., & Butte, A. J. (2024). Evaluating the use of large language models to provide clinical recommendations in the Emergency Department. *Nat Commun*, *15*(1), 8236. https://doi.org/10.1038/s41467-024-52415-1
42. Yang, H., Li, J., Liu, S., Du, L., Liu, X., Huang, Y., Shi, Q., & Liu, J. (2023). *Exploring the Potential of Large Language Models in Personalized Diabetes Treatment Strategies*.
43. Yang, X., & Li, J. (2025). A clustering-based federated deep learning approach for enhancing diabetes management with privacy-preserving edge artificial intelligence. *Healthcare Analytics*, *7*. https://doi.org/10.1016/j.health.2025.100392
44. Zou, Y., Zhao, L., Zhang, J., Wang, Y., Wu, Y., Ren, H., Wang, T., Zhang, R., Wang, J., Zhao, Y., Qin, C., Xu, H., Li, L., Chai, Z., Cooper, M. E., Tong, N., & Liu, F. (2022). Development and internal validation of machine learning algorithms for end-stage renal disease risk prediction model of people with type 2 diabetes mellitus and diabetic kidney disease. *Ren Fail*, *44*(1), 562-570. https://doi.org/10.1080/0886022x.2022.2056053