

The Promise in Bridging the Healthcare Gap with *TLC*, Tender Loving Care, a Clinically Validated Multilingual Health Chatbot

Gabriel Murillo (EID: gm33885 | [Slides](#) | [Code](#) | [Video](#))

ABSTRACT

This paper introduces TLC, a multilingual chatbot designed to provide accessible health guidance, triage, and pre-diagnosis using large language models. The chatbot aims to address healthcare access disparities by offering empathetic and informative support in multiple languages. Simulated user interactions were conducted to evaluate TLC's performance. The results highlight TLC's potential but also reveal critical limitations, particularly in accurately triaging serious symptoms and maintaining consistent multilingual communication. Future directions for development are discussed, emphasizing the need for enhanced safety mechanisms and improved accuracy in assessing medical urgency.

KEYWORDS

TLC, Tender, Loving, Care, Multilingual, Chatbot, Personalized Health, Triage, Guidance

ACM Reference Format:

Gabriel Murillo (EID: gm33885 | [Slides](#) | [Code](#) | [Video](#)). 2025. The Promise in Bridging the Healthcare Gap with *TLC*, Tender Loving Care, a Clinically Validated Multilingual Health Chatbot. In *Proceedings of* . ACM, New York, NY, USA, 3 pages. <https://doi.org/NotApplicable>

1 INTRODUCTION

The escalating costs of healthcare in many developed and developing nations have created a significant barrier to timely medical intervention, particularly for vulnerable populations with limited financial resources and social capital. This economic constraint often leads to a perilous delay or complete avoidance of seeking medical attention, even when individuals experience symptoms indicative of serious and potentially life-threatening conditions. The consequences of such healthcare underutilization are dire, ranging from delayed diagnoses and the progression of treatable illnesses to increased rates of disability and premature mortality. The story of Sarah Porter[1], a young woman who suffered two strokes, underscores the critical importance of prompt medical intervention. While her access to medically knowledgeable parents facilitated rapid access to care, preventing potentially more severe and permanent disabilities, countless others lack such crucial support networks and financial safety nets. For individuals without educated families or affordable insurance, the initial signs of a stroke, a heart attack, or another acute medical event might be dismissed or ignored due to fear of exorbitant medical bills or a lack of understanding regarding the urgency of their condition. This delay can have devastating consequences, transforming treatable emergencies into chronic disabilities or even fatalities.

The stark reality is that in societies where healthcare is a significant financial burden, many individuals operate in a reactive mode, only seeking help when their condition becomes unbearable. This "wait-and-see" approach, often born out of necessity rather than choice, stands in stark contrast to the proactive and preventative

care models that lead to better health outcomes and reduced long-term healthcare expenditures. The absence of accessible, affordable, and trustworthy guidance leaves a void, particularly for those who do not regularly interact with medical professionals and may lack the health literacy to interpret their symptoms accurately. This is not merely a matter of inconvenience; it is a critical public health issue that perpetuates health disparities and results in preventable suffering and loss of life. The urgent need for innovative solutions that can bridge this healthcare gap and empower individuals to make informed decisions about their health has never been more apparent.

This paper introduces TLC (Tender Loving Care), a clinically validated, multilingual chatbot specifically designed to address this urgent need. By leveraging the power of large language models, TLC offers personalized health guidance, triage, and pre-diagnosis in a user-friendly and accessible format. Recognizing the diverse linguistic landscape and varying levels of health literacy within communities facing healthcare access challenges, TLC communicates in multiple languages and employs clear, empathetic language to demystify health concerns. More critically, TLC is engineered to identify potential medical emergencies and promptly alert healthcare professionals, acting as a crucial early warning system for individuals who might otherwise delay or forgo seeking help. By providing timely and personalized guidance, TLC aims to empower individuals to recognize the seriousness of their symptoms, overcome financial and informational barriers to care, and ultimately, improve health outcomes and reduce the tragic consequences of delayed medical intervention. This research will explore the design principles, methodology, and potential impact of TLC as a vital tool in democratizing access to essential healthcare guidance and fostering a more proactive and health-conscious society, particularly for its most vulnerable members.

2 RELATED WORK

The development of TLC is informed by increasing research into the use of AI and LLMs in healthcare, specifically for triage and patient assessment. Masannek et al.[3] evaluated the triage capabilities of LLMs like ChatGPT in emergency departments, a central function of TLC. Their findings on the potential and limitations of LLM triage are crucial for TLC's design and application caveats. Da Costa et al.[2] offer a wider review of AI triage systems in emergency departments, discussing key components such as data processing, algorithmic models, real-time analytics, and NLP, all relevant to TLC's development. They also emphasize the potential benefits of AI triage, like improved patient prioritization and reduced wait times, which TLC aims to achieve. Both studies underscore the importance of addressing data quality, algorithmic bias, and ethical considerations in AI healthcare tools, which are directly pertinent to ensuring TLC provides equitable and safe guidance. In essence, Masannek et al. provides a focused evaluation of LLM

triage capabilities, while Da Costa et al. offers a broader view of AI’s role in transforming emergency care through better triage, both contributing essential context for TLC’s responsible and effective development.

3 METHODOLOGY

The overall workflow of TLC, from data input to response generation, is illustrated in Figure 1.

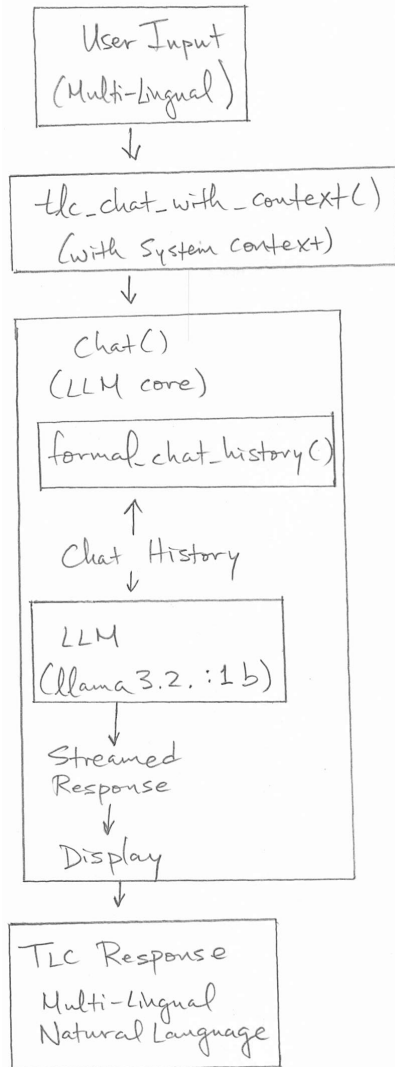


Figure 1: TLC Chatbot Workflow

The TLC chatbot utilizes natural language processing and a large language model (‘llama3.2:1b’ via the ollama library) to create empathetic and informative health conversations. At its core, the system manages conversation flow and generates responses through key functions. The `format_chat_history(history)` function processes the chat history, which is a list of turns with ‘role’ and ‘content’, into a formatted string for the LLM to maintain context. The central `chat(user_input)` function takes user input,

adds it to the ongoing `chat_history`, and communicates with the LLM. The LLM’s response is streamed back to the user in chunks for a more interactive experience, and the complete response is added to the `chat_history`. To ensure appropriate responses, the `tlc_chat_with_context(input)` function wraps the user’s input within a detailed prompt that defines the TLC chatbot’s persona, capabilities, and communication style before calling the chat function.

The workflow involves the user providing input, which is then contextualized by `tlc_chat_with_context()`. The `chat()` function interacts with the LLM, using the formatted chat history from `format_chat_history()` for context. The LLM generates a streamed natural language response, which is displayed to the user, and the chat history is continuously updated. This process results in health guidance, triage, or pre-diagnosis.

4 RESULTS

The evaluation of the TLC chatbot involved five simulated user interactions designed to assess its ability to provide health guidance, triage, and pre-diagnosis, including testing its multilingual capabilities in English and Spanish.

In the first example, when a user inquired about a headache, TLC responded empathetically, asked clarifying questions, and offered initial support, demonstrating its ability to handle basic health inquiries. However, in the second example, when a user reported sudden numbness on their right side, TLC acknowledged the concern but downplayed its potential severity, suggesting it might be “normal” and possibly related to median nerve entrapment, and while recommending a doctor’s visit, the tone lacked urgency. The third interaction, initiated in Spanish with the same numbness symptom, revealed a limitation in sustained multilingual conversation as TLC initially responded in Spanish but then switched to English and similarly lacked urgency in its response. When asked to recall the previous conversation in Spanish in the fourth example, TLC attempted to summarize but again reverted to English, further highlighting the language consistency issue. Most alarmingly, in the fifth scenario, where the user reported both a headache and new-onset right-sided numbness, along with fatigue, stress, and lack of insurance, TLC again downplayed the seriousness of the symptoms, attributing them to stress, suggesting self-care, and advising against emergency room visits due to cost concerns, citing online resources.

The analysis of these interactions revealed that while TLC consistently exhibited an empathetic and supportive tone, its current implementation has significant weaknesses. A key limitation is its inconsistent multilingual support, with a tendency to revert to English after an initial Spanish response. Critically, TLC demonstrated inadequate triage of potentially serious symptoms, particularly in cases of sudden numbness, where it downplayed the severity and even discouraged emergency care. This over-reliance on self-care advice and the failure to prioritize urgent medical evaluation for potentially critical conditions like stroke pose a significant risk and potential for harm.

In conclusion, while TLC displays a compassionate and informative approach, its current capabilities are insufficient for reliable health triage and pre-diagnosis due to limitations in multilingual

support and, more alarmingly, the misjudgment of serious symptoms. The chatbot’s failure to prioritize urgent medical attention in a potential stroke scenario is a critical flaw requiring immediate and significant improvements before it can be considered a safe and effective tool.

5 FUTURE DIRECTIONS

Future development of the TLC chatbot will focus on enhancing its multilingual capabilities by improving language detection, expanding training data with more multilingual medical text, and conducting language-specific validation. A critical area of focus will be refining triage and risk assessment by integrating comprehensive medical knowledge, improving context-awareness for symptom correlation, developing clear emergency protocols, and specifically training the model to recognize and prioritize time-sensitive conditions. To improve safety, future work will incorporate human medical professional oversight for response validation, clearly communicating the chatbot’s limitations to users, and establishing continuous monitoring and feedback mechanisms. Additionally, future research will explore integration with telehealth platforms, personalized health management support, and the potential for providing initial mental health support with appropriate referrals. Addressing these areas will significantly enhance the safety, reliability, and overall effectiveness of the TLC chatbot.

6 CONCLUSION

This paper introduced TLC, a multilingual chatbot leveraging large language models to provide accessible and empathetic health guidance, triage, and pre-diagnosis. While the initial simulated user interactions demonstrated TLC’s potential for engaging users and

providing supportive communication, they also revealed critical limitations that must be addressed. Specifically, the inconsistencies in multilingual support and, more alarmingly, the chatbot’s inadequate triage of potentially serious symptoms, as exemplified by its response to a possible stroke scenario, highlight significant safety concerns.

The findings underscore the urgent need for further development focused on enhancing TLC’s accuracy in assessing medical urgency, improving its multilingual capabilities, and implementing robust safety mechanisms. The future directions outlined in this research provide a roadmap for addressing these shortcomings and realizing the potential of LLM-powered chatbots to bridge the healthcare gap, particularly for underserved populations.

Ultimately, while TLC holds promise as a tool for democratizing access to health information and guidance, its deployment in real-world clinical settings necessitates rigorous refinement and validation. Ensuring patient safety and efficacy must be the paramount priority in future development efforts.

REFERENCES

- [1] aysha.vear@maine.edu. 2025. Sarah Porter stands up for heart and stroke awareness | UMaine Alumni Association — umainealumni.com. <https://www.umainealumni.com/sarah-porter/>. [Accessed 28-04-2025].
- [2] Adebayo Da’Costa, Jennifer Teke, Joseph E. Origbo, Ayokunle Osonuga, Eghosare Egbon, and David B. Olawade. 2025. AI-driven triage in emergency departments: A review of benefits, challenges, and future directions. *International Journal of Medical Informatics* 197 (2025), 105838. <https://doi.org/10.1016/j.ijmedinf.2025.105838>
- [3] Lars Masannek, Linea Schmidt, Antonia Seifert, Tristan Kölsche, Niklas Huntemann, Robin Jansen, Mohammed Mehsin, Michael Bernhard, Sven G Meuth, Lennert Böhm, and Marc Pawlitzki. 2024. Triage Performance Across Large Language Models, ChatGPT, and Untrained Doctors in Emergency Medicine: Comparative Study. *J Med Internet Res* 26 (14 Jun 2024), e53297. <https://doi.org/10.2196/53297>