OPTIMIZING THE MANAGEMENT OF TRANSPLANT PATIENT INQUIRIES THROUGH AUTOMATED ANSWERS: DEVELOPMENT AND EVALUATION OF A SAFE, PERSONALIZED AI CHATBOT

Cristian Jarry, MD, MSc; Ismael Peregrina, MBA; Miguel Blanco; Mauricio Barrón; Alejandra Cancino, RN;
Any Requena, RN; Sofía Abedrapo, MD; María E Vial, MD; Diego Sanhueza, MD; Pablo Achurra, MD, MSc.
Center for Simulation and Experimental Surgery, Pontificia Universidad Católica de Chile;
CHRISTUS Excellence and Innovation Center, Monterrey, Mexico;
Transplant Institute, UC-CHRISTUS Clinical Hospital, Chile

Background

Nearly 50,000 organ transplants were performed in the U.S. in 2023, with transplant patients requiring continuous care and frequent consultations with healthcare providers. While many of these interactions occur in formal settings, others often occur informally via WhatsApp, phone calls, or email. At our institution, which manages over 500 solid-organ transplant patients, nurses often serve as the first line of contact, receiving multiple informal inquiries daily. AI-driven tools could help address patient questions, but commercially available large language models (LLMs) are not designed for clinical use and may produce inaccurate or harmful information, raising concerns about patient safety and confidentiality. This study aims to develop and evaluate an AI-based chatbot to safely manage patient inquiries and reduce the workload on nursing staff.

Methods

A prospective study was conducted to evaluate an AI-based chatbot for transplant patient queries. A set of **150 frequently asked patient questions was curated by the transplant institute and narrowed to 80 essential questions.** Regarding the AI-tool, a JSONfile served as the knowledge base, and Term-Frequency-Inverse Document Frequency (TF-IDF) was applied to identify key concepts. It was developed in Python and deployed using Azure OpenAI GPT-4.0 Turbo, with a focus on data security. Using a Stage-Gate approach, random subsets of questions were tested, with a multidisciplinary team (surgeon, nurse, and researchers) assessing the chatbot's responses. Content accuracy, safety and language appropriateness were assessed and weighed in a summary "Approval rate (%)". Each Stage-Gate (SG 1-3) introduced new challenges, including typos, biased questions, multiple queries in a single submission, and attempts to "bypass" the Chatbot's safety mechanisms. A 90% approval threshold was set to progress to the next stage, with lower-performing configurations ("creativity levels") excluded from subsequent stages.

Results

Five users, evaluating 20 new questions on each stage, analysis of the "spontaneous answer" provided by the ChatBot and a Yes/no approval assessment regarding: **content accuracy, safety and language appropriateness**

STAGE-GATE 1: All 5 "Creativity levels" configuration

"Creativity levels" (0, 0.25, 0.5, 0.75, 1)
0.5 = 96% approval rate
0.75 = 92% approval rate

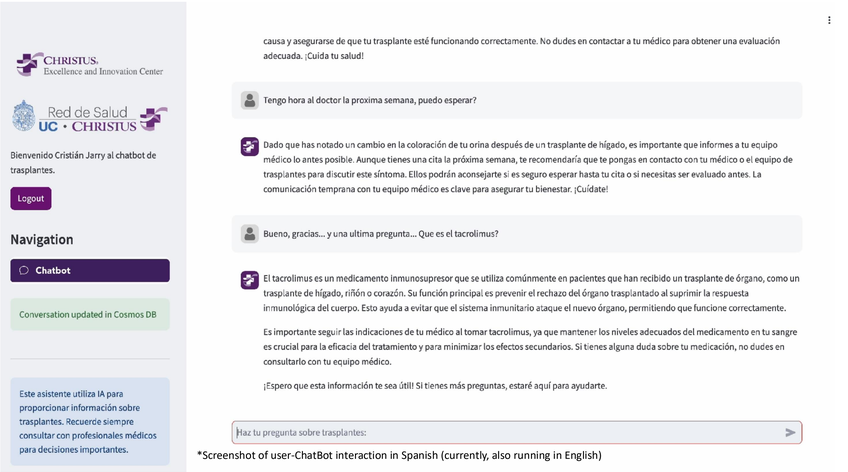
STAGE-GATE 2: "Creativity level validation", testing 0.5 against 0.75

0.75 = sustained 92% approval rate
Used onwards...

STAGE-GATE 3: Complex questions, typos, biased questions, multiple queries and bypass attempts

Overall 94.79% Approval rate

STAGE-GATE 4 (undergoing): preparing for patient use



*Screenshot of user-ChatBot interaction in Spanish (currently, also running in English)

Conclusion

Preliminary findings suggest that the AI-based chatbot can accurately and safely handle common transplant patient inquiries in a protected and confidential environment. Given its success in this controlled environment, future steps include testing with a selected group of patients under strict monitoring.

SAGES 2025