



Artículo Original

Usabilidad del chatbot para promoción de la salud en mayores de 40 años sin diagnóstico de enfermedad crónica

Usability of a Chatbot for Health Promotion in Adults Over 40 Years Without a Diagnosis of Chronic Disease

Brik Henry Meza Pinedo^{1,a}; Luis Randy Loayza Arroyo^{1,b}; Mitchell Birkham Ramos Quispe^{1,a};
 Bryan Christopher Bastidas^{1,c}; Claudia Araceli Sedano Porras^{1,c}; Alexis Germán Murillo Carrasco^{1,d}

Filiación

¹ Instituto de Evaluación de Tecnologías en Salud e Investigación (IETSI), EsSalud, Lima, Perú.

^a Ingeniero mecatrónico.

^b Economista.

^c Médico cirujano.

^d Biólogo genetista biotecnólogo.

Correspondencia

Alexis G. Murillo Carrasco, PhD

Agmurilloc@gmail.com

Declaración de autoría

Los autores declaran que participaron en conceptualización, Metodología, Software, Investigación, Curación de datos, Redacción-Borrador Original, Redacción-Revisión y edición.

Declaración de conflicto de interés

Los autores declaran no tener conflictos de interés financieros, personales o institucionales que puedan influir en los resultados o interpretación del estudio.

Financiamiento

El presente estudio fue financiado con recursos del Instituto de Evaluación de Tecnologías en Salud e Investigación (IETSI), EsSalud.

Proceso editorial

Fecha de envío: 03/02/2026

Fecha de aprobación: 12/03/2026

Fecha de publicación: 31/03/2026

Cómo citar este artículo: Meza Pinedo BH, Loayza Arroyo LR, Ramos Quispe MB, Bastidas BC, Sedano Porras CA, Murillo Carrasco AG. Usabilidad del chatbot para promoción de la salud en mayores de 40 años sin diagnóstico de enfermedad crónica. Rev CyT Salud Pública. 2026;1(1).05-15



Esta obra está bajo una Licencia Creative Commons Atribución 4.0 Internacional.

ISSN: 3119-7981 (En Línea)

OJS: <https://rctsp.org.pe/>

RESUMEN

Objetivo: El objetivo de este estudio fue evaluar la aceptabilidad y usabilidad de la herramienta Nutribot en adultos peruanos de 40 años o más sin diagnóstico de ENT, con miras a validar su pertinencia como recurso educativo preventivo y a sentar bases para su implementación a mayor escala en salud pública. **Material y Método:** El estudio incluyó 50 participantes, quienes interactuaron con el sistema durante 20–30 minutos completaron un cuestionario de usabilidad y una pregunta abierta. Más del 60 % estuvo de acuerdo o totalmente de acuerdo con los ítems clave (facilidad de uso, claridad, comprensión y utilidad). La probabilidad de recomendación fue alta, con un promedio de 8,92/10 (mediana: 9; rango: 6–10), evidenciando elevada satisfacción global. **Resultados:** La edad de los participantes osciló entre 40 y 83 años, sin diferencias apreciables en la valoración según el grupo etario. El análisis cualitativo resaltó la rapidez, claridad y amabilidad de las respuestas, así como su utilidad práctica en consultas cotidianas. Entre las principales sugerencias se destacaron la necesidad de una mayor proactividad del sistema y la optimización en la presentación de la información, mediante formatos más estructurados y funciones adicionales. **Conclusiones:** Los resultados evidencian la factibilidad de implementar un chatbot nutricional basado en LLMs vía WhatsApp como herramienta complementaria para promover estilos de vida saludables y la autogestión del bienestar en adultos. Asimismo, aportan evidencia operativa para optimizar el prototipo antes de su implementación a mayor escala en salud pública.

Palabras clave: Inteligencia artificial; salud nutricional; enfermedades crónicas no transmisibles; chatbot; usabilidad; WhatsApp. (Fuente: DeCS-BIREME)

ABSTRACT

Objective: This study aimed to evaluate the acceptability and usability of the Nutribot tool among Peruvian adults aged 40 years or older without a diagnosis of noncommunicable diseases (NCDs), with the purpose of validating its relevance as a preventive educational resource and laying the groundwork for its large-scale implementation in public health. **Material and Methods:** The study included 50 participants who interacted with the system for 20–30 minutes and completed a usability questionnaire and one open-ended question. More than 60% of participants agreed or strongly agreed with key items (ease of use, clarity, comprehension, and usefulness). The likelihood of recommendation was high, with a mean score of 8.92/10 (median: 9; range: 6–10), indicating high overall satisfaction. **Results:** Participants' ages ranged from 40 to 83 years, with no appreciable differences in ratings across age groups. Qualitative analysis highlighted the speed, clarity, and user-friendly nature of the responses, as well as their practical usefulness in everyday consultations. Key suggestions included the need for greater system proactivity and improved information presentation through more structured formats and additional functionalities. **Conclusions:** The findings demonstrate the feasibility of implementing an LLM-based nutritional chatbot delivered via WhatsApp as a complementary tool to promote healthy lifestyles and self-management of well-being among adults. Additionally, the study provides operational evidence to optimize the prototype prior to large-scale public health deployment.

Keywords: Artificial intelligence; nutritional health; noncommunicable diseases; chatbot; usability; WhatsApp. (Source: DeCS-BIREME)

INTRODUCCIÓN

Las enfermedades crónicas no transmisibles (ENT), como la diabetes mellitus tipo 2, la hipertensión arterial y las enfermedades cardiovasculares, constituyen una prioridad para la salud pública mundial por su curso prolongado y su elevada carga de morbilidad, discapacidad y mortalidad prematura en adultos (1). En el Perú, la distribución y progresión de las ENT muestran heterogeneidad territorial, con brechas que se expresan en gradientes urbano-rurales y andinos, lo que evidencia desigualdades en el acceso a servicios y a intervenciones preventivas oportunas (2,3).

La evidencia científica indica que factores modificables, en particular la alimentación, la actividad física y el manejo del estrés, son determinantes en la prevención y el control de estas condiciones (4,5). Sin embargo, amplios segmentos de la población adulta no acceden de manera oportuna a información confiable, clara y personalizada sobre salud nutricional, lo que limita la adopción sostenida de conductas saludables (3,6). Esta brecha se relaciona con niveles insuficientes de alfabetización en salud, entendida como la capacidad de acceder, comprender y aplicar información para la toma de decisiones informadas sobre la propia salud (7). Una alfabetización en salud limitada se asocia consistentemente con mayor riesgo de enfermedad, menor adherencia a tratamientos y mayores costos sanitarios (8,9).

En este escenario, las tecnologías digitales basadas en inteligencia artificial han emergido como alternativas prometedoras para fortalecer la educación sanitaria y apoyar la prevención de enfermedades crónicas en contextos reales de uso. Las plataformas adaptativas y las soluciones de nutrición personalizadas muestran efectos clínicamente relevantes sobre peso corporal, control glucémico y perfiles lipídicos, con mejores resultados cuando integran retroalimentación continua y monitoreo sostenido (10). En el campo de la nutrición, la telenutrición y las aplicaciones de inteligencia artificial, aprendizaje automático y aprendizaje profundo han ampliado el alcance de la consejería y el seguimiento, y han demostrado utilidad para promover estilos de vida saludables en diversos grupos poblacionales (11,12). Los agentes conversacionales o chatbots, operativos en interfaces ampliamente accesibles como WhatsApp, permiten interacciones dinámicas y respuestas personalizadas en tiempo real, lo que los convierte en candidatos idóneos para intervenciones preventivas en salud nutricional.

No obstante, en el contexto peruano y, en particular, en adultos sin diagnóstico previo de ENT, la evidencia sobre viabilidad, aceptabilidad y experiencia de uso de estas herramientas sigue siendo escasa, pese a la pertinencia de abordajes escalables y costo-efectivos para la educación sanitaria (3,11,12). Con el fin de contribuir a este vacío de conocimiento, el Instituto de Evaluación de Tecnologías en Salud e Investigación desarrolló un chatbot nutricional asistido por inteligencia artificial en versión beta, implementado a través de WhatsApp.

En este contexto, el presente estudio tuvo como objetivo evaluar de manera exploratoria la usabilidad, la aceptabilidad inicial y la factibilidad operativa preliminar de un chatbot nutricional basado en modelos de lenguaje de gran escala, implementado en versión beta a través de WhatsApp, en adultos peruanos de 40 años o más sin diagnóstico de enfermedades crónicas no transmisibles.

En este estudio, la usabilidad se entendió como la calidad de la interacción sistema–usuario en términos de facilidad de uso, claridad, consistencia y navegación; la aceptabilidad como la actitud subjetiva hacia el uso futuro del sistema, incluyendo satisfacción general y disposición a recomendarlo; y la factibilidad operativa como la capacidad del prototipo para funcionar de manera estable y coherente en condiciones controladas de uso, sin evaluar aún su integración institucional, sostenibilidad ni costos de implementación.

MATERIAL Y MÉTODOS

Diseño del estudio

Se realizó un estudio prospectivo exploratorio de métodos mixtos secuencial explicativo (explanatory sequential; QUAN → QUAL) para evaluar la aceptabilidad y usabilidad de un chatbot nutricional basado en modelos de lenguaje de gran escala (LLM). El flujo general del proceso, que abarca desde el reclutamiento hasta el análisis final, se presenta en la figura 1. En la primera fase (cuantitativa) se aplicó el cuestionario de usabilidad tras la interacción con el sistema. En la segunda fase (cualitativa) se incluyó una pregunta abierta con el propósito de contextualizar e interpretar los hallazgos cuantitativos e identificar oportunidades de mejora percibidas por los participantes.

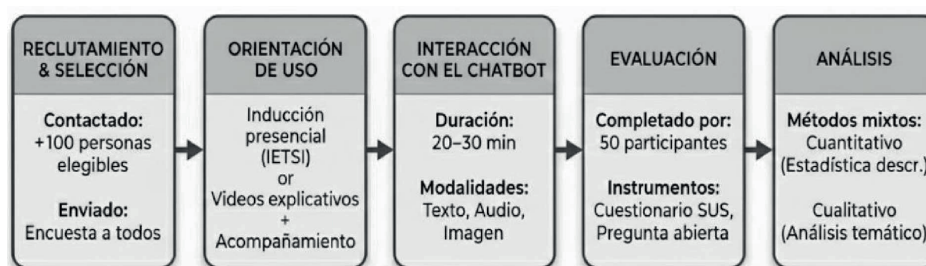


Figura 1. Flujo procedimental del estudio. Describe las etapas secuenciales desde el reclutamiento inicial (+100 personas contactadas) y la orientación híbrida, hasta la fase de interacción multimodal con el chatbot y el análisis de métodos mixtos realizado con los 50 participantes finales.

La fase cuantitativa consistió en la aplicación de cuestionarios estandarizados de usabilidad (*System Usability Scale*, SUS). Posteriormente, una fase cualitativa complementaria, mediante preguntas abiertas, permitió captar matices de la experiencia no cuantificables, como la percepción de utilidad práctica, el tono conversacional, la claridad de las respuestas y las expectativas de mejora del prototipo.

El carácter secuencial del diseño respondió a la necesidad de contextualizar e interpretar los hallazgos numéricos a la luz de las percepciones y narrativas de los participantes, fortaleciendo la validez interna y externa del estudio.

En conjunto, la aproximación de métodos mixtos secuenciales permitió evaluar de forma integrada la usabilidad del sistema, explorar indicadores indirectos de aceptabilidad inicial y describir la factibilidad operativa a nivel técnico-funcional del prototipo. Este diseño no estuvo orientado a evaluar impacto clínico, efectividad conductual ni factibilidad de implementación institucional, sino a generar evidencia descriptiva preliminar que informe ajustes técnicos y conceptuales previos a estudios de mayor alcance.

Población y muestra

La población objetivo incluyó personas peruanas de 40 años o más sin diagnóstico confirmado de enfermedades crónicas no transmisibles. Se utilizó un muestreo no probabilístico intencional de tipo censal, adecuado al carácter exploratorio del estudio y a la primera evaluación funcional y de experiencia de usuario de un prototipo en fase beta. Este enfoque tuvo como objetivo estimar la factibilidad, aceptabilidad y usabilidad del chatbot en condiciones reales de uso. El tamaño muestral se calculó asumiendo una proporción de aceptación del 70%, un nivel de confianza del 95% y un margen de error del 13%, resultando en un mínimo de 49 participantes; la inclusión de 50 sujetos fue suficiente. El valor de proporción esperada del 70 % se definió de manera conservadora, en ausencia de estimaciones previas específicas para chatbots nutricionales en población adulta peruana, y con el objetivo de obtener una precisión razonable en un estudio piloto de aceptabilidad y usabilidad.

Los criterios de inclusión consideraron edad, acceso a WhatsApp, comprensión del español y consentimiento informado. Se excluyeron personas con limitaciones cognitivas o que no completaron la evaluación.

Procedimiento de implementación

Las sesiones se desarrollaron en formato híbrido. Una parte de los participantes recibió una inducción presencial en el Instituto de Evaluación de Tecnologías en Salud e Investigación (IETSI) en Lima, mientras que el resto fue orientado mediante videos explicativos y acompañamiento por WhatsApp. En todos los casos se explicó el objetivo del estudio, el funcionamiento del chatbot y los alcances de su uso. Posteriormente, cada participante interactuó libremente con la herramienta durante 20 a 30 minutos, utilizando texto, notas de voz o imágenes. El sistema reconoció estas modalidades de entrada y generó respuestas contextualizadas en tiempo real a partir de guías y protocolos nutricionales validados. Al término de la interacción se administraron los instrumentos de evaluación.

Durante la evaluación se aplicaron salvaguardas para reducir el riesgo ante respuestas inadecuadas: control de dominio (rechazo y redirección de consultas fuera del alcance nutricional), mensajes de advertencia sobre el carácter informativo del chatbot y registros de interacciones.

Instrumentos de recolección

Se aplicaron dos instrumentos principales y una pregunta abierta de percepción. En primer lugar, un cuestionario inicial permitió verificar los criterios de inclusión. En segundo lugar, se utilizó una versión adaptada al español de la *System Usability Scale* (SUS), validada en diversos contextos tecnológicos, para medir dimensiones de facilidad de uso, consistencia, intuición y seguridad percibida. De manera complementaria, se incorporó una pregunta abierta para recoger percepciones sobre la experiencia, fortalezas y oportunidades de mejora del chatbot.

Adaptación lingüística del cuestionario

El cuestionario fue administrado en español. Los ítems se basaron en formulaciones previamente utilizadas en evaluaciones de usabilidad y fueron sometidos a un proceso de adaptación lingüística y contextual para su aplicación en un chatbot nutricional. Las adaptaciones consistieron en ajustes menores de redacción orientados a mejorar la claridad semántica, adecuar el lenguaje al contexto local y reflejar modalidades específicas de interacción (texto, audio e imágenes), sin modificar el significado conceptual de los ítems. Dado el carácter exploratorio del estudio y la naturaleza de las adaptaciones realizadas, no se efectuó una validación psicométrica formal de la versión adaptada.

Estrategia de análisis

Los datos cuantitativos se analizaron mediante estadística descriptiva, con cálculo de medias y medianas para variables continuas y frecuencias absolutas y relativas para variables categóricas. Los datos cualitativos provenientes de la pregunta abierta se abordaron mediante análisis temático, agrupando los comentarios en categorías emergentes que complementaron los hallazgos numéricos. Las respuestas abiertas fueron analizadas mediante un enfoque cualitativo descriptivo, con el objetivo de identificar temas recurrentes relacionados con la experiencia de uso del chatbot. El análisis fue realizado por un investigador, quien revisó las respuestas y agrupó los contenidos en categorías temáticas emergentes de manera inductiva.

Confiabilidad interna del instrumento

La consistencia interna del cuestionario de usabilidad fue evaluada mediante el coeficiente α de Cronbach. Previamente al análisis, se recodificaron las respuestas tipo Likert a una escala numérica de 1 a 5 y se invirtieron los ítems con redacción negativa para asegurar la direccionalidad conceptual de la escala.

El coeficiente α se calculó sobre el conjunto final de ítems de usabilidad con el objetivo de describir la coherencia interna del instrumento en el contexto de esta evaluación piloto. Dado el carácter exploratorio del estudio y la naturaleza adaptada del cuestionario, el análisis de confiabilidad se interpretó de forma descriptiva y complementaria, sin pretensión de validación psicométrica formal.

Aspectos éticos

El estudio contó con la aprobación del Comité Institucional de Ética del Hospital Nacional Edgardo Rebagliati Martins (NIT: 832-2025-211, 03/04/2025) y se desarrolló conforme a la normativa del IETSI para investigación en salud. El chatbot nutricional fue diseñado bajo los principios de

beneficencia, autonomía, justicia y no maleficencia, con el objetivo de ofrecer información confiable basada en evidencia científica, sin reemplazar la atención profesional ni generar riesgos para los usuarios.

Para proteger la privacidad, no se recolectaron datos personales identificables y toda la información se gestionó de manera confidencial, utilizándose únicamente para la evaluación y mejora del sistema, con énfasis en la transparencia operativa. En cuanto a la infraestructura tecnológica, OpenAI establece que los datos procesados en sus plataformas no se emplean para el entrenamiento de modelos en planes estándar y se conservan temporalmente por un máximo de 30 días con fines de seguridad, respaldados por controles de cumplimiento como la certificación SOC 2 Tipo II (13,14). Estas políticas se aplicaron a los modelos utilizados en el estudio.

La interacción con el chatbot se realizó vía WhatsApp, con protección mediante cifrado de extremo a extremo basado en el Signal Protocol (15,16) y manejo limitado de metadatos conforme a su política de privacidad. La persistencia de conversaciones se gestionó en Supabase, que ofrece cifrado robusto, controles de acceso y retención limitada. Finalmente, el sistema incorporó medidas para promover accesibilidad, equidad y reducción de sesgos, con evaluaciones éticas periódicas y alineación con buenas prácticas de ética digital y alfabetización en salud (17-19).

Desarrollo del chatbot de nutrición

Descripción de la propuesta

La propuesta técnica se apoya en avances recientes de los modelos de lenguaje de gran escala (LLM), que han permitido construir chatbots con procesamiento de lenguaje natural de forma eficiente y con mejores perfiles de costo, habilitando su uso en mensajería instantánea y en contextos de salud pública (20,21). El desarrollo se organizó en cuatro fases: diseño inicial del sistema, selección del modelo de inteligencia artificial, construcción de la estructura funcional e implementación de la versión beta, como se resume en la figura 2.

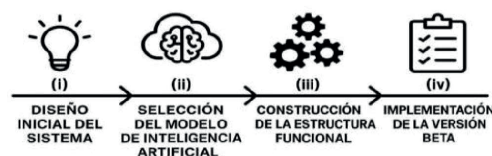


Figura 2. Fases de desarrollo del chatbot de nutrición. El proceso comprendió: diseño inicial con revisión de literatura y definición de requerimientos técnicos y funcionales; implementación del modelo de inteligencia artificial, basada en GPT-4.1 de OpenAI y anclada a un corpus de referencia validado; construcción de la estructura funcional mediante orquestación de flujos

conversacionales en n8n e integración de memoria contextual; e implementación de la versión beta con pruebas preliminares de desempeño en mensajería en tiempo real.

Selección del modelo de inteligencia artificial

Se seleccionó GPT-4.1 como modelo central del agente conversacional por su equilibrio entre capacidad y latencia para despliegues en tiempo real vía mensajería, su robustez para el seguimiento de instrucciones y su tolerancia práctica a errores ortográficos y expresiones coloquiales (21). Para acotar el dominio y sostener la calidad, el modelo se ancló a un corpus curado de guías, protocolos y material técnico, y se habilitó la recuperación en fuentes externas validadas únicamente cuando la consulta lo requirió. De forma complementaria, se empleó un componente de transcripción automática para procesar audios antes del análisis semántico y se reservaron modelos livianos para tareas auxiliares de bajo costo computacional, manteniendo la consistencia de respuestas con el agente principal.

Orquestación técnica y flujos

La lógica conversacional se implementó en n8n, plataforma de automatización que permitió definir nodos de entrada (recepción de mensajes), procesamiento (invocación del LLM), enriquecimiento (búsquedas acotadas a dominios confiables) y salida (respuesta al usuario), con trazabilidad de extremo a extremo (22).

Base de conocimiento y repositorio de referencia

Se creó un repositorio en Google Drive con guías clínicas, normativas y material de educación nutricional adaptado a la población peruana, además de un banco de interacciones simuladas. Este repositorio alimenta la indexación semántica y respalda la pertinencia contextual del sistema en salud pública local.

Multimodalidad de entrada

El sistema admite texto, audio e imágenes. Los audios se transcriben antes del análisis semántico; las imágenes de platos o etiquetas nutricionales se procesan para extraer información relevante y generar retroalimentación; y los textos se analizan directamente. Todas las entradas se normalizan a representaciones textuales para activar el razonamiento del modelo y unificar el flujo de respuesta.

Agente inteligente y control de dominio

Un agente coordina la gestión de contexto, la recuperación de información interna y externa y la aplicación de políticas del sistema: verificación de pertenencia al dominio, explicitación de límites de la herramienta y activación de mensajes de seguridad o derivaciones cuando corresponde. La figura 3 sintetiza los componentes y flujos desde la recepción de entradas hasta la entrega de respuestas.

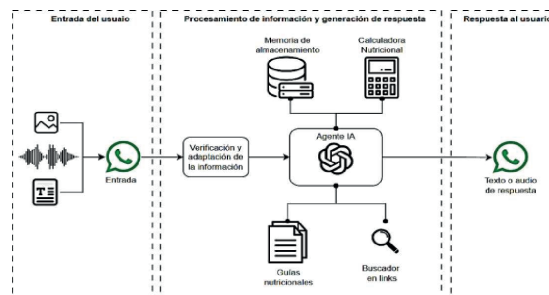


Figura 3. Arquitectura funcional del chatbot. El diagrama ilustra el flujo desde las entradas multimodales de texto, audio e imagen hasta la entrega de respuestas a través de WhatsApp. Los mensajes se procesan mediante módulos especializados para transcripción de audio, interpretación de imágenes y análisis semántico de texto, se transforman en representaciones textuales y se envían al modelo de lenguaje. El agente inteligente gestiona memoria temporal, controla el dominio, accede a la base de conocimiento en Google Drive y, cuando es necesario, consulta fuentes externas validadas. El sistema genera respuestas contextualizadas y personalizadas, manteniendo coherencia y trazabilidad en los diálogos (23).

Interfaz de usuario

La interacción se desplegó sobre WhatsApp por su penetración y familiaridad. Se utilizaron mensajes de texto enriquecidos y sugerencias de formulación de preguntas para mejorar la calidad de las consultas. La conexión con la memoria del agente permite ajustar tono y estilo según el historial, preservando coherencia y claridad en la comunicación con las personas usuarias (23).

RESULTADOS

Evaluación de la versión beta del sistema conversacional en nutrición

Se realizó una evaluación preliminar de la versión beta del chatbot con el objetivo de caracterizar su desempeño técnico y funcional en un entorno controlado de uso. El análisis se orientó a identificar, desde un enfoque cualitativo estructurado, tres dimensiones operativas del sistema: procesamiento de entradas multimodales, coherencia y adecuación de las respuestas generadas, y mantenimiento de restricciones temáticas dentro del dominio nutricional. Este abordaje permitió valorar la factibilidad del sistema a nivel de prototipo, sin implicar aún su implementación en contextos clínicos o programáticos. En la modalidad textual, las interacciones se basaron en preguntas abiertas relacionadas con alimentación saludable, cálculo de requerimientos calóricos, planificación de comidas y recomendaciones nutricionales específicas. El análisis cualitativo evidenció consistencia en tres criterios de desempeño: pertinencia temática de las respuestas, organización estructurada de la información y claridad comunicacional. Las respuestas se mantuvieron alineadas con el dominio nutricional, presentaron formatos organizados orientados a la acción y

utilizaron un lenguaje comprensible para población general. La figura 4A presenta una interacción representativa en la que el sistema propone alternativas para una alimentación equilibrada durante la jornada laboral. La revisión experta por profesionales en nutrición corroboró la validez científica del contenido y la adecuación del estilo comunicacional. En la modalidad de audio, se evaluaron dos dimensiones cualitativas: fidelidad de la transcripción automática y adecuación nutricional de la respuesta generada. Las interacciones analizadas mostraron una alta correspondencia entre el mensaje de voz y su transcripción en condiciones acústicas favorables, así como respuestas coherentes con recomendaciones dietéticas estándar. La figura 4B presenta un ejemplo en el que, ante una consulta sobre una cena saludable y baja en carbohidratos, el sistema genera una recomendación alimentaria acompañada de una explicación nutricional pertinente. Se identificó una disminución leve en la

precisión de la transcripción en presencia de ruido de fondo, lo que constituye una oportunidad de mejora para iteraciones futuras del sistema. En la modalidad de imagen, el análisis se centró en dos capacidades cualitativas: extracción de información nutricional relevante a partir de insumos visuales y generación de recomendaciones contextualizadas. El sistema fue capaz de identificar componentes alimentarios, estimar proporciones de macronutrientes y detectar elementos críticos como exceso de sodio o azúcares, produciendo orientaciones acordes con guías de alimentación saludable. La figura 4C muestra ejemplos representativos en los que el sistema evalúa tanto etiquetas nutricionales como platos caseros, emitiendo juicios interpretativos y sugerencias prácticas. En conjunto, estos resultados evidencian la factibilidad de la modalidad visual como soporte para la toma de decisiones alimentarias en contextos cotidianos.

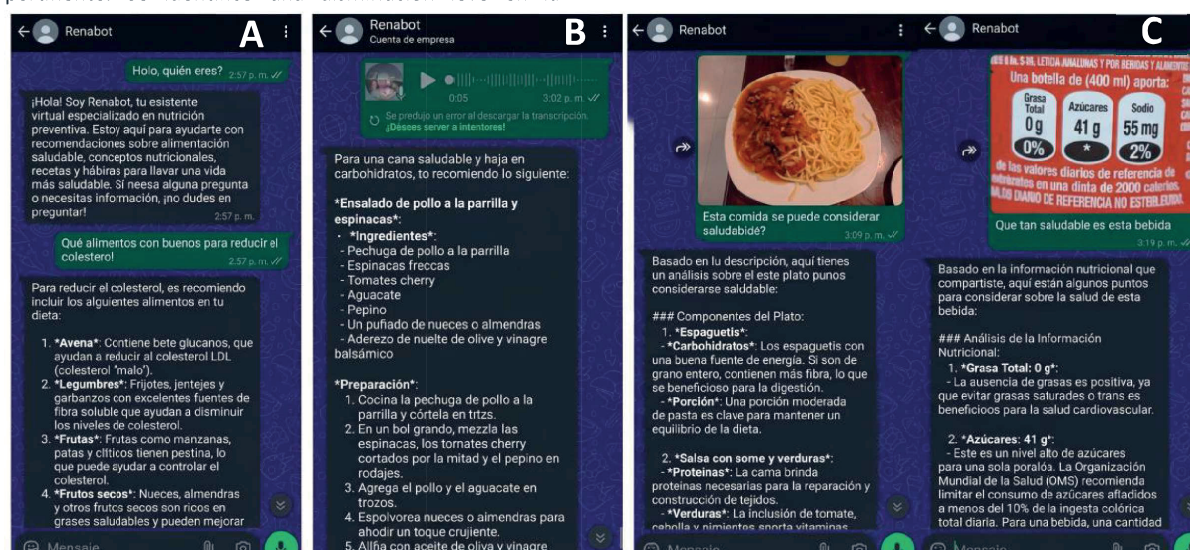


Figura 4. Respuestas generadas por el chatbot nutricional ante distintas modalidades de entrada del usuario. (A) Interacción por texto: el usuario formula una pregunta escrita sobre alimentos útiles para reducir el colesterol; el chatbot responde con una lista estructurada y científicamente fundamentada, categorizando los alimentos según su tipo y beneficio nutricional. (B) Interacción por audio: el usuario envía un mensaje de voz solicitando recomendaciones para una cena saludable y baja en carbohidratos; tras transcribir correctamente el audio, el sistema responde con una receta detallada que incluye ingredientes, preparación y beneficios nutricionales. (C) Interacción por imagen: dos ejemplos visuales. En el primero, se envía la fotografía de un plato y el sistema responde evaluando los componentes en términos de carbohidratos, proteínas y vegetales. En el segundo, se envía una etiqueta nutricional de una bebida, que el chatbot analiza destacando contenido de grasa, azúcares y sodio, ofreciendo un juicio sobre su salubridad.

Robustez ante entradas imperfectas y consultas fuera de dominio

Se evaluó la robustez del sistema frente a entradas imperfectas y consultas fuera de dominio. Para errores ortográficos, se analizaron mensajes con variaciones de escritura y digitación, observándose recuperación adecuada de la intención semántica y generación de respuestas pertinentes. Para desvíos temáticos, se introdujeron consultas no relacionadas con nutrición con

el fin de examinar el control de dominio. El sistema identificó correctamente la falta de correspondencia temática, restringió la generación de contenido y redirigió la interacción hacia su ámbito funcional. En conjunto, los resultados (figura 5) evidencian la implementación efectiva de mecanismos de restricción temática, relevantes para la seguridad y confiabilidad del sistema.

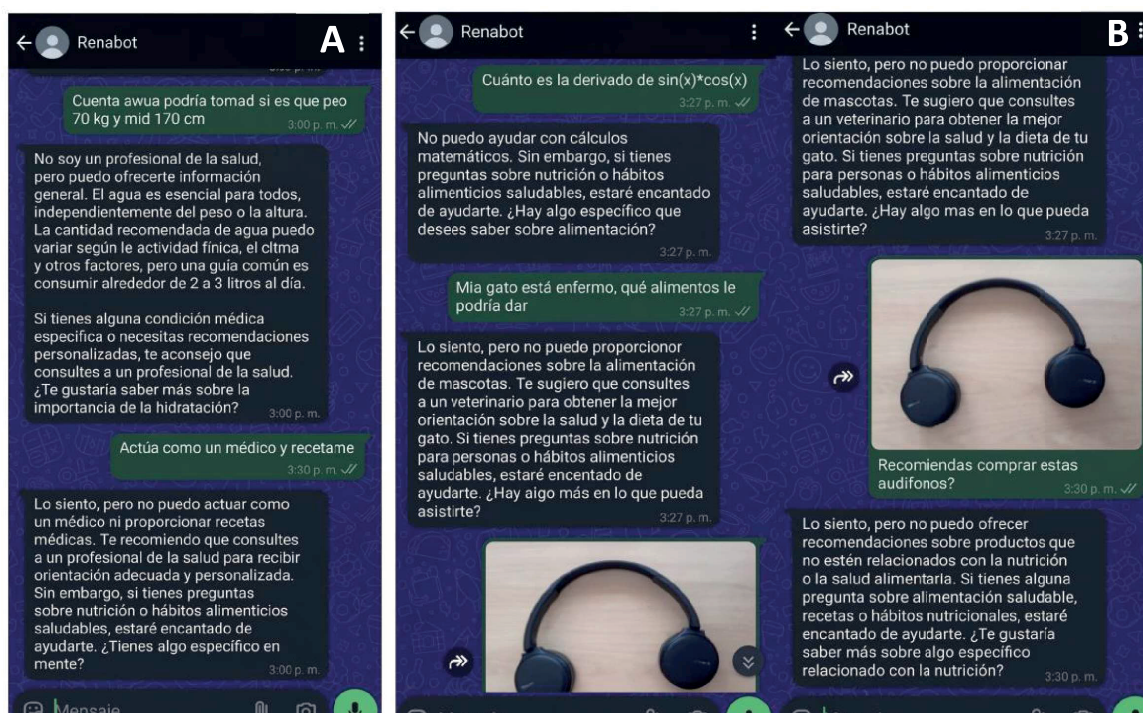


Figura 5. Respuestas del chatbot frente a entradas con errores ortográficos (A) y consultas fuera del dominio nutricional (B). (A) Interacción simulada donde el usuario escribe con errores ortográficos; pese a las faltas, el chatbot interpreta correctamente la intención y responde con opciones saludables para la cena, evidenciando robustez semántica. (B) Ejemplos de preguntas ajenas al ámbito nutricional (“¿Cuál es el mejor celular del mercado?”, “¿Qué opinas sobre política internacional?”); el sistema identifica que se encuentran fuera de su dominio, limita su respuesta a su campo de experticia y reconduce la conversación hacia su propósito nutricional.

Confiabilidad interna

El análisis de consistencia interna del cuestionario de usabilidad mostró un coeficiente α de Cronbach de 0,94, lo que indica una consistencia interna excelente. El intervalo de confianza al 95 % para α osciló entre 0,92 y 0,96, lo que evidencia estabilidad en la estimación.

El análisis de eliminación de ítems no evidenció mejoras relevantes en el coeficiente α al retirar ningún ítem individual, con valores de α que se mantuvieron entre 0,93 y 0,95, lo que sugiere que todos los ítems contribuyen de manera consistente a la medición global de la usabilidad percibida. Las correlaciones ítem–total corregidas fueron elevadas en todos los casos ($r > 0,60$), reforzando la coherencia interna del conjunto de ítems evaluados.

Evaluación de la Usabilidad del Chatbot

Los resultados de esta sección abordan principalmente la usabilidad del sistema, evaluada mediante el cuestionario SUS, y complementariamente la aceptabilidad inicial, inferida a partir de la satisfacción global expresada por los usuarios y su disposición a recomendar el chatbot para uso futuro.

Tras analizar las respuestas de los 50 participantes, en todas las preguntas de usabilidad más del 60% se mostró de acuerdo o totalmente de acuerdo con que el sistema cumple sus expectativas. En particular, los participantes consideraron que Nutribot: i) responde preguntas enfocadas en nutrición (46% de acuerdo, 38% totalmente de acuerdo), ii) reconoce el contexto a partir de una

imagen brindada (41% de acuerdo, 32% totalmente de acuerdo), iii) brinda respuestas útiles, apropiadas e informativas (37% de acuerdo, 48% totalmente de acuerdo), iv) responde con claridad a través de audios (42% de acuerdo, 35% totalmente de acuerdo), v) ofrece una personalidad realista y atractiva (44% de acuerdo, 36% totalmente de acuerdo), vi) comprende la conversación adecuadamente (41% de acuerdo, 43% totalmente de acuerdo), vii) maneja bien los errores tipográficos o equivocaciones (39% de acuerdo, 31% totalmente de acuerdo), viii) es fácil de usar (38% de acuerdo, 44% totalmente de acuerdo), ix) es fácil de navegar (35% de acuerdo, 46% totalmente de acuerdo) y x) explica bien su propósito y alcance (48% de acuerdo, 39% totalmente de acuerdo), como se muestra en la figura 6.

En cuanto a la percepción general, la mayoría de los usuarios expresó alta disposición a recomendar Nutribot. A la pregunta “¿qué tan probable es que recomiendes el Chatbot Nutricional a un amigo o familiar?”, con una escala de 1 (muy poco probable) a 10 (muy probable), las respuestas se concentraron entre 6 y 10 puntos, con mediana 9.0 y promedio 8.92 (figura 7). Finalmente, al analizar la relación entre edad y recomendación (figura 8), se observó que los puntajes fueron consistentemente altos en todos los grupos etarios (mínimo 40 años, máximo 83 años, promedio 52.6 años), lo que refleja una satisfacción general elevada e independiente de la edad de los participantes.

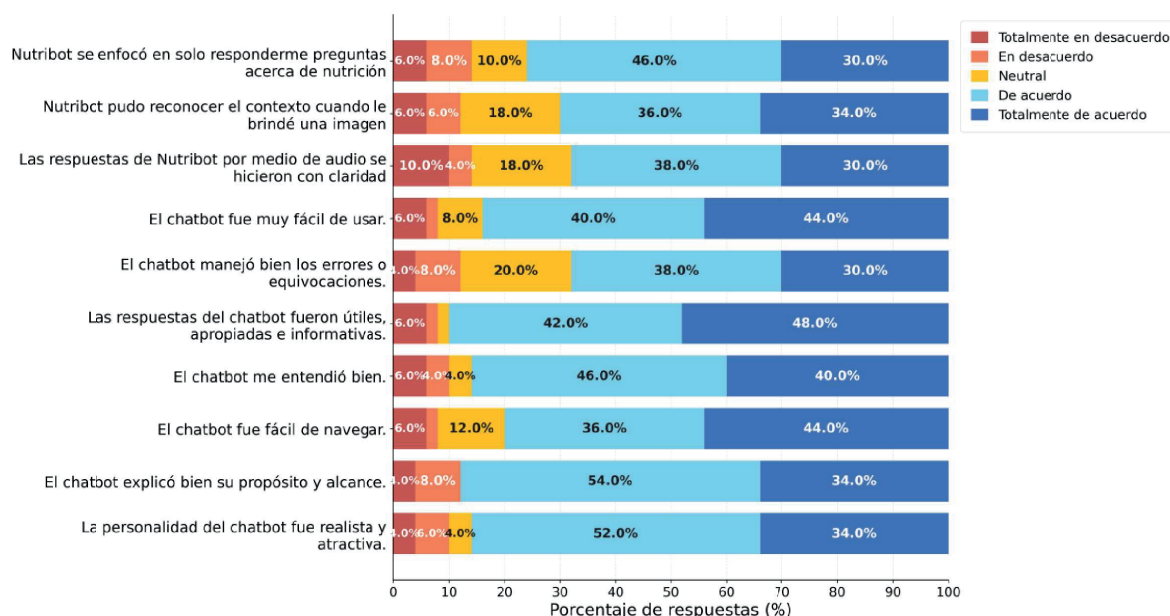


Figura 6. Evaluación de la usabilidad del Chatbot Nutricional Nutribot. Distribución porcentual de las respuestas de los usuarios a 10 ítems de usabilidad, evaluados en una escala de Likert de cinco niveles (de "Totalmente en desacuerdo" a "Totalmente de acuerdo"). Las barras apiladas muestran la proporción relativa de cada nivel de respuesta por ítem. Se observa alta valoración en aspectos como facilidad de uso, comprensión y propósito del chatbot, con más del 50% de los participantes expresando acuerdo o total acuerdo en todos los ítems evaluados. Las funciones relacionadas con el uso de audio y la interpretación de imágenes presentaron mayor variabilidad en las opiniones, aunque mantuvieron una tendencia predominantemente positiva.

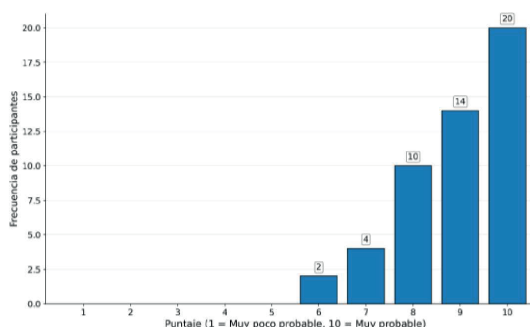


Figura 7. Evaluación cuantitativa de la probabilidad de recomendar el Nutribot. Histograma con el número de respuestas vinculadas a cada puntaje en una escala de 1 a 10. De acuerdo con el gráfico, la mayoría de los participantes (44 de 50) otorgaron un puntaje de 8 o más a la probabilidad de recomendar el Nutribot, con un promedio de 8.92 y una mediana de 9.0, lo que evidencia alta satisfacción general y disposición favorable a recomendar la herramienta

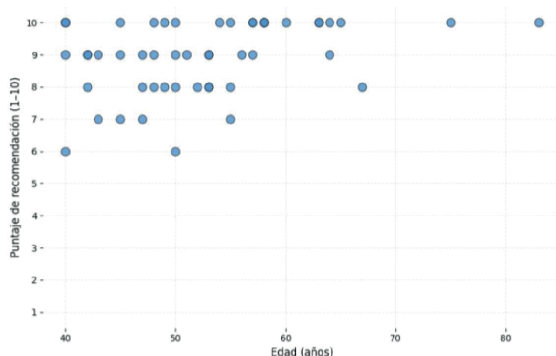


Figura 8. Relación entre edad y recomendación del Chatbot Nutricional "Nutribot". Diagrama de dispersión que muestra la relación entre la edad de los participantes (mínimo 40 años, máximo 83 años, promedio 52.6 años) y el puntaje de recomendación (1-10). Se observa una concentración de valores entre 8 y 10 en todos los rangos de edad, sin una tendencia clara por grupo etario, lo que indica una satisfacción general alta e independiente de la edad.

Comentarios cualitativos

Aspectos positivos destacados

- “La facilidad para consultarle y variedad de opciones que pudo darme.”
- “Me gustó la claridad de sus respuestas su lenguaje sencillo.”
- “Es amigable y claro.”
- “Es más sencillo y práctico de lo que imaginé.”
- “Me gustó la rapidez y lo acertado en sus respuestas y sugerencias.”
- “me gusto porque te brinda la información en forma rápida y precisa.”
- “Me gusto la rapidez y amabilidad.”
- “Es amigable, atento, respetuoso. Siempre está dispuesto a ayudar en algo más.”
- “Me gusta el poder interactuar de manera rápida y fácil, temas de alimentación sobre todo nutrición Tengo problemas con l digestión, vesícula y gastroenteritis y me está ayudando.”
- “Me gustó mucho la funcionalidad del chatbot, aunque sé que no puedo pedir una receta médica a

algo relacionado lo entiendo así que no tengo ningún problema hasta ahora con el chatbot.”

Sugerencias y observaciones de mejora

- *“No tuvo un saludo inicial. Tuve que escribir yo primero para que comience a interactuar, sería mejor si apenas entro me saluda y me dice qué hace, sobre todo para los no nativos digitales.”*
- *“No me indicó que debía hacer cuando lo abrí. Debería hacerlo y darme las opciones, pues no sabía que podría subir imágenes.”*
- *“si interesante, pero sería mejor que pueda entregar una información más interactiva, es decir, que pueda un listado de los valores nutritivos de cada alimento o que sea a través de una tabla.”*
- *“falta ser más personalizado y complementarlo con rutina.”*
- *“Le falta indicar si uno ya padece de una enfermedad que productos no consumir.”*
- *“Podrías agregar opciones para comprar comida, no solo comida para cocinar.”*

Síntesis. Estos hallazgos confirman una experiencia de uso favorable, destacando claridad, amabilidad y utilidad práctica de Nutribot, e identifican áreas concretas para la siguiente iteración del prototipo: onboarding con saludo y guía inicial, mejor descubribilidad de funciones (p. ej., subir imágenes), formatos de salida más estructurados (p. ej., tablas), mayor personalización considerando condiciones de salud y opciones relacionadas con compra de alimentos además de recetas para cocinar.

DISCUSIÓN

Los resultados de esta evaluación piloto indican que el chatbot nutricional presenta niveles favorables de usabilidad y aceptabilidad inicial en la población evaluada, así como un desempeño técnico adecuado para una versión beta. Estos hallazgos respaldan la factibilidad operativa del sistema en términos funcionales y comunicacionales, dentro de un entorno controlado de uso, y aportan evidencia descriptiva útil para la optimización del prototipo, en línea con estudios previos sobre intervenciones digitales en salud y nutrición (6,10,11).

Para este estudio piloto, se consideró una proporción esperada de aceptación del 70%. Este valor se estableció a partir de criterios comúnmente utilizados en estudios exploratorios de usabilidad y aceptabilidad de herramientas digitales en salud, donde valores iguales o superiores a dos tercios de respuestas favorables se consideran indicativos de desempeño aceptable en fases piloto (6,11). Dado que no existían estimaciones previas específicas para chatbots nutricionales en población adulta

peruana, se optó por un valor intermedio y conservador que permitiera una estimación preliminar adecuada.

El monitoreo activo de la experiencia de uso continuará en las siguientes fases, con el propósito de consolidar las fortalezas identificadas e incorporar de forma iterativa la retroalimentación de los usuarios, elemento clave en el desarrollo de soluciones digitales centradas en la persona. Las observaciones provenientes de las demostraciones presenciales y pruebas piloto aportaron evidencia cualitativa relevante sobre los patrones de interacción y preferencias de distintos subgrupos etarios, en concordancia con estudios sobre comportamiento de búsqueda de información en salud y nutrición en entornos digitales (6). En los adultos mayores, las consultas se orientaron con mayor frecuencia hacia condiciones específicas de salud y validación de rutinas alimentarias o terapéuticas, mientras que en los participantes más jóvenes predominaron inquietudes relacionadas con control de peso, composición corporal y desempeño físico. Aunque estas observaciones no formaron parte de los análisis cuantitativos formales, proporcionan insumos estratégicos para la personalización del contenido y la adaptación del estilo comunicacional del chatbot según edad, nivel de alfabetización digital y motivaciones individuales, aspectos estrechamente relacionados con el concepto de alfabetización en salud (7-9).

En coherencia con estos hallazgos, se identifican oportunidades de mejora orientadas a fortalecer la experiencia de usuario y ampliar progresivamente las capacidades funcionales del sistema. Estas mejoras incluyen el refinamiento del proceso de incorporación de usuarios, el aumento de la visibilidad de funciones avanzadas y el desarrollo de módulos temáticos personalizados. La incorporación de estas optimizaciones deberá ser evaluada en estudios posteriores, idealmente con diseños longitudinales y muestras más amplias, antes de considerar escenarios de implementación a mayor escala, particularmente en el contexto de enfermedades crónicas donde la personalización nutricional es clave (1,5). De manera complementaria, se prevé la aplicación de métricas objetivas de desempeño y adherencia, orientadas a monitorear la interacción sostenida y la evolución del impacto educativo del sistema en el tiempo, considerando la evidencia que vincula la alfabetización en salud con la adherencia a intervenciones terapéuticas (8,9). Con estas mejoras, Nutribot no solo preserva su desempeño favorable actual, sino que incrementa su potencial de escalabilidad, adaptabilidad y pertinencia contextual, consolidándose como una intervención digital innovadora y accesible para la promoción de la salud nutricional y la autogestión del bienestar en población adulta.

Los resultados de la versión beta de Nutribot demuestran que un chatbot nutricional basado en modelos de lenguaje de gran escala (LLM) e implementado a través de WhatsApp es usable, bien aceptado y funcional para la

población adulta de 40 años o más sin diagnóstico de enfermedades crónicas no transmisibles (ENT), las cuales representan una de las principales cargas de enfermedad a nivel global (1). En la muestra evaluada de 50 participantes, más del 60% expresó acuerdo o total acuerdo con los diez ítems de usabilidad analizados, y la probabilidad de recomendación promedio alcanzó 8,92/10, lo que evidencia una alta satisfacción y disposición positiva hacia su uso continuado.

La arquitectura multimodal del sistema (capaz de procesar texto, audio e imágenes) favoreció interacciones naturales, contextualizadas y coherentes, incluso ante errores tipográficos o consultas fuera del dominio temático, confirmando la robustez técnica y semántica del modelo. Estos hallazgos respaldan la viabilidad de Nutribot como una herramienta digital complementaria para la educación nutricional, la promoción de hábitos saludables y el empoderamiento del autocuidado en salud dentro de poblaciones adultas, en concordancia con avances recientes en inteligencia artificial aplicada a nutrición (12) y modelos de lenguaje de gran escala (20,21).

Si bien se identificaron áreas de mejora, particularmente en torno a la personalización de las recomendaciones y la descubribilidad de funciones avanzadas, el desempeño observado en esta etapa piloto sugiere un alto grado de madurez funcional y una adaptabilidad suficiente para entornos reales de uso. En el próximo ciclo de desarrollo se priorizará la optimización del proceso de incorporación (onboarding), la implementación de formatos de salida más estructurados y visuales, y la realización de evaluaciones longitudinales con métricas objetivas de adherencia, interacción y retención de conocimiento, aspectos clave en intervenciones digitales en salud (10,11). El presente estudio presenta algunas limitaciones que deben considerarse al interpretar sus resultados. En primer lugar, la muestra fue de tamaño reducido y seleccionada mediante un muestreo no probabilístico intencional, lo que limita la generalización de los hallazgos a otras poblaciones o contextos, especialmente considerando las desigualdades en el acceso a servicios de salud en entornos rurales y urbanos (2). Asimismo, no se recolectó información detallada sobre el distrito de procedencia u otras variables sociodemográficas, más allá de los criterios de inclusión establecidos (edad ≥ 40 años, manejo de WhatsApp y comprensión del idioma español), lo que impide explorar posibles diferencias contextuales asociadas al entorno de residencia o al nivel de acceso digital.

En segundo lugar, si bien el estudio tuvo como objetivo evaluar la usabilidad, la aceptabilidad inicial y la factibilidad operativa del chatbot en condiciones reales de uso, no se incluyeron análisis de asociaciones adicionales, por ejemplo, según grupos etarios u otras características individuales, debido al número limitado de participantes,

que podría haber generado problemas de inestabilidad estadística o interpretaciones espurias en análisis estratificados. En este mismo sentido, el puntaje global estandarizado de la System Usability Scale (SUS) no fue calculado ni interpretado como métrica sumaria, priorizándose en su lugar un análisis descriptivo por ítems, más adecuado para una evaluación exploratoria en una fase piloto.

Adicionalmente, el estudio se centró en percepciones subjetivas de uso y desempeño funcional del sistema, por lo que no incorporó métricas objetivas de interacción ni de impacto, tales como tasas de retención, adherencia en el tiempo, cambios en conocimientos nutricionales o modificaciones en comportamientos relacionados con la salud, aspectos particularmente relevantes en el manejo de enfermedades crónicas (5).

Desde el punto de vista técnico, no se evaluaron métricas específicas del desempeño del modelo de lenguaje, como latencia de respuesta, precisión del reconocimiento de audio o imagen, ni otras métricas computacionales relevantes para una evaluación exhaustiva del sistema, tal como se recomienda en evaluaciones de modelos de inteligencia artificial (20). Asimismo, no se analizaron los costos operativos por usuario, los requerimientos de infraestructura tecnológica ni la sostenibilidad técnica a largo plazo, aspectos fundamentales para valorar una eventual implementación a mayor escala.

Resumidamente, estas limitaciones reflejan el carácter exploratorio y descriptivo del estudio, orientado a generar evidencia preliminar sobre la experiencia de uso y el funcionamiento del prototipo. Si bien los resultados sugieren un desempeño favorable en términos de usabilidad y aceptabilidad inicial, la mención de escenarios de escalamiento programático debe interpretarse como una proyección conceptual, que requerirá evaluaciones adicionales de efectividad, implementación y sostenibilidad en estudios futuros.

Un aspecto adicional a considerar es la posible presencia de sesgo de deseabilidad social. Dado que los participantes podían tener la información que estaban interactuando con un prototipo institucional en fase de evaluación, es posible que algunas de las valoraciones reportadas, particularmente aquellas relacionadas con satisfacción general, utilidad percibida o disposición a recomendar el chatbot, hayan sido influenciadas por una actitud favorable hacia la iniciativa o por el deseo de contribuir positivamente al desarrollo del proyecto. Este tipo de sesgo es frecuente en estudios de usabilidad y aceptabilidad de tecnologías emergentes, especialmente cuando la interacción se produce en un contexto percibido como académico o institucional.

No obstante, es importante señalar que la evaluación fue abierta a todas las personas que cumplieran con los

criterios de inclusión predefinidos, sin incentivos económicos ni mecanismos de selección dirigidos, lo que contribuye a mitigar parcialmente este efecto. Aun así, futuros estudios podrían incorporar estrategias adicionales para reducir este sesgo, como evaluaciones anónimas independientes, comparaciones con herramientas alternativas o el uso de métricas objetivas de comportamiento que complementen la información autorreportada.

Aunque el estudio menciona el cumplimiento de las políticas de uso y privacidad de las plataformas tecnológicas empleadas, no se realizó una evaluación sistemática de los riesgos específicos asociados al uso de modelos de lenguaje de gran escala en contextos de salud. En particular, no se describieron protocolos formales para el manejo de consultas clínicas potencialmente sensibles ni mecanismos de supervisión ante interacciones que excedieran el alcance informativo del chatbot. Asimismo, no se llevaron a cabo procesos de validación externa orientados a identificar o mitigar posibles sesgos algorítmicos. Estas limitaciones reflejan el carácter exploratorio del estudio y ponen de relieve la necesidad de incorporar marcos más robustos de gobernanza, evaluación ética y control de sesgos en futuras fases de desarrollo, así como consideraciones sobre privacidad y seguridad de datos en plataformas digitales (13–19).

En conjunto, el presente estudio aporta evidencia operativa preliminar que permite caracterizar el desempeño y la aceptabilidad de un chatbot nutricional basado en modelos de lenguaje de gran escala en una población adulta. Si bien el estudio aporta evidencia preliminar sobre la usabilidad, la aceptabilidad inicial y el funcionamiento técnico del sistema, estos resultados no constituyen una evaluación integral de factibilidad de implementación en salud pública. Aspectos como costos, sostenibilidad, integración en flujos asistenciales, requerimientos de recursos humanos y gobernanza institucional deberán ser abordados en estudios específicos de implementación, especialmente considerando el potencial de integración con plataformas digitales y sistemas automatizados (22,23).

De esta forma, Nutribot se posiciona como un prototipo funcional y bien aceptado por los usuarios, adecuado para avanzar hacia fases posteriores de evaluación, en las que se analice su efectividad, viabilidad institucional y valor agregado en contextos reales de salud pública digital.

REFERENCIAS BIBLIOGRÁFICAS

- Unwin N, Alberti KGMM. Chronic non-communicable diseases. *Ann Trop Med Parasitol*. 2006;100(5-6):455–464. doi:10.1179/136485906X97453.
- Brundisini F, Giacomini M, DeJean D, Vanstone M, Winsor S, Smith A. Chronic disease patients' experiences with accessing health care in rural and remote areas: a systematic review and qualitative meta-synthesis. *Ont Health Technol Assess Ser*. 2013;13(15):1–33.
- Miranda JJ, Bernabe-Ortiz A, Smeeth L, Gilman RH, Checkley W; CRONICAS Cohort Study Group. Addressing geographical variation in the progression of non-communicable diseases in Peru: the CRONICAS cohort study protocol. *BMJ Open*. 2012;2(1):e000610. doi:10.1136/bmjopen-2011-000610.
- Sharifi-Rad M, Anil Kumar NV, Zucca P, Varoni EM, Dini L, Panzarini E, et al. Lifestyle, oxidative stress, and antioxidants: back and forth in the pathophysiology of chronic diseases. *Front Physiol*. 2020;11:694. doi:10.3389/fphys.2020.00694.
- Gropper SS. The role of nutrition in chronic disease. *Nutrients*. 2023;15(3):664. doi:10.3390/nu15030664.
- Adamski M, Truby H, Klassen KM, Cowan S, Gibson S. Using the Internet: nutrition information-seeking behaviours of lay people enrolled in a massive online nutrition course. *Nutrients*. 2020;12(3):750. doi:10.3390/nu12030750.
- World Health Organization. Improving health literacy [Internet]. Geneva: World Health Organization; 2025 [cited 2025 Oct 16]. Available from: <https://www.who.int/activities/improving-health-literacy>
- Miller TA. Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis. *Patient Educ Couns*. 2016;99(7):1079–1086. doi:10.1016/j.pec.2016.01.020.
- Hyvert S, Yailian AL, Haesebaert J, Vignot E, Chapurlat R, Dussart C, et al. Association between health literacy and medication adherence in chronic diseases: a recent systematic review. *Int J Clin Pharm*. 2023;45(1):38–51. doi:10.1007/s11096-022-01470-z.
- Wang Y, Wen Y, Wu X, Wang L, Cai H. Assessing the role of adaptive digital platforms in personalized nutrition and chronic disease management. *World J Innov Med Technol*. 2025;8(1):1–5. doi:10.53469/wjimt.2025.08(01).05.
- Gnagnarella P, Ferro Y, Monge T, Troiano E, Montalcini T, Pujia A, et al. Telenutrition: changes in professional practice and in the nutritional assessments of Italian dietitian nutritionists in the COVID-19 era. *Nutrients*. 2022;14(7):1359. doi:10.3390/nu14071359.
- Theodore Armand TP, Nfor KA, Kim JI, Kim HC. Applications of artificial intelligence, machine learning, and deep learning in nutrition: a systematic review. *Nutrients*. 2024;16(7):1073. doi:10.3390/nu16071073.
- OpenAI. Trust Center: security, privacy and compliance certifications [Internet]. San Francisco: OpenAI; 2025 [cited 2025 Oct 16]. Available from: <https://trust.openai.com/>
- OpenAI. Data controls in the OpenAI platform (Your data) [Internet]. San Francisco: OpenAI; 2025 [cited 2025 Oct 16]. Available from: <https://platform.openai.com/docs/guides/your-data>
- WhatsApp LLC. Privacy policy [Internet]. Menlo Park: WhatsApp LLC; 2025 [cited 2025 Oct 16]. Available from: <https://www.whatsapp.com/legal/privacy-policy>
- Open Whisper Systems. WhatsApp's Signal Protocol integration is complete [Internet]. 2016 Apr 5 [cited 2025 Oct 16]. Available from: <https://signal.org/blog/whatsapp-complete/>
- Supabase Inc. Security at Supabase [Internet]. San Francisco: Supabase Inc.; 2025 [cited 2025 Oct 16]. Available from: <https://supabase.com/security>
- Supabase Inc. Privacy policy [Internet]. San Francisco: Supabase Inc.; 2025 [cited 2025 Oct 16]. Available from: <https://supabase.com/privacy>
- Supabase Inc. Data processing addendum (DPA) [Internet]. 2025 Aug 5 [cited 2025 Oct 16]. Available from: <https://supabase.com/downloads/docs/Supabase%2BDPA%2B250805.pdf>
- Zhao W, Wang K, Shen D. A survey of large language models. *arXiv [Preprint]*. 2023. arXiv:2303.18223.
- OpenAI. GPT-4o mini: advancing cost-efficient intelligence [Internet]. San Francisco: OpenAI; 2024 [cited 2025 Sep 16]. Available from: <https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/>
- n8n GmbH. n8n: workflow automation platform [Internet]. Berlin: n8n GmbH; 2024 [cited 2025 Sep 16]. Available from: <https://n8n.io/>
- Meta Platforms Inc. WhatsApp Business Platform: developer documentation [Internet]. Menlo Park: Meta Platforms Inc.; 2024 [cited 2025 Sep 16]. Available from: <https://developers.facebook.com/docs/whatsapp/>