



Título: Traducción y comentarios sobre el artículo “ChatGPT accurately performs genetic counseling for gynecologic cancers”

(ChatGPT realiza con precisión asesoramiento genético para cánceres ginecológicos)

Nombre del revisor: Tamara Pradillo Aramendi. Hospital Universitario Infanta Elena.

1. - Artículo Original:

Patel JM, Hermann CE, Growdon WB, Aviki E, Stasenko M. ChatGPT accurately performs genetic counseling for gynecologic cancers. *Gynecol Oncol.* 2024 Apr 26;183:115-119. doi: 10.1016/j.ygyno.2024.04.006. Epub ahead of print. PMID: 38676973.

2.- Resumen del Artículo:

2.1 Introducción:

Los sistemas de inteligencia artificial (IA) como ChatGPT pueden realizar exámenes médicos y asesorar a los pacientes sobre diagnósticos médicos. Nuestro objetivo es cuantificar la precisión de ChatGPT V3.4 al responder

preguntas comunes relacionadas con las pruebas genéticas y el asesoramiento para cánceres ginecológicos.

2.2 Metodología

Se formuló una lista de 40 preguntas sobre pruebas genéticas y asesoramiento en lo que respecta a la oncología ginecológica en conjunto con oncólogos ginecológicos y se adaptó de los sitios web de sociedades profesionales. Esto incluyó el sitio web del Colegio Americano de Obstetricia y Ginecología, CancerCare. Organización y Cancer.Net. Las preguntas se formularon en un nivel de lectura de octavo grado en primera persona para emular a un paciente promedio que hace preguntas como lo haría en un entorno clínico.

Se utilizó la versión 3.5 de ChatGPT, la versión disponible al público. Las preguntas se dividieron en dos categorías: directrices de asesoramiento genético y preguntas relacionadas con trastornos genéticos específicos. Las respuestas fueron calificadas por dos oncólogos ginecológicos según la siguiente escala: 1) correcta y completa, 2) correcta pero no completa, 3) parcialmente correcta, parcialmente incorrecta y 4) completamente incorrecta. Las discrepancias en la calificación se resolvieron con un tercer revisor. Se calculó la proporción de respuestas que obtuvieron cada puntuación en general y dentro de cada categoría de preguntas.

2.3 Resultados

De 40 preguntas, la mitad ($n = 20$) de las preguntas se clasificaron como asesoramiento genético general y la otra mitad ($n = 20$) de las preguntas se referían al síndrome de Lynch y HBOC.

En general, ChatGPT proporcionó respuestas correctas y completas al 82,5 % ($n = 33/40$) de las preguntas, respuestas correctas pero no completas al 15% ($n = 6/40$) de las preguntas, respuestas parcialmente incorrectas al 2,5% ($n = 1/40$) de las preguntas y proporcionó 0% ($n = 0/40$) de respuestas incorrectas.

No hubo preguntas que requirieran que más de tres revisores asignaran una puntuación final. Las tasas de concordancia entre los revisores en general fueron del 92,5% ($n = 37/40$).

De las preguntas generales sobre asesoramiento genético, ChatGPT proporcionó respuestas correctas y completas al 100 % ($n = 20$) de las preguntas con una tasa de concordancia del 100 % entre los dos revisores.

En el subconjunto de preguntas que pertenecen a un síndrome específico (Lynch o HBOC), 17 preguntas fueron sobre el síndrome HBOC y 3 preguntas sobre el síndrome de Lynch. De las preguntas sobre el síndrome HBOC, ChatGPT proporcionó respuestas correctas y completas al 64,7% ($n = 11/17$) de preguntas, respuestas correctas pero no exhaustivas al 29,4% ($n = 5/17$) de las preguntas, y respuestas parcialmente incorrectas al 5,8% ($n = 1/17$) de las preguntas. No se informaron respuestas completamente incorrectas. Las tasas de concordancia entre los dos revisores alcanzaron

el 88,2 % (n = 15/17) y solo 2 preguntas requirieron un tercer revisor para su puntuación.

En preguntas sobre el síndrome de Lynch, ChatGPT proporcionó respuestas correctas y completas al 66,6 % (n = 2/3) de las preguntas y respuestas correctas pero no completas al 33,3 % (n = 1/3) de las preguntas. Las tasas de concordancia entre los dos revisores alcanzaron el 66 % y solo se envió una pregunta al tercer revisor para su puntuación.

3.- Comentario:

Nuestro estudio indica que ChatGPT puede brindar asesoramiento genético con precisión y responder preguntas sobre pruebas genéticas generales, pero es menos preciso en preguntas relacionadas con pruebas genéticas específicas de síndromes.

Los resultados son contradictorios respecto a los de otros estudios que analizan las respuestas de ChatGPT sobre otros temas médicos.

Aun así, se están llevando a cabo varios estudios que invierten sus aplicaciones en el análisis de imágenes, la selección de pacientes para ensayos clínicos y la ayuda en el diagnóstico médico. Nuestros hallazgos indican que ChatGPT es preciso y consistente al brindar asesoramiento general sobre pruebas genéticas, lo que indica que puede convertirse en una herramienta poderosa para ayudar a los médicos y asesores genéticos, especialmente a medida que el modelo continúa mejorando.

Los estudios futuros necesitarían evaluar la seguridad y precisión de diferentes modelos de ChatGPT y cómo las aportaciones de los médicos podrían mejorar la precisión de ChatGPT y su utilidad para los pacientes. A medida que evoluciona nuestra comprensión de la IA, debemos continuar evaluando cómo se puede aplicar para ayudar a mejorar la atención al paciente. Concluimos que ChatGPT puede proporcionar asesoramiento genético con precisión, pero se necesitan más estudios antes de que pueda recomendarse como recurso para los pacientes.