O3.32. ¿Podría Chat-GPT sustituir a los expertos en la interpretación de las cistouretrografías miccionales seriadas?

Mireia Batalla Vilacís[1], Mar Carbonell Pradas[1], Oriol Martín Solé[1], Sonia Pérez Bertólez[1], Jesús González Cayón[1], Xavier Tarrado Castellarnau[1], Luis Garcia Aparicio[1].
1. Hospital Sant Joan de Déu

Objetivos

Evaluar la fiabilidad de ChatGPT-5.1 en la interpretación de imágenes de cistouretrografía miccional seriada (CUMS), analizando la concordancia entre urólogos y ChatGPT.

Material y método

Se realizó un estudio retrospectivo de corte transversal. Dos urólogos pediátricos revisaron de forma independiente cuatro imágenes anonimizadas de CUMS (llenado vesical, oblicua derecha, oblicua izquierda y posmiccional) de 65 pacientes; de mediana de edad 8 meses (RIC: 4-28) y 52,3% varones. Se estudió la presencia de reflujo vesicoureteral (RVU), lateralidad, grado (I–V) y presencia de divertículos paraureterales. Las mismas imágenes fueron interpretadas con indicaciones idénticas en dos cuentas de ChatGPT-5.1. La concordancia interobservador se analizó mediante el coeficiente kappa de Cohen.

Resultados

La concordancia entre urólogos fue excelente para la detección y lateralidad del RVU (κ=0,98) y buena para la gradación (κ=0,76). La concordancia en la detección de divertículos paraureterales fue buena en el lado izquierdo (κ=0,73) y baja en el derecho (κ=0,30). La concordancia entre urólogos y ChatGPT-5.1 fue baja para la presencia y lateralidad del RVU (κ=0,27–0,38) y deficiente para la gradación (κ=0,09–0,20). La concordancia entre las dos cuentas de ChatGPT-5.1 fue moderada para la detección del RVU (κ=0,44) y deficiente para su gradación (κ=0,33–0,36). ChatGPT-5.1 no identificó ningún divertículo paraureteral.

Conclusiones

ChatGPT-5.1 mostró una consistencia limitada y una baja concordancia con los urólogos, especialmente en la gradación del RVU y en hallazgos morfológicos sutiles. Los modelos actuales generales de inteligencia artificial carecen de validación adecuada para la toma de decisiones clínicas en imagen urológica pediátrica.