



85 - IMPACTO DE LA INTELIGENCIA ARTIFICIAL EN EL APRENDIZAJE DE LAS CARACTERÍSTICAS ECOGRÁFICAS DE SOSPECHA DEL NÓDULO TIROIDEO (PÓSTER SELECCIONADO)

G. Díaz Soto^{1,2}, P. Fernández Velasco^{1,2}, B. Torres Torres^{1,2}, P. Pérez López^{1,2}, E. Delgado^{1,2} y D. de Luis^{1,2}

¹Servicio de Endocrinología y Nutrición, Hospital Clínico Universitario de Valladolid. ²Centro de Investigación en Endocrinología y Nutrición Clínica (CIENC), Universidad de Valladolid.

Resumen

Objetivos: Analizar el impacto de un programa basado en inteligencia artificial (IA) en el aprendizaje del endocrinólogo (E) de las características ecográficas sospechosas (CES) del nódulo tiroideo

Métodos: Estudio prospectivo de la imagen por ultrasonidos (US) de los 100 primeros nódulos analizados con IA. Se evaluó el aprendizaje en la caracterización de la US expresado como el grado de acuerdo obtenido por seis E sin y con experiencia previa. Se analizaron los Índices Kappa de Fleiss (?) obtenidos para cada CES del ACR-TIRADS de las primeras US con respecto de las últimas 20 evaluadas.

Resultados: Se evaluaron 100 pacientes (84% mujeres, edad media $55,8 \pm 18,5$ años) clasificados como ACR-TIRADS 3,4 y 5 el 43%, 28% y 16% de los nódulos analizados. El diámetro máximo nodular fue de $2,9 \pm 1,1$ cm, siendo el 7,8% malignos. Globalmente, las CES presentaron un grado elevado de concordancia en forma, composición, ecogenicidad y categoría ACR-TIRADS ($\kappa = 0,760$, $\kappa = 0,531$, $\kappa = 0,460$, $\kappa = 0,471$; $p < 0,001$, respectivamente), siendo aceptable en la evaluación de focos ecogénicos ($\kappa = 0,322$, $p < 0,001$) y pobre en márgenes nodulares ($\kappa = 0,187$, $p < 0,001$). Al comparar el aprendizaje en las primeras y últimas 20 US analizadas, se observó un incremento del grado de acuerdo en forma del 27,2% ($\kappa = 0,581 \pm 0,160$ vs. $0,739 \pm 0,127$), ecogenicidad 29,8% ($\kappa = 0,386 \pm 0,079$ vs. $0,501 \pm 0,070$) y categoría ACR-TIRADS 24,6% ($\kappa = 0,398 \pm 0,059$ vs. $0,496 \pm 0,057$), manteniéndose estable en el resto CES. La correlación entre la puntuación media ACR-TIRADS de los E previa y tras el uso de la IA mostró un incremento significativo del grado de acuerdo: $R = 0,741$ vs. $R = 0,981$, $p < 0,001$.

Conclusiones: La utilización de la IA se relacionó con una mejora en la capacidad de interpretación de la US con o sin experiencia previa. La ecogenicidad, forma y categoría ACR-TIRADS fueron las CES que mostraron un mayor incremento. Los sistemas de IA redujeron la variabilidad interobservador y aumentaron la reproducibilidad de las escalas de estratificación del riesgo de la US.