

# Precisión diagnóstica del *software* de cuantificación automática en pacientes con sospecha de COVID-19 del Instituto Nacional de Cardiología Ignacio Chávez

## Diagnostic accuracy of the automatic quantification software in patients with suspected COVID-19 of the Instituto Nacional de Cardiology Ignacio Chávez

Estefanía Gallego-Díaz<sup>1,2</sup>, César N. Cristancho-Rojas<sup>3</sup> y Sergio A. Criales-Vera<sup>4,5\*</sup>

<sup>1</sup>Programa de Imagenología Diagnóstica y Terapéutica, Departamento de Radiología, Instituto Nacional de Cardiología Ignacio Chávez; <sup>2</sup>Grupo CT Scanner, Universidad Nacional Autónoma de México; <sup>3</sup>Departamento de Radiología, Salud Digna; <sup>4</sup>Departamento de Radiología e Imagen, Instituto Nacional de Cardiología Ignacio Chávez; <sup>5</sup>Servicio de Tomografía y PET, Grupo CT Scanner. Ciudad de México, México

### Resumen

**Antecedentes y Objetivos:** Establecer la precisión diagnóstica por tomografía computarizada (TC) de la probabilidad de neumopatía por enfermedad por coronavirus 2019 (COVID-19), dada por el sistema de inteligencia artificial (IA) diseñado por Siemens, y el resultado de la evaluación cualitativa CO-RADS (COVID-19 Reporting and Data System) con el estándar de referencia reacción en cadena de la polimerasa transcriptasa inversa (RT-PCR), entregando así la experiencia de nuestra institución. **Métodos:** Se realizó un estudio observacional, comparativo y retrolectivo en 192 pacientes adultos con sospecha de infección por coronavirus 2 del síndrome respiratorio agudo grave (SARS-CoV-2) que contaban con prueba PCR. Se obtuvo la información de precisión diagnóstica luego de comparar el estándar de referencia (RT-PCR) con el CO-RADS realizado por los observadores y la probabilidad de COVID-19 que arrojaron las imágenes de TC mediante la IA. **Resultados:** La comparación de la probabilidad de COVID-19 obtenida por la IA vs. la RT-PCR para SARS-CoV-2 generó un AUC ROC de 0.774 (IC: 0.69-0.81) con  $p = 0.0001$ . La probabilidad de COVID-19 tuvo una precisión aceptable, con un buen valor predictivo positivo del 87.80%, pero con un pobre valor predictivo negativo del 58.80%. La variable CO-RADS vs. PCR obtuvo una mayor precisión con valores de sensibilidad y especificidad del 91.80 y 88.7% respectivamente. **Conclusión:** La comparación entre los resultados obtenidos por la IA y por la variable CO-RADS mostró mayor efectividad en esta última, sin embargo se logró documentar el alto impacto que tiene el sistema de cuantificación automática en la evaluación de estos pacientes, ya que permite agilizar la valoración del radiólogo y funciona como complemento en casos de dudas diagnósticas.

**Palabras clave:** Inteligencia artificial. Cuantificación automática. COVID-19. CO-RADS. RT-PCR.

### Abstract

**Background and Objectives:** To establish the diagnostic accuracy of the computed tomography (CT) comparing the probability of COVID-19 pneumopathy, obtained through artificial intelligence (AI) system designed by Siemens Healthineers, and the qualitative evaluation CO-RADS (COVID-19 Reporting and Data System) with the reference standard (RT-PCR), and thus providing the experience of our institution. **Methods:** An observational, comparative and retrolective study was performed on 280 adult patients with suspected SARS-CoV2 infection, 192 of whom had PCR testing. Diagnostic accuracy information was

**\*Correspondencia:**

Sergio A. Criales-Vera  
E-mail: scriales@gmail.com

Fecha de recepción: 10-05-2022

Fecha de aceptación: 13-07-2022

DOI: 10.24875/ACM.220001481

Disponible en internet: 25-10-2022

Arch Cardiol Mex. 2023;93(Supl 6):94-101

www.archivoscardiologia.com

1405-9940 / © 2022 Instituto Nacional de Cardiología Ignacio Chávez. Publicado por Permanyer. Este es un artículo *open access* bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

obtained after comparing the reference standard (RT-PCR) with the CO-RADS performed by observers and the probability of COVID-19 yielded by CT images through AI software. **Results:** The comparison of COVID-19 probability acquired by AI vs. SARS CoV-2 RT-PCR generated an AUC ROC 0.774 (CI 0.69-0.81) with  $p = 0.0001$ . The COVID-19 probability had an acceptable accuracy, with a good PPV 87.80%, but with a poor NPV of 58.80%. The CO-RADS vs. RCP variable got a higher accuracy with much higher sensitivity and specificity values, reaching 91.80% and 88.7% respectively. **Conclusion:** The comparison between the results obtained by the AI and those referring to the CO-RADS variable showed greater effectiveness in the latter for patients with suspected COVID-19 however, it was possible to document the high impact of the automatic quantification system in the evaluation of these patients since it allows speeding up the radiologist's assessment and works as a complement in cases of diagnostic doubts.

**Keywords:** Artificial intelligence. Automatic quantification. COVID-19. CO-RADS. RT-PCR.

## Introducción

El estándar de referencia para el diagnóstico de enfermedad por coronavirus 2019 (COVID-19) es la reacción en cadena de la polimerasa transcriptasa inversa (RT-PCR), la cual detecta el virus mediante muestras de hisopado nasofaríngeo, esputo o lavado bronquial<sup>1,2</sup>. Sin embargo, debido a la alta demanda de pruebas de detección, se han desarrollado otros métodos de laboratorio que se conocen como pruebas de antígeno y anticuerpos, caracterizadas por sus resultados a corto plazo, como son los anticuerpos inmunoglobulina (Ig) M e IgG, entre otros que son igualmente utilizados en el análisis de los casos (IgA, IgE, IgD), los cuales cuentan con una menor sensibilidad y especificidad<sup>2,3</sup>.

El método de imagen utilizado para la detección y estadificación de la neumopatía secundaria a COVID-19 es la tomografía computarizada (TC), que aunque no se recomienda para detección sistemática, ha demostrado ser una herramienta rápida complementaria en el diagnóstico, especialmente en aquellos lugares donde las pruebas no están ampliamente disponibles<sup>4-6</sup>. La TC posee una alta sensibilidad (98%) basándose en el sistema de puntuación CO-RADS (*COVID-19 Reporting and Data System*), que proporciona un nivel de sospecha etiológica<sup>7,8</sup>; así como datos pronósticos de la enfermedad, determinados con base en características y medición del compromiso pulmonar, aparte de la valoración vascular en pacientes con riesgo de tromboembolia pulmonar agregada<sup>9,10</sup>.

Ante la saturación de los servicios de radiología, herramientas auxiliares como la inteligencia artificial (IA) han adquirido relevancia con *software packages* como el CT Pneumonia Analysis (Siemens Healthineers), el cual se define como un sistema de procesamiento de imágenes que proporciona un análisis cualitativo y cuantitativo de imágenes DICOM provenientes de TC<sup>11-13</sup>. Este sistema se utiliza para asistir a radiólogos y médicos a la hora de evaluar y valorar patologías

pulmonares, lo que hace de él una herramienta útil de apoyo al diagnóstico y medición de los hallazgos tomográficos pulmonares<sup>14</sup>, sobre todo en una enfermedad como la COVID-19, que tiene patrones muy similares a otras patologías, agilizando así el análisis y la obtención de resultados<sup>15-17</sup>. Las ventajas del *software* en cuanto a servicios con prevalencia elevada de esta enfermedad y alta carga de trabajo son: mayor objetividad, cuantificación tomográfica de la neumopatía, reducción significativa del tiempo de interpretación, mejor abordaje en la supervisión con índices pronósticos y el seguimiento con impacto en el manejo médico.

La IA representa una herramienta prometedora en el campo de la imagen clínica<sup>18</sup>, sin embargo se debe tener presente que al ser un programa computacional que se codifica con algoritmos, tiene errores de precisión diagnóstica. Esto se da principalmente debido a patologías de base del paciente que pueden terminar confundiendo al *software*, a artefactos de respiración, o simplemente a errores aleatorios del sistema.

El Instituto Nacional de Cardiología Ignacio Chávez ha enfrentado el compromiso de dar atención a pacientes con COVID-19, contando con un manejo multidisciplinario en el que la imagen médica ha sido clave; aun así, no hay datos fiables que demuestren el valor diagnóstico de las herramientas humanas y de *software* aplicables a nuestra experiencia local en comparación con la evaluación por el radiólogo. Realizamos un estudio de prueba diagnóstica para la evaluación de la probabilidad de neumopatía por COVID-19 obtenida mediante el *software* de IA, con el resultado de la evaluación cualitativa CO-RADS, comparando esto con el estándar de referencia RT-PCR para coronavirus 2 del síndrome respiratorio agudo grave (SARS-CoV-2), y entregando así la experiencia de nuestra institución; al implementarse de esta manera, además de documentar la certeza diagnóstica permite evaluar la reproducibilidad de la IA documentando sus potenciales factores que contribuyen a errores en el *software*, así como sus ventajas y desventajas.

## Materiales y métodos

Se realizó un estudio observacional, comparativo y retroactivo en pacientes adultos con sospecha de infección por SARS-CoV-2 ingresados en el Instituto Nacional de Cardiología Ignacio Chávez durante el periodo comprendido entre diciembre de 2020 y enero del 2021, tiempo durante el cual el instituto se encontraba en fase COVID (centro híbrido), donde su principal objetivo era la atención de pacientes con diagnóstico o sospecha de esta infección. A estos sujetos se les realizó TC de tórax como parte de su evaluación clínica. La privacidad de los pacientes fue asegurada mediante la anonimización de la información, teniendo como base los principios de la declaración de Helsinki para el tratamiento de los datos.

## Población de estudio

Se incluyeron pacientes mayores de 18 años, registrados de forma consecutiva en nuestro sistema PACS-RIS, a quienes se les realizó una TC de tórax por sospecha clínica de COVID-19. Se tuvieron en cuenta datos de la historia clínica electrónica, incluyendo los registros de RT-PCR para SARS-CoV-2 para cada uno de los participantes, y se excluyó a aquellos a quienes no se les realizó esta prueba. Se registraron variables demográficas como la edad y el sexo, además de variables clínicas como comorbilidades y diagnósticos diferenciales.

## Protocolo de TC de tórax

La TC de tórax se realizó en fase inspiratoria por medio de un tomógrafo Somatom Definition Flash de 256 cortes Somatom Definition Flash (Siemens, Alemania). Se emplearon en todas las adquisiciones CARE Dose 4D, pitch mayor de 0.8, colimación adecuada para tórax (FoV aproximado de 350 mm) y tiempo de rotación de 0.28 s. Además, el kilovoltaje (Kv) con límite máximo de 120 de acuerdo con corrección de atenuación en el eje Z, los miliamperios por segundo (mAs) de 120 a 150 y dosis de radiación desde 120 a 250 mGy\*cm, con dosis promedio de 2 milisieverts (mSv). Las reconstrucciones se realizaron con corte de 1 mm utilizando un Kernel B50f (de alta frecuencia espacial), según las recomendaciones dadas por el proveedor del *software*.

## Análisis de imágenes y resultados del *software IA*

Las reconstrucciones fueron visualizadas y analizadas en una estación dedicada (SYNGO VIA versión

VA48A), mediante el sistema de IA (CT Pneumonia Analysis versión 2.0. Siemens Healthineers, Erlangen, Alemania). Este es un *software* que toma como entrada una TC de tórax simple e identifica y fracciona en 3D los pulmones y los lóbulos antes de segmentar las anomalías. Posteriormente, genera dos medidas combinadas de la gravedad de la afectación de pulmón/lóbulo, cuantificando la extensión de la presencia de opacidades altas (vidrio despulido y consolidación), de las que ya se ha demostrado su relación con síntomas graves y peor pronóstico (asistencia ventilatoria y mortalidad). La primera medida de gravedad de la enfermedad es global, mientras que la segunda se refiere específicamente a lóbulos<sup>6-12</sup>. El algoritmo de IA fragmenta de forma automática ambos pulmones y sus lóbulos, y muestra los contornos resultantes como una superposición gráfica. Con base en la atenuación de los vóxeles y de acuerdo con umbrales preestablecidos, este cuenta volúmenes y porcentajes entregando el cálculo de las zonas de ocupación acinar, las cuales están representadas por aquellos con una densidad mayor de  $-200$  U (*high average volumen*), mientras que el parénquima pulmonar normal tiene densidades de  $-750$  a  $-950$  UH<sup>12</sup>, un grosor de segmentación de 2 mm con un nivel de ventana de  $-600$  y un ancho de ventana de 1,200. Es importante resaltar que los valores generales de la probabilidad de COVID-19, la puntuación de opacidad total y los valores de porcentaje de opacidad que se muestran junto al gráfico corresponden a la totalidad del pulmón<sup>11</sup> (Fig. 1). Las tres variables analizadas en el presente estudio, el cual fue construido con base en cohortes de comprobación, son:

- Probabilidad de COVID-19: esta puntuación varía entre 0 y 1, y hace referencia a la probabilidad de que los hallazgos de la IA sean causados por COVID-19.
- Puntaje de opacidad total: su rango va de 0 a 20, representa la suma de los puntajes de opacidad de los cinco lóbulos pulmonares.
- Porcentaje de opacidad: este valor refiere el porcentaje de opacidad de todo el pulmón. El análisis tardó en promedio un minuto, y se registraron las tres variables previamente mencionadas.

Posteriormente se extrajeron los resultados de CO-RADS de las interpretaciones de radiólogos expertos en tomografía de tórax con más de cinco años de experiencia, quienes se basaron en la guía descrita por Prokop et al.<sup>7</sup>, con las siguientes definiciones:

- CO-RADS 0: cuando el resultado no encaja dentro de ninguna de las cinco categorías, debido a que las exploraciones son incompletas o de calidad insuficiente.



**Figura 1.** Ejemplo de los resultados arrojados por la inteligencia artificial. **A:** lista de los resultados de las mediciones que se muestran para cada lóbulo y pulmones. Los valores encasillados en las líneas blancas representan las constantes principales valoradas en este estudio. **B:** resultados de la segmentación y de la medición de la densidad pulmonar. **C:** imagen en volumen *rendering* de los pulmones y lesiones que se definieron en la configuración.

- CO-RADS 1: significa un nivel muy bajo de sospecha de afectación pulmonar por COVID-19, ya que los resultados de la TC son normales, o sus hallazgos son de origen inequívoco no infeccioso.
- CO-RADS 2: hace referencia a un bajo nivel de sospecha de afectación pulmonar por COVID-19 con base en hallazgos de TC en los pulmones típicos de origen infeccioso no COVID-19.
- CO-RADS 3: representa hallazgos indeterminados para la afectación pulmonar por COVID-19.
- CO-RADS 4: denota un alto nivel de sospecha de afectación pulmonar por COVID-19, determinando que los hallazgos de la TC son típicos de COVID-19. Sin embargo, también muestran cierta superposición con otras neumonías (virales).
- CO-RADS 5: nivel muy alto de sospecha de afectación pulmonar por COVID-19 por hallazgos típicos de TC.
- CO-RADS 6: indica resultado conocido de RT-PCR para SARS CoV-2 positivo. El índice de severidad por tomografía utilizado es el puntaje de severidad total (TSS, por sus siglas en inglés). En este los hallazgos encontrados se dividen por lóbulos y no por segmentos pulmonares, es un método visual cuantitativo que

consiste en puntuar los porcentajes de cada uno de los cinco lóbulos implicados (lóbulo superior derecho, lóbulo medio, lóbulo inferior derecho, lóbulo superior izquierdo y lóbulo inferior izquierdo): 1 punto, < 5% de afectación; 2 puntos, 5-25% de afectación; 3 puntos, 26-49% de afectación; 4 puntos, 50-75% de afectación, y 5 puntos, > 75% de afectación. La puntuación total de la TC es la suma de las puntuaciones lobares individuales, y puede oscilar entre 0 (sin afectación) y 25 (afectación máxima), la cual se da cuando los cinco lóbulos muestran más del 75% de afectación.

### RT-PCR para SARS-CoV-2

Finalmente, se cotejaron y se registraron en la base de datos los valores de IA, los reportes tomográficos realizados por los radiólogos y los resultados de RT-PCR para SARS-CoV-2. En nuestra institución se realiza la prueba RT-PCR para SARS-CoV-2 por medio de exudado del tracto respiratorio superior (nasofaringe), y consta de dos procesos automatizados: la extracción de ácidos nucleicos y la reacción de amplificación. El

**Tabla 1.** Variables calculadas de la edad y las otorgadas por la inteligencia artificial junto con la frecuencia en porcentajes de las variables diagnósticos diferenciales con respecto al total de la muestra analizada, la cual fue de un total de 192 pacientes distribuidos de la siguiente manera:

CO-RADS en categoría 1 alcanza 42 pacientes

CO-RADS en categoría 2 alcanza 2 pacientes

CO-RADS en categoría 3 alcanza 23 pacientes

CO-RADS en categoría 4 alcanza 0 pacientes

CO-RADS en categoría 5 alcanza 53 pacientes

CO-RADS en categoría 6 alcanza 72 pacientes.

CO-RADS 6 fue eliminado del análisis de resultados en la comparación CO-RADS vs. RT-PCR.

	Mediana (rango intercuartílico)	Rango mínimo- máximo	Diagnósticos diferenciales	Frecuencia n (porcentaje)
Edad	(38-62)	1-97	Sobrecarga hídrica/edema	20 (10.4%)
Probabilidad de COVID-19	1 (0.25-1)	0-1	Hipertensión pulmonar	8 (4.1%)
Puntaje de opacidad total	5 (1-10)	0-20	Neumopatía intersticial/enfisema	10 (5.2%)
Porcentaje de opacidad total	12 (1 –33)	0-88	Infeción por otros agentes	2 (1.04%)
			Infarto pulmonar	4 (2.08%)
			Total	44 (22.9%)
			Total de pacientes	192 (100%)

CO-RADS: COVID-19 Reporting and Data System; RT-PCR: reacción en cadena de la polimerasa transcriptasa inversa.

equipo que se utilizó es un Elite Ingenius serie 500. La tecnología es ELISA (ensayo por inmunoabsorción ligado a enzimas) y la determinación es del genoma viral SARS-CoV-2. El distribuidor de reactivos es Abalat y se encuentra autorizado por el InDRE (Instituto de Diagnóstico y Referencia Epidemiológico).

## Análisis estadístico

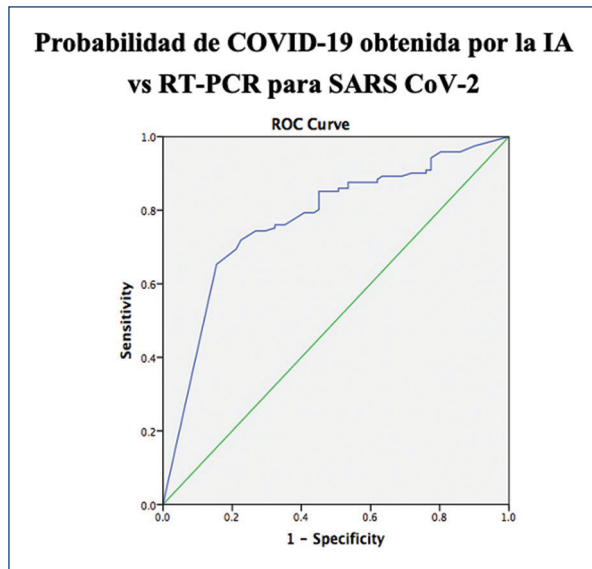
Se computaron resultados estadísticos de tendencia central y de dispersión para las variables continuas, mientras que las variables categóricas se expresaron como porcentajes y frecuencias. Se tomó como significancia estadística una  $p < 0.05$  (resultado clínicamente importante). Se obtuvo la información de precisión diagnóstica luego de comparar el estándar de referencia (RT-PCR) con el CO-RADS realizado por los observadores y la probabilidad de COVID-19 que arrojaron las imágenes de TC por medio del *software* de IA. Todo esto mediante un análisis basado en ROC (*Receiver-Operating-Characteristic*), con su respectiva área bajo la curva (AUC, por sus siglas en inglés) e intervalos de confianza del 95% (IC 95%). Los umbrales óptimos se obtuvieron utilizando el índice de Youden para describir la sensibilidad, la especificidad y los valores predictivos. El análisis se realizó en el *software* STATA SE versión 14.1.

## Resultados

Se consideraron en primera instancia 280 sujetos potenciales con estudios de TC almacenados en el PACS realizados por sospecha de SARS-CoV-2 en el periodo ya expuesto, de los cuales se analizaron 192 pacientes elegibles que contaban prueba RT-PCR para SARS-CoV-2 (107 hombres y 85 mujeres). En la comparación CO-RADS vs. RT-PCR para SARS-CoV-2 se omitieron 72 pacientes por presentar CO-RADS 6, ya que esto significaba prueba conocida de RT-PCR positivo al momento de la evaluación de las imágenes de TC por parte de los radiólogos que determinaron el CO-RADS. La edad media de los 192 pacientes es de 50.39 años y la mediana de 49 años ( $\pm 15.61$ ), con un mínimo de 18 años de edad y un máximo de 97 años (Tabla 1).

### Precisión diagnóstica de las variables tomográficas en relación con la RT-PCR para SARS-CoV-2

Se realizó la curva ROC entre la probabilidad de COVID-19 obtenida por la IA vs. la RT-PCR para SARS-CoV-2, encontrando un AUC de 0.774 (IC 95%: 0.69-0.81) con  $p = 0.0001$ , con sensibilidad del 65.3%, especificidad del 84.50%, área ROC de 0.75%, razón de verosimilitud positiva de 4.21 y negativa de 0.41,



**Figura 2.** Gráfico de la curva que muestra la precisión diagnóstica de la probabilidad de COVID-19 determinada por el *software* vs. la RT-PCR para SARS-CoV-2 (línea azul: probabilidad de COVID-19 obtenida por IA). COVID-19: enfermedad por coronavirus 2019; IA: inteligencia artificial; RT-PCR: reacción en cadena de la polimerasa transcriptasa inversa; SARS-CoV-2: coronavirus 2 del síndrome respiratorio agudo grave.

razón de momios diagnóstica de 10.26, valor predictivo negativo (VPN) del 58.8% y valor predictivo positivo (VPP) del 87.8% (Fig. 2). Gracias al índice de Youden se halló el punto de corte, siendo una probabilidad de 0.975 (IC 95%: 0.63-0.80) con sensibilidad y especificidad del 72 y 77%, respectivamente. Se dicotomizó esta variable respecto a dicho valor ( $> 0.97$ ), obteniendo una sensibilidad, especificidad, VPP y VPN del 67.3, 76.10, 66 y 77.1%, respectivamente (Tabla 2).

Se realizó la curva ROC entre CO-RADS vs. RT-PCR para SARS-CoV-2, obteniendo un AUC de 0.9 (IC 95%: 0.85-0.96), con sensibilidad del 91.80%, especificidad del 88.7%, VPP del 84.9 % y VPN del 94% (Tabla 3).

## Discusión

Esta investigación se enfocó en evaluar la precisión diagnóstica de la variable de probabilidad de COVID-19 brindada por el *software* de IA diseñado por Siemens y la valoración hecha por el radiólogo (CO-RADS), comparadas con la RT-PCR en una institución de alta complejidad con población mexicana. La probabilidad

**Tabla 2.** Dicotomización de la probabilidad de COVID-19 en  $> 0.97$

Parámetro	$> 0.97$ 95% - Intervalo de Confianza
Prevalencia	40.80% (32.0-50.2%)
Sensibilidad	67.30% (52.5-80.1%)
Especificidad	76.10% (64.5-85.4%)
Área ROC	0.72 (0.63-0.80)
Razón de verosimilitud (+)	2.81 (1.78-4.45)
Radio de verosimilitud (-)	0.43 (0.28-0.66)
Razón de momios diagnóstica	6.55 (2.94-14.62)
Valor predictivo positivo	66.00% (51.2-78.8%)
Valor predictivo negativo	77.10% (65.6-86.3%)

ROC: Receiver-Operating-Characteristic; COVID-19: enfermedad por coronavirus 2019.

**Tabla 3.** Variables otorgadas en la comparación CO-RADS vs. RT-PCR para SARS-CoV-2, con un AUC 95%

Parámetro	CO-RADS vs. RT-PCR para SARS CoV-2 95% - Intervalo de Confianza
Prevalencia	40.80% (32.0-50.2%)
Sensibilidad	91.80% (80.4-97.7%)
Especificidad	88.70% (79.0-95.0%)
Área ROC	0.9 (0.85-0.96)
Razón de verosimilitud (+)	8.15 (4.22-15.74)
Radio de verosimilitud (-)	0.09 (0.04-0.24)
Razón de momios diagnóstica	88.59 (25.80-301.20)
Valor predictivo positivo	84.90% (72.4-93.3%)
Valor predictivo negativo	94.00% (85.4-98.3%)

AUC: área bajo la curva; CO-RADS: COVID-19 Reporting and Data System; ROC: Receiver-Operating-Characteristic; RT-PCR: reacción en cadena de la polimerasa transcriptasa inversa; SARS-CoV-2: coronavirus 2 del síndrome respiratorio agudo grave.

de COVID-19 tuvo una precisión aceptable, con un buen VPP (87.80%), pero con un bajo VPN (58.80%). La variable CO-RADS vs. la RT-PCR para SARS CoV-2 obtuvo una mayor precisión con valores de sensibilidad y especificidad mucho más elevados, alcanzando el 91.80 y 88.7%, respectivamente.

La menor precisión del proceso automatizado de análisis se produjo debido a que existen constantes

que pueden provocar el error en la IA, como aquellos escenarios en los que existen otros procesos infecciosos y no infecciosos diferentes del COVID-19 (p. ej., CO-RADS 1, 2 y 3), los cuales en nuestra institución son muy frecuentes, por las enfermedades de base y diagnósticos diferenciales de la población analizada, representadas por sobrecarga hídrica/edema agudo de pulmón, hipertensión pulmonar, neumopatía intersticial/enfisema, infección por otros agentes e infarto pulmonar (Tabla 1). Estas condiciones producen fallas en el análisis del *software*, y en estas circunstancias es que el radiólogo al tener la correlación clínica adecuada y el patrón de los hallazgos, sumado a su experiencia, pueden ser clave en el diagnóstico diferencial puede determinar si la neumopatía proviene o no del COVID-19.

El metaanálisis de Moezzi et al.<sup>19</sup> demostró un rendimiento satisfactorio utilizando el algoritmo de IA para la identificación de COVID-19 por medio de TC asistida por IA frente a muestras sanas. Demostraron que la IA fue precisa en la afectación pulmonar durante el COVID-19, con una sensibilidad combinada de 0.90 (IC 95%: 0.90-0.91), una especificidad de 0.90 (IC 95%: 0.90-0.91) y un AUC de 0.96 (IC 95%: 0.91-0.98), lo que contrasta con nuestro estudio, pues este demostró una sensibilidad y especificidad de 0.65 y 0.84, respectivamente. Al igual que la investigación realizada por Mei<sup>20</sup>, donde indican el potencial valor de precisión diagnóstica de los sistemas de cuantificación tomográfica automática, con un 84.3% de sensibilidad y un 82.8% de especificidad. En comparación con nuestra investigación, su sensibilidad y especificidad fueron mayores, por lo tanto se debe realizar una correlación con la información clínica para tener un diagnóstico más apropiado, y lograr implementar algoritmos multivariados.

La variable de probabilidad COVID-19 otorgada por la IA nos permite salir de dudas frente a pacientes de los que no tengamos la certeza diagnóstica de COVID-19, pero debemos tener claro que el sistema no supera la lectura del radiólogo, sino que la complementa. Hay varias publicaciones sobre el apoyo de la IA con la severidad clínica, encontrándola acertada para el manejo de estos pacientes, y con adecuada comparación respecto a los pacientes de severidad elevada, con resultados cuantitativos reproducibles. En cuanto a la precisión diagnóstica, se ha validado en múltiples ocasiones el *software*, sin embargo actualmente ninguno ofrece la posibilidad de mejora para el sistema determinado por los diagnósticos diferenciales. Georgescu et al.<sup>18</sup> encontraron que si bien hay propiedades

dominantes que se pueden observar en la COVID-19, como la presencia de opacidades en vidrio despolido y distribución periférica y basal, estas características no se observan en todos los casos. Sin embargo, mencionaron que teniendo en cuenta que los resultados de RT-PCR para SARS-CoV-2 pueden tardar varias horas, el sistema se puede utilizar para la clasificación rápida de casos positivos, particularmente en entornos con recursos limitados donde la experiencia radiológica puede no estar disponible de inmediato. Además, demostraron un AUC de 0.90 con una sensibilidad del 86% y especificidad del 81%, siendo mucho mayor la sensibilidad que la adquirida en nuestro trabajo (65.3%), probablemente debido al tipo de pacientes que manejamos. Este sistema podría ayudar al radiólogo a priorizar la interpretación de TC en pacientes con COVID-19 mediante la detección de casos de menor probabilidad<sup>21</sup>.

En la investigación realizada por Machnicki et al.<sup>3</sup> se demostró que la apariencia típica de la TC de tórax tiene una sensibilidad y especificidad del 73.5 y 82.8% en relación con la precisión diagnóstica del radiólogo con respecto a la RT-PCR para SARS-CoV-2, así como la sensibilidad y especificidad de la apariencia indeterminada es del 92 y 41%. Para la IA se vuelve un limitante los pacientes con una morfología atípica, los cuales cada vez son más frecuentes.

Este estudio tiene varias limitaciones: la muestra poblacional es pequeña, la investigación tiene un diseño retrolectivo y un 22.9% de los pacientes tienen enfermedades de base (sobrecarga hídrica/edema agudo de pulmón, hipertensión pulmonar, neumopatía intersticial/enfisema, infección por otros agentes e infarto pulmonar), lo que puede disminuir la especificidad de la IA.

Invitamos a seguir investigando sobre la cuantificación tomográfica automática en pacientes con COVID-19, en especial en aquellos que presentan comorbilidades que pueden ser un desafío para los sistemas de cómputo. En un futuro se requiere implementar algoritmos avanzados que permitan la mejora de la especificidad en el diagnóstico de estos pacientes.

## Conclusión

La comparación entre los resultados obtenidos por el CO-RADS e IA vs. el estándar de referencia en pacientes con sospecha de COVID-19 mostró mayor efectividad diagnóstica del primero. Sin embargo, la IA tiene impacto en la evaluación de estos pacientes, ya que puede apoyar la sospecha de infección viral por su alto

VPP, además de que permite agilizar la valoración del radiólogo.

## Financiamiento

Ninguno.

## Conflicto de intereses

Ninguno.

## Responsabilidades éticas

**Protección de personas y animales.** Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

**Confidencialidad de los datos.** Los autores declaran que en este artículo no aparecen datos de pacientes.

**Derecho a la privacidad y consentimiento informado.** Los autores declaran que en este artículo no aparecen datos de pacientes.

## Bibliografía

- Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507-13.
- Machnicki S, Patel D, Singh A, Talwar A, Mina B, Oks M, et al. The usefulness of chest CT imaging in patients with suspected or diagnosed COVID-19. *Chest*. 2021;160(2):652-70.
- Gestoso-Pecellín L, García-Flores Y, González-Quintana P, Marro-Arencibia JL. Recommendations for use of the diverse tests for detection of SARS-COV-2 infection. *Enferm Clin*. 2021;31:S40-S48.
- Chaganti S, Grenier P, Balachandran A, Chabin G, Cohen S, Flohr T, et al. Automated Quantification of CT Patterns Associated with COVID-19 from Chest CT. *Radiol Artif Intell*. 2020;2(4):e200048.
- Weikert T, Rapaka S, Grbic S, Re T, Chaganti S, Winkel DJ, et al. Prediction of patient management in covid-19 using deep learning-based fully automated extraction of cardiothoracic ct metrics and laboratory findings. *Korean J Radiol*. 2021;22(6):994-1004.
- Li K, Fang Y, Li W, Pan C, Qin P, Zhong Y, et al. CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19). *Eur Radiol*. 2020;30(8):4407-16.
- Prokop M, van Everdingen W, van Rees Vellinga T, van Ufford HQ, Stöger L, Beenen L, et al. CO-RADS: A Categorical CT Assessment Scheme for Patients Suspected of Having COVID-19-Definition and Evaluation. *Radiology*. 2020;296(2):E97-E104.
- Pan F, Ye T, Sun P, Gui S, Liang B, Li L, et al. Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. *Radiology*. 2020;295(3):715-21.
- Poyiadji N, Cormier P, Patel PY, Haded MO, Bhargava P, Khanna K, et al. Acute pulmonary embolism and COVID-19. *Radiology*. 2020;297(3):E335-E338.
- Tang N, Li D, Wang X, Sun Z. Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. *J Thromb Haemost*. 2020;18:844-7.
- AI Rad Companion Research CT Pneumonia Analysis prototype. Application Manual [Internet]. Siemens Healthineers. Disponible en: [https://marketplace.teamplay.siemens.com/api/download/media/Siemens%20Healthcare%20GmbH/AI-Rad%20Companion%20Research%20CT/VA11A/ApplicationManual\\_AIResearch\\_CTPneumoniaAnalysisV2.pdf](https://marketplace.teamplay.siemens.com/api/download/media/Siemens%20Healthcare%20GmbH/AI-Rad%20Companion%20Research%20CT/VA11A/ApplicationManual_AIResearch_CTPneumoniaAnalysisV2.pdf)
- Okuma T, Hamamoto S, Maebayashi T, Taniguchi A, Hirakawa K, Matsushita S, et al. Quantitative evaluation of COVID-19 pneumonia severity by CT pneumonia analysis algorithm using deep learning technology and blood test results. *Jpn J Radiol*. 2021;39(10):956-65.
- Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J, et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *Eur Radiol*. 2021;31(8):6096-104.
- Lessmann N, Sánchez CI, Beenen L, Boulogne LH, Brink M, Calli E, et al. Automated assessment of COVID-19 reporting and data system and chest CT severity scores in patients suspected of having COVID-19 using artificial intelligence. *Radiology*. 2021;298(1):E18-E28.
- Smallwood NH. The potential role for CT in the diagnosis of coronavirus disease 2019. *Chest*. 2021;159(3):906-7.
- Homayounieh F, Bezerra Cavalcanti Rockenbach MA, Ebrahimi S, Doda Khara R, Roberts M, Driggs D, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence*. 2021;3(3):199-217.
- Bizzo BC, Buch V, Babaei R, Karimi Mobin H, Mohseni I, Mitschke M, et al. Multicenter assessment of CT pneumonia analysis prototype for predicting disease severity and patient outcome. *J Digit Imaging*. 2021;34(2):320-9.
- Mortani Barbosa Jr, E. J., Georgescu, B., Chaganti, S., Bastarrika Aleman, G., Broncano Cabrero, J., et al. Machine learning automatically detects COVID-19 using chest CTs in a large multicenter cohort. <https://doi.org/10.1007/s00330-021-07937-3> Published
- Moezzi M, Shirbandi K, Shahvandi HK, Arjmand B, Rahim F. The diagnostic accuracy of Artificial Intelligence-Assisted CT imaging in COVID-19 disease: A systematic review and meta-analysis. *Inform Med Unlocked*. 2021;24:100591.
- Mei X, Lee HC, Diao KY, Huang M, Lin B, Liu C, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat Med*. 2020;26(8):1224-8.
- Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. *Radiology*. 2020;296(2):E32-E40.