

Variabilidad interobservador en cinco escalas de severidad determinada por tomografía computarizada para la valoración de neumonía por COVID-19

Karina I. Holguín-Andrade,* Estefanía Murrieta-Peralta, Ana P. Chischitz-Condey, Dania G. Solís-Cano, Roberto Ríos-Muñoz, Mario Calva-Arcos y Héctor Murrieta-González

Centro Médico ABC, Departamento de Radiología e Imagen Molecular, Ciudad de México, México

Resumen

Introducción: A finales de 2019 se presentó un brote de neumonía causada por un nuevo coronavirus, enfermedad a la que se denominó COVID-19. La tomografía computarizada ha desempeñado un papel importante en el diagnóstico de los pacientes con COVID-19. **Objetivo:** Demostrar la variabilidad interobservador con cinco escalas propuestas para la medición de la extensión de la neumonía ocasionada por COVID-19 mediante tomografía. **Métodos:** Se analizaron 35 tomografías de tórax iniciales de pacientes que asistieron al triaje respiratorio por sospecha de neumonía por COVID-19. Tres radiólogos realizaron la clasificación de las imágenes tomográficas de acuerdo con las escalas de severidad propuestas por Yang (1), Yuan (2), Chun (3), Wang (4) e INER-Chung-Pan (5). Se calculó el porcentaje de concordancia entre los evaluadores para cada escala con el índice de correlación intraclase. **Resultados:** La mayoría de los pacientes presentó afección de cinco lóbulos pulmonares (77.1 % de los pacientes). Las escalas 1, 2, 4 y 5 mostraron una correlación intraclase > 0.91 , con $p < 0.0001$, por lo que la concordancia fue casi perfecta. **Conclusiones:** La escala 4 (de Wang) mostró la mejor concordancia interobservador, con un coeficiente de 0.964 ($p = 0.001$).

PALABRAS CLAVE: Neumonía por COVID-19. Variabilidad interobservador. Escala de severidad por tomografía.

Inter-observer variability with five computed tomography severity scales for COVID-19 pneumonia assessment

Abstract

Introduction: By the end 2019 there was an outbreak of pneumonia caused by a new coronavirus, a disease that was called COVID-19. Computed tomography has played an important role in the diagnosis of COVID-19 patients. **Objective:** To demonstrate inter-observer variability with five scales proposed for measuring the extent of COVID-19 pneumonia on tomography. **Methods:** 35 initial chest computed tomography scans of patients who attended respiratory triage for suspected COVID-19 pneumonia were analyzed. Three radiologists classified the tomographic images according to the severity scales proposed by Yang (1), Yuan (2), Chun (3), Wang (4) and INER-Chung-Pan (5). The percentage of agreement between the evaluators for each scale was calculated using the intra-class correlation index. **Results:** In most patients were five pulmonary lobes compromised (77.1 % of the patients). Scales 1, 2, 4 and 5 showed an intra-class correlation > 0.91 ($p < 0.0001$), with agreement thus being almost perfect. **Conclusions:** Scale 4 (proposed by Wang) showed the best inter-observer agreement, with a coefficient of 0.964 ($p = 0.001$).

KEY WORDS: COVID-19 pneumonia. Inter-observer variability. Tomography severity scale.

Correspondencia:

*Karina I. Holguín-Andrade

E-mail: iliana.holguina@gmail.com

0016-3813/© 2020 Academia Nacional de Medicina de México, A.C. 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/>).

Fecha de recepción: 14-11-2020

Fecha de aceptación: 10-12-2020

DOI: 10.24875/GMM.20000826

Gac Med Mex. 2021;157:405-410

Disponible en PubMed

www.gacetamedicademexico.com

Introducción

A finales de 2019 se presentó un brote de neumonía de etiología desconocida en la provincia de Wuhan, China. En enero de 2020 se determinó que estos casos eran secundarios a un nuevo virus, SARS-CoV-2.¹ La enfermedad causada por este nuevo virus ha sido denominada enfermedad por coronavirus 2019 (COVID-19).² Esta enfermedad ha tenido una rápida progresión en el mundo, por lo que fue declarada pandemia en marzo de 2020.³

Para el diagnóstico de la enfermedad por COVID-19, la reacción en cadena de polimerasa en tiempo real (RT-PCR) se ha utilizado como el estándar de referencia, con sensibilidad y especificidad reportadas actualmente de 42 a 71 % en fases iniciales de la enfermedad.⁴⁻⁶ Esta variabilidad y baja sensibilidad de la RT-PCR implica que muchos pacientes con COVID-19 podrían no ser identificados en etapas tempranas.⁴

La tomografía computarizada ha desempeñado un papel importante en el diagnóstico de la neumonía por COVID-19.⁷ Diversos estudios han descrito los patrones tomográficos característicos en pacientes con neumonía secundaria a COVID-19,^{6,8,9} entre los cuales destacan principalmente zonas de vidrio esmerilado, con o sin asociación con zonas de consolidación, de distribución periférica y con predominio en lóbulos inferiores.¹⁰ Asimismo, se ha estudiado la sensibilidad de la tomografía computarizada y la prueba RT-PCR positiva. Dichos estudios han demostrado una sensibilidad de hasta 97 % para la tomografía computarizada con hallazgos característicos en el contexto de una prueba de RT-PCR positiva.¹¹

Sin embargo, se ha planteado la utilidad de la imagen tomográfica como un mecanismo para estratificar la severidad de la enfermedad pulmonar. Múltiples autores han propuesto la aplicación de índices tomográficos de severidad para la medición semicuantitativa de la afección pulmonar en los que se evalúe la afección por lóbulos pulmonares, por segmentos pulmonares o por la división en ambos pulmones en tres tercios.¹²⁻¹⁵

Entre las escalas de severidad por tomografía se destacan cinco:

1. Escala de Yang *et al.*,¹² que evalúa el porcentaje de afección en cada segmento pulmonar y lo estratifica en tres categorías (0, 0 % afección; 1, < 50 % de afección; 2, > 50 % de afección) y que otorga una puntuación máxima de 40.

2. Escala tomográfica de mortalidad de Yuan,¹⁶ que divide cada pulmón en tres zonas y en cada una evalúa el patrón tomográfico de la neumonía (0, normal; 1, vidrio despulido; 2, consolidación), que se multiplica por el porcentaje de afección (0, normal; 1, < 25 %; 2, 25-50 %; 3, 50-75 %; 4, > 75 %) y otorga una puntuación máxima de 72.
3. Escala de involucramiento pulmonar de Wang.⁷ Aunque es similar a la anterior, elimina la variable del patrón tomográfico, por lo que la puntuación máxima es de 24 puntos.
4. Escala de severidad total de Chung,⁸ en la que se evalúa cada lóbulo y el porcentaje de afección de cada uno (0, 0 %; 1, 1-25 %; 2, 26-50 %; 3, 51-75 %; 4, 76-100 %) y otorga una puntuación máxima de 20.
5. Escala propuesta por varios autores como Chung, Pan y sus respectivos colaboradores.¹⁷ Esta escala, incorporada recientemente en guías propuestas por el Instituto Nacional de Enfermedad Respiratorias en México (INER),¹⁴ evalúa cada lóbulo pulmonar y lo multiplica por el porcentaje de afección, agregando una categoría adicional: 1, < 5 %; 2, 5-25 %; 3, 25-50 %; 4, 50-75 %; 5, > 75 %. Con ella se obtiene un máximo de 25 puntos.

A pesar de la utilidad de las escalas de severidad descritas, no se reporta la variabilidad interobservador; en las escalas propuestas por Yang¹² y Wang⁷ solo se indica la variabilidad intraobservador

El objetivo principal de este estudio es demostrar cuál es la variabilidad interobservador de las anteriores escalas para la medición por tomografía computarizada de la extensión de la neumonía por COVID-19.

Métodos

Se analizaron las tomografías de tórax iniciales de pacientes que asistieron al triaje respiratorio por sospecha de neumonía por COVID-19 y que disponían de prueba PCR positiva.

Para el cálculo del tamaño de la muestra se utilizó una concordancia esperada de 90 % (IC 95 % = 80-100 %), con lo que el número de pacientes estimado fue de 35, mismos que se incluyeron.

La tomografía computarizada de tórax fue realizada en posición supina siguiendo el protocolo para la exploración de tórax: 2.5 mm de grosor de corte con reconstrucción a 1.25 mm en ventana pulmonar, tiempo de rotación del tubo de 0.8 segundos, pitch de

Tabla 1. Descripción de las escalas para valoración de hallazgos tomográficos en neumonía por COVID-19

1. Yang R. <i>et al.</i> ¹²	Valora cada segmento pulmonar y confiere una puntuación de afección: 0, 0 %; 1, < 50 %; 2, > 50 %.	La puntuación máxima es de 40. El umbral es de > 19.5 puntos para enfermedad severa e ingreso hospitalario.
2. Yuan M. <i>et al.</i> ¹⁶	Divide ambos pulmones en tres zonas, con la carina y vena pulmonar inferior como límites para cada una. Valora cada zona y confiere una puntuación de afección: 0, 0 %; 1, < 25 %; 2, 25-50 %; 3, 50-75 %; 4, > 75 %. Posteriormente multiplica cada punto conferido por el patrón que predomina en cada zona (1, normal; 2, vidrio deslucido; 3, consolidación).	La puntuación máxima es de 72. El umbral para predecir mortalidad es > 24.5 puntos.
3. Chung M. <i>et al.</i> ⁸	Valora cada lóbulo pulmonar y confiere una puntuación de afección: 0, normal; 1-25 %, mínimo; 26-50 %, leve; 51-75 %, moderado; > 75 %, severo.	La puntuación máxima es de 20 puntos. No tiene un umbral para correlacionar con la severidad clínica.
4. Wang Y. <i>et al.</i> ⁷	Divide ambos pulmones en tres zonas, con la carina y vena pulmonar inferior como límites para cada una. Valora cada zona y confiere una puntuación de afección: 0, 0 %; 1, < 25 %; 2, 25-50 %; 3, 50-75 %; 4, > 75 %.	La puntuación máxima es de 24 puntos. No tiene un umbral para correlacionar con la severidad clínica.
5. INER ¹⁴ /Chung ⁹ /Pan <i>et al.</i> ¹⁷	Valora cada lóbulo pulmonar y confiere una puntuación de afección: 0, normal; 1, < 5 %; 2, 5-25 %; 3, 26-50 %; 4, 51-75 %; 5, > 75 %.	La puntuación máxima es de 25 puntos. Divide la enfermedad en leve (1-5 puntos), moderada (5-15 puntos) y severa (> 15 puntos).

1.375:155.00, tiempo de exposición de 3.8 segundos con 120 kV y 580 mA.

Tres radiólogos con experiencia superior a cinco años en el servicio de tomografía computarizada realizaron la clasificación de las imágenes tomográficas con los cinco escalas descritas (Tabla 1), una escala por semana. Las imágenes fueron distribuidas aleatoriamente cada semana entre los tres radiólogos, proceso que se continuó hasta contar con dos evaluaciones de cada escala por imagen radiográfica.

Se calculó el porcentaje de concordancia entre los evaluadores para cada tipo de escala mediante el índice de correlación intraclass.^{18,19} Esta es una medida de la diferencia entre el acuerdo observado y el acuerdo esperado, si bien se dispuso de la puntuación absoluta para cada escala. Existe una clasificación para el grado de acuerdo y se considera por estratos de la siguiente forma:¹⁸

- < 0.3, acuerdo malo o nulo.
- 0.31-0.50, acuerdo leve.
- 0.51-0.70, acuerdo moderado.
- 0.71-0.90, acuerdo bueno.
- > 0.91, acuerdo casi perfecto.

En todos los casos, el umbral para evaluar la significación estadística se estableció en un nivel alfa de 0.05.

Se obtuvo autorización por parte del Comité de Ética e Investigación del Centro Médico ABC. No existió ninguna interacción directa con los pacientes. La información se obtuvo a partir de los expedientes y se preservó la confidencialidad y el anonimato de los pacientes.

Resultados

Se estudiaron 35 pacientes, de los cuales 40 % fue del sexo femenino (14 pacientes) y 60 % del masculino (21 pacientes), con edad promedio de 54.14 años (Tabla 2).

La mayoría de los pacientes presentó afección de cinco lóbulos pulmonares (77.1 % de los pacientes). El patrón tomográfico que predominó fue vidrio esmerilado en 21 (60 %), consolidación en siete (20 %) y patrón empedrado en seis (17.1 %).

La descripción del método con el que se evaluó la severidad de la neumonía de acuerdo con cada escala se detalla en la Tabla 1.

La escala que mostró la mejor concordancia interobservador fue la propuesta por Wang *et al.*,⁷ con un coeficiente de 0.964 (IC 95 %, p = 0.001). Las siguientes en orden descendente de concordancia fue la propuesta por Chung, Pan y sus colaboradores,¹⁷ la cual fue adoptada por el INER¹⁴ y que en este estudio

Tabla 2. Características demográficas

Edad (años)	Media \pm desviación estándar 54.14 \pm 14.86	Rango 29-78
Sexo	Femenino	14 (40 %)
	Masculino	21 (60 %)
Hallazgos tomográficos	Vidrio esmerilado	21 (60 %)
	Consolidación	7 (20 %)
	Empedrado	6 (17.1 %)
	Negativo	1 (2.9 %)
Lóbulos pulmonares afectados	0	1 (2.9 %)
	1	0 (0 %)
	2	1 (2.9 %)
	3	3 (8.6 %)
	4	3 (8.6 %)
	5	27 (77.1 %)
Enfisema pulmonar	Negativo	29 (82.9 %)
	Positivo	6 (17.1 %)
Derrame pleural	Negativo	33 (94.3 %)
	Positivo	2 (5.7 %)
Manejo ambulatorio	Negativo	32 (91.4 %)
	Positivo	3 (8.6 %)
Ingreso a unidad de cuidados intensivos	Negativo	16 (45.7 %)
	Positivo	19 (54.3 %)
Intubación	Negativo	19 (54.3 %)
	Positivo	16 (45.7 %)

mostró un coeficiente de 0.935 (IC 95 %, $p = 0.001$), y la de Yuan *et al.*,¹⁶ con un coeficiente de 0.930 (IC 95 %, $p = 0.001$). La escala que mostró la concordancia más baja fue la propuesta por Chung *et al.*,⁸ con un coeficiente de -0.154 (IC 95 %, $p = 0.661$), como se muestra en la Tabla 3.

Cada escala tiene diferentes puntuaciones, que varían entre 0 y 40, 72, 20, 24 y 25 puntos, respectivamente, en cada una. En la Tabla 4 se muestra el promedio de la puntuación conferida de acuerdo con las dos lecturas de cada escala.

Discusión

La población estudiada representa la informada en la literatura. La mayoría de los pacientes con

neumonía por COVID-19 fueron hombres y presentaron enfermedad difusa con afección de los cinco lóbulos pulmonares.²⁰

Los patrones encontrados coincidieron con los reportados en series de casos más grandes,²¹ tanto en frecuencia como en orden; el más común fue el vidrio esmerilado, seguido por el patrón de consolidación y en empedrado. Tal como lo mencionan Zheng *et al.*,⁶ una de las manifestaciones topográficas más frecuentes es la presencia de zonas de ocupación con imagen de vidrio esmerilado, hasta en 98 % de los pacientes; seguida por consolidaciones hasta en 64 % de los pacientes y patrón en empedrado hasta en 36 %.

En algunos reportes también se han encontrado hallazgos adicionales como derrame pleural (8 %),²² que podrían conferir un mayor riesgo de neumonía severa con necesidad de intubación o, incluso, de muerte. Sin embargo, en este estudio encontramos solo dos pacientes con derrame pleural (5.7 %): uno presentó puntuaciones altas en las diferentes escalas de severidad, 36 en la de Yang *et al.*,¹² 52 en la de Yuan *et al.*,¹⁶ 15 en la de por Chung *et al.*,⁸ 21 en el de Wang *et al.*⁷ y 22 en la propuesta por Chung-Pan-INER, lo cual correlaciona con la observación de que los pacientes con derrame pleural tendrán una enfermedad severa y, por lo tanto, un desenlace distinto. El segundo paciente tuvo puntuaciones cercanas a la media (21 en la escala de Yang *et al.*,¹² 14 en la de Yuan M. *et al.*,¹⁶ 12 en la Chung *et al.*,⁸ 15 en la de Wang *et al.*⁷ y 12 en la propuesta por Chung-Pan-INER), por lo que esa observación no aplicaría para este paciente.

Este estudio se fundamenta en las observaciones y mediciones realizadas en 35 pacientes, y los intervalos de confianza muestran la precisión de las estimaciones.

Las mayores puntuaciones se observaron en pacientes ingresados en la unidad de cuidados intensivos y en pacientes intubados. De igual forma, las menores puntuaciones se registraron en pacientes manejados de forma ambulatoria. Lo anterior indica la utilidad de la clasificación de severidad pulmonar como indicador pronóstico.

Respecto a nuestro objetivo principal, encontramos que la mayoría de las escalas tienen poca variabilidad interobservador. Sin embargo, la escala 3 mostró una variabilidad interobservador significativa, con una correlación intraclase de -0.154 , con $p < 0.661$, por lo que la correlación es nula. El resto de las escalas (1, 2, 4 y 5) tuvieron una correlación intraclase > 0.91 ,

Tabla 3. Resultados de correlación intraclase

	Escala de severidad				
	1. Yang <i>et al.</i> ¹²	2. Yuan <i>et al.</i> ¹⁶	3. Chung <i>et al.</i> ⁸	4. Wang <i>et al.</i> ⁷	5. INER ¹⁴ /Chung <i>et al.</i> ⁹ /Pan <i>et al.</i> ¹⁷
Correlación intraclase	0.916	0.930	-0.154	0.964	0.935
Rango correlación intraclase	0.834-0.958	0.861-0.965	-1.286-0.417	0.928-0.982	0.872-0.967
p	0.001	0.001	0.661	0.001	0.001

Tabla 4. Promedio de puntuación en cada escala

		Media	Rango	Desviación estándar
Escala 1	Lectura 1	20.71	0-40	9.161
	Lectura 2	20.17	0-37	8.793
Escala 2	Lectura 1	27.83	0-66	15.775
	Lectura 2	27.43	0-60	14.250
Escala 3	Lectura 1	15.69	0-60	13.226
	Lectura 2	19.60	0-46	12.360
Escala 4	Lectura 1	12.46	0-24	5.972
	Lectura 2	13.14	0-22	5.281
Escala 5	Lectura 1	13.54	0-23	5.078
	Lectura 2	13.71	0-23	5.496

con $p < 0.0001$, casi perfecta. Con ello encontramos que, con excepción de la 3, las escalas pueden ser utilizadas en la práctica clínica con mayor confianza y menor variabilidad entre observadores.

Identificamos varios aspectos en cuanto al análisis más específico de cada escala. Uno es que las más prácticas son las escalas 1 y 2, ya que son las únicas que proponen una correlación de la puntuación con un umbral para predecir la admisión hospitalaria (escala 1) y la mortalidad (escala 2). La escala 5 confiere un grado de severidad de forma arbitraria y no traduce un desenlace o situación clínica específica.

La escala 2 es particular, ya que es la única que confiere una puntuación de acuerdo con la extensión de la neumonía, pero que también toma en cuenta el patrón (vidrio esmerilado o consolidación). En un principio pensábamos que el hecho de realizar un segundo paso, que no se efectúa en el resto de las escalas, podría incrementar la variabilidad interobservador, lo cual no ocurrió, si bien al momento los observadores consideraron esta escala como de mayor grado de dificultad.

Creímos que las escalas 3 y 5 tendrían menor variabilidad interobservador, ya que son las únicas que dividen las zonas evaluadas de acuerdo con límites anatómicos que observamos en la tomografía, ya que valoran la extensión de la neumonía por lóbulo pulmonar y las cisuras que los dividen son visibles en el estudio y que el resto de las escalas, al conferir un límite anatómico no visible (segmentos en la escala 1 y tres zonas por pulmón en las escalas 2 y 4), mostrarían una mayor variabilidad interobservador.

Conclusiones

La mayoría de las escalas evaluadas tuvieron una correlación intraclase casi perfecta, por lo cual pueden ser utilizadas en la práctica diaria. No recomendamos el uso de la propuesta por Chung *et al.*⁸ ya que presentó la mayor variabilidad interobservador.

Agradecimientos

Los autores agradecen la colaboración del doctor Juan Osvaldo Talavera Piña.

Conflicto de intereses

Los autores declaran no tener conflicto de intereses.

Financiamiento

Los autores declaran no haber recibido financiamiento para la realización de este estudio.

Responsabilidades éticas

Protección de personas y animales. Los autores declaran que para esta investigación no realizaron experimentos en seres humanos ni en animales.

Confidencialidad de los datos. Los autores declaran que siguieron los protocolos de su centro de trabajo sobre la publicación de datos de pacientes.

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

Bibliografía

1. He F, Deng Y, Li W. Coronavirus disease 2019: What we know? *J Med Virol.* 2020;1-7.
2. World Health Organization [Internet]. Suiza: Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected; 2020.
3. Simpson S, Kay FU, Abbara S, Bhalla S, Chung JH, Michael-Chung TSH, et al. Radiological Society of North America Expert Consensus Statement on Reporting Chest CT Findings Related to COVID-19. Endorsed by the Society of Thoracic Radiology, the American College of Radiology, and RSNA. *Radiol Cardiothorac Imaging.* 2020;2.
4. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology.* 2020;296:E32-E40.
5. Wen Z, Chi Y, Zhang L, Liu H, Du K, Li Z, et al. Coronavirus disease 2019: initial detection on chest CT in a retrospective multicenter study of 103 Chinese patients. *Radiol Cardiothorac Imaging.* 2020;2:e200092.
6. Zheng Y, Zhang Y, Wang Y, Zixiang H, Song B. Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review. *Eur Radiol.* 2020;30:4389-4389.
7. Wang Y, Dong C, Hu Y, Li C, Ren W, Zhang X, Shi H, et al. Temporal changes of CT findings in 90 patients with COVID-19 pneumonia: a longitudinal study. *Radiology.* 2020;296:E55-E64.
8. Chung M, Bernheim A, Mei X, Zhang N, Huang M, Zeng X, et al. CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology.* 2020;295:202-207.
9. Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raoof S, et al. The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society. *Chest.* 2020;158:106-116.
10. Pan Y, Guan H, Zhou S, Wang Y, Li Q, Zhu T, et al. Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV): a study of 63 patients in Wuhan, China. *Eur Radiol.* 2020;30:3306-3309.
11. Murrieta-Peralta E, Chischistz-Condey AP, Ramírez-Landero J, Moctezuma-Velasco CR, Murrieta-González H, Salazar-Segovia J. El reporte radiológico en pacientes con sospecha de COVID-19: nuestra experiencia en el Centro Médico ABC. *An Radiol Mex.* 2020;19:276-285.
12. Yang R, Li X, Liu H, Zhen Y, Zhang X, Xiong Q, et al. Chest CT severity score: an imaging tool for assessing severe COVID-19. *Radiol Cardiothorac Imaging.* 2020;30:e200047.
13. Kunwei L, 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). *European Radiology.* 2020;30:4407-4416.
14. Juárez-Hernández F. Presentación en línea de la página oficial INER [Internet]. México: Instituto Nacional de Enfermedades Respiratorias; 2020.
15. 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 with suspected COVID-19: definition and evaluation. *Radiology.* 2020;296:E97-E104.
16. Yuan M, Yin W, Tao Z, Tan W, Hu Y. Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China. *PLoS One.* 2020;15:e0230548.
17. Pan F, Ye T, Sun P, Gui S, Liang B, Li L, et al. Time course of lung changes at chest CT during recovery from coronavirus disease 2019 (COVID-19). *Radiology.* 2020;295:715-721.
18. Pita-Fernández S, Pértega-Díaz S, Rodríguez-Maseda E. La fiabilidad de las mediciones clínicas: el análisis de concordancia para variables numéricas. *Cad Aten Primaria.* 2003;10:290-296.
19. Camacho-Sandoval J. Coeficiente de concordancia para variables continuas. *AMC.* 2008;50:211-212.
20. Zhou S, Wang Y, Zhu T, Xia L. CT features of coronavirus disease 2019 (COVID-19) pneumonia in 62 patients in Wuhan, China. *AJR Am J Roentgenol.* 2020;214:1-8.
21. Salehi S, Abedi A, Balakrishnan S, Gholamrezanezhad A. Coronavirus disease 2019 (COVID-19): a systematic review of imaging findings in 919 patients. *AJR Am J Roentgenol.* 2020;215:1-7.
22. Carotti M, Salaffi F, Sarzi-Puttini P, Agostini A, Borgheresi A, Minorati D, et al. Chest CT features of coronavirus disease 2019 (COVID-19) pneumonia: key points for radiologists. *Radiol Med.* 2020:1-11.