

Artículo original

Evaluación de concordancia entre médicos, internos y estudiantes sobre gravedad COVID-19 mediante Kappa de Fleiss

Assessment of concordance between physicians, interns and students on COVID-19 severity using Fleiss Kappa

Daniela Abigail Cobo Álvarez¹ <https://orcid.org/0000-0002-3927-870X>

Carlos Escobar Suárez² <https://orcid.org/0000-0002-9280-1555>

Enrique Rodríguez Reyes¹ <https://orcid.org/0000-0002-5311-7312>

¹Universidad Regional Autónoma de los Andes (UNIANDES Ambato). Ecuador.

²Universidad Regional Autónoma de los Andes (UNIANDES). Ecuador.

Autor para la correspondencia: ua.danielacobo@uniandes.edu.ec

RESUMEN

Introducción: El uso del coeficiente Kappa de Fleiss se aplica en la evaluación de la concordancia entre evaluadores, lo que presenta ciertos desafíos dentro del ámbito de la práctica médica.

Objetivo: El objetivo del estudio fue evaluar la concordancia entre médicos, internos y estudiantes de la carrera de Medicina, sobre la gravedad de pacientes con COVID-19 en un hospital de Ambato, Ecuador, mediante el coeficiente Kappa de Fleiss.

Métodos: La concordancia entre evaluadores se evaluó mediante la prueba estadística Kappa de Fleiss. Fue un estudio de tipo analítico, observacional, transversal y prospectivo. La población de estudio se compuso de dos médicos, dos internos y dos estudiantes del último año de la carrera de Medicina, que laboraban en un hospital de Ambato, Ecuador.

Resultados: Se obtuvo un valor de Kappa global de 0,593 con un error estándar asintótico de 0,012, un valor Z de 50,952, y un p-valor de 0,000. Los intervalos de confianza asintóticos del 95% indicaron que el valor de Kappa se situó entre 0,571 y 0,616, lo que sugirió que la concordancia global estuvo en el rango de moderada-considerable. Los resultados de los coeficientes Kappa para categorías individuales reflejaron distintos niveles de concordancia entre los evaluadores.

Conclusiones: Si bien se observa un nivel de concordancia global moderado-considerable entre los evaluadores, al analizar categorías individuales, se evidencia una concordancia casi perfecta para pacientes clasificados como Leve, pero niveles de concordancia moderada y considerable para las categorías Moderado y Grave, respectivamente.

Palabras clave: COVID-19; Kappa de Fleiss; concordancia entre evaluadores; estudiantes de medicina; médicos internos residentes.

ABSTRACT

Introduction: The use of Fleiss' Kappa coefficient is applied in the evaluation of inter-rater concordance, which presents certain challenges within the medical practice setting.

Objective: The objective of this study was to evaluate the concordance between physicians, interns and medical students on the severity of patients with COVID-19

in a hospital in Ambato, Ecuador, using Fleiss' Kappa coefficient.

Methods: Inter-rater concordance was evaluated using Fleiss' Kappa statistical test. This was an analytical, observational, cross-sectional and prospective study. The study population consisted of two physicians, two interns and two final year medical students working in a hospital in Ambato, Ecuador.

Results: An overall Kappa value of 0.593 with an asymptotic standard error of 0.012, a Z value of 50.952, and a p-value of 0.000 were obtained. The asymptotic 95 % confidence intervals indicated that the Kappa value ranged from 0.571 to 0.616, which suggested that the overall agreement was in the moderate-considerable range. The results of Kappa coefficients for individual categories reflected different levels of agreement between raters.

Conclusions: Although a moderate-considerable level of overall agreement is observed among the evaluators, when analyzing individual categories, almost perfect agreement is evident for patients classified as Mild, but moderate and considerable levels of agreement for the Moderate and Severe categories, respectively.

Keywords: COVID-19; Fleiss' Kappa; inter-rater agreement; medical students; resident interns.

Recibido: 02/01/2024

Aceptado: 13/02/2024

Introducción

En el marco de Ecuador, la Constitución de la República que se promulga en 2008,

en su artículo 32, consagra el derecho a la salud como un derecho fundamental que el Gobierno Central está obligado a garantizar.⁽¹⁾ En el contexto de la pandemia de COVID-19, se ha establecido claramente en las Unidades de Cuidados Críticos (UCI) de Ecuador el criterio para seleccionar el sistema de triaje más idóneo en las unidades de atención de urgencias.⁽²⁾ Sin embargo, previo a este proceso, surge la necesidad de clasificar a los pacientes infectados con SARS-CoV-2 para una pronta atención y toma de decisiones médicas. En este punto, no siempre existe una completa coincidencia entre el personal médico que se encarga de su evaluación.

En este estudio, el principal problema identificado radica en la discrepancia de criterios para clasificar a los pacientes con COVID-19 entre médicos, internos y estudiantes de Medicina en un hospital de Ambato, Ecuador, durante la actual pandemia. Esta divergencia puede ser problemática, ya que podría resultar en una asignación ineficiente de recursos y generar disparidades en la atención médica otorgada a los pacientes afectados por el SARS-CoV-2.

La pregunta central que guía esta investigación es: ¿Cuál es el grado de concordancia existente entre médicos, internos y estudiantes de Medicina en la categorización de la gravedad de los pacientes con COVID-19 en un hospital de Ambato, Ecuador? El propósito de evaluar la concordancia mediante el coeficiente Kappa de Fleiss radica en analizar la coherencia en las decisiones de clasificación y comprender más a fondo la gestión de recursos y las necesidades de los pacientes en este contexto específico.

Este estudio es crucial ya que podría contribuir significativamente a la mejora de la precisión diagnóstica, la optimización de la atención médica, el fortalecimiento de la formación médica y la validación de métodos de evaluación clínica en el contexto específico de la pandemia de COVID-19 en el hospital de Ambato, Ecuador.

El objetivo del estudio es evaluar la concordancia entre médicos, internos y estudiantes de la carrera de Medicina, sobre la gravedad de pacientes con COVID-19 en un hospital de Ambato, Ecuador, mediante el coeficiente Kappa de Fleiss.

Métodos

El estudio llevado a cabo se enmarcó dentro de un enfoque analítico, observacional, prospectivo y transversal. Para evaluar la concordancia entre diferentes observadores, se empleó el coeficiente Kappa de Fleiss, siguiendo criterios predefinidos para la evaluación de la concordancia entre ellos.

Población de estudio y criterios éticos

La población de estudio consistió en dos médicos, dos internos y dos estudiantes de último año de la carrera de Medicina que estaban desempeñando sus labores en un hospital de Ambato, Ecuador, durante el mes de febrero de 2022.

Para ser considerados en la investigación, se incluyó a aquellos médicos, internos y estudiantes de último año de Medicina que contaban con experiencia en el manejo de pacientes con COVID-19 y estuvieron activamente involucrados en su atención durante ese mes. Por motivos éticos y en concordancia con los principios de la Declaración de Helsinki y sus posteriores revisiones que rigen la investigación con seres humanos, se excluyó a los participantes que no proporcionaron su consentimiento informado para participar en el estudio. Además, se obtuvo la aprobación del comité de ética de la institución correspondiente para llevar a cabo esta investigación.

Es pertinente aclarar que, dentro de la población bajo estudio, se definió la muestra mediante una técnica de muestreo por conveniencia, donde los investigadores seleccionaron a los participantes basándose en su accesibilidad y alto interés

demostrado por formar parte del estudio. Esta estrategia aseguró a los investigadores la disponibilidad de la muestra. Sin embargo, es importante destacar que esta selección no garantizó la representatividad de todos los médicos, internos y estudiantes de último año de Medicina en la población general. A pesar de ello, permitió obtener la información necesaria para llevar a cabo el estudio. Es relevante subrayar que los investigadores no buscaron extrapolar los resultados obtenidos de esta muestra a toda la población.

Procedimiento de estudio

El procedimiento del estudio implicó la evaluación de la concordancia entre dos médicos, dos internos y dos estudiantes de último año de Medicina que estaban trabajando en un hospital de Ambato, Ecuador, durante febrero de 2022. En total, se atendieron 247 pacientes con COVID-19, requiriendo una rápida y precisa clasificación de su estado de salud con respecto a la enfermedad.

Se aplicó el coeficiente Kappa de Fleiss para medir la concordancia entre este personal sanitario en la clasificación de los pacientes en tres niveles: Leve, Moderado y Grave. Cada evaluador ofreció su propia clasificación de la gravedad para cada paciente, lo que permitió determinar si existía un acuerdo significativo en esta categorización. El coeficiente Kappa de Fleiss se interpretó en una escala de 0 a 1, donde 0 señala falta de acuerdo entre los evaluadores y 1 indica una coincidencia perfecta.

Es esencial destacar que el Kappa de Fleiss no evaluó la precisión de las mediciones en términos de si reflejaban con precisión la gravedad real de la enfermedad. En cambio, midió la concordancia entre los evaluadores, garantizando consistencia en la toma de decisiones médicas.

Variables de estudio

- **Clasificación de los pacientes:** se refirió al proceso en el que el personal

sanitario asignó a cada paciente una categoría de gravedad, clasificándolos como Leve, Moderado o Grave, según su estado de salud debido a la COVID-19.

- **Nivel de concordancia entre evaluadores:** esta variable buscó medir el grado de acuerdo entre los diferentes miembros del personal sanitario que participaron en la clasificación de la gravedad de los pacientes. El coeficiente Kappa de Fleiss se empleó como medida para evaluar la concordancia interobservador. La interpretación de esta variabilidad se basó en los criterios establecidos por Landis y Koch en 1977:⁽³⁾
 - Kappa > 0,8: Indica un nivel de concordancia casi perfecto.
 - Kappa > 0,6: Indica un nivel de concordancia considerable.
 - Kappa > 0,4: Indica un nivel de concordancia moderado.
 - Kappa > 0,2: Indica un nivel de concordancia aceptable.
 - Kappa > 0: Indica un nivel de concordancia bajo.
 - Kappa < 0: Indica ausencia de concordancia.

Las tres categorías dentro del personal sanitario se conceptualizaron de la manera siguiente:

- **Médicos:** profesionales con experiencia en el manejo clínico de pacientes, capacitados y con conocimiento especializado en el diagnóstico y tratamiento de enfermedades.
- **Internos:** estudiantes de Medicina que habían completado su formación básica y se encontraban en una etapa avanzada de su educación, realizando prácticas clínicas bajo supervisión directa de médicos experimentados, también conocidos como médicos internos residentes (MIR) en algunos países.
- **Estudiantes de último año de Medicina:** estudiantes que estaban cursando el último año de su formación médica y se encontraban en la fase final de

su educación, adquiriendo experiencia práctica en entornos clínicos bajo supervisión directa.

Resultados

La Tabla 1 resume los resultados del coeficiente Kappa de Fleiss, un indicador fundamental para evaluar la concordancia entre el personal sanitario en la clasificación de la gravedad de los pacientes con COVID-19 en un hospital de Ambato, Ecuador. Este coeficiente proporciona información esencial sobre la consistencia de las clasificaciones realizadas por los evaluadores y su significancia estadística.

Tabla 1- Concordancia Global - Resultados del coeficiente Kappa de Fleiss

| Acuerdo global ^a | | | | | | |
|-----------------------------|-----------------|----------------|--------|---------------|---|-----------------|
| | Kappa de Fleiss | Asintótica | | | 95 % de intervalo de confianza asintótico | |
| | | Error estándar | Z | Significancia | Límite inferior | Límite superior |
| Acuerdo global | 0,593 | 0,012 | 50,952 | 0,000 | 0,571 | 0,616 |

a. Los datos de muestra contienen 247 sujetos eficaces y 6 evaluadores.

Los resultados de la Tabla 1 revelan un valor global de Kappa de 0,593, con un error estándar asintótico de 0,012. Este valor refleja la concordancia entre los evaluadores en la clasificación de la gravedad de los pacientes. El valor Z, 50,952, y el p-valor de 0,000 indican una concordancia significativa entre el personal sanitario en sus clasificaciones.

Los intervalos de confianza asintóticos del 95 % (Límite Inferior - LI y Límite Superior - LS) sugieren que el valor de Kappa se encuentra entre 0,571 y 0,616. Esto

implica una concordancia global moderada a considerable. En este contexto específico de la pandemia de COVID-19 en el hospital de Ambato, Ecuador, los evaluadores mostraron una consistencia aceptable en la clasificación de la gravedad de los pacientes. Este hallazgo es relevante para asegurar una atención médica efectiva y coherente en esta población de pacientes.

La Tabla 2 exhibe los resultados de los coeficientes Kappa, los cuales permiten una evaluación detallada de la concordancia entre los evaluadores en la clasificación de pacientes en distintos niveles de gravedad durante la pandemia de COVID-19 en el hospital de Ambato, Ecuador.

Tabla 2- Concordancia por categoría - Resultados de los coeficientes Kappa

| Acuerdo en categorías individuales ^a | | | | | | | |
|---|--------------------------|-------|----------------|--------|------|---|------|
| Categoría de puntuación | Probabilidad condicional | Kappa | Asintótica | | | 95 % de intervalo de confianza asintótico | |
| | | | Error estándar | Z | Sig. | LI | LS |
| Leve | 0,874 | 0,822 | 0,016 | 50,025 | 0,00 | 0,79 | 0,85 |
| Moderado | 0,595 | 0,393 | 0,016 | 23,903 | 0,00 | 0,36 | 0,42 |
| Grave | 0,739 | 0,582 | 0,016 | 35,428 | 0,00 | 0,55 | 0,61 |

a. Los datos de muestra contienen 247 sujetos eficaces y 6 evaluadores.

En la Tabla 2 se presentan los resultados de los coeficientes Kappa para cada una de las tres categorías de clasificación de pacientes. Se añade un error estándar asintótico de 0,016 para referencia.

- Leve: el coeficiente Kappa de 0,822, con un valor Z de 50,025 y un p-valor de 0,00, sugiere un nivel de concordancia casi perfecto en esta categoría.
- Moderado: la concordancia en esta categoría fue representada por un coeficiente Kappa de 0,393, un valor Z de 23,903 y un p-valor de 0,00, lo que indicó un nivel de concordancia moderado.
- Grave: en esta categoría, se observó un coeficiente Kappa de 0,582, un valor Z de 35,428 y un p-valor de 0,00, reflejando un nivel de concordancia considerable.

Los coeficientes Kappa para las categorías individuales reflejaron distintos niveles de concordancia entre los evaluadores. El valor de Kappa fue más alto para la categoría "Leve", mostrando una concordancia casi perfecta, mientras que las categorías "Moderado" y "Grave" exhibieron niveles de concordancia moderada y considerable, respectivamente.

Estos resultados resaltan la variabilidad en la concordancia de las clasificaciones de los evaluadores, lo cual sugiere áreas donde se puede mejorar la consistencia en la evaluación de la gravedad de los pacientes con COVID-19 en este contexto específico.

La Tabla 3 presenta los resultados de la Kappa ponderada de Cohen, que es una medida de concordancia ajustada que considera el grado de acuerdo entre los evaluadores ponderando las discrepancias entre las clasificaciones. Esta tabla refleja la concordancia entre los distintos tipos de evaluadores (médicos, internos y estudiantes) en la clasificación de la gravedad de los pacientes con COVID-19 en el hospital de Ambato, durante la pandemia.

Tabla 3- Concordancia mediante Kappa ponderada de Cohen entre los distintos evaluadores en la clasificación de gravedad de pacientes COVID-19

| Kappa ponderada de Cohen | | | | | | |
|---------------------------|-------------------------------|-----------------------------|----------------|-------|---------------------------------------|-----------------|
| Calificaciones | Kappa ponderados ^a | Asintótica | | | 95% Intervalo de confianza asintótica | |
| | | Error estándar ^b | Z ^c | Sig. | Límite inferior | Límite superior |
| Médico 1 – Médico 2 | 0,859 | 0,025 | 17,761 | 0,000 | 0,811 | 0,908 |
| Médico 1 - Interno1 | 0,781 | 0,029 | 15,931 | 0,000 | 0,724 | 0,839 |
| Médico 1 - Interno2 | 0,714 | 0,031 | 14,670 | 0,000 | 0,652 | 0,775 |
| Médico 1 - Estudiante1 | 0,519 | 0,038 | 11,375 | 0,000 | 0,444 | 0,593 |
| Médico 1 - Estudiante2 | 0,469 | 0,038 | 10,487 | 0,000 | 0,395 | 0,542 |
| Médico 2 - Interno1 | 0,647 | 0,036 | 13,093 | 0,000 | 0,577 | 0,718 |
| Médico 2 - Interno2 | 0,590 | 0,037 | 11,911 | 0,000 | 0,517 | 0,663 |
| Médico 2 - Estudiante1 | 0,482 | 0,042 | 10,094 | 0,000 | 0,399 | 0,565 |
| Médico 2 - Estudiante2 | 0,565 | 0,038 | 11,988 | 0,000 | 0,490 | 0,639 |
| Interno1 - Interno2 | 0,925 | 0,018 | 18,045 | 0,000 | 0,891 | 0,960 |
| Interno1 - Estudiante1 | 0,737 | 0,031 | 15,209 | 0,000 | 0,677 | 0,797 |
| Interno1 - Estudiante2 | 0,676 | 0,032 | 14,032 | 0,000 | 0,613 | 0,740 |
| Interno2 - Estudiante1 | 0,807 | 0,027 | 16,181 | 0,000 | 0,755 | 0,859 |
| Interno2 - Estudiante2 | 0,751 | 0,029 | 14,995 | 0,000 | 0,694 | 0,809 |
| Estudiante1 - Estudiante2 | 0,928 | 0,019 | 18,335 | 0,000 | 0,891 | 0,964 |

a. La estimación de la kappa ponderada utiliza ponderaciones lineales.

b. El valor no depende de hipótesis nulas o alternativas.

c. Estima el error estándar asintótico asumiendo la hipótesis nula de que la kappa ponderada es cero.

Los valores de Kappa ponderada para las diferentes comparaciones entre evaluadores muestran la concordancia en las calificaciones otorgadas. Las calificaciones más altas de Kappa ponderada, como en la comparación entre 'Interno 1 - Interno2 ' y 'Estudiante 1 – Estudiante 2', indican un nivel más alto de concordancia en la clasificación de gravedad de los pacientes.

Por otro lado, las comparaciones que muestran valores más bajos de Kappa ponderada, como en 'Médico 1 – Estudiante 1' y 'Médico 1 – Estudiante 2', sugieren un acuerdo más bajo en la clasificación. Estos resultados resaltan la variabilidad en la concordancia entre los diferentes tipos de evaluadores, lo que puede tener implicaciones en la consistencia de la atención médica brindada a los pacientes en función de quien realice la clasificación de gravedad.

Discusión

A partir de los resultados que se obtienen en este estudio, sus autores interpretan que existe una variabilidad significativa en la concordancia entre médicos, internos y estudiantes de la carrera Medicina al clasificar la gravedad de los pacientes con COVID-19 en el hospital de Ambato, Ecuador. Si bien se observa un nivel de concordancia global moderado-considerable entre los evaluadores, al analizar categorías individuales, se evidencia una concordancia casi perfecta para pacientes clasificados como 'Leve', pero niveles de concordancia moderada y considerable para las categorías 'Moderado' y 'Grave', respectivamente.

Estos hallazgos destacan la necesidad de estandarizar y mejorar la consistencia en la clasificación de gravedad entre el personal sanitario, ya que estas discrepancias podrían influir en la toma de decisiones clínicas y en la asignación de recursos, lo que a su vez podría afectar la atención y el manejo de los pacientes

con COVID-19.

Los autores consideran importante este estudio, dado que evaluar la concordancia entre médicos, internos y estudiantes en la evaluación de la gravedad del COVID-19 ayuda a mejorar la precisión en el diagnóstico y la clasificación de pacientes, lo que podría resultar en un tratamiento más eficaz y oportuno. Comprender la concordancia entre diferentes evaluadores permite identificar posibles discrepancias en la percepción de la gravedad de la enfermedad. Esto podría conducir a una atención médica más homogénea y coherente para los pacientes con COVID-19 en el hospital de Ambato.

Este estudio proporciona información valiosa sobre la consistencia de las evaluaciones clínicas entre médicos experimentados, internos y estudiantes de Medicina. Esto podría contribuir a mejorar la formación de los estudiantes y ayudar a los médicos en formación a comprender la importancia de la evaluación clínica en situaciones de emergencia, como la pandemia de COVID-19.

Paralelamente, la investigación de la concordancia entre diferentes grupos de evaluadores también valida los métodos utilizados para evaluar la gravedad del COVID-19 en el entorno clínico, identificando posibles áreas de mejora en la evaluación clínica y la necesidad de una mayor estandarización en los criterios de evaluación.

La principal prioridad en pacientes afectados por COVID-19 es atender sus crecientes necesidades de oxígeno, con el propósito de evitar el uso de ventilación mecánica, la cual se asocia con una tasa de mortalidad elevada que varía entre el 76,4 % y el 97,2 %. Se están explorando nuevas técnicas de suministro de oxígeno, como las campanas o cascos de oxígeno, que podrían ser beneficiosos al mejorar los niveles de saturación de oxígeno en la sangre (SaO₂), reducir la dependencia de la ventilación mecánica durante la hospitalización, disminuir las tasas de

mortalidad y acortar la estancia hospitalaria en pacientes hipóxicos con COVID-19 que no responden adecuadamente a los métodos convencionales de administración de oxígeno.⁽⁴⁾

En marzo de 2020, la Organización Mundial de la Salud (OMS) declara oficialmente la COVID-19 como una pandemia. La falta de información previa sobre esta enfermedad genera dificultades en la toma de decisiones para todos los profesionales de la salud, incluyendo a los departamentos quirúrgicos.⁽⁵⁾ El uso de oxígeno nasal de alto flujo se considera el tratamiento inicial recomendado para la insuficiencia respiratoria hipoxémica aguda y se emplea ampliamente en pacientes con COVID-19.⁽⁶⁾

Los pacientes con COVID-19 pueden experimentar una gama de síntomas y complicaciones que abarcan desde el síndrome de dificultad respiratoria aguda (SDRA) hasta miocardiopatía, disfunción de múltiples órganos y shock, los cuales pueden originarse debido a alteraciones en el sistema renina-angiotensina (RAS).⁽⁷⁾

Los médicos que trabajan en unidades de cuidados intensivos (UCI), al igual que aquellos involucrados en esta investigación, deben considerar que aún no se ha establecido de manera definitiva la efectividad de los glucocorticoides inhalados para reducir la duración de los síntomas o prevenir la hospitalización o la mortalidad en pacientes ambulatorios con COVID-19 de gravedad leve a moderada.⁽⁸⁾

Es crucial que el personal médico encargado de atender a pacientes con COVID-19, como los profesionales involucrados en el estudio que se examina acerca de la evaluación de concordancia entre evaluadores utilizando el coeficiente Kappa de Fleiss, consideren la integración de la inteligencia artificial (IA) junto con herramientas estadísticas y matemáticas en su enfoque de casos de urgencia.

Los modelos de IA demuestran una precisión destacable en la predicción de la

mortalidad en pacientes con COVID-19, proporcionando un valor pronóstico significativo. Estos modelos incluyen técnicas como KNN (K-Vecinos más Cercanos), SVM (Máquinas de Soporte Vectorial), ANN (Redes Neuronales Artificiales), RF (Bosques Aleatorios), XGBoost, entre otros, que muestran niveles altos de precisión.⁽⁹⁾

La incorporación de la IA y el análisis estadístico avanzado en las decisiones médicas puede resultar sumamente beneficiosa para mejorar la atención de pacientes críticos, permitiendo una evaluación más precisa de la gravedad de la enfermedad y una mejor asignación de recursos y tratamientos. Esta sinergia entre la tecnología y la práctica médica puede contribuir significativamente a tomar decisiones basadas en evidencia y optimizar los resultados clínicos, especialmente en situaciones críticas como la pandemia de COVID-19. Además, existen valiosos avances en la investigación que exploran con éxito el uso de la Neutrosfía para estos fines.^(10,11,12)

Los autores quieren destacar que el coeficiente Kappa de Fleiss se utiliza con frecuencia en diferentes estudios dentro del ámbito de la salud. Por ejemplo, se emplea en una investigación actual que se lleva a cabo en Colombia para evaluar la validez y fiabilidad del Cuestionario de Adherencia al Tratamiento para Pacientes con Hipertensión Arterial (TAQPH).⁽¹³⁾ Asimismo, se encuentra en un estudio realizado en Francia donde se observa que las tasas de concordancia entre evaluadores, en el caso de cuatro examinadores, son de 0,65 (Fleiss kappa, IC del 95% [0,46, 0,90]) para la presencia de estructuras linfoides terciarias maduras (mTLS, por sus siglas en inglés) y de 0,90 para la madurez (IC del 95 % [0,83, 0,99]).^(14,15)

Limitaciones del estudio

Los investigadores señalan como una limitación de este estudio el uso de una

muestra reducida de la población, así como la utilización de la técnica de muestreo por conveniencia.

Conclusiones

Se identifica un nivel general de concordancia moderado a considerable entre médicos, internos y estudiantes de la carrera de Medicina al evaluar la gravedad de los pacientes con COVID-19 en el hospital de Ambato, Ecuador, mediante el coeficiente Kappa de Fleiss.

Aunque se observó un grado de acuerdo aceptable en la clasificación global, al considerar categorías específicas de gravedad, se detectaron diferencias significativas entre los evaluadores. Se destacó una alta concordancia, casi perfecta, para pacientes clasificados como Leve, mientras que para las categorías Moderado y Grave se evidenciaron niveles de concordancia moderada y considerable, respectivamente.

Estos resultados enfatizaron la importancia de establecer criterios de clasificación más precisos y uniformes entre los profesionales de la salud que participan en la evaluación de pacientes con COVID-19. La discrepancia encontrada en la evaluación de la gravedad puede tener implicaciones cruciales en la toma de decisiones clínicas y en la asignación de recursos, lo que resalta la necesidad de implementar estrategias para mejorar la consistencia en la evaluación de pacientes en contextos similares.

Referencias bibliográficas

1. Chérrez-Anguizaca J, León-Micheli E. La aplicación del Triage, en la prestación del servicio de salud en el Ecuador. CM [Internet]. 1 dic.2021 [citado

30sep.2023];7(3):98-23.

Available

from:

<https://www.cienciamatriarevista.org.ve/index.php/cm/article/view/572>

2. Jaramillo MN, Chuga ZN, Hernández CP, Lits RT. Análisis multicriterio en el ámbito sanitario: selección del sistema de triaje más adecuado para las unidades de atención de urgencias en Ecuador. *Rev Investig Oper.* 2022;43(3):316-324

3. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977 Mar;33(1):159-74. PMID: 843571.

4. Dayya D, O'Neill OJ, Feiertag TD, Tuazon-Boer R, Sullivan J, Perez L, Gurash S, Eaton M, Bodley T, Marker J, Smykowski E, Hall T. The use of oxygen hoods in patients failing on conventional high-flow oxygen delivery systems, the effects on oxygenation, mechanical ventilation and mortality rates in hypoxic patients with COVID-19. A Prospective Controlled Cohort Study. *Respir Med.* 2021 Apr;179:106312. <https://10.1016/j.rmed.2021.106312>.

5. Bresadola V, Biddau C, Puggioni A, Tel A, Robiony M, Hodgkinson J, Leo CA. General surgery and COVID-19: review of practical recommendations in the first pandemic phase. *Surg Today.* 2020 Oct;50(10):1159-1167. <https://10.1007/s00595-020-02086-4>.

6. Grieco DL, Menga LS, Cesarano M, Rosà T, Spadaro S, Bitondo MM, et al. Effect of Helmet Noninvasive Ventilation vs High-Flow Nasal Oxygen on Days Free of Respiratory Support in Patients With COVID-19 and Moderate to Severe Hypoxemic Respiratory Failure: The HENIVOT Randomized Clinical Trial. *JAMA.* 2021 May 4;325(17):1731-1743. <https://10.1001/jama.2021.4682>.

7. Diah W, Chabane M, Tourette C, Azbekyan A, Morelot-Panzini C, Hajjar LA, et al. Testing the efficacy and safety of BIO101, for the prevention of respiratory deterioration, in patients with COVID-19 pneumonia (COVA study): a structured summary of a study protocol for a randomised controlled trial. *Trials.* 2021 Jan

- 11;22(1):42. <https://10.1186/s13063-020-04998-5>.
8. Boulware DR, Lindsell CJ, Stewart TG, Hernandez AF, Collins S, McCarthy MW, et al. Inhaled Fluticasone Furoate for Outpatient Treatment of Covid-19. *N Engl J Med*. 2023 Sep 21;389(12):1085-1095. <https://10.1056/NEJMoa2209421>.
9. Xin Y, Li H, Zhou Y, Yang Q, Mu W, Xiao H, et al. The accuracy of artificial intelligence in predicting COVID-19 patient mortality: a systematic review and meta-analysis. *BMC Med Inform Decis Mak*. 2023 Aug 9;23(1):155. <https://10.1186/s12911-023-02256-7>.
10. Valenzuela-Chicaiza CV, Arciniegas-Paspuel OG, Carrera-Cuesta PY, Álvarez-Hernández SR. Neutrosophic Psychology for Emotional Intelligence Analysis in Students of the Autonomous University of Los Andes, Ecuador. *Neutrosophic Sets and Systems*. 2020; 34(1). https://digitalrepository.unm.edu/nss_journal/vol34/iss1/1
11. González Chico MG, Hernández Bandera N, Blacksmith Loop S, Laica Sailema N. Evaluación de la Relevancia de la Atención Médica Intercultural. Muestreo Neutrosófico. *Neutrosophic Sets and Systems*. 2021;44(1):46. Disponible en: https://digitalrepository.unm.edu/nss_journal/vol44/iss1/46
12. Jaramillo MN, Chuga ZN, Hernández CP, Lits RT. Análisis multicriterio en el ámbito sanitario: selección del sistema de triaje más adecuado para las unidades de atención de urgencias en Ecuador. *Rev Investig Oper*. 2022;43(3):316-324.
13. Esquivel Garzón N, Díaz Heredia LP. Validity and Reliability of the Treatment Adherence Questionnaire for Patients with Hypertension. *Invest Educ Enferm*. 2019 Sep;37(3):e09. <https://10.17533/udea.iee.v37n3e09>.
14. Vanhersecke L, Bougouin A, Crombé A, Brunet M, Sofeu C, Parrens M, et al. Standardized Pathology Screening of Mature Tertiary Lymphoid Structures in Cancers. *Lab Invest*. 2023 May;103(5):100063. <https://10.1016/j.labinv.2023.100063>.

15. Ortiz-Rosales A, Vázquez-Céspedes M, Cedeño-Perdomo MY, Rodas-Belmonte SA, Araujo-Alvarez A. Incremento de conocimientos sobre COVID-19 en una comunidad del municipio Caroní, Venezuela. Rev Inf Cient [Internet]. 2022;101(2). Disponible en: <https://revinfcientifica.sld.cu/index.php/ric/article/view/3845>