

Universidad de Ciencias Médicas de La Habana

Escuela Nacional de Salud Pública

**Modelos predictivos para la estimación y clasificación del
riesgo de muerte en COVID-19**

Tesis para optar por el Grado Científico de Doctor en Ciencias Médicas

Dr. Maicel Eugenio Monzón Pérez, M. Sc.

Autor

La Habana 2024

Universidad de Ciencias Médicas de La Habana

Escuela Nacional de Salud Pública

**Modelos predictivos para la estimación y clasificación del
riesgo de muerte en COVID-19**

Tesis para optar por el Grado Científico de Doctor en Ciencias Médicas

Dr. Maicel Eugenio Monzón Pérez, M. Sc.

Autor

Tutores

Dra. Cs. Lizet Sánchez Valdés

Dr. Cs. Agustín Lage Dávila

Dr. C. Frank Daniel Martos Benítez

La Habana 2024

"La simplicidad es la máxima sofisticación"

Leonardo da Vinci

Agradecimientos

Mi más profundo agradecimiento a mis tutores de tesis, el Dr. Cs. Agustín Lage Dávila y a la Dra. Cs. Lizet Sánchez Valdés, también al Dr. C. Frank Daniel Martos. Su paciencia, conocimiento y dedicación han sido invaluable para este trabajo. Agradezco a mis profesoras la Dra. C. María del Carmen Pria Barros, la Dra. C. María Esther Álvarez Lauzarique y la Dra. C. Ileana Castañeda Abascal por su gentileza y aportes para la mejora de mi investigación. A mi amiga la MCs. Tania Sánchez por la revisión de este documento. A mis colegas y amigos del grupo de WhatsApp “Aula de Bioestadística” que me han ayudado en la investigación, como el Dr. Cs. Jorge Bacallao Gallestey, el Dr. C. Héctor Bayarre Vea, quienes me ofrecieron sus recomendaciones para mejorar este trabajo. A mis antiguos y nuevos compañeros de trabajo por su apoyo y colaboración, MCs. Gladys Giménez, Dra. C. María Acelia Miragaya, MCs. Liliam Quelle Santana y, en especial, a mi amiga la MCs. Amaylid Arteaga García.

Mi familia merece un reconocimiento especial, a mi padre Eugenio, mi madre Maira, mi hermano Maikel, mis hijos Melanie y Gabriel y mi esposa Ivette. Su amor incondicional, apoyo constante y fe inquebrantable en mis capacidades han sido mi fortaleza durante este viaje.

Finalmente, quiero agradecer a todas las instituciones y personas que han hecho posible esta investigación.

Este logro es tanto mío como de todos los que me han apoyado en este camino.

A todos ustedes, mi más sincero agradecimiento.



Dr. Maicel Eugenio Monzón Pérez

SÍNTESIS

Introducción: Los modelos predictivos mejoran el control de la transición del estado del paciente a la gravedad en la COVID-19. Esta investigación aborda la ausencia de modelos para la estimación y la clasificación del riesgo de muerte en la población cubana, y el desarrollo y validación de estos modelos.

Objetivos: Desarrollar modelos predictivos de muerte para pacientes con COVID-19 atendidos en la comunidad y el hospital, validar los modelos en pacientes cubanos y diseñar un nomograma y una aplicación para móvil con el modelo seleccionado.

Diseño metodológico: Se realizó un estudio de cohorte retrospectiva para desarrollar un modelo para el escenario comunitario, así como un estudio de casos y controles para diseñar dos modelos para hospitales. Se realizó una validación de los modelos a través de métricas de calibración y discriminación. Se comparó el desempeño de los modelos propuestos con modelos internacionales. Finalmente se diseñó un nomograma y programó una aplicación para ejecución del modelo seleccionado en JavaScript, HTML y CSS.

Resultados: Los modelos propuestos demostraron buena fiabilidad de las predicciones con cifras cercanas a cero en el intercepto, y uno en la pendiente de la recta de calibración, así como capacidad de diferenciación del resultado con valores superiores a 0,80 en el área bajo la curva.

Conclusiones: Los modelos diseñados son válidos para estimar y clasificar el riesgo de muerte por COVID-19 y mostraron buena capacidad de diferenciación del resultado y fiabilidad de las predicciones. Se diseñó un nomograma y una aplicación para facilitar el uso de COVIDCENCEC.

Tabla de contenido

INTRODUCCIÓN	1
Contexto histórico	1
Problema científico	6
Preguntas científicas	6
Objetivos	6
Novedad científica	7
Actualidad	8
Pertinencia.....	8
Aportes	8
Metodología general del estudio	10
Socialización de los resultados de la tesis.....	10
CAPÍTULO I. ESTADO EN LA INVESTIGACIÓN SOBRE MODELOS PREDICTIVOS PARA LA ESTIMACIÓN Y CLASIFICACIÓN DEL RIESGO DE MUERTE EN COVID-19.....	11
I.1 Generalidades	11
I.2. Producción científica, fuentes clave, autores y documentos relevantes .	12
I.3 Principales modelos de predicción clínica para la COVID-19	17
I.3.1 Sesgos en los modelos de predicción clínica para la COVID-19	18
I.3.2 Modelos de mayor calidad	19
I.3.3 Variables de resultado.....	21
I.3.4 Variables predictoras.....	22
I.3.5 Tipo de modelo	24
I.4 Estrategias de modelado	26
1.4.1 Definición del problema e inspección de datos	27
1.4.2 Codificación de las variables predictoras	28
1.4.3 Especificación del tipo de modelo	29
1.4.4 Estimación del modelo	31
1.4.5 Evaluación del rendimiento del modelo	34
1.4.6 Evaluación de la validez del modelo	35
1.4.7 Presentación del modelo	38
I.5 Consideraciones sobre el desarrollo y validación del software	39
I.6 Generalización y extrapolación de los modelos a escenarios de atención médica en Cuba.....	40
I.6.1 Limitaciones para la generalización de modelos	40
I.6.2 Fuentes de datos disponibles para la investigación	42
I.6.3 El contexto cubano como escenario para la generalización.....	43
I.7 Consideraciones finales.....	44

CAPÍTULO II. DISEÑO METODOLÓGICO	45
II.1 Generalidades	45
II.2 Tipo de estudio y contexto	46
II.2.1 Primera etapa del estudio	47
II.2.2 Segunda etapa del estudio.....	48
II.3 Universo y muestra	49
II.3.1 Muestreo para primera etapa de la investigación.....	49
II.3.2 Muestreo para la segunda etapa de la investigación	49
II.4 Operacionalización de variables	51
II.4.1 Procedimientos para la confección del Índice de riesgo ponderado de síntomas (IRPS).....	52
II.5 Técnicas y procedimientos.....	52
II.5.1 Técnicas y procedimientos de obtención de información	52
II.5.2 Técnicas de procesamiento y análisis de la información.....	53
II.5.3 Técnicas y procedimientos de presentación de los resultados	57
II.6 Aspectos éticos.....	58
II.7 Fortalezas y limitaciones metodológicas.....	58
II.8 Consideraciones finales.....	60
CAPÍTULO III. RESULTADOS	62
III.1 Desarrollo de modelos para escenarios de atención médica de la comunidad	62
III.2 Desarrollo de modelos para escenarios hospitalarios	67
III.2.2.1 Desarrollo del modelo árbol de alto riesgo AAR.....	70
III.2.2.2 Desarrollo del modelo COVIDCENCEC	73
III.3 Validación de modelos propuestos en distintos escenarios.....	74
III.3.1 Validación de modelos para pacientes de la comunidad.....	74
III.3.2 Validación de modelos para escenarios hospitalarios.....	75
III.4 Contraste entre COVIDCENCEC y modelos de mayor calidad	77
III.5. Presentación e implementación en la práctica de los modelos.....	79
III.5.2.1 Nomograma	79
III.5.2.2 Diseño de COVIDCENCEC-APK	79
III. 6 Consideraciones finales.....	80
CAPÍTULO IV. DISCUSIÓN	81
IV.1 Modelos para escenarios de atención médica de la comunidad	81
IV.3 Modelos para escenarios de atención médica en hospitales	85
IV.4 Estrategia de aplicación de modelos de predicción clínica en la gestión de la transición a la gravedad en Cuba	94
IV.5 Consideraciones finales	98

CONCLUSIONES.....	99
RECOMENDACIONES	100
REFERENCIAS BIBLIOGRÁFICAS	101
ANEXOS	i
Anexo 1. Participación en eventos científicos y publicaciones a partir de los resultados de la tesis	i
Anexo 2. Principales recursos de información, autores y artículos sobre el tema según la base de datos Dimension	ii
Anexo 3. Principales conceptos de los trabajos seleccionados	iii
Anexo 4. Escenarios de atención médica en Cuba para el enfrentamiento a la COVID-19 y su relación con la estratificación del riesgo	iv
Anexo 5. Relación entre modelo de mayor calidad y sus predictores.....	v
Anexo 6. Indicadores de estrato	vi
Anexo 7. Variaciones temporales a lo largo de la pandemia en Cuba.....	ix
Anexo 8. Versiones del PAN y frecuencia de casos por versiones.....	x
Anexo 9. Variables disponibles en RNPC.....	xi
Anexo 10. Variables disponibles en RICPA	xii
Anexo 11. Extrapolación y comparación de los modelos.....	xiii
Anexo 12. Estrategias para la identificación y mitigación del sesgo	xiv
Anexo 13. Determinación del tamaño de muestra según criterios y parámetros seleccionados	xvi
Anexo 14. Operacionalización de las variables según etapas de la investigación	xvii
Anexo 15. Índice de riesgo ponderado de síntomas (IRPS)	xxii
Anexo 16. Frecuencia de datos perdidos en pacientes hospitalizados.....	xxiii
Anexo 17. Factor de inflación de varianza del ajuste para el modelo COVIDCENCEC	xxiv
Anexo 18. Estrategia para controlar las interacciones.....	xxv
Anexo 19. Ecuación del modelo COVIDCENCEC	xxvi
Anexo 20. Métricas de rendimiento del modelo ABR según partición de datos	xxvii
Anexo 21. Métricas de calibración y discriminación del modelo AAR según partición de datos	xxviii
Anexo 22. Métricas de calibración y discriminación del modelo COVIDCENCEC según partición de datos	xxix
Anexo 23. Rendimiento de los modelos propuestos en el estudio e identificados como de mayor calidad.....	xxx
Anexo 24. Nomograma del modelo COVIDCENCEC	xxxi
Anexo 25. Interfaz de la aplicación COVIDCENCEC-APK	xxxii
Anexo 26. Informe final de validación del software COVIDCENCEC-APK	xxxiii

GLOSARIO DE TÉRMINOS Y ABREVIATURAS

COVIDPRECISE: es un consorcio internacional de expertos que evalúa sistemáticamente el estado de las investigaciones sobre modelos de predicción clínica para la COVID-19. La base de datos del sitio incluye información sobre artículos publicados y pre-impresos que abordan el desarrollo o validación de modelos diagnósticos y pronósticos específicos para la COVID-19 o las actualizaciones de modelos predictivos conocidos para otras enfermedades aplicados a este problema de salud. La información se encuentra disponible en el sitio web <https://www.covprecise.org> .

COVIDCENCEC-APK: se trata de una calculadora de riesgo para ejecutar el modelo COVIDCENCEC. Esta herramienta fue desarrollada para el sistema operativo Android y se adhiere al paradigma Open Surce. El código fuente de este software está disponible en el perfil de Reserch Gate de este autor y se puede descargar a través de la siguiente dirección electrónica: DOI:[10.13140/RG.2.2.12842.84160](https://doi.org/10.13140/RG.2.2.12842.84160) .

Escenarios de atención médica: se trata de los diferentes escenarios de atención médica en que se organizó el Sistema Nacional de Salud cubano para hacer frente a la COVID-19. Se distinguen dos tipos de escenarios fundamentales, el escenario comunitario o de APS que incluye el hogar y los centros de aislamientos. En el escenario hospitalario se incluyen hospitales destinados o no al manejo de la COVID-19, Unidades de Vigilancia Intensiva, Unidades de Cuidados Intermedios y Unidades de Cuidados Intensivos. El tipo y distribución de escenarios durante toda la pandemia sufrió modificaciones en el tiempo y en distintos entornos geográficos.

Modelos de mayor calidad: son modelos de predicción clínica identificados por las revisiones sistemáticas llevada a cabo por el grupo COVIDPRECISE como de mayor calidad. Los criterios para definir la calidad de un modelo se basan en atributos como la buena calibración, fuerte discriminación, utilidad clínica, facilidad de generalización, transparencia en sus publicaciones y disponibilidad.

Paciente de bajo riesgo: son pacientes que tienen un riesgo bajo de transición de la gravedad. Este grupo puede ser manejado de forma ambulatoria o en escenarios comunitarios. Estos pacientes suelen requerir seguimiento clínico, aislamiento en domicilio y medidas de apoyo sintomático, sin necesidad de hospitalización. También es posible definir pacientes de bajo riesgo en escenarios hospitalarios. En este caso, son pacientes con baja probabilidad de muerte.

Pacientes de alto riesgo: son pacientes que tienen un riesgo alto de transición en su estado clínico a la gravedad. Según el contexto, estos pacientes pueden requerir hospitalización y manejo en unidades especializadas. En el contexto hospitalario, es posible que estas personas puedan necesitar cuidados intensivos, monitorización continua y tratamiento específico para complicaciones asociadas a la COVID-19.

Pacientes de la comunidad: pacientes que en el momento del estudio recibieron atención en escenarios comunitarios.

Pacientes hospitalizados: pacientes que en el momento del estudio recibieron atención en escenarios hospitalarios.

Patrones de atención médica: modelo de tratamiento que se regulariza en un grupo de pacientes de características similares en el Protocolo de actuación

nacional para la COVID-19 en Cuba (PAN). Se incluyen las regularidades diagnósticas y las terapéuticas.

Protocolo de actuación nacional: es un documento oficial elaborado por el Ministerio de Salud Pública de Cuba (Minsap), que tiene como objetivo establecer pautas, procedimientos y recomendaciones para la prevención, el control y manejo de la COVID-19. Este protocolo sirve como una guía integral para los profesionales de la salud, autoridades sanitarias y la población en general, con el fin de coordinar acciones para hacer frente a la pandemia en Cuba.

Transición hacia la gravedad: proceso en el que un paciente contagiado desarrolla formas graves de la enfermedad. Se considera la muerte como desenlace de mayor gravedad.

Triage: es el proceso utilizado para evaluar, clasificar y priorizar a los pacientes según el riesgo de muerte y la urgencia de su tratamiento. El objetivo del triaje es identificar a los pacientes de alto riesgo que necesitan mayor atención y luego intensificar la conducta médica, al mismo tiempo, detectar a aquellos con bajo riesgo que pueden beneficiarse de una atención más conservadora. Además, el triaje permite hacer un uso eficiente de los recursos de salud limitados en momentos de crisis sanitaria.

Abreviaturas en orden alfabético

Abreviaturas	Significado
4C MORTALITY	Coronavirus Clinical Characterisation Consortium Mortality
AAR	Árbol clasificador de pacientes de alto riesgo de muerte por COVID-19
ABR	Árbol clasificador de pacientes de bajo riesgo de muerte por COVID-19
APP	Antecedentes patológicos personales
AUC	Area under the curve
CCA	Estudio de casos y controles anidados
CENCEC	Centro Nacional Coordinador de Ensayos Clínicos
CIA	Criterio de información de Akaike
Cont.	Conteo absoluto de células
COVIDCENCEC	Modelo para la estimación y clasificación del riesgo de muerte en pacientes hospitalizados con COVID-19 del CENCEC
Cr	Creatinina sérica
CRD	Cuaderno de recogida de datos
CSS	Cascading Style Sheets
CTREE	Conditional Inference Trees
DE	Desviación estándar
DSS	Determinantes sociales de la salud
Dxy	Correlación de rango de D de Somers entre x e y
EC	Enfermedades cardiovasculares
ECV	Enfermedad cerebrovascular
EE	Error estándar
EIC	Enfermedad isquémica del corazón
ENSAP	Escuela Nacional de Salud Pública
EPOC	Enfermedad pulmonar obstructiva crónica
EPP	Eventos por parámetro predictivo
ERC	Enfermedad renal crónica
FC	Frecuencia cardiaca
FIV	Factor de inflación de varianza
FR	Frecuencia respiratoria
GCS	Glasgow Coma Scale
HTML	HyperText Markup Language
IA	Inteligencia artificial
IAAS	Infección asociada a la asistencia de salud
IC	Intervalo de confianza
IEC	International Electrotechnical Commission
IMC	Índice de masa corporal
INL	Índice neutrófilo linfocito
IRPS	Índice de riesgo ponderado de síntomas
ISO	International Organization for Standardization
LASSO	Least Absolute Shrinkage and Selection Operator

LAT	Latido
LDH	Lactato deshidrogenasa
LR-	Razón de verosimilitud negativa
LR+	Razón de verosimilitud positiva
MC	Media de citas
Me	Mediana
ME	Mortalidad explicada
MIN	Minutos
Minsap	Ministerio de Salud Pública
NEWS 2	National Early Warning Score 2
OMS	Organización Mundial de la Salud
OR	Odds Ratio
PAM	Presión arterial media
PAN	Protocolo de actuación nacional para la COVID-19 en Cuba
PCP	Proporción de citas de campo
PCR	Proteína C reactiva
PLT	Conteo de plaquetas en células por 10 ⁹
PRIEST	Pandemic Respiratory Infection Emergency System Triage
R ²	Coeficiente de determinación
RCP	Reacción en cadena a la polimerasa
RESP	Respiraciones
RF	Random Forest
RIC	Rango intercuartílico
RICPA	Registro del proyecto del Minsap “Impacto de la COVID-19 en pacientes adultos en Cuba (2020-2023)”
RL	Regresión logística binaria múltiple
RM	Riesgo de mortalidad
RNPC	Registro nacional de pacientes confirmados de COVID-19 del Minsap
ROC	Receiver Operating Characteristic Curve
RPART	Recursive Partitioning and Regression Trees
RS	Riesgo de severidad
SDRA	Síndrome de distrés respiratorio del adulto
SE	Severidad explicada
SEG	Segundos
SNS	Sistema Nacional de Salud
SpO2	Saturación de oxígeno capilar periférico
T	Temperatura corporal
TA	Tensión arterial
TRE	Tamaño relativo del estrato
UCI	Unidades de Cuidados Intensivos
VPN	Valor predictivo negativo
VPP	Valor predictivo positivo
VSG	Velocidad de sedimentación globular

INTRODUCCIÓN

La pandemia de COVID-19, iniciada a finales de 2019, se convirtió en uno de los mayores desafíos de salud pública del siglo XXI. La COVID-19 ha tenido efectos negativos en la sociedad, la economía y la salud ⁽¹⁾. Según la Organización Mundial de la Salud (OMS), se han registrado aproximadamente 5,42 millones de muertes atribuibles directamente a la COVID-19 hasta diciembre de 2021. Además, el exceso total de mortalidad asociado a la enfermedad se estima en 14,83 millones a nivel global ⁽²⁾.

Durante la pandemia, los tres componentes más importantes para la reducción de la mortalidad fueron el control de la transmisión del virus, la prevención de la transición hacia la gravedad y el tratamiento efectivo de la enfermedad grave ^(2,3). La calidad en la estimación y clasificación del riesgo es fundamental para el control de la transición hacia la gravedad. Una metodología diseñada para este propósito genera y utiliza conocimientos que resultan difíciles de obtener completamente a través de la literatura científica. La razón principal de esta dificultad radica en el carácter contextual de la información. Por ejemplo, se conoce que durante la epidemia de H1N1 ⁽⁴⁾, escalas de uso común para el triaje de pacientes en unidades de cuidados intensivos (UCI) como SOFA ⁽⁵⁾ y APACHE II ⁽⁶⁾ resultaron inútiles para predecir la muerte en estos pacientes ⁽⁷⁾. En este caso, se aplicó a una enfermedad que en su momento presentaba características nuevas. Además, en Cuba, la conducta médica que se adoptó fue diferente a la del resto del mundo ⁽²⁾.

Contexto histórico

La COVID-19 fue una enfermedad nueva porque, a pesar de los antecedentes de infecciones humanas por otros coronavirus como el SARS-CoV ⁽⁸⁾, en esta

zoonosis causada por el virus SARS-CoV-2 ⁽⁹⁾, la presentación y el curso clínico de la enfermedad, así como las posibles complicaciones y la gravedad se desconocían en sus primeras etapas ⁽¹⁰⁾. Es importante destacar que, al inicio de la pandemia, había gran incertidumbre acerca de cómo responderían los pacientes a la conducta médica en general y, en particular, a los tratamientos farmacológicos producidos por la industria biotecnológica cubana ⁽¹¹⁾. En la COVID-19, el cuadro clínico fue muy heterogéneo, con una evolución muy variable que abarcó desde pacientes asintomáticos hasta aquellos que evolucionaron hacia la gravedad ⁽¹²⁾. El grupo con peor pronóstico, al que se debía orientar el tratamiento más intensivo, solía ser reducido ⁽³⁾. No fue conveniente tratar a todos los pacientes de la misma manera, ya que solo una pequeña fracción de los confirmados (aproximadamente el 2 %) evolucionaron hacia la gravedad ⁽²⁾. Además, las decisiones para el triaje se debían tomar lo más rápido posible, ya que los pacientes con COVID-19 podía evolucionar a formas graves en un corto tiempo ^(3,13). La heterogeneidad clínica de la COVID-19 es el resultado de un proceso sindémico complejo, donde factores biológicos y estructurales interactúan sinérgicamente ⁽¹⁴⁾. Los determinantes económicos, sociales, culturales y demográficos, junto con las condiciones de vida, comportamientos individuales y la organización de los sistemas de salud, también desempeñan un papel fundamental. Las desigualdades sociales aumentan la vulnerabilidad de ciertos grupos ⁽¹⁵⁾. La evolución de la enfermedad está influenciada por decisiones clínicas, características del tratamiento y acceso a los servicios de salud ^(14,15). Estos elementos pueden variar significativamente de un país a otro o incluso dentro de una misma zona geográfica ⁽¹⁶⁾.

En Cuba, las estrategias de enfrentamiento a la enfermedad tuvieron características diferentes en comparación con el resto del mundo. Por ejemplo, se abogó por el ingreso temprano de todos los pacientes, se confirmó la enfermedad por la prueba de Reacción en Cadena de la Polimerasa (RCP) o el diagnosticador de fabricación cubana Umelisa SARS-COV-2 IgG, y se utilizó el índice neutrófilo linfocito (INL) para clasificar a los enfermos ⁽²⁾. A lo largo del tiempo, el manejo de la COVID-19 sufrió transformaciones en los patrones de atención médica y no todas las etapas de la epidemia fueron iguales. Hasta enero de 2021, las autoridades realizaron seis actualizaciones del Protocolo de Actuación Nacional para la COVID-19 en Cuba (PAN) ⁽¹⁷⁾, lo que pautó cambios en la conducta médica ⁽¹⁷⁻¹⁹⁾.

Antecedentes

La pandemia de la COVID-19 sobrepasó los servicios de salud a nivel global. De acuerdo con la OMS ⁽²⁰⁾, más de la mitad de los países experimentaron disrupciones significativas en la atención médica. Cuba no fue la excepción en el enfrentamiento a esta enfermedad, por lo que fue necesario implementar soluciones para evitar el colapso de los servicios, especialmente en la UCI. Hasta el 29 de octubre de 2022, en el país se confirmaron 1 111 281 pacientes y se reportaron 8 530 muertes ⁽²⁾. Con la implementación de la versión 1.5 del PAN en junio de 2020 ⁽¹⁸⁾, inició el empleo de modelos de predicción clínica para la estratificación del riesgo de transición hacia la gravedad ⁽³⁾. Dichos modelos perfeccionaron los procedimientos de derivación de los pacientes hacia los distintos escenarios de atención médica. Una versión anterior de uno de los modelos propuestos en este estudio, conocido como el Árbol de bajo riesgo (ABR), se presentó ante el Comité de Innovación del Minsap. Este modelo, entre

otros análisis, contribuyó a la definición de los pacientes de bajo a de alto riesgo en el PAN y a la toma de decisiones para el enfrentamiento de la pandemia ^(3,17,18). Los modelos de predicción clínica sirvieron como base para el desarrollo de una metodología de clasificación de riesgo de transición hacia la gravedad en el contexto de la atención médica en Cuba durante la COVID-19. Estos modelos proporcionaron una clasificación del riesgo de muerte y, tras ser sometidos a pruebas rigurosas, evidenciaron su robustez y precisión ⁽³⁾. Durante la epidemia de COVID-19 en Cuba, se desarrollaron otros instrumentos para el pronóstico de pacientes con la enfermedad. Herrera y colaboradores ⁽²¹⁾ propusieron un índice pronóstico de gravedad basado en una cohorte de 150 pacientes en Villa Clara entre marzo y junio de 2020. Además, García Álvarez y colaboradores aplicaron un modelo cubano para predecir la muerte en pacientes graves por COVID-19 en Lombardía, Italia ⁽²²⁾. Sin embargo, estos estudios solo tuvieron en cuenta los datos del inicio de la pandemia, se realizaron con un número reducido de pacientes y no fueron sometidos a una validación externa rigurosa.

El interés en los modelos de predicción clínica creció de manera notable durante la pandemia de COVID-19 ⁽²³⁾. En 2019, se registraron 10 753 publicaciones sobre este tema en la base de datos “Dimensions” ⁽²⁴⁾ que incluían los términos “prediction model” o “prognostic model” en el título o el resumen. En 2020 y 2021, se publicaron diariamente un promedio de 41 y 55 artículos, respectivamente. El año 2022 marcó un hito con el mayor número de publicaciones hasta la fecha ⁽²⁵⁾. Este crecimiento reflejó la importancia de los modelos de predicción clínica en la lucha contra la COVID-19. El grupo COVIDPRECISE ⁽²⁶⁾ (acrónimo de, Precise Risk Estimation to optimise COVID-19 Care for Infected or Suspected patients in diverse sEttings) llevó a cabo una revisión sistemática de más de

126 000 artículos relacionados con modelos predictivos para la COVID-19, en la cual se evaluó la calidad de artículos que daban a conocer modelos de predicción clínica para la COVID-19 ⁽²⁷⁾. El estudio determinó que la mayoría de los modelos publicados no eran recomendables para un uso generalizado en la práctica clínica debido a la falta de información y un alto riesgo de sesgo en sus publicaciones ^(25,27). Sin embargo, como resultado de la revisión de COVIDPRECISE ⁽²⁵⁾ se identificaron cinco modelos destacados por poseer un conjunto de atributos de calidad (modelos de mayor calidad, en lo adelante) ^(25,27). En sus orientaciones del 25 de enero de 2021 para el manejo clínico de la COVID-19 ⁽²⁸⁾, la OMS no recomendó la utilización de modelos predictivos en la práctica clínica. Sin embargo, esta institución estimuló la validación en el contexto de cada país de los modelos de mayor calidad identificados por COVIDPRECISE ⁽²⁵⁾.

Justificación

Los modelos de predicción clínica son una tecnología valiosa porque hace más eficiente la atención del paciente y optimiza la utilización de recursos de salud ^(3,25,27). Aunque se han identificado modelos de mayor calidad ^(25,27), su aplicación generalizada requiere una validación externa en Cuba y la consideración de las circunstancias individuales y contextuales ^(29,30). También se podría proponer el desarrollo de modelos cubanos contextualizados a la organización de los sistemas y servicios de salud del país como los que ya se venían aplicando a partir de la versión 1.5 del PAN ^(3,18). Ante la persistencia de condiciones como el deterioro ambiental, la superpoblación y la alta movilidad que produce la globalización, se reconoce la posibilidad de una nueva pandemia y la aparición de enfermedades emergentes y reemergentes ⁽³¹⁾. Por ende, es esencial equipar

al país con metodologías que se hayan desarrollado, debatido y validado previamente, en previsión de la potencial aparición de una nueva crisis sanitaria provocada por coronavirus u otro agente biológico ⁽³²⁾. Además, es necesario implantar modelos de predicción clínica en el marco de la atención médica en Cuba para perfeccionar el PAN ⁽¹⁹⁾. Adicionalmente, esta investigación puede servir como base para la extrapolación de enfoques metodológicos para el triaje en otros problemas de salud.

Problema científico

Hasta donde conoce el autor, no existen modelos de predicción clínica para la estimación y clasificación del riesgo de muerte por COVID-19 específicos para la población cubana en los distintos escenarios de atención médica, por lo que es necesario desarrollar y validar estos modelos en Cuba. Esta acción permitiría una atención eficiente y personalizada desde el inicio de la enfermedad, así como optimizar los recursos de salud.

Preguntas científicas

¿Qué modelos de predicción clínica se podrían utilizar como fundamento para estimar y clasificar el riesgo de muerte por COVID-19 en los distintos escenarios de atención médica en Cuba?

¿Cuál es la validez de los modelos predictivos propuestos?

Objetivos

- Desarrollar modelos predictivos de muerte para pacientes confirmados de COVID-19 atendidos en escenarios de atención médica comunitarios y hospitalarios.
- Validar los modelos predictivos desarrollados.
- Diseñar un nomograma y una aplicación móvil para la predicción del riesgo de muerte en pacientes hospitalizados.

Novedad científica

La novedad radica en el desarrollo de modelos para la estimación y clasificación del riesgo de muerte específicos para la COVID-19. El primer modelo es una actualización del árbol de bajo riesgo (ABR) para escenario de atención médica comunitaria. Una versión anterior de esta herramienta se utilizó para la estratificación del riesgo en el PAN y para la derivación de pacientes hacia los distintos escenarios de atención médica. En este estudio, además de la actualización de ABR, se proponen otros dos modelos para su uso en escenarios hospitalarios (Árbol de alto riesgo [AAR] y COVIDCENCEC). En general, el sistema pronóstico que se desarrolló puede favorecer la reducción de la mortalidad a través de la prevención de la transición hacia la gravedad con herramientas que mejoran la calidad de la estratificación.

Se realizó una innovación que consistió en añadir una nueva capa de análisis a las reglas de clasificación que se obtienen de los modelos de árboles de clasificación. Esta metodología consiste en incorporar tres nuevos indicadores que se pueden emplear para realizar intervenciones en función del riesgo de muerte en grupos de pacientes específicos. Este método ofrece una forma novedosa de interpretar y aplicar los resultados en los modelos de predicción clínica que, al parecer, no se ha descrito en la literatura hasta este momento.

Si bien el modelado predictivo mediante algoritmos de aprendizaje automático se utiliza desde hace tiempo, la estrategia y los resultados específicos para el enfrentamiento a la COVID-19 en Cuba son inéditos porque dependen del contexto. Este enfoque puede ser provechoso para países en vías de desarrollo, donde la adaptación o creación de modelos de predicción clínica se enfrenta a desafíos adicionales derivados de las limitaciones de recursos.

Actualidad

Este trabajo, de relevancia por satisfacer una necesidad social, desarrolla modelos de predicción clínica para la COVID-19 en respuesta a una prioridad nacional. Los modelos y enfoque metodológico que se emplean en este estudio resultan útiles para hacer frente a futuras crisis sanitarias, abordar otros problemas de salud y pueden contribuir al perfeccionamiento del PAN.

Pertinencia

La investigación se inscribe en el proyecto del Minsap titulado ⁽³³⁾: “Impacto de la COVID-19 en pacientes adultos en Cuba. 2020-2023” (ICPA, por sus siglas). Cumple con el Lineamiento 96 de la Política Económica y Social del Partido y la Revolución ⁽³⁴⁾, que busca mejorar la atención médica y la eficiencia de los recursos.

Aportes

Los aportes del estudio son:

1) Sistema de modelos de predicción clínica para la estimación y clasificación del riesgo de muerte en COVID-19 (ABR, AAR y COVIDCENCE). Los modelos propuestos favorecen una mejor atención del paciente, y contribuyen a proteger al sistema de salud de un eventual colapso por saturación de la atención médica. El modelo ABR para escenario de atención médica comunitaria puede contribuir a brindar una atención personalizada a cada paciente y remitirlos a un hospital solo si presentan un alto riesgo de muerte al inicio de la enfermedad. Asimismo, ABR se puede utilizar para guiar el seguimiento clínico por la APS y las medidas de apoyo en casa. Los modelos AAR y COVIDCENCE tienen la capacidad de potenciar el triaje en escenarios de atención médica hospitalaria, identificar a pacientes de alto riesgo que podrían requerir cuidados especiales y tratamiento específico contra la COVID-19.

2) Herramientas de implementación de modelos de predicción clínica (nomograma y software COVIDCENCEC-APK). Se aportan herramientas de implementación inmediata y que podría utilizar el personal de salud para estimar el riesgo de muerte del paciente a partir del modelo COVIDCENCEC. El uso de esta tecnología permite la accesibilidad y su facilidad de uso.

4) Indicadores para la estratificación de riesgo a partir de árboles de clasificación. Se aporta un conjunto de indicadores que se adicionan a las reglas de clasificación de los árboles de clasificación. El aporte permite realizar intervenciones en función de las características específicas de los diferentes grupos de pacientes. Las intervenciones pueden consistir en la derivación de pacientes hacia los distintos escenarios de atención médica, la modulación de la conducta y el tratamiento médico, así como la gestión de recursos de salud.

4) Criterios sobre la validez de modelos foráneos en la población cubana. Se establecieron criterios para evaluar la “transportabilidad” de los modelos a escenarios de atención médica hospitalarios en Cuba. El ajuste de los modelos a los datos de este estudio permitió evaluar la generalización de los modelos de mayor calidad. La comparación de modelos ayudó a elegir aquellos que mejor se adaptan a los escenarios de atención médica en Cuba.

6) Aplicación de la inteligencia artificial en el ámbito de las ciencias médicas y las ciencias de la salud pública en Cuba. La utilización de algoritmos de aprendizaje supervisado como RPART, CTREE y la regresión logística se puede considerar como una aplicación válida y relevante de la inteligencia artificial para el enfrentamiento de la pandemia de COVID-19 en Cuba. Este estudio se alinea con los principios del aprendizaje automático y

contribuye a la aplicación de la inteligencia artificial en la práctica diaria de las ciencias médicas y en la salud pública.

7) Metodología para el enfrentamiento a futuras pandemias o aplicación a otros problemas de salud. Se propone una metodología para el triaje que puede contribuir a preparar al país ante futuras crisis sanitarias o ser aplicada a distintos problemas de salud. El enfoque metodológico de este estudio se puede aplicar en circunstancias donde la demanda de intervenciones de salud excede la oferta disponible. En particular, se pudiera utilizar en situaciones donde se necesita la estandarización de procesos, realizar una clasificación rápida y eficiente para establecer prioridades, en la optimización del flujo de pacientes, en la priorización de recursos limitados o dar respuestas a situaciones de emergencia como epidemias, pandemias o desastres naturales o causados por el hombre.

Metodología general del estudio

Este estudio se realiza en dos fases: en ambas se desarrollan y validan modelos predictivos para la estratificación del riesgo de muerte en pacientes con COVID-19. La primera fase se enfoca en los pacientes confirmados en escenarios comunitarios y la segunda etapa en pacientes ingresados en hospitales. Los modelos resultantes demuestran ser eficaces en la predicción de muerte en pacientes con COVID-19. Se diseña un nomograma según la metodología propuesta por Harrell ⁽³⁰⁾ y una aplicación para teléfonos móviles o tabletas para la estimación del riesgo de muerte por COVID-19 a partir de un modelo seleccionado.

Socialización de los resultados de la tesis

Los resultados del proceso investigativo se presentaron en diversos eventos y en publicaciones científicas que se relacionan en el anexo 1.

CAPÍTULO I. ESTADO EN LA INVESTIGACIÓN SOBRE MODELOS PREDICTIVOS PARA LA ESTIMACIÓN Y CLASIFICACIÓN DEL RIESGO DE MUERTE EN COVID-19

El presente capítulo tiene como objetivo describir la fundamentación teórica que permita entender y justificar cada uno de los objetivos planteados en la tesis. Se presentan las cuestiones fundamentales sobre el estado de la investigación en el desarrollo y validación de modelos de predicción clínica para la estimación y clasificación del riesgo de muerte en pacientes con COVID-19.

Además, se expone el soporte teórico que sustenta el diseño metodológico empleado en el estudio, así como las justificaciones de las decisiones metodológicas adoptadas a lo largo de este. Finalmente, se realiza una valoración crítica sobre la manera en que estos modelos se aplican en la práctica y su generalización en el contexto de los escenarios de atención médica en Cuba.

I.1 Generalidades

Los modelos de predicción clínica son modelos matemáticos que establecen una relación entre las características del paciente y un evento de interés, y que se utilizan para estimar la probabilidad de que dicho evento ocurra en pacientes similares ^(29,30,35). Estas herramientas se pueden emplear para el diagnóstico o el pronóstico de un problema de salud ⁽²⁹⁾. A pesar de que la estructura de estos modelos puede ser relativamente simple, el proceso que conduce a modelos válidos y extrapolables a una población determinada es un desafío metodológico complejo ^(29,30). Para el desarrollo y validación de un modelo de predicción clínica, la utilización del conocimiento previo disponible en la literatura científica es de suma importancia ⁽²⁹⁾, dado que muchos de los desafíos futuros que puedan

surgir podrían haber sido ya estudiados por otros investigadores ^(25,29,30). En un análisis cienciométrico realizado por este autor y colaboradores ⁽²⁵⁾, cuyos resultados se encuentran publicados en la Revista Habanera de Ciencias de Médicas se describen algunos aspectos sobre este tema. El estudio abarca la producción científica entre los años 2019 y 2021 en el ámbito de los modelos de predicción clínica empleados en la COVID-19.

I.2. Producción científica, fuentes clave, autores y documentos relevantes

Luego de la aparición de la pandemia de COVID-19, el mundo experimentó una explosión de literatura científica sobre esta temática ^(25,27). El tema abarcó el 2,12 % del total de la producción científica mundial anual en el año 2020, lo que resultó un hecho sin precedentes en la historia ⁽²⁵⁾. Una consulta en la base de datos Dimensions ⁽²⁴⁾ que consistió en la unión de los términos “prediction model” o “prognostic model” en el título o el resumen, más las palabras “COVID-19” o “SARS-CoV-2” en artículos originales publicados entre los años 2019 y 2021 en los campos de investigación Ciencias Biomédicas y Clínicas o Ciencias de la Salud devolvió un total de 519 documentos. De estas publicaciones, el 86,21 % recibió al menos una cita ⁽²⁵⁾. La tabla muestra la lista de los —cinco más importantes— recursos de información, autores y documentos (anexo 2, tabla 1) ⁽²⁵⁾.

El recurso de información que más aportó al tema fue Scientific Reports ⁽³⁶⁾ del portafolio de Nature ⁽³⁷⁾ y PLOS ONE ⁽³⁸⁾. Ambas revistas son de acceso abierto y están entre las de más alto nivel de impacto en el campo de la salud. Con una diferencia marcada respecto al resto de publicaciones, el artículo de mayor interés correspondió a la revisión sistemática viva titulada “Prediction Models for Diagnosis and Prognosis of Covid-19: Systematic Review and Critical Appraisal.” publicada en The British Medical Journal (BMJ) ^(27,39) por Wynants y

colaboradores ⁽²⁷⁾. En esta revisión auspiciada por COVIDPRECISE ⁽²⁶⁾ (revisión de COVIDPRECISE para abreviar) se realizó una evaluación crítica sobre la validez y utilidad de los informes publicados o preimpresos de modelos para el diagnóstico y el pronóstico de pacientes con COVID-19 ^(25,27).

El autor más importante fue Martin van Smeden del Centro Julius de Ciencias de la Salud y Atención Primaria en Utrecht ⁽⁴⁰⁾, Países Bajos, coautor además de la revisión sistemática mencionada ⁽²⁷⁾ y líder del proyecto COVIDPRECISE ⁽²⁶⁾.

La versión tres de la revisión de COVIDPRECISE ⁽²⁶⁾ y cuya autora principal es Laure Wynants ⁽²⁷⁾ constituyó la obra de referencia más importante dentro de la investigación en este campo, según los indicadores bibliométricos analizados ⁽²⁵⁾. Esta obra es una referencia importante en esta tesis porque sirve como base para identificar los modelos de mayor calidad, la variable de resultado, el conjunto de variables predictoras candidatas, así como los tipos de algoritmos usados para el desarrollo de los modelos propuestos ^(25,27).

El mapa temático muestra los temas más relevantes en este frente de investigación (anexo 3, figura 1). La figura visualiza conglomerados a partir de los títulos de las publicaciones que se obtuvieron mediante la técnica de escalamiento multidimensional. Adicionalmente, se identifican los artículos más populares dentro de cada grupo ⁽²⁵⁾. En el mapa temático se distinguen tres ejes de temas. Dos de ellos son bastante homogéneos, mientras que uno presenta mayor heterogeneidad. A continuación, de arriba a abajo, se describe cada uno de los conglomerados:

Conglomerado verde (superior): Incluyó las palabras unidades de cuidados (care unit) y cuidados intensivos (intensive care). Los artículos del grupo se enfocaron en temas como el desarrollo de modelos en escenarios hospitalarios

de muy alto riesgo de transición a la gravedad. El trabajo más relevante (162 citas) dentro del clúster fue: “Prediction model and risk scores of ICU admission and mortality in COVID-19” de Zirun Zhao y colaboradores ⁽⁴¹⁾ publicado en la revista PLOS ONE ⁽³⁸⁾.

Conglomerado rojo (medio): Resulta el grupo más amplio y variable. El término principal lo constituye desarrollo de modelos (model development), acompañado de palabras como pacientes hospitalizados (patients hospitalized, hospitalized covid), que podrían indicar el contexto hospitalario como escenario principal para la aplicación de estos modelos. En el conglomerado también se encuentran palabras que podrían ser asociadas con la estimación del riesgo de enfermar (predicting COVID-19, COVID-19 infection, COVID-19 disease), de transición a formas graves de la enfermedad (disease progression), en particular, de admisión en unidades de cuidados intensivos (icu admission) o necesidad de ventilación mecánica (mechanical ventilation), así como del riesgo de muerte (predict mortality, mortality risk). Este conglomerado incluye la revisión de COVIDPRECISE ⁽²⁶⁾ realizada por Wynants y colaboradores ⁽²⁷⁾ que fue el artículo más citado de esta colección en general ⁽²⁵⁾.

Conglomerado azul (inferior): El grupo contiene palabras como pacientes críticos (critically ill patients) o pacientes críticos con COVID-19 (critically ill covid patients) que denota un interés de la aplicación de modelos para explicar o predecir la enfermedad grave, así como el uso de técnicas de inteligencia artificial ⁽⁴²⁾. El artículo de mayor relevancia (27 citas) fue publicado por Mohammad M. Banoei, en la revista Critical Care ⁽⁴³⁾ y se titula: “Machine-learning-based COVID-19 mortality prediction model and identification of patients at low and high risk of dying” ⁽⁴⁴⁾.

Un elemento que se deriva del análisis anterior es que la contextualización de los modelos en diferentes escenarios de la atención médica es un tema fundamental ^(29,30,45). Es importante destacar que los modelos de predicción clínica se pueden clasificar en función del contexto poblacional al que se dirigen ⁽²⁷⁾. Esta clasificación se usa con frecuencia en la literatura para agrupar los modelos predictivos de acuerdo con su población objetivo ⁽⁴⁵⁾. Por ejemplo, en la revisión de COVIDPRECISE, los modelos pronósticos para la COVID-19 se agruparon en tres categorías fundamentales: modelos para población general, modelos para la atención médica en la comunidad y modelos para pacientes ingresados en el hospital ⁽²⁷⁾.

De esta manera, se distingue que se desarrollaron modelos específicos para pacientes atendidos en escenarios de atención médica comunitaria como para pacientes hospitalizados. En opinión de este autor, esto respalda los términos utilizados en los objetivos y en la organización de este estudio, está en línea con la literatura científica existente y es congruente con los distintos escenarios de atención médica descritos en el PAN (anexo 4, figura 2) ^(17,18).

Un tema de atención por parte de la comunidad científica fue el uso de algoritmos de aprendizaje automático para la construcción de modelos predictivos ^(25,27). Precisamente, los algoritmos de aprendizaje supervisado como RPART ⁽⁴⁶⁾, CTREE ⁽⁴⁷⁾ y la regresión logística binaria múltiple ⁽⁴⁸⁾ se usan en este estudio para el desarrollo de los modelos cubanos propuestos. La capacidad de los algoritmos para aprender patrones y tomar decisiones informadas ante nuevas situaciones (aprendizaje automático) ⁽⁴⁹⁾ es una de las subdisciplinas más importantes dentro de la Inteligencia Artificial (IA), junto al Procesamiento de Lenguaje Natural (PLN), el procesamiento de imágenes, la generación de

contenidos y la robótica ^(50,51). La IA está revolucionando la medicina moderna. Su capacidad para mejorar la precisión diagnóstica, personalizar tratamientos, utilizar grandes volúmenes de datos y de variados tipos, así como acelerar la investigación puede transformar la atención médica al paciente ⁽⁴⁹⁾. Se reconoce el interés de los desarrolladores de modelos por la integración de múltiples fuentes de datos para construir modelos más precisos ⁽⁵²⁾. En este sentido, la incorporación de información que ofrecen sensores biológicos instalados en dispositivos de la vida diaria como pulseras y relojes inteligentes, datos genómicos, biomarcadores e imágenes médicas, datos de determinantes sociales y ambientales a los modelos también resultó de gran interés ⁽²⁵⁾.

La creciente generación de grandes cantidades de información provenientes de múltiples fuentes de datos dio origen al concepto de “Big Data” (del inglés, grandes datos). Este enfoque se centra en la gestión y análisis de grandes volúmenes de información diversa que son de difícil manejo por métodos tradicionales. El Big Data puede ofrecer una visión profunda de los sistemas complejos, lo que a su vez permite tomar decisiones más efectivas bajo un enfoque integrador ^(29,53).

Otros enfoques holísticos como el de “una salud” también son atractivos para los desarrolladores de modelos predictivos. Una razón es porque la integración de predictores relacionados con varias dimensiones de la salud como la ambiental, los estilos de vida, los determinantes sociales de la salud (DSS), la organización de los servicios médicos ⁽¹⁵⁾, puede mejorar la exactitud de este tipo de herramientas ⁽²⁹⁾.

La medicina de precisión fue otro frente de investigación relacionado con los modelos predictivos para la COVID-19. Este paradigma es un nicho para la

aplicación de modelos de predicción clínica porque estos modelos son útiles para personalizar el tratamiento, identificar riesgos de manera oportuna, optimizar recursos, aumentar la precisión diagnóstica y mejorar la calidad de vida de los pacientes ^(29,49). La medicina de precisión es un enfoque de la terapéutica que adapta el tratamiento en función de las características individuales del paciente, se utilizan datos de características individuales, perfil genético, entorno y estilo de vida para mejorar la eficacia y seguridad de los tratamientos ⁽⁵⁴⁾.

I.3 Principales modelos de predicción clínica para la COVID-19

Hasta el 9 de agosto de 2022, se registró un total de 232 modelos de predicción clínica ⁽²⁷⁾. De estos, siete se utilizaron para pronosticar la susceptibilidad a la COVID-19 en la población general, 118 para realizar clasificaciones diagnósticas como la presencia de COVID-19, COVID-19 severo y diagnóstico por imágenes ^(25,27). Además, 107 se emplearon para estimar el riesgo de transición a la gravedad en pacientes confirmados, incluyeron 28 para predecir el estado crítico y 39 para la muerte ⁽²⁷⁾. Hasta la fecha del análisis realizado por este autor ⁽²⁵⁾, la base de datos de COVIDPRECISE ⁽²⁶⁾ ha registrado 731 publicaciones ^(25,27). Asimismo, la revisión de COVIDPRECISE ⁽²⁶⁾ tiene cuatro actualizaciones ⁽²⁷⁾. El artículo se publicó originalmente el 7 de abril de 2020 y se realizó una actualización posterior el 3 de febrero de 2021 ^(25,27). Durante el periodo de estudio, algunas versiones ya no estaban disponibles públicamente. Esto limitó la capacidad para obtener referencias completas y precisas para todas las actualizaciones. Sin embargo, se hizo una búsqueda exhaustiva para incluir la información más relevante y actualizada disponible en el momento de la redacción de esta tesis.

En opinión de este autor, este trabajo se erige como una referencia imprescindible para investigadores y profesionales del área, al identificar y

analizar en profundidad los principales sesgos que afectan a estos modelos. Esta obra proporciona una guía práctica para desarrollar y evaluar modelos más robustos y confiables.

I.3.1 Sesgos en los modelos de predicción clínica para la COVID-19

En revisión sistemática de COVIDPRECISE ⁽²⁶⁾ se reporta que los modelos descritos en los artículos frecuentemente exhibieron problemas de sobreajuste, causados por tamaños de muestra insuficientes y una validación inadecuada. Esto incluyó la ausencia de validación externa y poca atención a la calibración. Además, se reportó una deficiencia en la transparencia de las metodologías que se describieron, las elecciones de modelado y los resultados; muchas publicaciones no presentaron ecuaciones de modelos subyacentes o gráficas de calibración. Se hizo un uso inapropiado de datos, existió inconstancia en la medición de las variables de resultados y en ocasiones los modelos se implementaron de manera anticipada y sin documentación ni validación suficiente. Estas limitaciones hicieron que los modelos fueran poco confiables para su uso clínico ^(25,27). Muchas de las reglas pronósticas no se aplicaron en la práctica de la asistencia médica diaria, sino que su uso se circunscribió al periodo durante el cual se llevó a cabo la investigación.

En opinión de este autor, la identificación y manejo transparente de los sesgos resultan fundamentales para la credibilidad y aplicabilidad de los hallazgos de la investigación. En este estudio se identificaron situaciones que incrementaron el riesgo de sesgos, como los sesgos de selección, información y confusión ⁽⁵⁵⁾. Sin embargo, se implementaron medidas de mitigación. Por ejemplo, la aleatorización por pacientes y hospitales resultó clave, ya que permitió una distribución equitativa de las características observables y no observables de los

participantes, redujo el efecto de sitio y aseguró resultados más representativos y generalizables a diferentes escenarios de atención médica ⁽⁵⁶⁾. Además, la aleatorización minimiza la influencia de factores externos que podrían distorsionar los resultados, y ofrece una base sólida para la validez interna del estudio con el uso de técnicas como la validación cruzada ⁽²⁹⁾. A pesar de estas medidas, se reconoció la presencia de ciertos sesgos que limitaron la generalización de los resultados, lo cual se consideró al interpretar las conclusiones del estudio.

I.3.2 Modelos de mayor calidad

La revisión de COVIDPRECISE ⁽²⁶⁾ identificó en la versión original cinco modelos de mayor calidad que se podrían adaptar a cada uno de los escenarios de atención médica del Sistema Nacional de Salud (SNS) para el manejo de la COVID-19 ^(25,27).

Los modelos de mayor calidad son los siguientes: el modelo QCOVID ⁽⁵⁷⁾ que se podría adaptar para identificar grupos vulnerables a enfermar y morir a causa de la COVID-19 en la población general; el modelo PRIEST ⁽⁵⁸⁾, útil para clasificar a los pacientes con COVID-19 que acuden a consultas de urgencias; los modelos de deterioro 4C MORTALITY ⁽⁵⁹⁾; el modelo NEWS 2 ⁽⁶⁰⁾; y el modelo de Xie ⁽⁶¹⁾, que podrían servir como complemento para tomar las decisiones, según el riesgo de transición a la gravedad en pacientes hospitalizados ⁽²⁷⁾.

A continuación, se ofrece una descripción más detallada de los modelos de mayor calidad.

1. **NEWS 2** ⁽⁶⁰⁾: El National Early Warning Score 2 (NEWS 2, por sus siglas en inglés) es un sistema de puntuación que se utiliza para evaluar la gravedad de la enfermedad de un paciente y para identificar signos tempranos de

deterioro clínico. La herramienta no es específica para la COVID-19, sino que se usa para predecir la mortalidad en un conjunto amplio de enfermedades. El modelo NEWS 2 es una versión actualizada del modelo NEWS, que se produjo por primera vez en 2012 y se actualizó en diciembre de 2017. Esta regla de predicción se basa en una puntuación en la que se asignan puntos a las mediciones fisiológicas ya registradas en la práctica rutinaria, cuando los pacientes se presentan o son atendidos en el hospital.

2. **4C MORTALITY** ⁽⁶²⁾: El modelo Coronavirus Clinical Characterisation Consortium Mortality (4C MORTALITY, por sus siglas en inglés) es un modelo de estratificación de riesgo desarrollado en el Reino Unido por el consorcio ISARIC 4C que predice la mortalidad hospitalaria específicamente para la COVID-19 a partir de datos demográficos, clínicos y parámetros de laboratorio.
3. **PRIEST** ⁽⁵⁸⁾: El modelo Pandemic Respiratory Infection Emergency System Triage (PRIEST, por sus siglas en inglés) se utiliza para estimar la gravedad clínica en pacientes con COVID-19 basado en la puntuación NEWS 2 ⁽⁶⁰⁾. Se evaluó en 22 485 pacientes en el Reino Unido. El sistema PRIEST ⁽⁵⁸⁾ proporciona una evaluación de riesgo más completa y específica para la COVID-19 en comparación con NEWS 2 ⁽⁶⁰⁾. El modelo PRIEST ⁽⁵⁸⁾ mejora el modelo NEWS 2 ⁽⁶⁰⁾ al agregar puntos adicionales para las variables predictoras como la edad, el sexo y la capacidad de un paciente para realizar actividades cotidianas y cuidar de sí mismo (estado funcional).
4. **Modelo de Xie** ⁽⁶¹⁾: Se trata de un índice para predecir la mortalidad intrahospitalaria en pacientes con COVID-19. La herramienta se desarrolló en la República Popular China y fue utilizada para predecir la muerte en

pacientes con esta enfermedad. En la revisión de COVIDPRECISE ⁽²⁶⁾, se menciona que se están esperando más detalles sobre el modelo de Xie ⁽²⁷⁾. En la primera versión de la revisión se incluye este modelo dentro de los de mayor calidad, pero a partir de esa versión ya no se vuelve a hacer referencia a esta herramienta.

5. QCOVID ⁽⁵⁷⁾: Se trata de un modelo desarrollado por la Universidad de Oxford para estimar el riesgo de hospitalización y muerte por COVID-19 en pacientes adultos en el Reino Unido en la población general. El modelo se desarrolló en una población de 6,08 millones de pacientes adultos. Este modelo es muy interesante porque tiene en cuenta determinantes sociales de la salud entre sus variables predictoras y se puede aplicar a personas que aún no se han contagiado de la COVID-19.

Los modelos de mayor calidad se basan en algoritmos de regresión logística binaria múltiple o regresión LASSO ⁽⁶³⁾. Solo se tiene referencia que NEWS 2 ⁽⁶⁰⁾ y 4C MORTALITY ⁽⁶²⁾ se usaron fuera del contexto de la investigación y en escenarios de atención médica reales.

I.3.3 Variables de resultado

En la revisión de COVIDPRECISE ⁽²⁶⁾, la variable de resultado más importante resultó ser la muerte dentro de los primeros 30 días (mortalidad). La mortalidad estuvo presente en el 44,8 % de las publicaciones. Le siguieron la progresión a la gravedad y la admisión en la UCI, presentes en el 14,2 % y 9,0 % de los casos, respectivamente. Las variables de resultados como la aparición de neumonía, la estadía hospitalaria y el tiempo hasta el alta hospitalaria tenían menor frecuencia, en solo el 1,4 %, 1,7 % y 1,0 % de los casos, respectivamente. Además, se

identificaron otras variables que estaban compuestas por distintas variables de manera integrada en el 13,7 % de los casos.

Resulta interesante la alta frecuencia de la mortalidad a los 30 días como variable de respuesta fundamental en los modelos reportados en la revisión de COVIDPRECISE ⁽²⁶⁾. Este autor considera que esto podría indicar un enfoque en los resultados a corto plazo en la investigación de COVID-19. Sin embargo, las variables como la estadía hospitalaria y el tiempo hasta el alta en el hospital, que son indicadores importantes de la carga del sistema de salud se usaron con menos frecuencia ⁽²⁷⁾.

I.3.4 Variables predictoras

En relación con el conjunto de los predictores que corresponden a los modelos de mayor calidad ^(25,27), se observó que se incluía la edad, el sexo, la presencia de comorbilidades, el nivel saturación de oxígeno (SpO₂), la frecuencia respiratoria (FR), la presión arterial sistólica y la creatinina sérica (Cr). El modelo PRIEST ⁽⁵⁸⁾ se añadió a estas variables el valor obtenido por la escala APACHE II ⁽⁶⁾ y el conteo de plaquetas (PLT). Además, el modelo QCOVID ⁽⁵⁷⁾ incorporó el índice de masa corporal (IMC), el grupo étnico, la pobreza y el antecedente de ser fumador. Los modelos 4C MORTALITY ⁽⁶²⁾ y el NEWS 2 ⁽⁶⁰⁾ incluyeron la cuantificación de la enzima lactato deshidrogenasa (LDH) y los niveles de troponina. Finalmente, el modelo de Xie ⁽⁶¹⁾ añadió a estos predictores, el conteo de linfocitos.

Se muestra un gráfico aluvial, que permite visualizar relaciones entre categorías ⁽⁶⁴⁾. Los flujos en el gráfico representan la relación entre los predictores de los modelos (a la izquierda) y los modelos seleccionados (a la derecha). El ancho

de los flujos indica la frecuencia o importancia de cada predictor en los diferentes modelos (anexo 5, figura 3).

Se consideró la edad, y las comorbilidades como factores de riesgo relevantes para la muerte por COVID-19 en la mayoría de los modelos estudiados ^(27,65,66).

Es una práctica común en la investigación pronóstica que los modelos resuman un conjunto de variables clínicas en un solo predictor ^(29,30). Por ejemplo, para el caso de las comorbilidades se utiliza el Índice de Comorbilidad de Charlson (ICC) como única variable ^(67,68). El ICC se emplea para estimar la mortalidad de pacientes con múltiples comorbilidades en distintos contextos ⁽⁶⁸⁾. Entre las enfermedades que se incluyen en este índice se encuentran el asma bronquial, diabetes mellitus, EPOC, cardiopatía isquémica, enfermedad renal crónica, neoplasias, enfermedad cerebrovascular e inmunodeficiencias ^(67,67,69). En el contexto de la COVID-19, se incorporaron a las variables originales del ICC, otros antecedentes como la hipertensión arterial (HTA) y la obesidad, dada su asociación con la mortalidad por COVID-19 ⁽⁷⁰⁾.

Entre las comorbilidades más importantes en el pronóstico de la COVID-19 se incluyen trastornos metabólicos como la diabetes tipo 2 y la obesidad, que están asociados con procesos inflamatorios que alteran el sistema inmunitario ⁽⁷¹⁾. La hipertensión arterial también está relacionada con la inflamación de bajo grado, que puede ser tanto causa como consecuencia de la enfermedad ⁽⁷²⁾. La oncogénesis, especialmente en pacientes con cáncer de pulmón, aumenta el riesgo de complicaciones graves debido a la función pulmonar deteriorada y la inflamación crónica ⁽⁷³⁾. La edad es un factor de riesgo significativo, ya que el envejecimiento se asocia con una disminución de la función inmunitaria ⁽⁷⁴⁾.

En este estudio se emplea como punto de corte para la variable edad 65 y más años. La razón es que los 65 años son un estándar epidemiológico reconocido para diferenciar entre poblaciones más jóvenes y más viejas ⁽⁷⁵⁾.

Síntomas como la disnea ⁽⁷⁶⁾, la confusión mental y la tos son indicadores de mayor riesgo de mortalidad ⁽⁷⁷⁾. La frecuencia respiratoria (FR) y la presión arterial media (PAM) son importantes para evaluar la gravedad de la enfermedad ⁽⁷⁸⁾. Parámetros de laboratorio como niveles elevados de LDH y Cr, así como cifras bajas de PLT ⁽⁷⁹⁾, son útiles para predecir la gravedad de la COVID-19. El índice neutrófilo-linfocito (INL) y las plaquetas (PLT) son marcadores de inflamación y coagulación ⁽⁸⁰⁾, respectivamente, y un INL elevado o un PLT bajo se asocian con peores resultados ⁽⁸¹⁾.

En opinión del autor, estos hallazgos subrayan la multicausalidad en la progresión hacia la gravedad de la COVID-19 y la necesidad de adoptar enfoques integrales para la selección de predictores. Además, es importante destacar que los predictores candidatos de los modelos propuestos en este estudio están respaldados por la literatura científica y se encuentran validados por el conocimiento sobre la fisiopatología de la enfermedad.

1.3.5 Tipo de modelo

Los algoritmos utilizados en la modelación se pueden clasificar por el tipo de modelo utilizado en modelos paramétricos, semiparamétricos o modelos no paramétricos ⁽⁸²⁾. Por regla general, los algoritmos paramétricos se basan en la estimación de parámetros y los no paramétricos no hacen suposiciones estrictas sobre la forma funcional de la relación entre la variable de resultado y las variables predictoras ^(29,49).

En el ámbito del aprendizaje automático, una clasificación muy utilizada es la de modelos de aprendizaje supervisado o no supervisado ^(49,83). Los algoritmos de aprendizaje supervisado utilizan datos etiquetados (pares de entrada-salida). A su vez, los algoritmos de aprendizaje supervisado se dividen en modelos de regresión y de clasificación si la variable de respuesta es cuantitativa o cualitativa, respectivamente ^(49,83). Por su parte, los modelos de aprendizaje no supervisado exploran datos no etiquetados y su objetivo es descubrir patrones, relaciones o estructuras en los datos. Además, dentro de la clasificación de los algoritmos de aprendizaje automático se incluye los de aprendizaje por refuerzo y los de aprendizaje semisupervisado ⁽⁴⁹⁾.

Según el estudio realizado por COVIDPRECISE ⁽²⁶⁾, se observó que la regresión logística binaria múltiple ⁽⁴⁸⁾ fue el tipo de modelo más empleado, presente en 215 artículos (38,2 %). En el resto de las publicaciones se incluyeron otros algoritmos para especificar el modelo (126; 22,4 %), como las redes neuronales (85; 15,1 %), los árboles de clasificación (55; 9,8 %), la regresión de Cox (45; 8,0 %), la máquina de soporte de vectores (17; 3,0 %), la regresión lineal, multinomial u ordinal (19; 6,4 %) y los modelos de supervivencia (1; 0,2 %) ⁽²⁷⁾.

Este autor considera que es importante subrayar el uso predominante de la regresión logística binaria múltiple ⁽⁴⁸⁾, lo que podría reflejar una preferencia de la comunidad científica por los métodos estadísticos tradicionales en la investigación sobre modelos de predicción clínica en la COVID-19. La regresión logística binaria múltiple se considera también un algoritmo paramétrico y de clasificación supervisada en el ámbito del aprendizaje automático. Las técnicas no paramétricas como las redes neuronales y los árboles de clasificación son pilares del enfoque de aprendizaje automático, pero se emplearon en menor

medida en la revisión de COVIDPRECISE ⁽²⁷⁾. Una explicación podría ser que, si el equipo de desarrollo tiene una experiencia significativa con modelos tradicionales, puede ser más cómodo utilizar este tipo de modelos. Además, los equipos de desarrollo pueden ser reacios a adoptar nuevas tecnologías, especialmente si ya tienen procesos establecidos y funcionan bien con los modelos tradicionales.

Todos los algoritmos utilizados para el desarrollo de los modelos cubanos son utilizados dentro del campo del aprendizaje automático. Estos tienen la capacidad de aprender patrones a partir de datos, y almacenar este conocimiento en la forma de árboles de clasificación para el caso de RPART ⁽⁴⁶⁾ y CTREE ⁽⁴⁷⁾ y en forma de una ecuación matemática para el caso de la regresión logística binaria múltiple ⁽⁴⁸⁾. Esta propiedad les permite realizar predicciones en nuevos conjuntos de datos a partir de pautas aprendidas con anterioridad en la partición de entrenamiento ^(29,49). Es por ello que se puede considerar este estudio como una aplicación práctica de la IA ⁽⁸³⁾ en el afrontamiento de la COVID-19 en Cuba ⁽³⁾.

I.4 Estrategias de modelado

Contar con una estrategia de modelado correcta es esencial para desarrollar y validar modelos de predicción clínica que sea fiables, precisos, interpretables, éticos, eficientes y adaptables. En la literatura se han propuestos varios enfoques ^(29,30,35). En esencia, las acciones realizadas durante el proceso son muy similares, pero varían en algunos elementos dentro de la secuencia de etapas. En el ámbito de la bioestadística, se destacan las propuestas de Frank E. Harrell de la Universidad de Carolina del Norte en EE. UU. y de Ewout Steyerberg de la Universidad de Leiden en Países Bajos. La estrategia de modelado de este

estudio se basa en el enfoque de siete pasos propuesto por Steyerberg, por resultar práctico ^(29,35). Sin embargo, se le realizan algunas adecuaciones. A continuación, se describen cada uno de esos pasos.

1.4.1 Definición del problema e inspección de datos

El proceso se inicia con el análisis de las preguntas de investigación ^(29,30). La elección de la variable de resultado está directamente relacionada con la pregunta de investigación, por lo que se operacionaliza en esta etapa. Se recopilan los datos y se verifica la integridad de la información. Se realiza el análisis exploratorio de datos EDA (del inglés, Exploratory Data Analysis) para comprender las características de las variables y dar solución a problemas que pueden surgir en etapas posteriores. Por ejemplo, se identifican y tratan los datos atípicos y faltantes.

En opinión de este autor, la elección del estado al egreso como variable de resultado es adecuada, porque se conecta directamente con las preguntas científicas planteadas en este estudio. Además, se encontraba accesible en las fuentes de datos a las que se tuvo acceso en este estudio y porque ofrece ventajas metodológicas adicionales. Estas ventajas consisten en que la muerte es un evento inequívoco, con escasas posibilidades de errores de medición y sesgos. Además, facilita la comparabilidad entre diferentes poblaciones, como la que se realizó en este estudio. La relevancia clínica de esta variable es indiscutible, dado que es de interés para pacientes, médicos y responsables de la toma de decisiones en salud ⁽²⁹⁾.

1.4.2 Codificación de las variables predictoras

Para la codificación de predictores se definen con precisión las variables medidas en escalas cualitativas, se crean nuevas clases o agrupan las categorías poco frecuentes ^(29,30).

En el caso de las variables cuantitativas, a menudo se necesitan realizar transformaciones. Por ejemplo, la aplicación de splines cúbicos restringidos o polinomios fraccionarios si no se cumple el supuesto de linealidad entre la variable de resultado y los predictores ^(29,30).

En ocasiones, como en el caso de este estudio, se construyen nuevos predictores resúmenes para condensar información en una sola variable sintética. Esta técnica puede ser útil cuando los predictores están fuertemente correlacionados o cuando se busca simplificar el modelo ⁽²⁹⁾. Según Harrell, la reducción de dimensionalidad es una alternativa a la pérdida de poder predictivo por parsimonia del modelo. Harrell no está de acuerdo con la idea generalizada de que los modelos con menor número de predictores son preferibles a aquellos que incluyen un mayor número de variables ⁽³⁰⁾.

El criterio de este autor es que, aunque la parsimonia no debe ser el único criterio en la elección de predictores. En determinados contextos, la simplificación mediante predictores resumidos o modelos más simples puede ser útil en la práctica. Por ejemplo, en el contexto de este estudio fue crucial encontrar un equilibrio entre la precisión de los modelos y la operatividad exigida en una investigación ejecutada en el curso de una crisis sanitaria, donde las condiciones no eran ideales y no se disponía de mucha información en las primeras etapas de la pandemia.

1.4.3 Especificación del tipo de modelo

En esta etapa se selecciona el algoritmo o tipo de modelo más adecuado para establecer la relación entre las variables de resultado y los predictores. Durante este paso se deben diferenciar los efectos principales de las interacciones entre las variables. Se debe evaluar en profundidad los supuestos como la aditividad (con términos de interacción) y la linealidad (con términos no lineales) ⁽³⁰⁾.

En algunos casos como los de este estudio, la relación entre la variable estado al egreso y el predictor no cumple el principio de linealidad, lo que representa una violación de uno de los supuestos de la regresión logística binaria múltiple ⁽⁴⁸⁾. Una solución a esta situación consiste en realizar transformaciones como la aplicación de splines cúbicos ⁽²⁹⁾. Un spline cúbico es una herramienta matemática que permite representar una función mediante la unión de polinomios cúbicos. En la regresión logística binaria múltiple ⁽³⁰⁾, los spline cúbicos se utilizan para modelar relaciones no lineales entre la variable de respuesta y los predictores ^(29,30). Al ajustar un polinomio cúbico en cada intervalo entre dos puntos consecutivos, los spline cúbicos pueden capturar patrones complejos en los datos y proporcionar un ajuste más flexible y preciso que un modelo lineal ⁽³⁰⁾.

La selección de variable predictoras representa un paso esencial en la etapa de especificación del tipo de modelo ^(29,35). Por su parte, los predictores deben tener relevancia para la variable de resultado y estar contrastados por su plausibilidad biológica ⁽³⁰⁾.

El tema de la selección de predictores finales en el contexto de la regresión logística ⁽⁴⁸⁾ es ampliamente debatido por la comunidad científica ⁽⁸⁴⁾. Autores como Silva ⁽⁸⁵⁾ argumentan la improcedencia del uso de métodos algorítmicos

como la “selección de variables paso a paso” para establecer los predictores finales. La selección automática de predictores a partir de algoritmos puede ser cuestionada porque el método no se basa en la comprensión teórica de las relaciones entre las variables, tampoco puede distinguir entre las asociaciones causales y aquellas que resultan de confusión o sesgo. Los métodos por pasos son sensibles a variaciones aleatorias e inestabilidades en los datos. Además, muchos de los algoritmos de “selección paso a paso” emplean como criterio de selección en sus etapas los valores de p de las pruebas de significación estadística ^(84,85). El uso de los valores p se ha criticado por diversas razones, una de ellas es que el rechazo de la hipótesis nula depende del tamaño de la muestra ⁽⁸⁶⁾. Además, un resultado estadísticamente significativo puede tener o no relevancia clínica y viceversa. Las pruebas de hipótesis se basan en el rechazo de la hipótesis nula y no en evaluar la fuerza de la evidencia en otros contextos. Por otra parte, la regla para la decisión de estas pruebas se sustenta exclusivamente en el valor de “ p ” y en otros parámetros como los errores tipo I y tipo II que se pueden manipular para obtener los resultados deseados. Esta es una de las razones por las cuales, en este estudio, se sustituye el criterio de significación estadística a partir del intervalo de confianza de los OR en lugar del criterio tradicional que usa los valores de “ p ”. Adicionalmente, otro argumento que apoya esta decisión es que los OR son una medida del tamaño del efecto y la amplitud del intervalo de confianza es también un indicador de precisión de la estimación.

En trabajos como los de Heinze, se ofrecen recomendaciones prácticas para la selección de predictores finales ⁽⁸⁷⁾. La presente investigación sigue algunas de estas recomendaciones para abordar el diseño metodológico de los modelos

propuestos. Por ejemplo, se considera la selección de un conjunto de predictores candidatos contrastados en la literatura como una buena práctica ^(29,87). En el caso de este estudio, se trata de los predictores de los modelos de mayor calidad que estaban disponibles en las fuentes de datos y tenían información suficiente para ajustar los modelos ^(25,27).

Este autor considera que emplear un enfoque integral para la selección de predictores que combina el juicio clínico con procedimientos estadísticos o computacionales es una buena práctica. En los tres modelos propuestos, la selección final de los predictores se basó en la plausibilidad biológica, el respaldo de la literatura científica, los métodos de selección de atributos de los algoritmos de aprendizaje automático basados en árboles, así como el criterio de significación estadística definido en este estudio.

1.4.4 Estimación del modelo

En la etapa de estimación de parámetros de un modelo, el objetivo es encontrar los mejores valores para los coeficientes en la relación entre la variable de resultado y las variables predictoras ^(29,48). Este proceso está condicionado por el tipo de modelo seleccionado y por las características de los datos. En el caso de la regresión logística binaria múltiple, para ajustar los coeficientes de la distribución Logit, se utilizan algoritmos como el de Walker-Duncan ⁽⁸⁸⁾ y el Método de Máxima Verosimilitud (MLE, por sus siglas en inglés) ⁽⁸⁹⁾. Estos dos procedimientos se pueden utilizar indistintamente porque producen coeficientes similares ^(48,88,90). Sin embargo, para obtener COVIDCENCEC, se seleccionó el método de MLE debido a su mayor versatilidad, ya que permite ajustar diversas distribuciones además de la Logit. Este método es el más popular dentro de la comunidad científica por su eficiencia computacional ⁽³⁹⁾.

En la etapa de especificación del modelo, una consideración clave es cómo manejar las interacciones entre las variables predictoras ^(29,30,35). Las interacciones en modelos de regresión, y de forma particular en la regresión logística, ocurren cuando el efecto de un predictor sobre la variable de resultado varía en dependencia del valor de otro predictor ⁽³⁰⁾. No tener en cuenta el efecto de las interacciones puede ser una causa de pérdida de presión y validez de los modelos. Sin embargo, la adición de un nuevo término para la interacción en la ecuación no siempre es necesario. La decisión de incluir un término de interacción se debe basar en la evidencia empírica y teórica, así como en el análisis realizado en el EDA. Es crucial evaluar si la interacción mejora el ajuste del modelo y considerar su simplicidad e interpretabilidad. Es importante evitar una complejidad innecesaria que pueda llevar al sobreajuste. Sin embargo, la inclusión de términos de interacción en la regresión logística puede mejorar la capacidad del modelo para capturar relaciones complejas ^(29,30).

En opinión de este autor, un modelo más simple resulta ser más robusto y fácil de interpretar, mientras que los modelos con múltiples interacciones muestran una mayor susceptibilidad al sobreajuste. En este estudio, las decisiones sobre la inclusión de los términos de interacción en el modelo COVIDCENCE consideran la significación estadística y la plausibilidad clínica de la relación entre las variables. Es importante destacar que esto solo aplica al modelo COVIDCENCEC, dado que en los modelos basados en árboles de clasificación no se toman en cuenta interacciones porque estos algoritmos eliminan intrínsecamente este efecto como parte del procedimiento ⁽²⁹⁾.

Los modelos de árboles de clasificación ABR y AAR aportan una clasificación del riesgo de muerte por COVID-19 de la cual se derivan los indicadores de

estrato propuestos (anexo 6, figura 4). Los estratos se definen a partir de las variables y divisiones desde el nodo principal hasta los nodos terminales. La clasificación resultante cumple con el principio de ser exhaustiva y mutuamente excluyente, además clasifica al paciente según el riesgo individual de muerte en diferentes niveles de riesgo. Esta solución es un aporte teórico al enfoque clásico de aplicación de árboles de clasificación en el manejo de problemas de salud que no se había descrito con anterioridad (anexo 6) ⁽⁴⁾.

En opinión de este autor, la elección de utilizar simultáneamente la regresión logística binaria múltiple ⁽⁴⁸⁾ como los árboles de clasificación ⁽²⁹⁾ en el desarrollo de modelos de predicción clínica propuestos se justifica por su complementariedad. La regresión logística binaria múltiple ⁽⁴⁸⁾ y los árboles de clasificación ⁽²⁹⁾ poseen fortalezas y debilidades complementarias, lo que permite capturar diferentes aspectos de los datos ⁽⁹¹⁾. Mientras que la regresión logística ⁽⁴⁸⁾ es útil para modelar relaciones lineales y proporciona resultados en términos de probabilidad, los árboles de clasificación pueden capturar relaciones no lineales y ofrecen un marco para establecer los indicadores de estratos propuestos ⁽²⁹⁾. Además, la utilización de varios tipos de modelos puede aumentar la robustez del análisis, por ejemplo, la elección de predictores a partir de los métodos de selección de atributos de los árboles de clasificación podría resolver las críticas a la selección algorítmicas en la regresión logística ⁽⁸⁵⁾. Por último, un argumento importante es que los algoritmos empleados en este estudio aprenden. Este enfoque tiene ventajas frente a otros tipos de modelos más tradicionales, dado que pueden adaptarse a la naturaleza cambiante de la COVID-19 o a modificaciones del contexto a medida que se recopilan nuevos datos.

1.4.5 Evaluación del rendimiento del modelo

El rendimiento del modelo generalmente se evalúa durante la etapa de desarrollo a partir de métricas de desempeño como la calibración y la discriminación en la partición de entrenamiento. La calibración se refiere a la concordancia entre las predicciones del modelo y los resultados observados ⁽⁹²⁾. Un modelo bien calibrado es aquel cuyas predicciones estimadas coinciden con los resultados observados ^(29,35). Por su parte, la discriminación se refiere a la capacidad del modelo para distinguir entre las diferentes categorías de la variable de resultado una vez ejecutado el pronóstico. En el contexto de este estudio, si el modelo tiene una buena discriminación será capaz de distinguir con exactitud entre pacientes que murieron a aquellos que egresaron vivos ⁽²⁹⁾.

Una herramienta muy popular para la evaluación de la calibración son las rectas de calibración ^(35,92). En los gráficos que proyectan estas rectas resultan de interés la pendiente y el intercepto de calibración ⁽⁹²⁾. Un intercepto cercano a cero y una pendiente cercana a uno indican un buen ajuste del modelo a los resultados observados ^(29,92). Este tipo de gráficos se seleccionaron para evaluar la calibración en la validación interna y externa de los modelos propuestos.

Para la validación interna en este estudio, se utilizó la validación cruzada de 10 pliegues ⁽²⁹⁾. Este método implica la división de los datos en 10 subconjuntos seguido de 10 iteraciones de entrenamiento y validación, en las que se emplea un subconjunto distinto para la validación en cada iteración ^(29,82). En esta técnica las métricas de desempeño del modelo se determinan a partir de la media de las puntuaciones obtenidas en estas iteraciones.

1.4.6 Evaluación de la validez del modelo

La validez del modelo se evalúa típicamente después que se ha desarrollado. Para ello se utilizan dos enfoques, la validación interna y la validación externa ^(29,30). En la validación interna se realiza en la partición de datos que se usaron para el desarrollo del modelo (partición de entrenamiento) ⁽²⁹⁾. Este tipo de validación permite comprobar cómo el modelo se ajusta a los datos y puede ayudar a detectar problemas como el sobreajuste. Los métodos de validación interna incluyen la validación aparente, la validación por división de datos, el remuestreo y la validación cruzada ^(30,92). Por su parte, la validación externa se realiza en un conjunto de datos independientes que no se haya utilizado en el entrenamiento (partición de prueba) ⁽²⁹⁾. La validación externa se puede llevar a cabo a través de una muestra independiente. Para seleccionar este conjunto independiente de datos se puede usar un criterio temporal, espacial o geográfico. Los métodos de remuestreo y validación cruzada se pueden usar también para hacer más robusta la validación a partir de múltiples conjuntos de datos generados por varias iteraciones ⁽³⁰⁾.

El punto de corte para establecer la proporción de pacientes en los conjuntos de entrenamiento y prueba es un tema discutido ^(29,30). La partición 70:30 (aproximadamente 2:1) para establecer la partición se adopta como una práctica habitual en el ámbito del aprendizaje automático, aunque se han usado otros enfoques como la razón 90:10, 80:20, 60:40, 50:50. Un equilibrio entre las proporciones de las particiones es fundamental para el desarrollo de un modelo preciso. Una proporción mayor en el conjunto de entrenamiento permite que el modelo aprenda mejor los patrones de los datos. Sin embargo, un entrenamiento excesivo puede provocar sobreajuste ⁽⁹³⁾. Por el contrario, un conjunto de

pruebas mayor mejora la evaluación del rendimiento, pero restringe los datos de entrenamiento lo que podría provocar un ajuste deficiente ⁽⁹⁴⁾.

Harrel expresa que la “verdadera validación externa” se refiere a la validación en un tiempo o lugar diferente, o sea, la validación temporal o la geográfica, y que no solo consiste en establecer una partición independiente para el conjunto de prueba. Adicionalmente, afirma que, si bien la validación externa ofrece mayor información sobre el poder de generalización, no es garantía de total fiabilidad ⁽³⁰⁾. Esto se debe a que pueden aparecer discrepancias temporales o variaciones geográficas en las poblaciones ^(29,30). Es por ello que Harrell recomienda modelar de manera explícita estas discrepancias y usar técnicas robustas como el remuestreo o la validación cruzada en todos los datos, o sea, sin definir particiones de entrenamiento y prueba ⁽³⁰⁾.

En opinión de este autor, la elección de la partición temporal en la primera etapa del estudio se justifica por varias razones. Primero, refleja la realidad de cómo se utilizan estos modelos, ya que se entrenan con datos del pasado y se utilizan para predecir el futuro. Segundo, evita el sesgo de mirar hacia el futuro que puede surgir con una partición aleatoria. Tercero, permite evaluar la robustez del modelo a lo largo del tiempo, un indicador clave de la calidad de las reglas de predicción propuestas. Por último, permite evaluar la capacidad del modelo de predicción clínica para adaptarse a cambios en el tiempo, como los que ocurrieron durante la COVID-19, un aspecto crítico en el contexto clínico donde las prácticas y tecnologías pueden evolucionar rápidamente. Por lo tanto, la partición temporal proporciona una evaluación más rigurosa y realista del rendimiento del modelo sobre otras alternativas como la aleatoria ⁽²⁹⁾.

Por su parte, la elección de un enfoque de partición geográfica para la segunda etapa del estudio se justifica porque también refleja cómo se utilizan estos modelos en la vida real. Evita el sesgo geográfico, ya que los pacientes representan a hospitales de diferentes regiones y se consideran las posibles variaciones en la atención médica en las instituciones de salud. Además, permite evaluar la robustez del modelo y su capacidad para adaptarse a las diferencias entre hospitales ^(29,30).

En el proceso de validación de un modelo de predicción clínica se usan varias métricas de rendimiento, como la precisión, la sensibilidad, la especificidad, el valor predictivo positivo (VPP), el valor predictivo negativo (VPN), el coeficiente de determinación (R^2), el AUC (por sus siglas en inglés, "Area Under the Curve"), el índice de discriminación (D_{xy}), el intercepto, la pendiente de calibración y el índice de Brier. Además, se usan otras medidas como las razones de verosimilitud (LR) negativas (LR+) y positivas (LR-) para evaluar el desempeño de los modelos ^(29,30,35,92). Las métricas D_{xy} y AUC miden la capacidad del modelo para diferenciar entre diferentes resultados, con valores más altos que indican una mayor capacidad de discriminación ⁽³⁰⁾. El R^2 indica cuánta variación en los datos puede explicar el modelo. Un valor cercano a uno indica un alto nivel de explicación. El índice de Brier, además, mide la precisión de las predicciones probabilísticas, con valores más bajos que indican mejores predicciones ^(29,30).

En las escalas clínicas, la elección de los puntos de corte para establecer un paciente de alto riesgo es un procedimiento fundamental ⁽²⁹⁾. Existen varios enfoques para seleccionar los puntos de corte óptimos ^(29,30). Algunos de los procedimientos más usados son el empleo de la media y la desviación estándar, el método de cuartiles o percentiles, el uso de análisis discriminante, los criterios

basados en importancia clínica, los métodos basados en la AUC. El índice de Youden es un método común para seleccionar el punto de corte óptimo porque permite obtener un solo punto de corte y priorizar la sensibilidad sobre la especificidad del modelo, para proporcionar la mejor clasificación de los pacientes si estos tienen alto riesgo de morir ⁽²⁹⁾.

1.4.7 Presentación del modelo

La presentación del modelo implica la comunicación clara y efectiva de los resultados de los modelos a los usuarios finales de las herramientas ^(29,35). Este paso es esencial para asegurar que los hallazgos sean accesibles y útiles para otros investigadores y profesionales de la salud ⁽²⁹⁾. Una solución es aplicar el modelo a través de un nomograma impreso en papel o la codificación de las reglas de pronóstico en un software ⁽²⁹⁾. Una tendencia para la aplicación de modelos es ejecutar las ecuaciones pronósticas en calculadoras basadas en la web o como aplicaciones para teléfonos móviles y tabletas. Algunos modelos predictivos se pueden incorporar en la práctica clínica como parte de los sistemas de gestión hospitalaria o en la vida diaria, como parte de dispositivos como relojes inteligentes y pulseras que registran parámetros fisiológicos como la frecuencia cardíaca, frecuencia respiratoria, presión arterial, SpO₂, nivel de glucosa en sangre, temperatura corporal, actividad física, calidad del sueño y postura ^(29,49,82).

Un elemento importante en la aplicación de un modelo en escenarios reales es que en la estrategia de modelado se debe tomar en cuenta para la evaluación periódica del rendimiento de los modelos y la realización de actualizaciones periódicas ^(29,30). El modelo puede quedar obsoleto con el tiempo debido a cambios en la dinámica de la enfermedad y los pacientes. Por ello, la evaluación

sistemática del ajuste del modelo y la precisión predictiva es cardinal para identificar la necesidad de actualizaciones ⁽³⁰⁾.

En opinión de este autor, una vez terminada una investigación es necesario transferir el conocimiento resultante a la práctica a través de la creación de nuevas tecnologías. Para el caso de este estudio, una tecnología factible es el desarrollo de un software como COVIDCENCEC-APK. Sin embargo, este tipo de aplicaciones deben ser fiables y responder a los propósitos de la investigación. A continuación, se describen algunas consideraciones sobre el desarrollo y validación de software que todo diseño debe tomar en cuenta.

I.5 Consideraciones sobre el desarrollo y validación del software

Una vez desarrollada una tecnología de software, es necesario realizar acciones para mejorar la capacidad del personal médico para ejecutar el modelo y hacer estimaciones y clasificaciones del riesgo de muerte en un nuevo paciente. También es importante ofrecer la posibilidad a otros investigadores de reutilizar el código fuente, modificar los parámetros del modelo, así como rediseñar la herramienta en sentido general. Una vez terminada la aplicación, es imprescindible evaluar si la herramienta se comporta de manera adecuada en las condiciones previstas. Las pruebas de validación de software evalúan si el producto es fiable y seguro. Asimismo, es necesario ofrecer un informe final de validación del software para mayor transparencia ⁽⁹⁵⁾. Como marco normativo para la confección del informe final se tomó en cuenta la guía ética y de gobernanza para el uso de la inteligencia artificial (IA) en el ámbito de la salud, desarrollada por la Organización Mundial de la Salud (OMS) ⁽⁹⁶⁾.

I.6 Generalización y extrapolación de los modelos a escenarios de atención médica en Cuba

A continuación, se abordarán algunos desafíos para la generalización de modelos predictivos en Cuba, tales como la complejidad de adaptar modelos foráneos a distintos escenarios de atención médica debido a diferencias contextuales. También se discutirá el uso de las fuentes de datos en la investigación y se destacará la combinación de distintas fuentes de datos. Finalmente, se examinará el contexto cubano y las directrices para el triaje de pacientes y se enfatiza la importancia de adaptar los modelos a entornos con recursos médicos limitados y las particularidades del sistema de salud cubano.

I.6.1 Limitaciones para la generalización de modelos

Una limitación que se debe tomar en cuenta para aplicar los modelos foráneos en Cuba y compararlos con los propuestos en este estudio como COVIDCENCEC, es que algunos algoritmos como el caso de la regresión logística binaria múltiple no son extrapolables a diferentes poblaciones de manera directa. Por ejemplo, cuando se comparan dos modelos predictivos que utilizan la regresión logística, es importante tener en cuenta que el término independiente de la ecuación logística puede variar según la prevalencia de la condición que se está modelando. La razón fundamental es que la función logit ⁽⁴⁸⁾ no se puede usar para estimar probabilidades a partir de un perfil dado cuando el modelo no se ha obtenido en el mismo contexto donde se quieren extrapolar los resultados. Los coeficientes de la regresión son válidos, pero el término independiente necesita corrección porque depende de las prevalencias ⁽²⁹⁾.

Los cambios en la conducta médica a lo largo del tiempo, así como las modificaciones en las definiciones de variables, pueden alterar el impacto de

ciertos predictores en los resultados y afectar el rendimiento de los modelos de predicción. La inestabilidad de la efectividad del tratamiento en la población objetivo también puede afectar la precisión. Si la gravedad de la enfermedad varía entre las poblaciones, es probable que el modelo requiera ajustes para reflejar las diferencias en el efecto del tratamiento. Además, el impacto del tratamiento puede diferir entre distintas instituciones de salud debido a variaciones en las características de los pacientes, prácticas médicas, disponibilidad de medicamentos y otros factores contextuales que influyen en el pronóstico de la COVID-19 ⁽⁹⁷⁾.

En este estudio se podría anticipar diferencias geográficas entre las poblaciones de pacientes de diferentes instituciones de salud y otras variaciones temporales que podrían ser causas de sesgos.

El riesgo de muerte experimentó cambios en el tiempo (anexo 7, figura 5), es posible también que algunos predictores podrían ser inestables durante esta crisis sanitaria. Por ejemplo, la edad se relaciona con la gravedad de la enfermedad, la susceptibilidad a comorbilidades, la respuesta a los tratamientos, incluso con variaciones en la conducta médica.

La conducta médica sufrió modificaciones debido a seis actualizaciones del PAN (ver anexo 8, figura 6), lo que pudo generar diferencias entre las particiones de entrenamiento y prueba en el estudio. Además, no se puede descartar variabilidad en los tratamientos administrados en diferentes hospitales, a pesar de las pautas establecidas por el PAN. Por ejemplo, en algunos hospitales solo se aplicó Jusvinza, y en otros Itolizumab, Nimotizuman u otro antiinflamatorio según la disponibilidad en el tratamiento a los pacientes graves.

La heterogeneidad en los tratamientos puede actuar como un modificador del efecto y provocar sesgo de confusión. Asimismo, el riesgo de sesgo de selección e información se debe tomar en cuenta en este estudio.

En opinión de este autor, la adecuación de los modelos de mayor calidad en Cuba sería posible; sin embargo, también resultaría un proceso complejo, dado que, además de la actualización, se debería tomar en cuenta la aplicación en entornos reales. Por ejemplo, los modelos concebidos en países desarrollados podrían tener dificultades para ser aplicados en entornos de recursos médicos limitados como Cuba. También es importante señalar que algunos de estos modelos no tenían una descripción clara de su implementación en los artículos de origen o no se aplicaron con regularidad en escenarios de atención médica reales ⁽²⁷⁾. Por ello, el desarrollo de modelos cubanos que tengan en cuenta la organización del sistema de salud puede ser la opción más viable.

Es importante destacar, que gran parte de este estudio se realizó en dinámica de la emergencia sanitaria con resultados que se implementaron en tiempo real, esto añadió aún más complejidad a la etapa de ejecución de la investigación.

Una limitación que aumenta el riesgo de sesgo de información es que los datos utilizados en este estudio no provinieron de bases de datos diseñadas específicamente para dar salida a los objetivos de este trabajo, sino que fueron adecuadas a estos.

I.6.2 Fuentes de datos disponibles para la investigación

En el estudio se tuvo acceso a dos fuentes de datos: el Registro Nacional de Pacientes confirmados de COVID-19 (RNPC) y el Registro de la base de datos del proyecto Impacto de la COVID-19 en pacientes adultos en Cuba. Años 2020-2023. (RICPA). El registro RNPC ofrece una visión amplia de los pacientes

confirmados de COVID-19 en Cuba. A pesar del número limitado de variables disponibles en este conjunto, su cobertura nacional proporcionó información valiosa (anexo 9, figura 7). El registro RICPA contenía información de las historias clínicas de pacientes hospitalizados incluidos en el proyecto ICPA ⁽³³⁾. Para la confección de este se utilizó el programa de encuestas Xavia Sidec 3.0 que es una aplicación para manejo de datos de ensayos clínicos que se usó para codificar el cuaderno de recogida de datos (CRD) del proyecto ICPA (anexo 10, figura 8) ⁽³³⁾. Esta herramienta se seleccionó por la posibilidad de entrada de datos desde el sitio donde se encontraban las historias clínicas. Si bien, esta base datos tenía una representación más limitada de la población que RNPC, ofrece un mayor número de predictores clínicos.

En opinión de este autor, la combinación de estos conjuntos de datos permitió aprovechar las fortalezas de ambas fuentes. La investigación pudo beneficiarse de la amplia cobertura de RNPC y de la riqueza de información clínica de RICPA.

I.6.3 El contexto cubano como escenario para la generalización

En Cuba, las directrices para el triaje de pacientes a partir de la estratificación del riesgo de transición a la gravedad se establecieron en el PAN ⁽¹⁷⁻¹⁹⁾.

La versión 1.6 del PAN enfatiza la estratificación del riesgo como un proceso importante para identificar a los individuos con mayor riesgo de transición a la severidad ⁽¹⁷⁾.

El PAN establece una relación entre la estratificación del riesgo, el escenario de atención médica y los patrones de tratamiento ^(17,18). Por ejemplo, los pacientes clasificados como de Alto Riesgo (AR), ya sea por su edad (65 años o más) o por la presencia de comorbilidades, tienen una mayor probabilidad de ser admitidos en la UCI, especialmente si presentan síntomas graves. Asimismo, la

presencia de síntomas moderados o severos aumentan las probabilidades de admisión a la UCI, ya que condiciones como neumonía o síndrome de dificultad respiratoria aguda (SDRA) requieren atención intensiva. El protocolo detalla patrones de tratamientos específicos, por ejemplo, se recomienda el uso de antiinflamatorios para pacientes con signos de inflamación y la administración de heparina profiláctica para todos los pacientes en la UCI, así como dosis terapéuticas para aquellos con signos de enfermedad severa, de esta manera se prescribe una adecuada terapia anticoagulante para prevenir eventos tromboembólicos ⁽¹⁷⁾.

I.7 Consideraciones finales

Se explora la literatura científica sobre el desarrollo y validación de modelos de predicción clínica para estimar y clasificar el riesgo de muerte en pacientes con COVID-19. Se analizan fuentes clave, autores y publicaciones relevantes, así como los componentes de los modelos, que incluyen las variables de resultado, predictores y los distintos tipos de algoritmos. Además, se presentan diferentes tipos de modelos de predicción clínica usados en la COVID-19, en particular los que se consideran de mayor calidad. Se fundamentan las decisiones metodológicas para el diseño de la investigación en los escenarios comunitarios y hospitalarios. También se discute la estrategia de modelado, incluidas las métricas para evaluar la validez de los modelos y cómo se comparan los modelos propuestos con los de mayor calidad. Como parte de la estrategia de modelado, se presentan las bases teóricas para el desarrollo de una aplicación de software móvil llamada COVIDCENCEC-APK. Finalmente, se abordan los esfuerzos para mitigar el sesgo, así como los desafíos para la generalización de los modelos propuestos en los distintos escenarios de atención médica en Cuba.

CAPÍTULO II. DISEÑO METODOLÓGICO

El presente capítulo tiene como objetivo presentar el diseño metodológico que permitirá cumplir los objetivos del estudio. Se inicia con una síntesis de los elementos metodológicos más relevantes. A continuación, se describe el tipo de estudio y su contexto, así como la definición del universo y la muestra. Se aborda la operacionalización de las variables y se detallan las técnicas y procedimientos que se utilizan, incluidas las que se emplean en la estrategia de modelado. Además, se consideran los aspectos éticos, así como las fortalezas y limitaciones metodológicas. El capítulo concluye con las consideraciones finales.

II.1 Generalidades

El estudio se llevó a cabo en dos etapas, en las cuales se abordaron los objetivos contextualizados en los escenarios de atención médica comunitaria y hospitalaria. A continuación, se resumen algunos de los elementos clave en el diseño de los modelos propuestos según las etapas de la investigación (tabla 2).

Tabla 2. Síntesis de los elementos metodológicos fundamentales para el desarrollo y validación de los modelos según la etapa del estudio

Características	Etapas	
	Etapla 1 (Modelos para escenario comunitario)	Etapla 2 (Modelos para escenario hospitalario)
Diseño	Cohorte retrospectiva	Casos y controles anidados
Participantes	Confirmados en la comunidad	Hospitalizados
Fuentes	RNPC	RICPA
Muestra	770 957	3683
Resultado	Estado al egreso	Estado al egreso
Variables predictoras	Edad y comorbilidades	Edad, comorbilidades, IRP, frecuencia respiratoria, presión arterial media, INL, el conteo de plaquetas
Datos perdidos	Casos completos	Imputación múltiple
Tipo de modelo	Árbol de clasificación (CTREE)	Árbol de clasificación (RPART), regresión logística binaria múltiple
Validación	Temporal	Geográfica
Validación interna	Validación cruzada (10 pliegues)	Validación cruzada (10 pliegues)

En la etapa uno, se diseñó un modelo para escenarios de atención médica de la comunidad. Este modelo se basó en un diseño de cohorte retrospectiva y se aplicó un algoritmo de árbol de clasificación CTREE ⁽⁴⁷⁾. Los predictores utilizados fueron la edad y las comorbilidades. Como técnica para la validación externa se usó la validación temporal.

En la etapa dos, se crearon dos modelos para pacientes hospitalizados. Estos modelos se basaron en un diseño de casos y controles anidados y se aplicó un árbol de clasificación RPART ⁽⁴⁶⁾ y una regresión logística binaria múltiple ⁽⁴⁸⁾.

Los predictores que se incorporaron para estos modelos fueron el IRPS, FR, PAM, INL y el PLT. Como técnica para la validación externa se usó la validación geográfica.

Para comparar el modelo COVIDCENCEC con los modelos de mayor calidad, ^(25,27) se transformaron los términos independientes de los modelos de mayor calidad basados en el algoritmo de regresión logística binaria múltiple (anexo 11, tabla 3).

Para todos los modelos la variable de resultado fue el estado al egreso, se realizó validación cruzada como técnica para la validación interna.

Durante todo el proceso de investigación científica, se realizaron acciones para identificar y mitigar el sesgo (anexo 12, tabla 4).

A continuación, se desarrollan con más detalles cada una de las etapas de la investigación.

II.2 Tipo de estudio y contexto

Se realizó un estudio metacéntrico con cobertura para todas las instituciones del Sistema Nacional de Salud en Cuba destinadas a la lucha contra la COVID-19 y como parte del proyecto ICPA ⁽³³⁾. El periodo de evaluación de los pacientes

comprendió entre las fechas del 1 de noviembre de 2019 y el 29 de noviembre de 2021. En ambas etapas antes mencionadas, se dividió la totalidad de los casos en dos subconjuntos de datos, uno de derivación (entrenamiento) y otro de validación (prueba). En este proceso, se incluyó el 70 % de los datos en el conjunto de entrenamiento, mientras que el restante 30 % se asignó al conjunto de prueba.

II.2.1 Primera etapa del estudio

La primera etapa se definió como un estudio de cohorte retrospectiva ⁽²⁹⁾. En la figura 9 se describe la distribución de participantes según un criterio temporal para la definición de la partición de entrenamiento y de prueba ^(29,30).

Este estudio incluyó a 770 957 pacientes en total. Los datos de los predictores se recogieron en la captación por el SNS y la variable de resultado fue el estado al egreso. Según el PAN versión 1.5, el egreso debe ocurrir cuando se obtenga un resultado negativo de los exámenes RCP, diagnosticador o prueba rápida o desaparezcan los síntomas o la muerte del paciente ⁽¹⁸⁾.

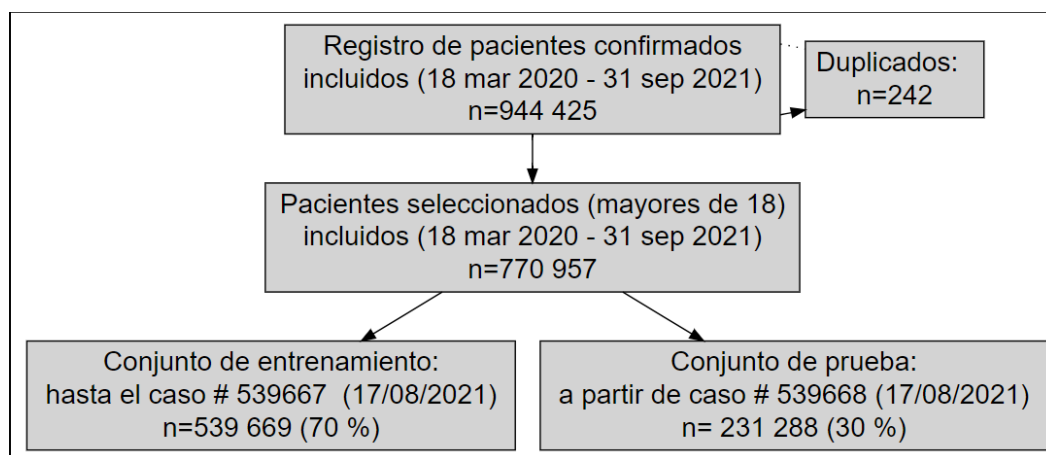


Figura 9. Participantes confirmados de COVID-19 según subconjunto para la modelación y criterio temporal para la definición de los grupos.

Se seleccionaron los primeros 539 669 pacientes confirmados para el conjunto de entrenamiento (validación temporal) y a partir de estos casos el subconjunto

de prueba ⁽²⁹⁾. El subconjunto de prueba incluyó a los restantes 231 288 pacientes.

II.2.2 Segunda etapa del estudio

La segunda etapa se concibió como un estudio de casos y controles anidados (CCA) ^(29,98). En la figura 10 se muestra la distribución de pacientes por hospitales, así como la partición de entrenamiento y prueba. Este análisis incluyó a 3683 pacientes de 37 hospitales de nivel secundario y terciario de todo el país.

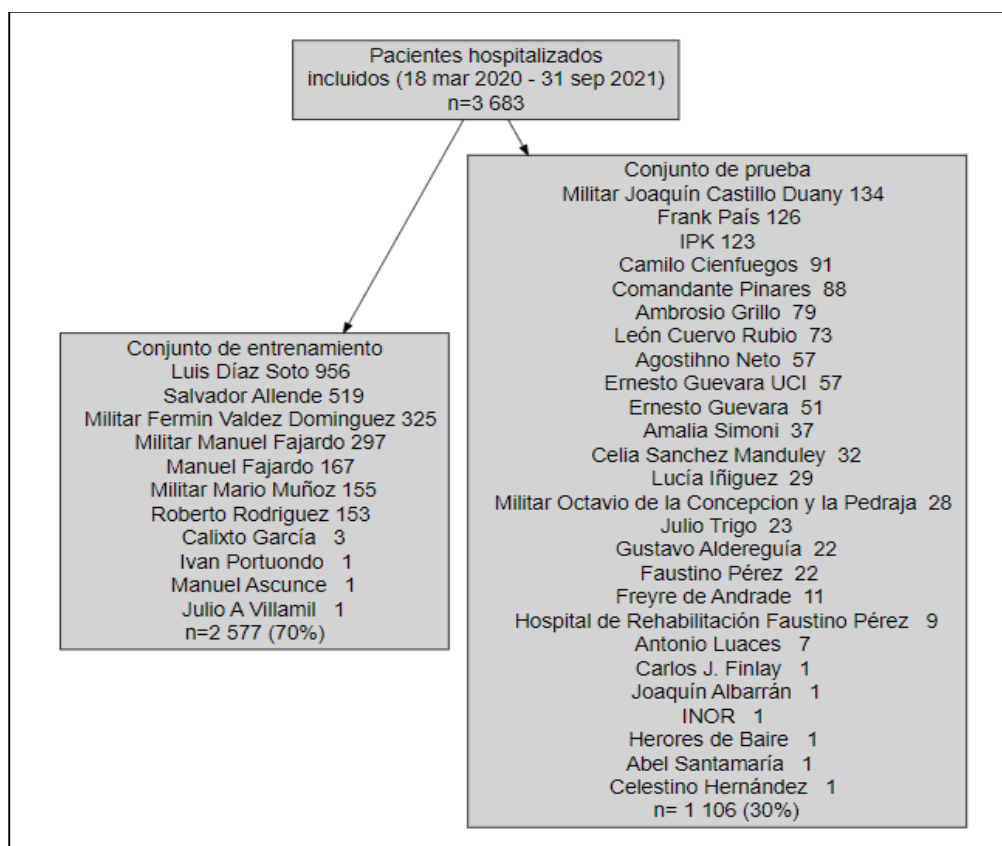


Figura 10. Participantes según hospitales y subconjuntos para la modelación.

En el subconjunto de entrenamiento se incluyeron 2577 pacientes ingresados, mientras que el subconjunto de prueba quedó conformado por 1106 participantes. En esta etapa se llevó a cabo una asignación aleatoria de los hospitales (validación geográfica) para definir los conjuntos de entrenamiento y prueba. A excepción de la variable de resultado que se recogió al egreso, el resto

de las variables predictivas se recogieron al ingreso en el hospital. En esta etapa, la razón de asignación es de tres controles por casos (1:3), o sea, tres vivos por cada fallecido.

II.3 Universo y muestra

En ambas etapas los criterios de inclusión y de exclusión fueron los siguientes:

Criterios de inclusión:

- Pacientes adultos de más de 18 años confirmados de COVID-19 por RCP o diagnosticadores cubanos y registrados en la base de datos del RNPC.

Criterios de exclusión:

- Pacientes con historias clínicas o registros duplicados.

Los pacientes se siguieron desde la confirmación de la enfermedad hasta que se cumplieran al menos uno de los siguientes criterios: negativización de la prueba RCP, resultado negativo del diagnosticador o desaparición de los síntomas. Se tomó un periodo de seguimiento máximo de 30 días a partir de la fecha de confirmación de la COVID-19 para identificar el acontecimiento de interés (muerte). Más allá de este periodo se consideró al paciente como libre de ocurrencia de este evento.

II.3.1 Muestreo para primera etapa de la investigación

En la primera etapa del estudio se analizaron a todos los pacientes que cumplieran los criterios de inclusión y exclusión durante el periodo de estudio.

II.3.2 Muestreo para la segunda etapa de la investigación

A continuación, se describe el proceso de cálculo del tamaño de muestra y selección de las unidades de análisis para la segunda etapa.

II.3.2.1 Cálculo del tamaño de muestra

Los parámetros para el cálculo del tamaño de muestra (TM) de pacientes hospitalizados se obtuvieron de la base del RNPC. Para determinar el número de sujetos requeridos para el desarrollo de un nuevo modelo de predicción clínica, se utilizó la función pmsampsize de la biblioteca de R del mismo nombre ⁽⁹⁹⁾. Esta función calcula el tamaño mínimo de la muestra necesario para el desarrollo de un nuevo modelo de predicción en el que se tienen en cuenta criterios que permiten minimizar el sobreajuste y garantizar una estimación precisa de los parámetros del modelo de predicción. Como parámetros se asumió el AUC de 0,9 del modelo de clasificación de la primera etapa (C), 10 predictores candidatos (PP).

La tasa de letalidad por COVID-19 en pacientes hospitalizados asumida fue de 7,4 % (IC del 95 % = del 7,4 % al 11,2 %) según se refiere en un metanálisis publicado por Sepadi y colaboradores ⁽¹⁰⁰⁾. Se establecen tres criterios para determinar el tamaño de muestra óptimo, el primero toma como base la diferencia entre el R-cuadrado aparente y ajustado, el segundo el margen de error en la estimación del intercepto y el tercero se basa en los eventos por parámetro del predictor (EPP). El criterio final se define como la muestra de mayor tamaño para garantizar que se cumplan todos los requerimientos (anexo 13, tabla 5).

El tamaño mínimo de muestra requerido para el desarrollo de nuevos modelos fue de 2930 pacientes. Finalmente se obtuvo información de 3683 sujetos, aproximadamente un 25 % más de sujetos de lo que se requería inicialmente, lo que se podría justificar como una previsión para posibles pérdidas.

II.3.2.2 Selección muestral

La selección de la muestra para la segunda etapa se realizó a partir de la lista completamente numerada de la base de datos RNPC. Solo se eligieron aquellos pacientes que se ingresaron en hospitales. Luego se seleccionaron 779 fallecidos (casos) y 2904 egresados vivos (controles), aproximadamente 3 controles por caso para un total de 3683 pacientes.

II.4 Operacionalización de variables

La variable dependiente fue la muerte (estado al egreso), definida como muerte confirmada o presunta por COVID-19 según el certificado de defunción o muerte que ocurrió en un individuo con infección confirmada por SARS-CoV-2 en cualquier momento desde el 1 de noviembre de 2019 hasta el 29 de noviembre del 2021.

Una de las ventajas de los algoritmos de árboles es su capacidad para identificar un punto de corte óptimo. En el ajuste del modelo ABR, se decidió a priori utilizar 65 años como punto de corte para la variable de edad. Esta decisión se fundamenta en varios factores. En primer lugar, el algoritmo CTREE ⁽⁴⁷⁾ exige dicotomizar a priori las variables cuantitativas, pero el algoritmo RPART ⁽⁴⁶⁾ puede definir el punto de corte óptimo. En este caso RPART ⁽⁴⁶⁾ definió esta división de manera automática ^(29,82).

En la segunda etapa, se modificó la operacionalización de la variable comorbilidades. A esta variable se le transformó su escala para evitar la pérdida de precisión que puede provocar la dicotomización ⁽¹⁰¹⁾. Aunque la dicotomización puede resultar en pérdida de información, en determinadas circunstancias puede facilitar un ajuste más adecuado de los datos al modelo predictivo ⁽³⁰⁾.

A continuación, se resume la operacionalización de las variables del estudio por etapas de la investigación (anexo 14, tabla 6). Todas los predictores son variables utilizadas al menos una vez en los modelos de mayor calidad ^(25,27), a excepción de IRPS que es una incorporación de este autor.

II.4.1 Procedimientos para la confección del Índice de riesgo ponderado de síntomas (IRPS)

Para confeccionar la variable sintética IRPS se realizó una suma ponderada de síntomas en función de su gravedad (anexo 15, tabla 7). Esta ponderación es el resultado del ajuste de un modelo de regresión logística a los antecedentes de síntomas o signos como variables independientes y la variable de resultado, estado al egreso. Posteriormente, los coeficientes se redondean al entero más próximo para obtener un valor para el síntoma en particular ⁽³⁰⁾. Se seleccionaron los coeficientes que eran mayores de uno. Esta variable está acotada de cero a cinco puntos, donde la disnea y la alteración del estado mental suman dos puntos cada una, y la tos suma un punto. El IRPS comienza en cero porque es el valor que representa la ausencia de síntomas (pacientes asintomáticos) y tiene su valor máximo en cinco porque es la suma de todos los puntos posibles de estar presentes todos los síntomas seleccionados.

II.5 Técnicas y procedimientos

Para el desarrollo y validación de los modelos en cada una de las etapas se aplicó el siguiente conjunto de técnicas y procedimientos.

II.5.1 Técnicas y procedimientos de obtención de información

Se importaron los ficheros de los registros RNPC y RICPA en formato de texto separado por comas (csv). Luego se cargaron en el software R versión 4.4.1.

Para ejecutar la estrategia de modelado se utilizaron dos metapaquetes, las bibliotecas tidyverse y tidymodels ⁽¹⁰²⁾.

II.5.2 Técnicas de procesamiento y análisis de la información

La estrategia de modelado de este estudio se basó en la metodología propuesta por Steyerberg ^(29,35). A continuación, se describen las técnicas y procedimientos empleados para el desarrollo y validación de los modelos propuestos según este enfoque.

II.5.2.1 Técnicas empleadas para la definición del problema e inspección de datos

Una vez analizado el problema de investigación y realizado el proceso de revisión documental, se seleccionó el estado al egreso como variable de resultado porque se relaciona directamente con el problema de la investigación. El evento muerte que corresponde a la categoría fallecido al egreso representó el evento de mayor interés por parte del sistema de salud durante la pandemia. Se realizó el proceso de EDA ⁽¹⁰²⁾ para explorar los tipos de datos presentes en base de datos y examinar sus distribuciones. Además, se calcularon medidas resúmenes como la mediana y el rango intercuartílico para las variables cuantitativas y la frecuencia absoluta y porcentajes para describir las características principales de los datos. Se analizaron las relaciones existentes entre los predictores y las asociaciones entre estos, así como con la variable de resultado. Se utilizó la técnica de winsorizing en el 1 y 99 percentil para todas las variables cuantitativas ⁽²⁹⁾, como estrategia para el manejo de datos atípicos. Se imputaron los valores ausentes en todas las variables seleccionadas que presentaban datos perdidos ⁽³⁰⁾. Las variables con un porcentaje inferior al 50 % de datos perdidos se excluyeron del estudio porque su imputación puede ser

improcedente (anexo 16, tabla 8). Una vez verificado el mecanismo de pérdida de datos, se llevó a cabo la imputación múltiple por medio de la biblioteca missRanger de R ⁽¹⁰³⁾. En esta biblioteca se implementa un método de imputación múltiple basado en el procedimiento 'MissForest', que permite la imputación de conjuntos de datos mixtos mediante el uso de algoritmos de bosques aleatorios ⁽¹⁰⁴⁾.

II.5.2.2 Técnicas empleadas para la codificación de variables predictoras

Para el caso del modelo COVIDCENCE que se basa en la regresión logística binaria múltiple, ⁽⁴⁸⁾ se verificó que existía una relación lineal entre la variable dependiente y las variables predictores de los modelos ⁽³⁰⁾. Para modelar las relaciones no lineales entre variables, se utilizó la técnica de spline cúbico con tres nodos ⁽¹⁰⁵⁾. Esta elección se basó en el equilibrio entre la flexibilidad del modelo y la prevención del sobreajuste ⁽³⁰⁾.

Se comprobó el supuesto de ausencia de multicolinealidad a través del factor de inflación de varianza (FIV). Para todas las variables, el valor de esta medida es bajo (anexo 17, figura 11). El hecho de que el FIV sea muy inferior a 5 se considera un indicador de no colinealidad en la regresión logística ^(30,82).

II.5.2.3 Técnicas empleadas para la especificación del tipo de modelo

Se desarrollaron y validaron tres modelos: el modelo de árbol CTREE ⁽⁴⁷⁾ nombrado ABR para el triaje en la primera etapa del estudio y los modelos AAR basados en árboles RPART, ⁽⁴⁶⁾ así como el modelo de regresión logística binaria múltiple nombrado como COVIDCENCEC para la segunda etapa.

El algoritmo que se utilizó para la estimación de los coeficientes de los predictores en COVIDCENCEC fue la estimación por máxima verosimilitud ⁽⁸⁹⁾. Para formular la clasificación del riesgo de muerte por COVID-19 a partir de los

modelos de árboles de clasificación ABR y AAR se calcularon los indicadores de estrato propuestos por este autor, el Tamaño relativo del estrato (TRE), el Riesgo de muerte (RM) y el Riesgo de muerte explicada (ME) (anexo 6).

En el caso del modelo COVIDCENCEC, se evaluaron los términos de interacción de la regresión logística; sin embargo, estos elementos no resultaron significativos (anexo 18, tabla 9). Se optó por un modelo más simple para evitar el sobreajuste.

La elección del conjunto de predictores candidatos y de los predictores finales de los modelos se basó en criterios de plausibilidad biológica, respaldo en la literatura y de significación estadística. En los modelos ABR y AAR, se utilizó el algoritmo de selección de atributos que forma parte del ajuste al árbol ⁽¹⁰⁶⁾.

II.5.2.4 Técnicas empleadas para la estimación de los modelos

Se utilizó la regresión logística binaria múltiple para ajustar el modelo COVIDCENCEC. Se aplicaron los algoritmos de árboles de clasificación binaria CTREE y RPART para los modelos ABR y AAR. La selección de predictores en los algoritmos basados en árboles se realizó mediante métodos de selección de atributos específicos para árboles de clasificación, lo que permitió identificar variables relevantes para el modelo ⁽²⁹⁾.

Elección del punto de corte para la escala COVIDCENCEC: Para establecer el punto del corte del modelo COVIDCENCEC, se utilizó el índice de Youden en la curva ROC ⁽²⁹⁾. Más allá de este valor definido se asumió que el paciente tenía alto riesgo de muerte.

II.5.2.5 Técnicas empleadas para la evaluación del rendimiento de los modelos

Se utilizó el paquete rms ⁽¹⁰⁷⁾ del programa R para validar los modelos de predicción clínica ⁽³⁰⁾. Se evaluó la calidad de los pronósticos a través de varias métricas de calibración y discriminación ⁽³⁵⁾. En las métricas de calibración se incluyó el intercepto y la pendiente de la recta de calibración ^(29,30,35), así como el índice de Brier ⁽³⁰⁾. Las métricas de discriminación fueron el índice de discriminación de Somers' D (Dxy), el R^2 y el AUC. Adicionalmente se utilizaron otras medidas de desempeño como la precisión, la sensibilidad, la especificidad, el VPP, el VPN, la LR+ y la LR- para evaluar el rendimiento de los modelos ⁽²⁹⁾.

II.5.2.6 Técnicas empleadas para la validación de los modelos

Para la validación interna de todos los modelos, se utilizó la validación cruzada de 10 pliegues ⁽²⁹⁾. En cuanto a la validación externa, se evaluó el rendimiento de los modelos entrenados a partir de un conjunto de datos independientes seleccionada por un criterio temporal para el caso de ABR y geográfico para los modelos AAR y COVIDCENCEC.

II.5.2.7 Técnicas empleadas para la presentación del modelo

Las reglas de decisión del modelo ABR se presentaron e integraron en la estratificación del PAN ⁽¹⁸⁾. Por su parte, para facilitar el uso de COVIDCENCEC se diseñó un nomograma y una aplicación para celulares o tabletas denominada COVIDCENCEC-APK para favorecer la generalización, la transparencia en su funcionamiento, así como su disponibilidad.

II.5.2.8 Técnicas empleadas para el desarrollo y validación de COVIDCENCEC-APK

El desarrollo de la aplicación COVIDCENCEC-APK para teléfonos móviles y tabletas implicó varias etapas. Primero, se diseñó la interfaz de usuario en HTML versión 5.3 y CSS3, para asegurar que fuera intuitiva y fácil de usar. Luego, se implementó la lógica de la aplicación a partir de JAVASCRIPT (versión ECMAScript 2021), que incluía la ejecución del modelo de regresión logística para calcular el riesgo basado en los datos del paciente ingresados por el usuario (anexo 19). Finalmente, se realizó una serie de pruebas para validar el desempeño de la aplicación y se realizó un informe final (anexo 26).

II.5.3 Técnicas y procedimientos de presentación de los resultados

Los resultados se presentaron en tablas y gráficos. Para el caso de las tablas se utilizó medidas resúmenes para las variables cualitativas como los porcentajes y se describió de la siguiente forma “frecuencia absoluta (porcentaje)”. Para el caso de las variables cuantitativas se utilizó como medidas resúmenes la mediana (Me) y el rango intercuartílico (RIC) de la siguiente forma “Me(RIC)”. El cálculo de las frecuencias en las tablas se realizó en función de los totales marginales de las columnas. Estas cifras se muestran en los encabezados de algunas tablas. En la comparación de algunas tablas se utiliza la diferencia porcentual. En el contexto de este estudio se usa la diferencia de los logaritmos naturales ⁽¹⁰⁸⁾. En las tablas donde se presentan los OR y su intervalo de confianza se utiliza el método de Bootstrap ⁽³⁰⁾. Para el caso de los OR ajustados se usa la regresión logística, dado que pueden obtener estimaciones del OR junto con sus intervalos de confianza directamente a partir de los exponenciales de los coeficientes ajustados. La ausencia del valor uno en el rango de los

intervalos de confianza para el OR constituye el criterio de significación estadística de este estudio.

II.6 Aspectos éticos

El proyecto de investigación recibió la aprobación del Consejo Científico del Centro Nacional Coordinador de Ensayos Clínicos. Esta investigación se incorporó al proyecto encargado por el Ministerio de Salud Pública de Cuba, titulado ⁽³³⁾ “Impacto del Programa de Actuación Nacional de la COVID-19 en pacientes adultos. Cuba, 2020-2023”. El proyecto fue la primera investigación aprobada por el Comité de Innovación del Ministerio de Salud Pública en marzo de 2020, y su ejecución comenzó en julio de 2020. Además, el estudio fue aprobado por el Comité de Ética del CENCEC. Los datos se utilizaron de forma anónima sin hacer referencia a los datos personales de los pacientes por lo que se garantizó la confidencialidad. Por las características del diseño del estudio no se evaluó directamente al paciente, sino a través de datos de registros e historias clínicas. Además, no se realizó experimentación en humanos.

II.7 Fortalezas y limitaciones metodológicas

La tesis propone un enfoque integral para la estratificación del riesgo de muerte por COVID-19, donde se combina el análisis de la literatura, varios algoritmos clásicos y técnicas de inteligencia artificial, diversas fuentes de datos y la colaboración de instituciones de todos los niveles del Sistema Nacional de Salud. El diseño multicéntrico, con un amplio número de participantes, proporciona mayor robustez y contextualización a los modelos propuestos. Se dio prioridad a la relevancia clínica, dado que se seleccionó a la variable “muerte” (estado al egreso) como variable de resultado y se seleccionó a predictores empleados por los modelos de mayor calidad reconocidos en la literatura ⁽²⁷⁾.

Los modelos propuestos son simples y fácilmente comprensibles, esto facilita su implementación en diversos escenarios de atención médica y su comprensión por parte del personal médico no familiarizado con la tecnología digital. Además, uno de los modelos propuestos fue utilizado en la práctica, lo que confiere un mayor prestigio al estudio. Asimismo, la implementación de uno de los modelos en un nomograma y en una aplicación para teléfonos móviles o tabletas facilita su uso, disponibilidad y capacidad de generalización en la atención médica.

En la segunda etapa del estudio, se utilizó el registro RICPA, con entrada de datos directamente desde los hospitales y verificación a partir de la historia clínica. Esta metodología es una fortaleza de este estudio y un enfoque moderno de recolección de datos en estudios multicéntricos. Además, la recopilación de datos durante los momentos de epidemia también se considera una fortaleza, ya que proporciona información valiosa sobre el comportamiento de la enfermedad en condiciones reales y en tiempo de ejecución.

El estudio presenta varias limitaciones metodológicas. En primer lugar, no fue posible comparar uno de los modelos seleccionados en la literatura con el modelo propuesto QCOVID ⁽⁵⁷⁾ para pacientes confirmados debido a la falta de disponibilidad de datos sobre sus predictores. En segundo lugar, se tuvo que descartar algunas variables importantes debido a que presentaban un alto porcentaje de datos perdidos. Además, no contar con registros de datos diseñados para la modelación predictiva y ajustados específicamente a los objetivos de la investigación también atenta contra la calidad de los modelos.

En la primera etapa del estudio, la base de datos del RNPC fue una fuente de gran utilidad. Sin embargo, es posible que contenga errores e inconsistencias. Por ejemplo, puede haber pacientes que se infectaron, pero no fueron

registrados. Además, la definición de muerte por COVID-19 puede haber sido atribuida o no a la enfermedad, lo que puede generar ambigüedad y afectar el desempeño de los modelos.

Otra limitación potencial es el periodo de seguimiento de 30 días para los pacientes. Si un paciente fallece después de este periodo, no se reflejaba en los datos del estudio, podría resultar en una representación inadecuada de la población estudiada.

En el estudio se llevaron a cabo un conjunto de estrategias para identificación y mitigación del sesgo (anexo 12, tabla 4). Sin embargo, siempre existirá un grado de incertidumbre inherente a cualquier estudio clínico. La transparencia en la identificación y manejo de estos sesgos resulta fundamental para la credibilidad y la aplicabilidad de los hallazgos. El estudio emplea criterios estrictos de inclusión y exclusión basados en medidas objetivas, como el estado al egreso a 30 días. Utiliza medidas estandarizadas y objetivas para los predictores, considera también el momento de las mediciones, y emplea múltiples fuentes de datos y técnicas de imputación múltiple para manejar los datos faltantes. Se aplican técnicas para gestionar los valores extremos. Además, adopta un enfoque multicéntrico junto con técnicas de validación cruzada para evaluar tanto la validez interna como externa del modelo, y favorecer su poder de generalización. Asimismo, el diseño multicéntrico y la aleatorización por pacientes y hospitales ayudan a mitigar el riesgo de sesgo de confusión.

II.8 Consideraciones finales

El diseño metodológico que se llevó a cabo en dos etapas permitió la creación de modelos predictivos para pacientes con COVID-19 confirmados en la comunidad y atendidos en hospitales. Se empleó un enfoque mixto al aplicar algoritmos clásicos y de aprendizaje automático. En opinión de este autor, el

enfoque utilizado de usar varios algoritmos y compararlos con otros modelos puede dar mayor confianza en los resultados. Finalmente, la presentación de uno de los modelos en forma de nomograma o aplicaciones para teléfonos móviles es factible con la tecnología disponible y facilita su generalización.

CAPÍTULO III. RESULTADOS

Se desarrolló el modelo ABR para la estimación y clasificación del riesgo de muerte por COVID-19 en escenarios de atención médica comunitaria, así como ABR y COVIDCENCEC para hospitales. Los modelos propuestos presentaron mejor rendimiento que los modelos de mayor calidad en escenarios de atención médica de Cuba. Finalmente se diseñó un nomograma y la aplicación COVIDCENCEC-APK para facilitar la generalización del modelo COVIDCENCEC.

III.1 Desarrollo de modelos para escenarios de atención médica de la comunidad

Para los 770 957 pacientes confirmados, la tasa de letalidad fue del 1 % (6619 fallecidos). La mediana de la edad para todos los pacientes estudiados fue de 48 años. Se observó un ligero predominio de mujeres (53 %) entre los confirmados. La mayoría de los pacientes fueron asintomáticos en el momento de la captación por el SNS (87 %). Por su parte, el 8,89 % de los pacientes (68 516) presentó al menos una comorbilidad.

La proporción de confirmados que presentaron comorbilidades en la partición de entrenamiento fue prácticamente la misma en el conjunto de entrenamiento (8,80 %) que en el de prueba (9,10 %). Se observó homogeneidad entre la mayoría de los predictores de ambas particiones. Todas las variables presentaron diez unidades o menos de variación porcentual entre los conjuntos de entrenamiento y el de prueba (tabla 10).

Tabla 10. Características clínicas de pacientes confirmados de COVID-19 según partición de los datos

Características	Global N = 770 957	Entrenamiento N = 539 669	Prueba N = 231 288
Fallecidos, <i>n</i> (%)	6 619 (0,86)	4 881 (0,90)	1 738 (0,75)
Edad años, <i>Me</i> (<i>RIC</i>)	48 (33; 59)	47 (32; 58)	50 (35; 62)
Sexo, <i>n</i> (%)			
<i>Femenino</i>	410 301 (53,22)	284 182 (52,66)	126 119 (54,53)
<i>Masculino</i>	360 648 (46,78)	255 479 (47,34)	105 169 (45,47)
Asintomáticos, <i>n</i> (%)	672 037 (87,17)	451 125 (83,59)	220 912 (95,51)
Comorbilidades¹, <i>n</i> (%)	68 516 (8,89)	47 469 (9,80)	21 047 (9,10)
HTA, <i>n</i> (%)	57 169 (7,42)	37 278 (6,95)	19 891 (6,60)
Asma bronquial, <i>n</i> (%)	10 924 (1,42)	2 766 (0,52)	8 158 (3,53)
Diabetes mellitus, <i>n</i> (%)	9 556 (1,24)	2 231 (0,42)	7 325 (3,17)
EPOC, <i>n</i> (%)	1 543 (0,20)	430 (0,08)	1 113 (14,15)
EIC, <i>n</i> (%)	7 203 (0,93)	1 711 (0,32)	5 492 (2,37)
Obesidad, <i>n</i> (%)	2 469 (0,32)	473 (0,09)	1 996 (0,86)
ERC, <i>n</i> (%)	824 (0,11)	286 (0,05)	538 (0,23)
Neoplasias, <i>n</i> (%)	2 404 (0,31)	751 (0,14)	1 653 (0,71)
ECV, <i>n</i> (%)	4 110 (0,53)	1 116 (0,21)	2 994 (1,29)
Inmunodeficiencias, <i>n</i> (%)	86 (0,01)	18 ($\approx$ 0,00)	68 (0,03)

Nota: ¹ comorbilidades = al menos una comorbilidad.

Se comprobó la asociación entre todos los predictores seleccionados y el estado al egreso fallecido en el conjunto de entrenamiento (tabla 11). A excepción de las inmunodeficiencias, todas las asociaciones con el estado al egreso cumplieron el criterio de significación estadística definido en el diseño metodológico de este estudio.

Tabla 11. Características clínicas y demográficas de pacientes confirmados de COVID-19 según el estado al egreso en el conjunto de entrenamiento

Características	Vivo, N = 534 788	Fallecido, N = 4 881	OR (95% IC)
Edad (años), Me (RIC)	47 (32; 58)	71 (59; 80)	1,08 (1,07; 1,08)
Sexo masculino, n (%)	252 692 (47,25)	2 787 (57,10)	1,49 (1,40; 1,57)
Sintomático, n (%)	446 577 (83,51)	4 548 (93,18)	2,73 (2,44; 3,06)
Comorbilidades, n (%)	42 783 (8,00)	4 686 (96,00)	1,22 (1,18; 1,25)
HTA, n (%)	33 959 (6,35)	3 319 (68,00)	1,07 (1,05; 1,11)
Asma bronquial, n (%)	2 273 (0,43)	493 (10,10)	0,15 (0,14; 0,16)
Diabetes mellitus, n (%)	547 (0,10)	1 684 (34,50)	34,85(31,63;38,3)
EPOC, n (%)	88 (0,02)	342 (7,00)	42,81(33,86;54,2)
EIC, n (%)	1 223 (0,23)	488 (10,00)	4,40(3,96;4,89)
Obesidad, n (%)	131 (0,02)	342 (7,00)	4,40(3,96;4,89)
ERC, n (%)	151 (0,03)	135 (2,77)	9,81(7,76;12,36)
Neoplasias, n (%)	369 (0,07)	382 (7,83)	11,41(9,88;13,17)
ECV, n (%)	818 (0,15)	298 (6,10)	3,13(1,03;9,50)
Inmunodeficiencias, n (%)	14 (≈0,00)	4 (0,08)	3,13(1,03;9,50)

Nota: ¹ comorbilidades = al menos una comorbilidad.

La edad fue un predictor de muerte a considerar, con un OR puntual de 1,08 y un intervalo de confianza del 95 % entre 1,07 y 1,08. La evidencia apoya el criterio de que los hombres, una vez confirmados, tenían una mayor probabilidad de morir en comparación con las mujeres (OR: 1,49; IC 95 %: 1,40-1,57). Por su parte, los pacientes con síntomas en el momento de la confirmación tenían un riesgo significativamente mayor de fallecer en comparación con los que en ese instante eran asintomáticos (OR: 2,73; IC 95 %: 2,44-3,06).

La presencia de comorbilidades aumentó el riesgo de muerte en comparación con los individuos que no tuvieron ese antecedente. En este sentido, se estimó un OR puntual de 1,22 y un intervalo de confianza del 95 % entre 1,18 y 1,25 para este antecedente. Varias comorbilidades específicas se asociaron a la mortalidad, las de mayor relevancia fueron la EPOC (OR: 42,81; IC 95 %: 33,86-54,2), la diabetes mellitus (OR: 34,85; IC 95 %: 31,63-38,3) y las neoplasias (OR: 11,41; IC 95 %: 9,88-13,17).

Todas las variables mencionadas en la tabla anterior se incorporaron al ajuste del algoritmo de árbol de clasificación para obtener el modelo ABR (figura 12). El nodo raíz de ABR estuvo constituido por la variable edad. Luego se definió el primer criterio de decisión a partir del antecedente del paciente de ser menor de 65 años. En los nodos terminales se incluyó la variable comorbilidades y en sus ramas se tuvo en cuenta la presencia o no del antecedente de al menos una comorbilidad. Para construir la clasificación del riesgo se emplearon los criterios de todas las ramas en su camino desde cada nodo terminal hasta el nodo raíz.

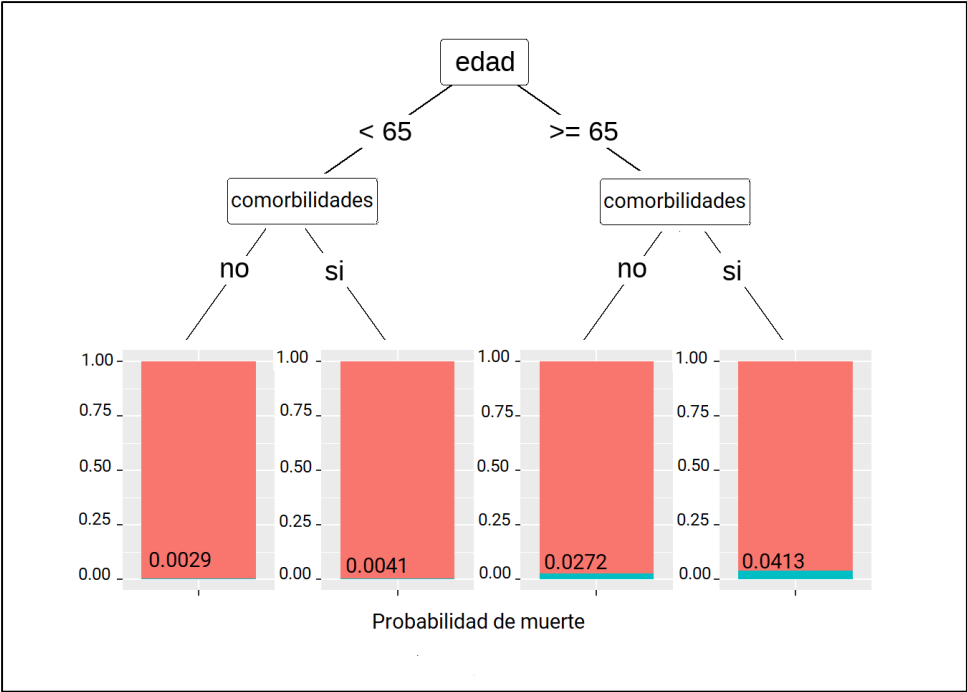


Figura 12. Especificación del modelo ABR para la estratificación de riesgo de muerte en pacientes confirmados por COVID-19.

Para definir la clasificación del riesgo se utilizaron los indicadores de estrato que se derivan del modelo ABR y que se describen a continuación (tabla 12).

Tabla 12. Clasificación de riesgo de transición a la gravedad en pacientes confirmados de COVID-19 según el clasificador ABR en la partición de desarrollo

Estrato	n	muertes	Indicadores de estrato		
			TRE (%)	RM (%)	ME (%)
Edad < 65 y sin comorbilidades	150 779	431	27,94	0,29	8,83
Edad < 65 y comorbilidades	301 204	1229	55,81	0,41	25,18
Edad ≥ 65 y sin comorbilidades	28 338	771	5,25	2,72	15,80
Edad ≥ 65 y comorbilidades	59 348	2450	11,00	4,13	50,19
Total	539 669	4881	100	100	100

El estrato de “paciente de bajo riesgo (BR)” se seleccionó como el estrato de menor probabilidad de muerte (RM=0,29 %). Este grupo se conformó con los pacientes confirmados menores de 65 años y sin comorbilidades.

Según el análisis, el riesgo individual para una persona sin comorbilidades y con edad menor de 65 años es apenas del tres por mil (RM=0,3 %), por lo que sería muy raro que en un grupo de cien pacientes confirmados muriera algún individuo a causa de la COVID-19. Sin embargo, este grupo incluía casi una tercera parte (TRE=27,94%) de todo el volumen de pacientes. Los resultados tienen implicaciones prácticas porque tratar a todos los pacientes de la misma manera o en el mismo escenario de atención médica (por ejemplo, si en lugar de ingresar al paciente en el hogar con tratamiento sintomático y realizar el seguimiento por la APS se hubieran atendido en un escenario hospitalario) hubiera resultado en un estrés innecesario al sistema de salud con pacientes de muy bajas probabilidades de muerte (ME=8,83%).

Por su parte, los pacientes de alto riesgo en el escenario comunitario se podrían definir como las personas mayores de 65 años con antecedentes de comorbilidades. En este grupo, el riesgo de mortalidad es del 4,13 % y es donde ocurre alrededor de la mitad de las defunciones (ME=50,19 %), por lo que se puede considerar intensificar las medidas terapéuticas. Sin embargo, se trata de

un grupo bastante reducido de individuos (TRE=11 %). Es posible que muchos de estos pacientes hayan sido atendidos en escenarios de atención médica hospitalaria.

Durante la pandemia de COVID-19, las reglas de clasificación del árbol y los indicadores de estrato de ABR, se utilizaron junto con otras evidencias, para definir criterios de estratificación de riesgo a partir de la versión 1.5 del PAN (figura 13) ⁽¹⁸⁾.

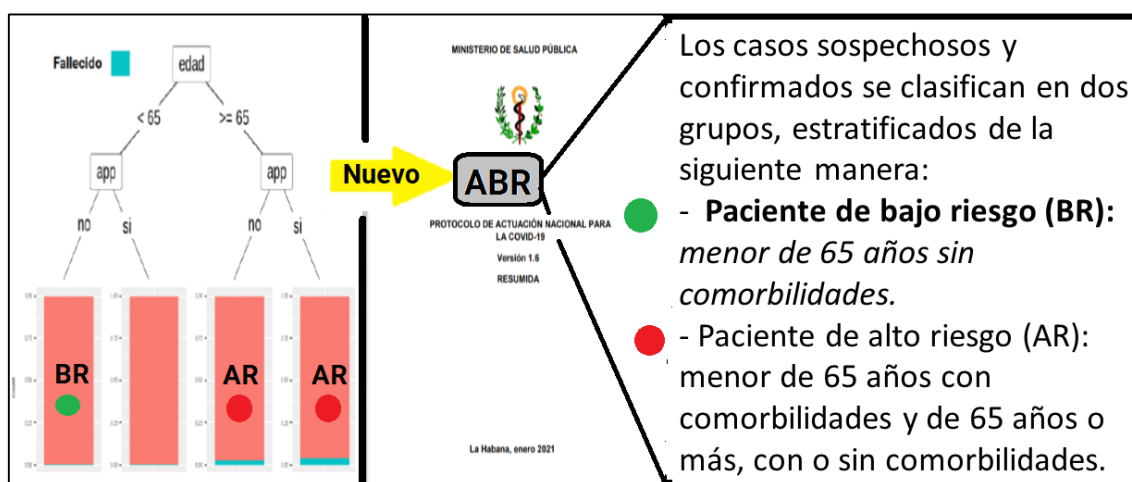


Figura 13. Relación de ABR con la estratificación del riesgo en el PAN.

Por ejemplo, según la estimación y clasificación del riesgo de muerte del modelo ABR y la conducta terapéutica pautada en el PAN ^(17,18), los pacientes menores de 65 años y sin comorbilidades se clasificaron como de bajo riesgo (BR) y se derivaron a centros de aislamiento no hospitalarios si el paciente era asintomático, o a un hospital o institución con condiciones básicas de atención si presentaba síntomas. En ambos casos, se tomó la decisión de prescribir el tratamiento con Heberferon ⁽¹⁷⁾.

III.2 Desarrollo de modelos para escenarios hospitalarios

Se analizaron 3683 pacientes hospitalizados, con una edad mediana de 56 años.

El 64,2 % estaban aparentemente sanos, el 25,71 % tenían una comorbilidad y 10,07 % padecía de dos o más enfermedades (tabla 13).

Las características de los pacientes en el escenario de atención médica hospitalaria fueron bastante similares entre el conjunto de entrenamiento y el de prueba. Los predictores con mayor heterogeneidad entre ambas particiones fueron el antecedente de HTA y los valores de LDH. Algunos predictores tenían más del 50 % de datos perdidos y se excluyeron de la modelación. Todas las variables cuantitativas se acotaron mediante winsorizing para reducir la influencia de los datos atípicos y mitigar el sesgo de información.

Tabla 13. Características demográficas, clínicas y de laboratorio de pacientes hospitalizados con COVID-19 según partición de datos

Características	Global N = 3683	Entrenamiento N = 2577	Prueba N = 1106
Fallecidos, <i>n</i> (%)	779 (21,15)	547 (21,23)	232 (20,98)
Edad, años, <i>Me (RIC)</i>	56 (40; 70)	56 (41; 71)	55 (38; 70)
Sexo, <i>n</i> (%)			
<i>Femenino</i>	1701 (46,18)	1275 (49,49)	529 (47,85)
<i>Masculino</i>	1982 (53,81)	1302 (50,51)	577 (52,15)
Comorbilidades, <i>n</i> (%)			
<i>Sin comorbilidades</i>	2365 (64,21)	1663 (64,53)	702 (63)
<i>1 comorbilidad</i>	947 (25,71)	669 (25,96)	278 (25)
<i>2 o más comorbilidades</i>	371 (10,07)	245 (9,51)	126 (11)
HTA, <i>n</i> (%)	1572 (42,68)	1121 (43,50)	451 (41)
Asma bronquial, <i>n</i> (%)	317 (8,61)	220 (8,54)	97 (8,8)
Diabetes mellitus, <i>n</i> (%)	599 (16,26)	411 (15,95)	188 (17)
EPOC, <i>n</i> (%)	110 (2,99)	80 (3,10)	30 (2,7)
EIC, <i>n</i> (%)	227 (6,16)	165 (6,40)	62 (5,6)
EC, <i>n</i> (%)	267 (7,25)	191 (7,41)	76 (6,9)
Obesidad, <i>n</i> (%)	115 (3,12)	74 (2,87)	41 (3,7)
ERC, <i>n</i> (%)	106 (2,88)	80 (3,10)	26 (2,4)
Neoplasias, <i>n</i> (%)	104 (2,82)	65 (2,52)	39 (3,5)
ECV, <i>n</i> (%)	35 (0,95)	25 (0,97)	10 (0,9)
Inmunodeficiencias, <i>n</i> (%)	32 (0,97)	24 (0,93)	8 (0,7)
IRPS, <i>Me (RIC)</i>	0 (0; 2)	0 (0; 2)	0 (0; 2)
FC (Lat. x min.), <i>Me (RIC)</i>	80 (76; 86)	80 (77; 86)	80 (76; 86)
FR (x min.), <i>Me (RIC)</i>	18 (18; 20)	18 (18; 20)	18 (18; 20)
TAM (mm/Hg), <i>Me (RIC)</i>	90 (83; 93)	90 (83; 93)	90 (83; 93)
VSG, <i>Me (RIC)</i>	23 (12; 45)	25 (12; 46)	20 (10; 43)
INL, <i>Me (RIC)</i>	2,21 (1,41; 3,82)	2,20 (1,38; 3,71)	2,23 (1,51; 4)
PLT, ($\times 10^9/\text{ml}$), <i>Me (RIC)</i>	230 (190; 278)	230 (193; 277)	230 (188; 279)
PCR, <i>Me (RIC)</i>	11 (3; 36)	11 (3; 37)	11 (3; 32)
Cr, ($\mu\text{mol/L}$), <i>Me (RIC)</i>	89 (73; 109)	89 (74; 109)	90 (71; 109)
Urea (mg/100 ml), <i>Me (RIC)</i>	6 (4; 11)	6 (4; 14)	5 (4; 10)

LDH (u/l), Me (RIC)	284 (186; 424)	284 (192; 421)	291 (167; 440)
SpO2, Me (RIC)	95 (90; 98)	95 (90; 98)	95 (91; 98)

Nota: ¹ suma de comorbilidades

Se identificaron varios predictores significativos del riesgo de muerte por COVID-19. La edad fue un predictor de muerte a considerar: con cada año adicional de edad aumenta el riesgo de muerte en un 1,08 %. Se encontró que la presencia de comorbilidades, como la hipertensión arterial, la diabetes mellitus y la enfermedad pulmonar obstructiva crónica, se asociaba con un mayor riesgo de morir.

El índice de riesgo ponderado de síntomas (IRPS), que resume la sintomatología del enfermo, también parecía desempeñar un papel importante en el pronóstico de los pacientes con COVID-19 (OR: 2,96, IC 95 %: 2,68-3,27). A medida que aumentaba el IRPS, el riesgo de muerte también se incrementó.

Los datos del examen físico como la frecuencia cardíaca y la FR resultaron ser predictores significativos del estado al egreso fallecido.

Por su parte, los exámenes complementarios como el INL, el PLT, la Cr y la urea fueron predictores importantes. El INL, un indicador de la respuesta inmunitaria del organismo, mostró que a medida que aumentaba, también lo hacía el riesgo de muerte (OR: 1,36, IC 95 %: 1,29-1,43). El OR para el PLT era cercano a dos, por lo que las plaquetas podrían tener un papel en el pronóstico de los pacientes con COVID-19. En los indicadores de la función renal y de la disfunción multiorgánica como la Cr y la urea, se describe que a medida que aumentaban sus cifras, también lo hacía el riesgo de muerte. En cuanto a la saturación de oxígeno periférico fue un factor protector (OR: 0,94, IC 95 %: 0,92-0,97) para el deterioro del estado de salud de estos pacientes (tabla 14). Estos resultados subrayan la complejidad de la COVID-19 y la necesidad de considerar una

amplia gama de hallazgos del interrogatorio, el examen físico y los exámenes complementarios al predecir la transición a la gravedad.

Tabla 14. Características clínicas según el estado al egreso de pacientes hospitalizados con COVID-19 en el conjunto de entrenamiento

Características	Vivo N = 1955	Fallecido N = 622	OR (IC 95%)
Edad (años), Me (RIC)	51 (35; 63)	73 (62; 82)	1,08 (1,07 ; 1,09)*
Sexo masculino, n (%)	970 (49,59)	332 (53,45)	1,16 (0,97 ; 1,39)
Comorbilidades ¹ , Me (RIC)	0 (0; 1)	1 (0; 1)	2,49 (2,17 ; 2,86)*
HTA, N (%)	763 (39,03)	386 (62,06)	2,55 (2,12 ; 3,07)*
Asma bronquial, n (%)	98 (5,01)	62 (9,97)	1,35 (0,96 ; 1,88)
Diabetes mellitus, n (%)	235 (11,78)	193 (31,03)	3,29 (2,65 ; 4,09)*
EPOC, N (%)	41 (2,06)	47 (7,56)	3,80 (2,49 ; 5,83)*
EIC, n (%)	94 (4,71)	87 (13,99)	3,21 (2,37 ; 4,37)*
Obesidad, n (%)	41 (2,06)	40 (6,43)	3,20 (2,06 ; 4,99)*
ERC, n (%)	35 (1,75)	54 (8,68)	5,18 (3,36 ; 7,99)*
Neoplasias, n (%)	29 (1,45)	40 (6,43)	4,54 (2,80 ; 7,36)*
ECV, n (%)	14 (0,70)	15 (2,41)	3,41 (1,66 ; 7,03)*
Inmunodeficiencias, n (%)	22 (1,10)	3 (0,48)	0,48 (0,15 ; 1,50)
IRPS, Me (RIC)	0 (0; 0)	2 (0; 2)	2,96 (2,68 ; 3,27)*
FC (Lat. x min), Me (RIC)	80 (76; 83)	86 (80; 96)	1,08 (1,07 ; 1,09)*
FR (Resp. x min), Me (RIC)	18 (17;20)	22 (19; 26)	1,34 (1,30 ; 1,39)*
TAM, (mm/Hg), Me (RIC)	90 (83, 93)	92 (83, 97)	1,00 (0,99 ; 1,01)
VSG, (mm/h), Me (RIC)	20 (10; 40)	55 (26; 80)	1,02 (1,02 ; 1,03)*
INL, Me (RIC)	1,90 (1,24, 2,85)	4,00 (2,33, 7,32)	1,36 (1,29 ; 1,43)*
PLT (x10 ⁹ /ml), Me (RIC)	230 (200, 277)	209 (170, 270)	2,00 (1,01 ; 3,00)*
PCR, Me (RIC)	10 (2, 36)	17 (11, 76)	1,02 (1,00 ; 1,03)*
Cr (μmol/L), Me (RIC)	85 (72; 103)	117 (87; 184)	1,01 (1,01 ; 1,01)*
UREA mg/100 ml, Me (RIC)	5 (4; 9)	19 (7; 31)	1,03 (1,02 ; 1,05)*
LDH (u/l), Me (RIC)	270 (190; 370)	432 (221; 639)	1,00 (1,00 ; 1,01)
SpO ₂ , Me (RIC)	97 (93; 98)	93 (83; 97)	0,94 (0,92 ; 0,97)*

Nota: 1 comorbilidades = Suma de comorbilidades

* resultado estadísticamente significativo

III.2.2.1 Desarrollo del modelo árbol de alto riesgo AAR

El modelo AAR es un árbol de clasificación de cuatro niveles de profundidad. En el algoritmo RPART ⁽⁴⁶⁾ se utilizó como nodo raíz el IRPS y como criterio de decisión de sus dos ramas valores de cero (asintomático) y uno o más (síntomas). Durante el ajuste también se seleccionaron predictores como la edad, la FR, el INL y las comorbilidades. La figura 14 muestra la estructura del árbol con los detalles de sus nodos internos, ramas con sus correspondientes

reglas de decisión, así como los nodos terminales u hojas. Nótese en la representación de AAR que para cada hoja se describe el estado al egreso pronosticado, así como un conjunto de valores que coinciden con dos de los indicadores de estrato propuestos por este autor (RM y TRE).

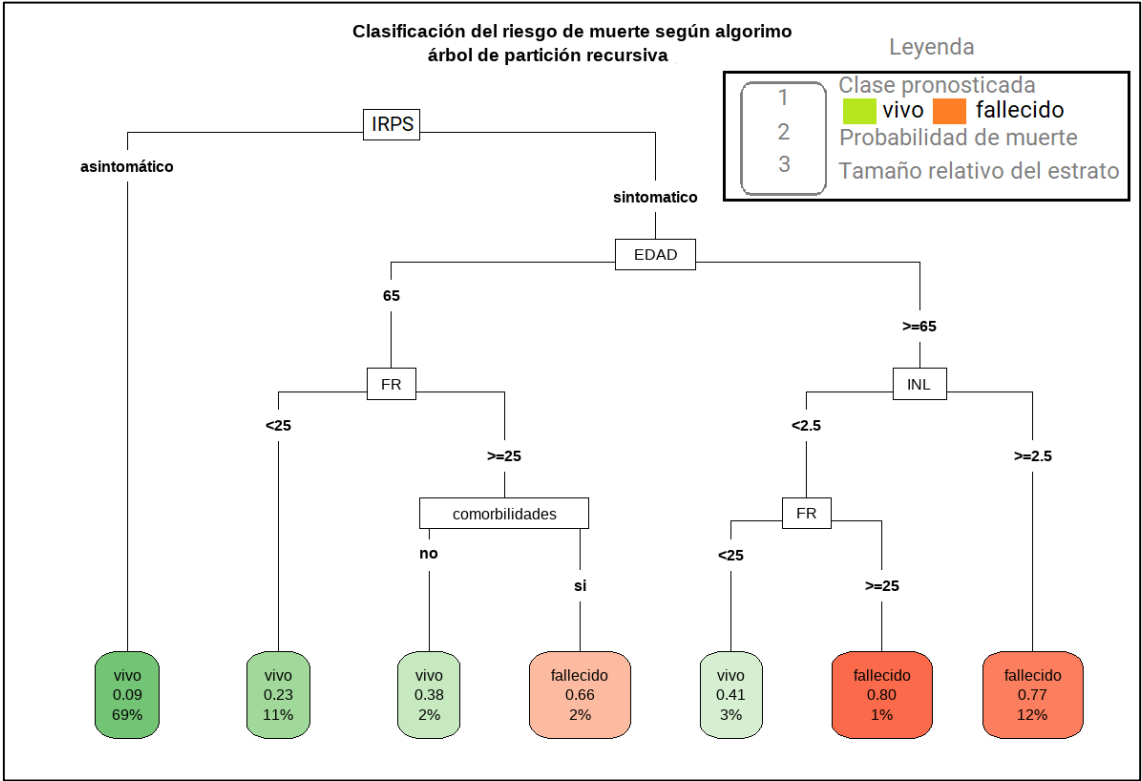


Figura 14. Modelo AAR ajustado al conjunto de entrenamiento.

Se describe la clasificación del riesgo y los indicadores que se derivan de AAR ajustado al conjunto de entrenamiento (tabla 15). Es importante señalar que estos valores corresponden a un estudio de casos y controles y que solo tienen un fin ilustrativo. A diferencia del ABR, estos no guardan relación con las tasas de prevalencia reales; sin embargo, este modelo se puede ajustar para la estratificación de pacientes en escenarios hospitalarios.

Por ejemplo, un individuo con una puntuación del IRPS menor que uno o igual tiene 0,09 probabilidades de morir, mientras que un individuo con IRPS de dos en lo adelante (al menos disnea o confusión mental), mayor de 65 años, INL

menor de 2,5 unidades, pero con una FR igual o superior a 25 respiraciones por minuto tiene una probabilidad de muerte de 0,8 y sería un paciente de muy alto riesgo. En contraste, una persona con un IRPS de uno o menos, sería de muy bajo riesgo en la estratificación en el escenario de atención médica hospitalaria. Por su parte, el grupo de personas con IRPS de uno o menos es el de más bajo riesgo (RM = 9 %) y mayor volumen de pacientes (TRE = 69 %). En cambio, el grupo de pacientes con IRPS superior a uno, con 65 años o más, INL mayor de 2,5 y FR mayor de 25 respiraciones por minuto o igual constituye la clasificación de mayor riesgo de muerte (RM = 80 %), pero el menor volumen de pacientes (TRE = 1 %). En este grupo de alto riesgo es precisamente donde se debería intensificar aún más la conducta médica.

Tabla 15. Clasificación de riesgo transición a la gravedad en pacientes hospitalizados por COVID-19 según el clasificador ABR en la partición de entrenamiento

Estratos	n	Fallecidos	Indicadores de estrato		
			TRE (%)	RM (%)	ME (%)
IRPS ≤ 1	1 778	160	68,99	9	28,12
IRPS > 1 Edad < 65 años FR < 25	283	65	10,98	22,97	11,42
IRPS > 1 Edad < 65 años FR ≥ 25 Sin comorbilidades	52	20	2,02	38,46	3,51
IRPS > 1 Edad < 65 años FR ≥ 25 Con comorbilidades	52	34	2,02	65,38	5,98
IRPS > 1 Edad ≥ 65 años INL $< 2,5$ FR < 25	77	32	2,99	41,56	5,62
IRPS > 1 Edad ≥ 65 años INL $< 2,5$ FR ≥ 25	26	21	1,01	80,77	3,69

IRPS <1, Edad ≥ 65 años INL ≥ 2,5	309	238	11,99	77,02	41,83
Total	2577	569	100	100	100

III.2.2.2 Desarrollo del modelo COVIDCENCEC

Se ajustó un modelo de regresión logística con el objetivo de predecir el estado de egreso de los pacientes a partir de los predictores seleccionados (tabla 16).

La edad presentó un OR de 1,11 (IC 95 %: 1,07; 1,15), y el INL, con un OR de 1,86 (IC 95 %: 1,48; 2,34). La suma de comorbilidades y la variable IRPS también resultaron ser predictores de muerte estadísticamente significativos. Los predictores se ajustaron con un spline cúbico ⁽³⁰⁾ para relajar el incumplimiento de linealidad entre la variable dependiente y el logaritmo de la variable independiente. En el anexo 19 se muestra la ecuación del modelo COVIDCENCEC.

Tabla 16. Ajuste de modelo de regresión logística a la mortalidad de pacientes hospitalizados con COVID-19

Características	OR (95 % CI)¹
Intercepto	0,00 (0,00; 0,09)
Edad (años)	1,11 (1,07; 1,15)
Edad (años) ' 	0,96 (0,93; 0,99)
FR (Resp. x min)	1,13 (1,03; 1,30)
FR (Resp. x min) ' 	1,00 (0,69; 1,45)
TAM (mm/Hg)	0,96 (0,92; 1,00)
TAM (mm/Hg) ' 	1,05 (0,99; 1,10)
INL	1,86 (1,48; 2,34)
INL ' 	0,46 (0,27; 0,77)
PLT (x10⁹/ml)	1,00 (0,99; 1,00)
PLT (x10⁹/ml) ' 	2,00 (1,15; 3,01)
Suma de comorbilidades	1,92 (1,49; 2,46)
IRPS	1,94 (1,61; 2,34)

Nota: “ ‘ ” señala la transformación de la variable original con un spline cúbico

En la selección de los predictores finales se combinaron tres criterios: en primer lugar, el respaldo por la evidencia, dado que todos los predictores son partes de al menos uno de los modelos de mayor calidad, el criterio de selección de

variables aportado por los algoritmos de árboles de clasificación y el criterio de significación estadística definido en este estudio.

En relación con la inclusión de predictores no significativos en el modelo final, es importante señalar que la significación estadística no es el único criterio para la selección de predictores en un modelo pronóstico. Los predictores pueden ser clínicamente relevantes y proporcionar información valiosa sobre la relación entre las características del paciente y el resultado. Además, la inclusión de estos predictores puede mejorar la precisión del modelo. Por lo tanto, la decisión de incluir estos predictores en el modelo final puede estar justificada desde una perspectiva clínica y metodológica.

III.3 Validación de modelos propuestos en distintos escenarios

Se validaron modelos para la estimación y clasificación del riesgo de muerte por COVID-19 de pacientes confirmados en la comunidad y atendidos en el hospital. Es importante destacar que, si bien los modelos para escenarios hospitalarios solo se validaron en este estudio, el modelo propuesto para el escenario comunitario sí fue utilizado en el triaje de pacientes durante la pandemia, por lo que se podría considerar adicionalmente una validación práctica.

III.3.1 Validación de modelos para pacientes de la comunidad

Se evaluó el rendimiento del modelo ABR en los conjuntos de entrenamiento y prueba. Aunque el R^2 no fue muy alto, el modelo se desempeñó bien frente a datos desconocidos. La precisión, el valor predictivo positivo y negativo y el área bajo la curva fueron más altos en el conjunto de prueba, e indicó un rendimiento del modelo aceptable (tabla 17). Sin embargo, la sensibilidad fue más baja en el conjunto de prueba, lo que se podría tomar en cuenta como una limitación al predecir la muerte (los falsos negativos serían fallecidos que escaparían al triaje inicial).

Tabla 17. Métricas de desempeño del modelo ABR según partición de datos

Métrica	Entrenamiento	Prueba
Precisión	0,62	0,65
Sensibilidad	0,59	0,41
Especificidad	0,66	0,89
VPP	0,63	0,79
VPN	0,62	0,60
Dxy	0,31	0,40
AUC	0,66	0,70
R2	0,10	0,14
Intercepto	0,00	-0,36
Pendiente	1,00	1,42

El modelo tendió ligeramente a subestimar las probabilidades predichas. La curva de calibración de ABR desciende ligeramente por debajo de la bisectriz en el conjunto de prueba (anexo 20, figura 15).

El modelo ABR puede ser de utilidad para clasificar en grupos de alto y bajo riesgo de muerte en pacientes confirmados en la comunidad. Esta estratificación puede contribuir a mejorar la eficacia de las acciones terapéuticas, lograr mayor beneficio para los pacientes y proteger el sistema de salud de la saturación al mismo tiempo que mejora el confort del paciente. En especial ABR es muy útil para identificar pacientes que podrían ser ingresados en el hogar con tratamiento sintomático y seguimiento por la APS, o sea, para la identificación de grupos de muy bajo riesgo.

III.3.2 Validación de modelos para escenarios hospitalarios

III.3.2.1 Validación del modelo Árbol de alto riesgo (AAR)

El modelo AAR presenta buenos resultados en relación con las métricas de calibración y discriminación (anexo 21, figura 16). Dentro de la figura, las curvas de calibración A y B muestran que las probabilidades predichas y observadas exhiben intersecciones cercanas a cero y pendientes aproximadamente iguales a uno. Esto indica que AAR tuvo una alta precisión y estaba bien calibrado.

Además, la alineación estrecha de la recta de calibración con la línea diagonal sugirió que no hay evidencia de sobreajuste o subajuste. En relación con la discriminación, las curvas ROC que se muestran en C y D describen valores altos del AUC. Se pudo comprobar que tanto para el conjunto de entrenamiento como para el de prueba los valores son aproximadamente 0,9.

III.3.2.2 Validación del modelo COVIDCENCEC y definición de puntos de corte

El modelo COVIDCENCEC muestra un buen desempeño en términos de calibración y discriminación (anexo 22, figura 17). Nótese que la recta de calibración es prácticamente ideal, con un intercepto de la recta de calibración (Intercepto, para abreviar) con valor cero tanto en el conjunto de entrenamiento como en el de prueba. Además, la pendiente de la recta de calibración (pendiente, para abreviar) es prácticamente uno, lo que demuestra que el modelo está bien calibrado. En relación con la discriminación, los valores de la curva ROC para ambos conjuntos son iguales a 0,92, lo que denota muy buena capacidad de discriminación.

Se definieron puntos de corte para agrupar las probabilidades pronosticadas en categorías de riesgo. En la subfigura C se ilustra la capacidad de discriminación de COVIDCENCEC y el valor máximo del índice de Youden ⁽²⁹⁾. Este índice se empleó para determinar el punto de corte de la escala, establecido en 0,44 (equivalente a 136 puntos en el nomograma). Se consideró que cualquier valor superior a este indicaba un estrato de ‘pacientes de alto riesgo’.

En el análisis que sigue se compara el modelo COVIDCENCEC con los modelos de mayor calidad en relación con su desempeño en el escenario de atención médica hospitalaria. La razón fundamental de seleccionar este modelo entre los

propuestos fue por la similitud entre los algoritmos, el propósito y los escenarios de atención médica donde podrían ser utilizados.

III.4 Contraste entre COVIDCENCEC y modelos de mayor calidad

Como parte de la validación se decidió contrastar el desempeño del modelo COVIDCENCEC con los modelos identificados en la literatura como modelos de mayor calidad ^(3,27). En la tabla 18 se describen las métricas de la matriz de confusión para los modelos seleccionados.

Tabla 18. Métricas de la matriz de confusión según modelos seleccionados

Métrica	COVIDCENCEC	NEWS2	PRIEST	Xie	4C MORTALITY
Precisión	0,88	0,81	0,66	0,70	0,82
Sensibilidad	0,87	0,84	0,67	0,70	0,85
Especificidad	0,85	0,78	0,65	0,71	0,79
VPP	0,86	0,75	0,61	0,71	0,76
VPN	0,87	0,87	0,71	0,69	0,87
LR+	5,80	3,82	1,91	2,41	4,05
LR-	0,15	0,21	0,51	0,42	0,19

Se destaca el modelo COVIDCENCEC, que superó a los demás en todas las métricas. Con una precisión de 0,88, COVIDCENCEC mostró la mayor proporción de verdaderos positivos sobre el total de resultados positivos. En términos de sensibilidad, COVIDCENCEC y 4C MORTALITY ⁽⁶²⁾ presentaron valores de 0,87 y 0,85, respectivamente, lo que reflejó su eficacia en la detección de verdaderos positivos. COVIDCENCEC también mostró una alta especificidad con un valor de 0,85, lo que reflejó su habilidad para evitar falsos positivos.

En cuanto al valor predictivo positivo, COVIDCENCEC, con un valor de 0,86, mostró una alta probabilidad de acierto cuando predijo un caso positivo. Finalmente, en el valor predictivo negativo, COVIDCENCEC, 4C MORTALITY ⁽⁶²⁾ y NEWS 2 ⁽⁶⁰⁾ presentaron valores de 0,87, respectivamente, lo que indicó una alta probabilidad de acierto cuando predijeron un caso negativo. Estos

resultados evidencian la superioridad del modelo COVIDCENCEC en comparación con los modelos seleccionados.

La tabla 19 muestra las métricas de desempeño para los modelos seleccionados.

Tabla 19. Métricas de rendimiento según modelos seleccionados

Métricas	COVIDCENCEC	NEWS2	PRIEST	Xie	4C Mortality
R ²	0,61 (0,01)	0,69 (0,01)	0,27 (0,01)	0,36 (0,03)	0,42 (0,01)
AUC	0,94 (0,01)	0,91 (0,01)	0,76 (0,02)	0,79 (0,04)	0,85 (0,01)
Dxy	0,62 (0,01)	0,85 (0,02)	0,27 (0,01)	0,57 (0,02)	0,53 (0,01)
Intercepto	1,64 (0,02)	13,5 (0,02)	-0,07 (0,01)	0,45 (0,03)	1,26 (0,01)
Pendiente	0,13 (0,01)	14,1 (1,10)	0,58 (0,03)	2,83 (0,07)	3,62 (0,08)
Brier	0,17 (0,02)	0,15 (0,01)	0,24 (0,01)	0,2 (0,02)	0,19 (0,01)

Se destacó que el modelo COVIDCENCEC mostró un rendimiento superior en comparación con los demás modelos (anexo 23, figura 18). Con una media de 0,61 y una desviación estándar de 0,01, el R² para COVIDCENCEC indicó que este modelo fue capaz de explicar una gran proporción de la variabilidad en los datos. En términos de la métrica AUC, COVIDCENCEC lideró con una media de 0,94 y una desviación estándar de 0,01, lo que reflejó su alta capacidad para distinguir entre las clases vivo y fallecido. En cuanto al índice Dxy, COVIDCENCEC obtuvo una media de 0,62 y una desviación estándar de 0,01, lo que sugirió su eficacia en la discriminación entre resultados positivos y negativos.

Con respecto a la calibración, el intercepto y la pendiente para COVIDCENCEC fueron de 1,64 y 0,13, respectivamente, lo que indicó un menor sesgo sistemático y una mayor precisión en sus predicciones en comparación con otros modelos.

Finalmente, en el Índice de Brier, COVIDCENCEC obtuvo una media de 0,17 y una desviación estándar de 0,02, lo que indicó que sus predicciones probabilísticas fueron precisas. Estos resultados evidenciaron la eficacia del

modelo COVIDCENCEC y su superioridad en la predicción de casos fallecidos de COVID-19.

III.5. Presentación e implementación en la práctica de los modelos

La ejecución del cálculo de la estimación del riesgo de muerte por COVID-19 a partir de un nomograma o aplicación para teléfonos móviles facilita la generalización y simplifica el uso del modelo por parte del personal de salud. A continuación, se dan detalles sobre el diseño de estas herramientas.

III.5.2.1 Nomograma

Se presentó un nomograma para visualizar el modelo de regresión logística desarrollado (anexo 24, figura 19). En el nomograma se proporcionó una representación gráfica intuitiva del pronóstico, donde se estimó la probabilidad de egreso fallecido de los pacientes basados en el modelo COVIDCENCEC.

En el nomograma, cada predictor se representa en una escala propia. Para utilizar el nomograma, se traza una línea vertical desde el valor del predictor hasta la escala de puntos en la parte superior del nomograma. Esto se hace para que en cada predictor exista una puntuación, y luego se suman los puntos correspondientes. Finalmente, se traza una línea desde la suma total de puntos hasta la escala de probabilidad en la parte inferior del nomograma para obtener la probabilidad estimada de muerte. Los coeficientes del modelo COVIDCENCEC se utilizaron para determinar la longitud de las escalas de los predictores en el nomograma.

III.5.2.2 Diseño de COVIDCENCEC-APK

De manera alternativa, este autor programó una aplicación nativa para el sistema operativo Android. Esta puede ser instalada en cualquier teléfono móvil o tableta para calcular las probabilidades de muerte para un paciente. Una vez introducido

los datos, el sistema calculará el riesgo de morir del paciente y lo muestra en una barra de progresos, si este valor sobrepasa el punto de corte establecido, la barra inferior cambia de color verde al rojo para indicar que se trata de un paciente de alto riesgo de muerte (anexo 25, figura 20). Los resultados de las pruebas para validar el desempeño de la aplicación demostraron que la aplicación era capaz de ejecutar con precisión el modelo COVIDCENCEC y proporcionar estimaciones de riesgo confiables para los pacientes, esto se describe en el informe final de validación de software (anexo 26).

III. 6 Consideraciones finales

En este capítulo se abordó el desarrollo de modelos predictivos de muerte para pacientes con COVID-19 de bajo y alto riesgo, la validación de estos modelos en pacientes cubanos y la comparación de los modelos de mayor calidad. Para el manejo de pacientes en la comunidad con bajo riesgo de mortalidad, se desarrolló el modelo ABR. Este modelo proporcionó un conjunto de indicadores útiles para la toma de decisiones, como la planificación de recursos y la gestión de la transición a la gravedad. En cuanto al manejo de pacientes hospitalizados con alto riesgo de mortalidad, se desarrollaron y validaron los modelos AAR y COVIDCENCEC. La escala COVIDCENCEC se puede aplicar a los pacientes mediante un nomograma o una aplicación para móviles.

CAPÍTULO IV. DISCUSIÓN

El empleo de modelos de predicción clínica para la estimación y clasificación del riesgo favorece la calidad en la estratificación. Esta tecnología tiene un impacto positivo en la prevención de la transición a la gravedad y, por consiguiente, en la reducción de las muertes por COVID-19. Como resultado de la falta de modelos de predicción clínica específicos en la población cubana, se desarrolló y validó el modelo ABR para la estratificación del riesgo en la comunidad, así como los modelos AAR y COVIDCENCEC para el triaje de pacientes en hospitales.

Los hallazgos del estudio apoyan la noción de que los modelos propuestos se adaptan mejor a los escenarios de atención médica en Cuba que los modelos de mayor calidad reportados en la literatura. En general, los resultados de este estudio son extrapolables a nuevas emergencias sanitarias y al abordaje de diversos problemas de salud. Por lo tanto, este estudio concluye con una propuesta de una estrategia para la implementación de modelos de predicción clínica en nuevos contextos.

IV.1 Modelos para escenarios de atención médica de la comunidad

En los escenarios de atención médica de la comunidad es donde se deben tomar las primeras decisiones médicas. Por ello, es necesario contar con herramientas que mejoren la calidad de la clasificación en ese contexto.

El modelo ABR se presentó como un instrumento práctico para la estratificación del riesgo de muerte por COVID-19 ⁽³⁾. En especial este modelo es muy recomendable para identificar de manera sencilla y económica a pacientes cuya conducta médica se puede atenuar.

Si bien ABR es un modelo bastante simple, su rendimiento se puede considerar como aceptable. Además, ABR es versátil, ya que tiene la capacidad de clasificar el riesgo de muerte a nivel individual como en conglomerados de individuos.

Este estudio se enfocó en el uso de modelos para el control de la transición a la gravedad en pacientes con confirmación de COVID-19. Estas herramientas se pueden emplear además para tomar decisiones en el tratamiento de la enfermedad grave. Sin embargo, el uso de modelos predictivos en el control de la transmisión como uno de los tres pilares fundamentales para reducir la mortalidad por COVID-19 no se aborda en este estudio. Por tanto, no se propusieron modelos para identificar personas con alto riesgo de contagio por el SARS-CoV-2, tampoco para estimar el riesgo de enfermar por la COVID-19. Se desconoce si el modelo ABR se puede utilizar con este fin o qué modelos de predicción clínica pueden ser válidos para mejorar las estrategias de control de la transmisión en Cuba. Esta laguna en el conocimiento puede sugerir direcciones para futuras investigaciones que aborden el tema de los modelos predictivos para la estimación y clasificación del riesgo de muerte en la población general. Precisamente, el modelo QCOVID ⁽⁵⁷⁾ se diseñó para el triaje en la población general y no en pacientes confirmados de la COVID-19 como ABR ⁽³⁾. El modelo ABR presenta algunos elementos comunes con QCOVID ⁽⁵⁷⁾; sin embargo, esta comparación en el marco de este estudio fue limitada. A diferencia del modelo ABR ⁽³⁾, QCOVID ⁽⁵⁷⁾ tiene como variable de resultado el tiempo hasta la muerte por COVID-19 y el tiempo hasta la admisión hospitalaria con infección confirmada por SARS-CoV-2. Como describe Clift ⁽⁵⁷⁾ en su publicación, el modelo está bien calibrado y tiene una buena capacidad de discriminación. La variable de resultado de ABR y de QCOVID ⁽⁵⁷⁾ son diferentes, por lo que en principio no se deben comparar directamente sus desempeños predictivos ⁽²⁷⁾. El estudio Clift ⁽⁵⁷⁾ y colaboradores que dio a conocer QCOVID incluyó una validación en dos periodos, con un total de 2 173 056 de pacientes en la corte

de validación ⁽²⁷⁾. En este estudio se utilizó datos de la población de Reino Unido que se obtuvo de la fuente de datos QResearch del Servicio Nacional de Salud (NHS) de ese país ⁽⁵⁷⁾.

La razón fundamental por la cual la utilización del modelo QCOVID ⁽⁵⁷⁾ en este estudio es limitada, se debió a que no se disponía de información sobre algunos de sus predictores. Cabe destacar que la construcción de ABR se realizó a partir de la base de datos de vigilancia RNPC, una fuente con fines diferentes al de la construcción de modelos de predicción clínica. Por su parte, en el modelo QCOVID ⁽⁵⁷⁾ se utilizó información específica como el índice de masa corporal, el puntaje de Townsend para medir pobreza ⁽¹⁰⁹⁾, el grupo étnico, así como otras condiciones y tratamientos médicos que son muy contextualizados para la población de Reino Unido. Los predictores edad y comorbilidades empleados en ABR también están incluidos entre los predictores del modelo QCOVID ⁽⁵⁷⁾.

Es conocido que incorporar información más amplia que caracteriza a la salud humana, además de la biología humana, como predictores sociales, del entorno ambiental, los determinantes socioeconómicos y de la organización de servicios médicos pueden mejorar la precisión de los modelos predictivos ⁽²⁹⁾.

Es importante reconocer las limitaciones de este estudio, en relación con la evaluación de predictores que abarquen una gama más amplia de determinantes de la salud. Se justifica una investigación adicional para explorar si la incorporación de predictores relacionados con el proceso de determinación social de la COVID-19 en Cuba mejoran el rendimiento del modelo ABR o qué modelos se podrían proponer para este fin.

Sin embargo, es importante destacar que ABR y QCOVID ⁽⁵⁷⁾ tienen propósitos diferentes. El modelo QCOVID ⁽⁵⁷⁾ se propuso para ser utilizado como apoyo en

la selección de individuos para ensayos clínicos, priorización para la vacunación, y discusiones entre pacientes y médicos sobre la mitigación del riesgo laboral o de salud ⁽²⁷⁾. En cambio, ABR tuvo como objetivo clasificar a los pacientes para establecer el escenario de atención médica más adecuado según su riesgo de transición a la gravedad y tomar conducta de la manera más simple y oportuna posible.

Si bien el modelo ABR está compuesto por solo dos predictores como la edad y las comorbilidades, esta sencillez también se puede considerar una fortaleza ⁽³⁾. Aunque la literatura describe un grupo considerable de variables con alto valor predictivo para la transición a la gravedad y en particular para el riesgo de muerte, algunos de estos predictores no suelen estar disponibles en escenarios comunitarios o en hospitales con limitaciones de recursos. Esto puede dificultar la obtención de datos para el diseño de los modelos, así como importantes limitaciones para su posterior implementación en escenarios de atención médica reales ⁽²⁹⁾.

No siempre los modelos más complejos, que incluyen más predictores y de alta tecnología son más efectivos. En una revisión sistemática de modelos predictivos para la COVID-19 que realizó Gupta y colaboradores se expresa que ⁽¹¹⁰⁾: *“La saturación de oxígeno y la edad del paciente son fuertes predictores de deterioro y mortalidad en adultos hospitalizados con COVID-19. Ninguno de los 22 modelos pronósticos evaluados en la revisión agregan mayor valor de predicción que esos predictores univariados.”*

Este autor opina que, más que la complejidad del modelo o la disponibilidad de información de predictores de alta tecnología, que siempre serán limitados en el contexto de una crisis sanitaria, las claves para un buen triaje radican en la

implementación de un sistema de clasificación de calidad, la adherencia a protocolos de actuación nacional con patrones de tratamientos de reconocida eficacia y seguridad, así como un marco ético adecuado para guiar la toma de decisiones.

IV.3 Modelos para escenarios de atención médica en hospitales

Los modelos AAR y COVIDCENCEC se pueden aplicar en escenarios hospitalarios para tomar decisiones de triaje, reevaluar a los pacientes hospitalizados y enfocar las medidas terapéuticas en aquellos identificados como de alto riesgo ⁽¹⁷⁾.

En ambos modelos se utilizan datos que son el resultado de la aplicación del método clínico a pacientes con COVID-19. En este sentido, utiliza la edad y los antecedentes de comorbilidades. Además, se incluyen otras informaciones que se pueden obtener del interrogatorio del paciente como la sintomatología (representada por el predictor IRPS). Asimismo, se recoge información del examen físico como la FR y la PAM. Por último, se usan datos de exámenes complementarios, en particular el INL y el PLT.

Los modelos propuestos para escenarios de atención médica hospitalaria demostraron su buena discriminación y calibración. La escala COVIDCENCEC se comparó con los modelos de mayor calidad reportados en la literatura ^(25,27), pero su rendimiento en el contexto de los escenarios de atención médica en Cuba fue superior. A continuación, se describen más detalles de la comparación de COVIDCENCEC con los modelos de mayor calidad.

El modelo NEWS 2 ⁽⁶⁰⁾ se recomendó y utilizó en el Reino Unido para la estratificación del riesgo de los pacientes con COVID-19. NEWS 2 ⁽⁶⁰⁾ recibió el respaldo formal del Servicio Nacional de Salud ⁽¹¹¹⁾ para convertirse en el sistema de alerta temprana para identificar a pacientes con cuadros clínicos agudos, en

hospitales de ese país. En la validación externa realizada en este estudio, el modelo NEWS 2 ⁽⁶⁰⁾ demostró una alta discriminación con un AUC de 0,92, superior a la reportada en el estudio original (AUC de 0,74) ^(27,62) y en una validación externa posterior realizada por Jie Ma y colaboradores, se describió un valor de AUC superior al descrito en su presentación original (AUC de 0,82) ⁽¹¹²⁾. Sin embargo, en este estudio se observó un posible sobreajuste del modelo. Se identificó un intercepto elevado con un valor de 13,5 en la recta de calibración. Además, una pendiente de 14,1 que sugiere que las probabilidades predichas por el modelo NEWS 2 ⁽⁶⁰⁾ pueden estar subestimadas o sobreestimadas en los pacientes cubanos.

El modelo 4C MORTALITY ⁽⁶²⁾ se basó en un estudio que usó datos de 260 hospitales del Reino Unido, además se incluyó una representación bastante amplia de pacientes, en concreto 35 463 adultos hospitalizados por COVID-19. Aunque el modelo se desarrolló en una cohorte de pacientes del Reino Unido, su uso se ha extendido a otros contextos. Por ejemplo, se utilizó para predecir el riesgo de mortalidad en pacientes con COVID-19 en los Estados Unidos ⁽⁵⁹⁾. El modelo se evaluó en un total de 1276 sujetos admitidos en el King's College Hospital National Health Service en el Reino Unido, además en 6237 pacientes de cinco hospitales del NHS Trusts, un hospital en Noruega y dos hospitales en Wuhan, China. Este modelo mostró una alta discriminación para la mortalidad con AUC de 0,79 en la cohorte de derivación y 0,77 en la cohorte de validación. En este estudio, el valor del AUC fue de 0,85, se describió una calibración con valores del intercepto en la recta de 1,26 y una pendiente de 3,62, lo que es un indicador de sobreajuste.

El modelo PRIEST ⁽⁵⁸⁾ se basó en datos de 22 445 personas que asistieron a 70 departamentos de emergencia en el Reino Unido entre marzo y mayo de 2020. Su aplicación se ha extendido a otros contextos fuera de ese país como por ejemplo Grecia ⁽¹¹³⁾. En este modelo se reporta un valor del AUC de 0,80, prácticamente el mismo valor obtenido en este estudio (AUC = 0,85). Sin embargo, aunque el modelo muestra una buena discriminación con los datos de pacientes cubanos (como se evidencia por un AUC alto), los parámetros de calibración sugieren que puede no estimar con precisión las probabilidades de los resultados. Un intercepto de -0,07 indica que el modelo está sobreestimando ligeramente el riesgo. Por otro lado, una pendiente de 0,58 muestra que las probabilidades predichas por el modelo aumentan más lentamente que las observadas a medida que aumenta el riesgo. Esto puede sugerir que el modelo está subestimando las diferencias en el riesgo de muerte entre los individuos.

El modelo de Xie ⁽⁶¹⁾ se desarrolló y validó con datos que se recopilaron de 299 pacientes del Hospital Tongji en Wuhan. La validación externa se realizó en una cohorte retrospectiva de 145 pacientes del Hospital Jinyintan en esta misma localidad. En un estudio realizado por Wickstrom y colaboradores ⁽¹¹⁴⁾, se validaron cuatro modelos de predicción, en el que se incluyó el modelo de Xie ⁽⁶¹⁾. Para ello se usaron datos de pacientes admitidos con síntomas de COVID-19 en el Kings Cross Hospital (KCH) en Londres, Reino Unido, y en el Oslo University Hospital (OUH) en Oslo, Noruega. Según estos autores, el rendimiento de los modelos fue variable. Para la predicción de la mortalidad, el modelo de Xie tuvo una buena discriminación en OUH con un AUC de 0,87 (IC del 95 %: 0,79; -0,95). En este estudio, el valor del AUC fue de 0,8, resultado que es bastante similar al reportado por Wickstron ⁽¹¹⁴⁾. Sin embargo, en relación con la

calibración, el valor del intercepto de 0,45 sugiere que, en promedio, las probabilidades predichas por el modelo son sistemáticamente más bajas que las observadas. Esto puede indicar que el modelo está subestimando el riesgo en los datos de este estudio. Por otro lado, una pendiente de 2,83 indica que las probabilidades predichas por el modelo aumentan más rápidamente que las observadas a medida que se incrementa el riesgo de muerte. Esto puede sugerir que el modelo no está bien calibrado.

En los modelos, NEWS 2 ⁽⁶⁰⁾, 4C Mortality ⁽⁶²⁾, PRIEST ⁽⁵⁸⁾ y Xie ⁽⁶¹⁾ se describe un rendimiento inferior en los datos de pacientes cubanos hospitalizados. Esto puede ser explicado por las diferencias contextuales de los escenarios de atención médica en Cuba en comparación con las poblaciones donde se desarrollaron originalmente los modelos de mayor calidad.

Una implicación de los resultados obtenidos en este estudio es que los modelos ABR, AAR y COVIDCENCEC se adaptan mejor a los escenarios de atención médica en Cuba. Por tanto, se puede considerar que resulta más conveniente utilizar los modelos cubanos propuestos que los modelos de mayor calidad reportados en la literatura. Además, como se ha mencionado, ABR ya se probó en escenarios de atención médica reales durante la pandemia en Cuba.

La figura 21 muestra la potencial utilización de los modelos de predicción clínica en los escenarios de atención médica del SNS durante la pandemia.

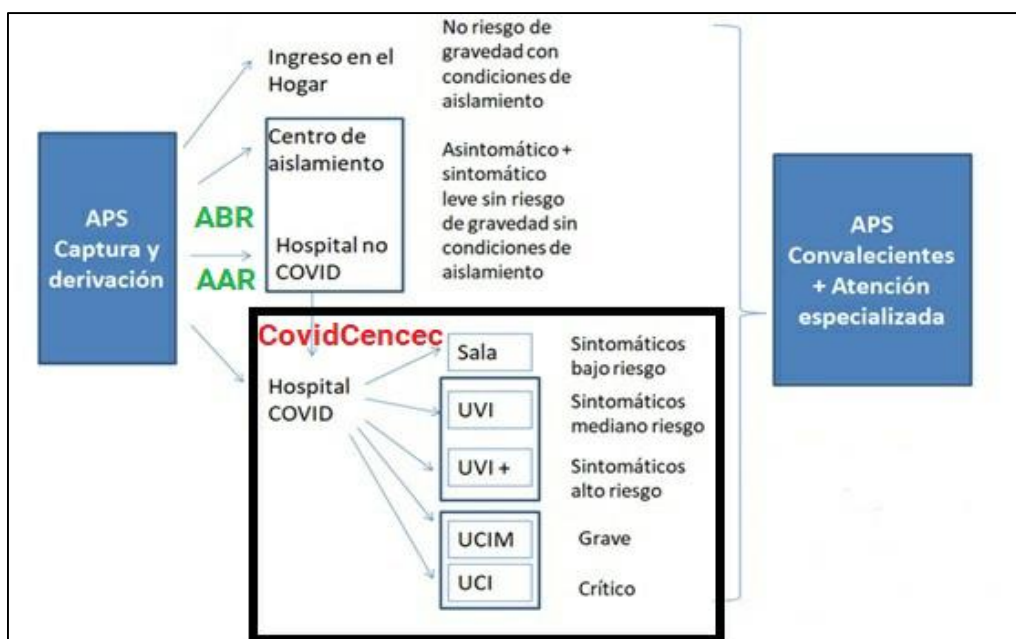


Figura 21. Aplicación de modelos de predicción clínica a escenarios de atención médica.

Si bien la literatura reconoce el valor de los modelos de mayor calidad algunos de ellos no se utilizaron en otros escenarios que no fueran los seleccionados para la ejecución de la investigación.

Sin embargo, en los modelos propuestos se utilizan fortalezas de los modelos de mayor calidad. Por ejemplo, todos los predictores incluidos se pueden considerar respaldados por la literatura porque fueron predictores incluidos y evaluados previamente por los modelos de mayor calidad ^(25,27). Además, todas las relaciones de los predictores con el resultado se sustentan por la plausibilidad biológica.

En un estudio realizado por Cruz y colaboradores, se afirma que la edad avanzada aumenta significativamente el riesgo de muerte, y que los coeficientes de la regresión logística indican un aumento del 1,08 por año ⁽¹¹⁵⁾. Estos valores de OR son muy similares a los que se obtienen en este estudio. Por su parte, Masudah H y colaboradores identificaron que la diabetes mellitus, la hipertensión y las enfermedades cardíacas son factores de riesgo importantes para la muerte

por COVID-19 ⁽¹¹⁶⁾. La FR, un indicador crucial de la función pulmonar, puede estar vinculada a la gravedad de enfermedades respiratorias como la COVID-19, donde un incremento significativo puede señalar un deterioro respiratorio agudo y un mayor riesgo de mortalidad. En un metanálisis realizado por Tredinnick-Rowe y colaboradores, se describe que la FR es un predictor del deterioro de la COVID-19 ⁽¹¹⁷⁾. Asimismo, la TAM, que refleja la función cardiovascular y la perfusión de órganos, variaciones extremas de este parámetro se pueden relacionar con disfunción cardiovascular aguda en caso de complicaciones graves de la COVID-19. En un estudio realizado por Yang Chen y colaboradores ⁽¹¹⁸⁾ en pacientes ingresados en el Hospital Universitario de Stony Brook de Nueva York, se arribó a la conclusión de que la PAM puede ayudar a identificar a los pacientes con COVID-19 con mayor riesgo de muerte. Además, el INL, un biomarcador de la respuesta inmunitaria y la inflamación importante, puede reflejar una respuesta inflamatoria exagerada en pacientes con COVID-19, y se puede asociar a formas más graves de la enfermedad. En un estudio realizado por Callejas se describe que, en los pacientes fallecidos, los valores del INL fueron notablemente más altos en comparación con los de los supervivientes, lo que sugiere su valor predictivo en la transición hacia la gravedad ⁽¹¹⁹⁾. Según una publicación realizada por Bi ⁽¹²⁰⁾ y una revisión sistemática realizada por Malik y colaboradores ⁽¹²¹⁾, los cambios en el PLT se relacionan con la activación de la coagulación, así como la respuesta inflamatoria, lo que puede desencadenar complicaciones trombóticas e incrementar el riesgo de muerte por COVID-19. Es importante subrayar que, para el caso de los modelos basados en árboles de clasificación como ABR y AAR se propusieron una serie de indicadores de estratos que se pueden usar para mejorar la atención médica, disminuir costos

y asignar pacientes a diferentes escenarios de atención médica. Se considera que el desarrollo de la metodología descrita podría servir como base para la toma de decisiones de políticas de salud bajo un enfoque de medicina personalizada. Una ventaja importante es que su costo de implementación es bajo, dado que los estratos pueden ser definidos a priori como parte de un flujograma de toma de decisiones.

Es creciente la inserción de la inteligencia artificial (IA) en la práctica diaria en el ámbito de las ciencias médicas y la salud pública.

Por ejemplo, durante la COVID-19 Chung y colaboradores emplearon la IA para analizar datos de la variabilidad de la FC obtenida de sensores disponibles en relojes y pulseras inteligentes para predecir la aparición de la enfermedad antes de que comenzaran los primeros síntomas ⁽¹²²⁾.

El sistema NYUtron representa un avance significativo en la aplicación de la IA en la atención médica, particularmente en el análisis de notas clínicas no estructuradas para predecir varios resultados clínicos. Al aprovechar millones de registros de pacientes, NYUtron puede pronosticar con precisión riesgos como la mortalidad, los reingresos y la duración de las estadías hospitalarias ⁽¹²³⁾. En un estudio realizado por Viderman y colaboradores ⁽⁵³⁾, se usó esta herramienta para predecir la gravedad de la COVID-19, la disfunción de órganos, la mortalidad, así como realizar estimaciones en la utilización de los recursos hospitalarios.

En Cuba, el uso de modelos de aprendizaje automático como ABR en el contexto de la pandemia es un ejemplo de su utilidad práctica. Los modelos basados en IA se pueden ajustar a diferentes poblaciones y contextos clínicos mediante el aprendizaje continuo. Esto significa que pueden mejorar con el tiempo a medida

que se recopilan más datos y se refinan las técnicas. En opinión de este autor, esta es una fortaleza de este tipo de algoritmos ante la pérdida de rendimiento de los modelos debido a inestabilidad de los predictores por cambios contextuales.

La aplicación de modelos para definir la conducta médica puede tener limitaciones de orden práctico por su complejidad, dado que el médico puede presentar dificultades para aplicar este tipo de herramientas en el momento en que tiene que tratar a un paciente. Por ello, estos modelos se deben concebir de la forma más simple posible ⁽³⁾.

Por otro lado, los profesionales de la salud pueden dudar en usar modelos de predicción clínica si no se comprende cómo funcionan, lo que puede llevar a resistencia para su adopción. Para tener un impacto real, los modelos predictivos se deben integrar sin problemas en la práctica médica. Esta integración requiere abordar cuestiones prácticas como usabilidad, accesibilidad e interoperabilidad con dinámica de trabajo existentes.

En consecuencia, se consideran ABR y AAR como modelos adecuados para el manejo de la COVID-19 por su simplicidad y transparencia. Los algoritmos de árboles de clasificación son considerados como unos de los más intuitivos e interpretables por parte de un público no especializado en estadística o aprendizaje automático. La regresión logística binaria múltiple empleada para el desarrollo de COVIDCENCEC, tiene una larga historia de utilización en el pronóstico de pacientes desde un enfoque estadístico más tradicional en ámbitos como la epidemiología o la bioestadística ⁽⁴⁸⁾.

La comunicación clara y efectiva de los resultados de la investigación es esencial para asegurar que los hallazgos sean accesibles y útiles para los pacientes como

para los profesionales de la salud ⁽²⁹⁾. Precisamente con este fin, se construyó un nomograma y COVIDCENCEC-APK, una aplicación para teléfonos móviles o tabletas que se puede utilizar para estimar y clasificar el riesgo de muerte a través del modelo COVIDCENCEC.

Es importante subrayar que informar a los pacientes sobre el uso de modelos de predicción clínica como herramientas pronóstico es esencial por varias razones éticas y prácticas. Primero, garantiza el consentimiento informado, es importante que los pacientes entiendan cómo se toman las decisiones sobre su salud. La transparencia en el uso de estos modelos fomenta la confianza entre pacientes y el personal médico y, asimismo, la comprensión de posibles sesgos inherentes a los datos ayuda a asegurar decisiones más justas. Además, al involucrar a los pacientes en el proceso, se facilita la toma de decisiones compartida, donde se pueden expresar preocupaciones y contribuir a la mejora continua de la tecnología propuesta.

El empleo de un marco ético adecuado para guiar la toma de decisiones en las que se emplean los modelos de predicción clínica tiene una gran connotación.

En su libro de texto “Explanatory Model Analysis” los autores Biecek y Burzykowski ⁽⁸²⁾ se inspiraron en las Leyes de la Robótica de Isaac Asimov para proponer tres requisitos clave que cualquier modelo predictivo debería cumplir. En el contexto de este estudio, las aplicaciones de estos principios se postulan de la siguiente manera: en primer lugar, las predicciones se deben comprender por las personas implicadas o sea para los pacientes y el personal médico, no como una caja negra. En segundo lugar, los modelos no deben ser engañosos, es decir su funcionamiento interno debe ser transparente y verificable. Por último,

la tercera Ley de la Robótica exige que los creadores del modelo sean responsables de su comportamiento.

En este mismo sentido, Frank Harrel, uno de los autores más importantes en el ámbito de los modelos de predicción clínica expresa que no solo se deben validar los modelos predictivos, sino también sus desarrolladores. Adicionalmente sugiere que un comité independiente reproduzca los resultados de la investigación para mayor confianza y transparencia de los modelos ⁽³⁰⁾.

IV.4 Estrategia de aplicación de modelos de predicción clínica en la gestión de la transición a la gravedad en Cuba

Si bien el objetivo de este trabajo no es proponer una metodología definitiva para la aplicación de modelos en gestión de la transición a la severidad, se recomiendan algunos criterios de cómo podrían ser empleadas estas herramientas en los escenarios de atención médica en el futuro.

La posibilidad de nuevas pandemias, ya sea del SARS-CoV-2 o de otros patógenos emergentes, es un riesgo que requiere atención. La OMS ha identificado múltiples virus que podrían causar futuros brotes ⁽¹²⁴⁾, y la mutación continua del SARS-CoV-2 plantea amenazas adicionales ⁽¹²⁵⁾. Factores como la degradación de ecosistemas aumentan la probabilidad de transmisión de virus zoonóticos ⁽¹²⁶⁾. Por ello, es crucial tener estrategias previamente discutidas para enfrentar emergencias sanitarias. La preparación proactiva permite respuestas más rápidas y efectivas, así como disminuir el impacto en la salud pública y la economía.

Los hallazgos de este estudio pueden contribuir a preparar al sistema de salud frente a la potencial amenaza de una nueva pandemia con metodologías previamente evaluadas y discutidas. Además, el enfoque que se adoptó en este estudio es extrapolable a otros problemas de salud, no solo en el triaje de

enfermedades infecciosas, sino también en el manejo de enfermedades no transmisibles, como el cáncer. Se justifican investigaciones adicionales para explorar estas aplicaciones.

Estas herramientas tienen la particularidad de ser fáciles de implementar y son poco costosas en relación con los insumos médicos. Estos modelos, por ejemplo, los basados en árboles (ABR y AAR) tienen la capacidad de ayudar a planificar estrategias de asignación de recursos de salud en función de los indicadores del estrato. Hasta donde conoce el autor, esta metodología es novedosa y no se ha descrito en otro sitio con anterioridad.

Las herramientas propuestas están diseñadas para apoyar, y no remplazar, al médico asistente, quien es el actor principal en la toma de decisiones. Estos modelos proporcionan información adicional que puede ser utilizada para informar y guiar el proceso, pero la decisión final siempre la tiene el profesional de la salud y, en primer lugar, el paciente.

Es importante enfatizar que muchas de las variables y mecanismos causales subyacentes suelen ser inestables en el tiempo y resultan susceptibles a cambios contextuales. Por ello, se requiere una vigilancia del rendimiento de los modelos de manera sistemática.

La posibilidad de que los modelos propuestos ya no sean útiles podría ser posible, por tanto, se requeriría una actualización o la implementación de nuevos modelos. A continuación, se describe una estrategia que se podría implementar ante una nueva emergencia. En la figura 22 se muestra una infografía que resume una metodología para la implementación de modelos de predicción clínica en el control de la transición a la gravedad en escenarios de atención médica.



Figura 22. Infografía sobre la estrategia para la incorporación de modelos de predicción clínica a la práctica.

Primer paso (identificación de factores de riesgo): El proceso comienza con la identificación de factores de riesgo relevantes para la enfermedad o condición clínica de interés. Para ello se pueden utilizar fuentes como la literatura científica, registros clínicos electrónicos, investigaciones previas y bases de datos de cohortes de pacientes. Además, se deben considerar definiciones estandarizadas, incluir múltiples variables basadas en evidencia.

Es importante contar con datos del interrogatorio, examen físico y exámenes complementarios para identificar predictores candidatos. Para ello es importante garantizar la infraestructura informática, en especial un software que permita obtener un conjunto de datos en cantidad y calidad suficiente para la modelación.

La aplicación XAVIA-SIDEC empleada para este estudio se podría reutilizar ya que contiene un CRD con información clínica suficiente para generar una gran variedad de modelos de predicción clínica a partir de datos de las historias clínicas.

Segundo paso (caracterización del riesgo): Una vez obtenidos los datos, se evalúa y cuantifica el riesgo asociado a cada factor identificado con la variable de resultado. Para ello se pueden utilizar técnicas y procedimientos más tradicionales como la bioestadística u otros enfoques como la aplicación de la IA.

Tercer paso (construcción de los modelos): El tercer paso incluye la identificación de modelos candidatos o el desarrollo y validación de nuevas propuestas de herramientas. Los tres modelos que se han propuesto podrían ser incorporados en esta etapa si se trata de una nueva ola de COVID-19, pues cada uno proporciona perspectivas únicas y valiosas para el trabajo inicial. Sin embargo, dada la naturaleza cambiante de los factores contextuales, la opción más recomendable es el desarrollo de nuevos modelos, en especial si se tratan de nuevos problemas de salud.

Cuarto paso (validación de los modelos): Una vez desarrollados los modelos, las herramientas deben ser validadas a partir de conjuntos de datos independientes o utilizar técnicas robustas como el remuestreo o la validación cruzada.

Quinto paso (implementación): Para integrar los modelos seleccionados se puede utilizar un protocolo que pauté la conducta y donde se discutan los resultados con otros criterios. Por ejemplo, los resultados se pueden integrar de manera racional a las pautas normativas como la del PAN ⁽¹⁷⁾.

Quinto paso (Capacitación): Es importante que se proporcione capacitación a los profesionales de la salud sobre cómo interpretar y aplicar las predicciones del modelo en la toma de decisiones clínicas. El diseño de software como COVIDCENCEC-APK puede facilitar este proceso.

Séptimo paso (evaluación continua): Finalmente, es imprescindible una evaluación continua de la utilidad clínica por si se requiere la realización de ajustes de los modelos en ejecución, la incorporación de nuevos modelos o su supresión.

IV.5 Consideraciones finales

El uso de modelos de predicción clínica desarrollados y validados en la población cubana contribuye a la calidad de la gestión de la transición a la gravedad, lo que impacta directamente en la reducción de las muertes por COVID-19. Además, pone en manos de la salud pública cubana una información valiosa que puede contribuir a preparar al sistema de salud ante futuras amenazas.

CONCLUSIONES

1. Un modelo de predicción clínica basado en la edad y la presencia de comorbilidades tiene un adecuado rendimiento para predecir la muerte de los pacientes confirmados de COVID-19 en los escenarios de atención médica comunitaria.
2. Dos modelos predictivos basados en variables clínicas como la edad, la presencia de comorbilidades, síntomas seleccionados, la frecuencia respiratoria, la presión arterial media, el índice neutrófilo-linfocitos, así como el conteo de plaquetas poseen un alto rendimiento para predecir la muerte de los pacientes ingresados con COVID-19 en los escenarios de atención médica hospitalarios.
3. El nomograma y la aplicación para móviles o tabletas pueden predecir el riesgo de muerte en pacientes hospitalizados.

RECOMENDACIONES

- Dado que una vez finalizada la pandemia, la COVID-19 se quedó como una endemia, se puede usar estos modelos para estratificar el riesgo y perfeccionar la conducta a tomar con los casos diagnosticados.
- En caso de una nueva pandemia, se recomienda implementar los modelos propuestos para gestionar la transición a la severidad y optimizar los recursos sanitarios en el contexto de los servicios de atención médica.
- Se propone la proyección de esta metodología para la estratificación del riesgo de gravedad de otras enfermedades, tanto transmisibles como crónicas no transmisibles. Dado que es generalizable, puede extenderse a otros problemas de salud. Este enfoque puede ser de utilidad en la organización de los servicios de atención médica, la investigación y la docencia.
- Actualizar el Protocolo de actuación nacional para la COVID-19 en Cuba con los resultados de los modelos propuestos para los escenarios de atención médica considerados.
- Capacitar al personal de la APS y de los hospitales en el uso de estos modelos y sus modos de empleo como el nomograma y la aplicación para móviles.
- Continuar la investigación para estimar el costo social y económico ahorrado por aumento de la eficiencia en los servicios de salud.
- Divulgar los resultados e informar a directivos del Sistema Nacional de Salud para promover su utilización y generalización en la atención médica del país.

REFERENCIAS BIBLIOGRÁFICAS

1. Msemburi W, Karlinsky A, Knutson V, Aleshin-Guendel S, Chatterji S, Wakefield J. The WHO estimates of excess mortality associated with the COVID-19 pandemic. *Nature* [Internet]. 5 de enero de 2023 [citado 1 de abril de 2024];613(7942):130-7. Disponible en: <https://doi.org/10.1038/s41586-022-05522-2>
2. Miranda-Portal JA. Respuesta de Cuba para enfrentar la COVID-19 y sus retos. *Revista Cubana de Medicina Tropical* [Internet]. 14 de diciembre de 2022 [citado 25 de marzo de 2022];74(3). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0375-07602022000300022&lng=es
3. Monzon-Perez M, Sanchez-Valdés L, Lage-Dávila A. Identificación de pacientes de muy bajo riesgo de transición a la severidad en confirmados de Covid-19. Cuba. Marzo 2020-Mayo 2021. *RHCM* [Internet]. 2 de febrero de 2022 [citado 3 de marzo de 2022];22(1). Disponible en: <https://revhabanera.sld.cu/index.php/rhab/article/view/4943/3276>
4. Farley MM. 2009 H1N1 influenza: a twenty-first century pandemic with roots in the early twentieth century. *Am J Med Sci* [Internet]. septiembre de 2010 [citado 6 de diciembre de 2024];340(3):202-8. Disponible en: <https://doi.org/10.1097/maj.0b013e3181e937b0>
5. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure: On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine (see contributors to the project in the appendix). *Intensive Care Med* [Internet]. julio de 1996 [citado 6 de diciembre de 2024];22(7):707-10. Disponible en: <https://doi.org/10.1007/BF01709751>
6. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* [Internet]. octubre de 1985 [citado 7 de diciembre de 2024];13(10):818-29. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/3928249/>
7. Enfield K, Miller R, Rice T, Thompson BT, Truitt J. Limited Utility of SOFA and APACHE II Prediction Models for ICU Triage in Pandemic Influenza. *Chest* [Internet]. octubre de 2011 [citado 30 de octubre de 2024];140(4):913A. Disponible en: <https://doi.org/10.1378/chest.1118087>
8. Zhang Y, Zhang H, Zhang W. SARS-CoV-2 variants, immune escape, and countermeasures. *Front Med* [Internet]. abril de 2022 [citado 2 de junio de 2024];16(2):196-207. Disponible en: <https://doi.org/10.1007/s11684-021-0906-x>
9. Pagani I, Ghezzi S, Alberti S, Poli G, Vicenzi E. Origin and evolution of SARS-CoV-2. *Eur Phys J Plus* [Internet]. 16 de febrero de 2023 [citado 30 de marzo

- de 2024];138(2):157. Disponible en: <https://doi.org/10.1140/epjp/s13360-023-03719-6>
10. Wagh PW, Mathankar HS, Kumbhalkar S, Rangari A. The Novel Virus: Covid-19. ECS Trans [Internet]. 24 de abril de 2022 [citado 18 de marzo de 2024];107(1):17371-80. Disponible en: <https://doi.org/10.1149/10701.17371ecst>
 11. Bastian H. Cuba's Pandemic Crisis. Current History [Internet]. 1 de febrero de 2023 [citado 2 de junio de 2024];122(841):56-62. Disponible en: <https://doi.org/10.1525/curh.2023.122.841.56>
 12. Mr. Kachare Vishal, Prof. Waghmare S. U., Poonam B. Kodage. Covid-19 A Comprehensive Review of Signs, Symptoms, Diagnosis, and Treatment Strategies. IJAR SCT [Internet]. 8 de mayo de 2024 [citado 8 de enero de 2025];51-68. Disponible en: <https://www.doi.org/10.48175/ijarsct-18110>
 13. Cherkashin M, Berezin N, Berezina N, Nikolaev A, Kuplevatskaya D, Kuplevatsky V, et al. Primary pre-hospital triage of patients with COVID-19. MES [Internet]. mayo de 2022 [citado 7 de diciembre de 2024];(2022(2)). Disponible en: <https://doi.org/10.47183/mes.2022.015>
 14. Álvarez Pérez A. Manejo de la epidemia de COVID-19 en Cuba con enfoques sindémico y de determinación social de la salud. 9 de octubre de 2023 [citado 31 de octubre de 2024];49(2). Disponible en: <https://revsaludpublica.sld.cu/index.php/spu/article/view/3947>
 15. Álvarez Pérez A. La determinación social de la salud [Internet]. 1.^a ed. Vol. 1. La Habana: ECIMED; 2023 [citado 7 de diciembre de 2024]. Disponible en: <http://www.bvscuba.sld.cu/libro/la-determinacion-social-de-la-salud-una-antologia/>
 16. Bouktif S, K N. Exploring Epidemiological Determinants of COVID-19: Clustering and Correlation of Multifaceted Factors. En: 2023 IEEE 11th Joint International Information Technology and Artificial Intelligence Conference (ITAIC) [Internet]. Chongqing, China: IEEE; 2023 [citado 7 de diciembre de 2024]. p. 1962-6. Disponible en: <https://doi.org/10.1109/itaic58329.2023.10409110>
 17. Minsap. Protocolo de actuación nacional para la Covid-19, (Versión 1.6) [Internet]. ECIMED; 2021 [citado 19 de marzo de 2024]. Disponible en: <https://instituciones.sld.cu/pdvedado/files/2021/02/PROTOCOLO-V.6-FEB.pdf>
 18. Minsap. Protocolo de Actuación Nacional para la COVID-19 (versión 1.5) [Internet]. ECIMED; 2020 [citado 6 de junio de 2023]. Disponible en: <https://covid19cubadata.github.io/protocolos/protocolo-version-5.pdf>
 19. Sanchez YM, Suárez IM, García AA, Sánchez IA. Estructuración del Protocolo Cubano de Actuación para la Atención de Casos COVID-19. Revista Cubana de Salud Pública [Internet]. 2021 [citado 7 de diciembre de 2024].

- 2024];47(3). Disponible en: <https://revsaludpublica.sld.cu/index.php/spu/article/view/2922>
20. WHO. Pulse survey on continuity of essential health services during the COVID-19 pandemic: interim report, 27 August 2020 [Internet]. 2023 [citado 1 de junio de 2022]. Disponible en: https://www.who.int/publications-detail-redirect/WHO-2019-nCoV-EHS_continuity-survey-2020.1
21. Herrera C, Lage Dávila A, Betancourt Cervantes J, Barreto Fiu E, Sánchez Valdés L, Crombet Ramos T. Nomograma de predicción para la estratificación del riesgo en pacientes con COVID-19. EJHR [Internet]. 6 de julio de 2021 [citado 10 de marzo de 2023];1-19. Disponible en: <https://revistas.uautonoma.cl/index.php/ejhr/article/view/1592/1144>
22. García Álvarez PJ, Morejón Ramos L, Grasso Leyva F. Aplicación de un modelo cubano predictivo de mortalidad en pacientes graves por covid-19 en Lombardía, Italia. Revista Médica Electrónica [Internet]. 2021 [citado 7 de diciembre de 2024];43(2):3047-60. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1684-18242021000203047
23. Buttia C, Llanaj E, Raeisi-Dehkordi H, Kastrati L, Amiri M, Meçani R, et al. Prognostic models in COVID-19 infection that predict severity: a systematic review. Eur J Epidemiol [Internet]. abril de 2023 [citado 31 de marzo de 2024];38(4):355-72. Disponible en: <https://doi.org/10.1007/s10654-023-00973-x>
24. Digital Science. Dimensions. 2022 [citado 16 de diciembre de 2022]. Dimensions. Disponible en: <https://app.dimensions.ai/>
25. Monzón-Pérez M, Sánchez-Valdés L, Lage-Dávila A. Estado en la investigación sobre modelos de predicción de la severidad en confirmados de la Covid-19. RHCM [Internet]. 22 de marzo de 2023 [citado 22 de marzo de 2023];22(1). Disponible en: <https://revhabanera.sld.cu/index.php/rhab/article/view/5257/3273>
26. Van Smeden M, Moons C, Van Calster B. COVID PRECISE. [citado 29 de marzo de 2023]. COVID Precise. Disponible en: <https://www.covprecise.org/>
27. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of Covid-19: systematic review and critical appraisal [Internet]. Vol. 369, BMJ. 2020 [citado 22 de marzo de 2023]. p. m1328. Disponible en: <https://doi.org/10.1136/bmj.m1328>
28. WHO. Living guidance for clinical management of COVID-19: living guidance, 23 November 2021 [Internet]. World Health Organization; 2021 [citado 21 de abril de 2022]. Disponible en: <https://apps.who.int/iris/bitstream/handle/10665/349321/WHO-2019-nCoV-clinical-2021.2-eng.pdf>
29. Steyerberg E. Clinical Prediction Models. USA: Springer; 2009. (Statistics for Biology and Health).

30. Harrell Jr FE. Regression Modeling Strategies. New York: Springer; 2015.
31. Tabish SA, Nabil S. An Age of Emerging and Reemerging Pandemic Threats. Health [Internet]. 2022 [citado 28 de enero de 2024];14(10):1021-37. Disponible en: <https://www.scirp.org/journal/doi.aspx?doi=10.4236/health.2022.1410073>
32. Frieden TR, Buissonnière M, McClelland A. The world must prepare now for the next pandemic. BMJ Glob Health [Internet]. marzo de 2021 [citado 31 de marzo de 2024];6(3):e005184. Disponible en: <https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2021-005184>
33. Amaylid Arteaga García, Maicel Monzón. Proyecto Impacto de la COVID-19 en pacientes adultos en Cuba. 2020-2023 (ICPA). 2024 [citado 1 de diciembre de 2024]; Disponible en: <https://rgdoi.net/10.13140/RG.2.2.29975.18087>
34. Escandell-Sosa CVE. Lineamientos de la política económica y social del partido y la revolución aprobados en el VI Congreso del Partido Comunista de Cuba: una visión desde la economía política. Disponible en: <https://anuarioeco.uo.edu.cu/index.php/aeco/article/view/642/615>
35. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. European Heart Journal [Internet]. 1 de agosto de 2014 [citado 9 de mayo de 2021];35(29):1925-31. Disponible en: <https://academic.oup.com/eurheartj/article-lookup/doi/10.1093/eurheartj/ehu207>
36. Scientific Reports. Scientific reports [Internet]. 2023 [citado 16 de junio de 2023]. Disponible en: <https://www.nature.com/srep/>
37. Nature. Nature [Internet]. 2023. Disponible en: <https://www.nature.com/>
38. PLOS ONE. PLOS ONE [Internet]. 2023 [citado 16 de junio de 2023]. Disponible en: <https://plos.org/>
39. British Medical Journal. British medical journal [Internet]. 2023. Disponible en: <https://www.bmj.com/>
40. Julius Center for Health Sciences and Primary Care. Julius center for health sciences and primary care [Internet]. 2023 [citado 16 de junio de 2023]. Disponible en: <https://juliuscentrum.umcutrecht.nl/en/>
41. Zhao Z, Chen A, Hou W, Graham JM, Li H, Richman PS, et al. Prediction model and risk scores of ICU admission and mortality in COVID-19. Adrish M, editor. PLoS ONE [Internet]. 30 de julio de 2020 [citado 6 de enero de 2023];15(7):e0236618. Disponible en: <https://dx.plos.org/10.1371/journal.pone.0236618>
42. He J, Liu X, Lu X, Zhong M, Jia C, Lucero-Prisno DE, et al. The impact of COVID-19 on global health journals: an analysis of impact factor and publication trends. BMJ Glob Health [Internet]. abril de 2023 [citado 31 de

- marzo de 2024];8(4):e011514. Disponible en: <https://gh.bmj.com/lookup/doi/10.1136/bmigh-2022-011514>
43. BMC. Critical Care. 2023 [citado 16 de junio de 2023]. Critical care. Disponible en: <https://ccforum.biomedcentral.com/>
 44. Banoei MM, Dinparastisaleh R, Zadeh AV, Mirsaeidi M. Machine-learning-based COVID-19 mortality prediction model and identification of patients at low and high risk of dying. Critical Care [Internet]. 8 de septiembre de 2021 [citado 16 de abril de 2023];25(1):328. Disponible en: <https://doi.org/10.1186/s13054-021-03749-5>
 45. Panda NR, Mahanta KL, Pati JK, Varanasi PR, Bhuyan R. Comparison of Some Prediction Models and their Relevance in the Clinical Research. ijsmr [Internet]. 8 de marzo de 2023 [citado 31 de marzo de 2024];12:12-9. Disponible en: <https://lifescienceglobal.com/pms/index.php/ijsmr/article/view/9043>
 46. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: A conditional inference framework. Journal of Computational and Graphical statistics [Internet]. 2006;15(3):651-74. Disponible en: <https://doi.org/10.1198/106186006X133933>
 47. Hothorn T, Hornik K, Zeileis A. ctree: Conditional inference trees. The comprehensive R archive network [Internet]. 2015;8. Disponible en: <http://cran.salud.gob.sv/web/packages/partykit/vignettes/ctree.pdf>
 48. Agresti, A. Categorical Data Analysis. 2da ed. Florida: Wiley; 2002.
 49. Panesar A. Machine Learning and AI for Healthcare [Internet]. 1.^a ed. UK: apress; 2018 [citado 12 de julio de 2024]. Disponible en: <https://doi.org/10.1007/978-1-4842-3799-1>
 50. Munde A. The Machine Learning Pipeline. En: Deep Learning Concepts in Operations Research [Internet]. 1.^a ed. New York: Auerbach Publications; 2024 [citado 26 de noviembre de 2024]. p. 226-43. Disponible en: <https://www.taylorfrancis.com/books/9781003433309/chapters/10.1201/9781003433309-18>
 51. Consoli S. Data Science for Healthcare Methodologies and Applications [Internet]. 1era ed. Vol. 1. Switzerland: Springer; 2019 [citado 12 de julio de 2024]. Disponible en: <https://doi.org/10.1007/978-3-030-05249-2>
 52. Ventresca M, Schünemann HJ, Macbeth F, Clarke M, Thabane L, Griffiths G, et al. Obtaining and managing data sets for individual participant data meta-analysis: scoping review and practical guide. BMC Med Res Methodol [Internet]. diciembre de 2020 [citado 5 de noviembre de 2023];20(1):113. Disponible en: <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-020-00964-6>

53. Viderman D, Kotov A, Popov M, Abdildin Y. Machine and deep learning methods for clinical outcome prediction based on physiological data of COVID-19 patients: a scoping review. *International Journal of Medical Informatics* [Internet]. febrero de 2024 [citado 3 de diciembre de 2024];182:105308. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/S138650562300326X>
54. Akhoon N. Precision medicine: A new paradigm in therapeutics. *Int J Prev Med* [Internet]. 2021 [citado 1 de diciembre de 2024];12(1):12. Disponible en: https://journals.lww.com/10.4103/ijpvm.IJPVM_375_19
55. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection Bias and Information Bias in Clinical Research. *Nephron Clin Pract* [Internet]. 21 de abril de 2010 [citado 19 de enero de 2025];115(2):c94-9. Disponible en: <https://www.doi.org/10.1159/000312871>
56. Raymond J, Valvassori L, Darsaut TE. Understanding the role of randomization in clinical research and practice. *Neurochirurgie* [Internet]. noviembre de 2023 [citado 19 de enero de 2025];69(6):101492. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/S0028377023000905>
57. Clift AK, Coupland CAC, Keogh RH, Diaz-Ordaz K, Williamson E, Harrison EM, et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. *BMJ* [Internet]. 20 de octubre de 2020 [citado 13 de junio de 2022];m3731. Disponible en: <https://doi.org/10.1136/bmj.m3731>
58. Goodacre S, Thomas B, Sutton L, Burnsall M, Lee E, Bradburn M, et al. Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: The PRIEST observational cohort study. Ashkenazi I, editor. *PLoS ONE* [Internet]. 22 de enero de 2021 [citado 13 de junio de 2022];16(1):e0245840. Disponible en: <https://dx.plos.org/10.1371/journal.pone.0245840>
59. Gordon AJ, Govindarajan P, Bennett CL, Matheson L, Kohn MA, Camargo C, et al. External validation of the 4C Mortality Score for hospitalised patients with COVID-19 in the RECOVER network. *BMJ open* [Internet]. 2022 [citado 26 de febrero de 2024];12(4):e054700. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/35450898/>
60. Carr E, Bendayan R, Bean D, Stammers M, Wang W, Zhang H, et al. Evaluation and improvement of the National Early Warning Score (NEWS2) for COVID-19: a multi-hospital study. *BMC Med* [Internet]. diciembre de 2021 [citado 13 de junio de 2022];19(1):23. Disponible en: <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-020-01893-3>
61. Xie J, Hungerford D, Chen H, Abrams ST, Li S, Wang G, et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19 [Internet]. *Infectious Diseases (except*

- HIV/AIDS); 2020 mar [citado 13 de junio de 2022]. Disponible en: <http://medrxiv.org/lookup/doi/10.1101/2020.03.28.20045997>
62. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ* [Internet]. 9 de septiembre de 2020 [citado 31 de octubre de 2021];m3339. Disponible en: <https://www.bmj.com/lookup/doi/10.1136/bmj.m3339>
 63. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* [Internet]. 2010;33(1):1-22. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/20808728>
 64. Bojanowski M, Edwards R. Alluvial: Alluvial Diagrams [Internet]. 2016 [citado 9 de diciembre de 2024]. p. 0.1-2. Disponible en: <https://CRAN.R-project.org/package=alluvial>
 65. Dessie ZG, Zewotir T. Mortality-related risk factors of COVID-19: a systematic review and meta-analysis of 42 studies and 423,117 patients. *BMC Infect Dis* [Internet]. 21 de agosto de 2021 [citado 6 de noviembre de 2023];21(1):855. Disponible en: <https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-021-06536-3>
 66. Romero Starke K, Reissig D, Petereit-Haack G, Schmauder S, Nienhaus A, Seidler A. The isolated effect of age on the risk of COVID-19 severe outcomes: a systematic review with meta-analysis. *BMJ Glob Health* [Internet]. diciembre de 2021 [citado 6 de noviembre de 2023];6(12):e006434. Disponible en: <https://gh.bmj.com/lookup/doi/10.1136/bmigh-2021-006434>
 67. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases* [Internet]. enero de 1987 [citado 14 de octubre de 2021];40(5):373-83. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/0021968187901718>
 68. Charlson ME, Charlson RE, Peterson JC, Marinopoulos SS, Briggs WM, Hollenberg JP. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *Journal of clinical epidemiology*. 2008;61(12):1234-40. Disponible en: <https://10.1016/j.jclinepi.2008.01.006>
 69. Chenchula S, Vidyasagar K, Pathan S, Sharma S, Chavan MR, Bhagavathula AS, et al. Global prevalence and effect of comorbidities and smoking status on severity and mortality of COVID-19 in association with age and gender: a systematic review, meta-analysis and meta-regression. *Sci Rep* [Internet]. 19 de abril de 2023 [citado 31 de mayo de 2024];13(1):6415. Disponible en: <https://www.nature.com/articles/s41598-023-33314-9>
 70. Al-Qudimat AR, Ameen A, Sabir DM, Alkharraz H, Elaarag M, Althani A, et al. The Association of Hypertension with Increased Mortality Rate during the COVID-19 Pandemic: An Update with Meta-analysis. *J Epidemiol Glob Health*

- [Internet]. 15 de junio de 2023 [citado 31 de mayo de 2024];13(3):495-503. Disponible en: <https://link.springer.com/10.1007/s44197-023-00130-3>
71. Charles-Messance H, Mitchelson KAJ, De Marco Castro E, Sheedy FJ, Roche HM. Regulating metabolic inflammation by nutritional modulation. *Journal of Allergy and Clinical Immunology* [Internet]. octubre de 2020 [citado 9 de diciembre de 2024];146(4):706-20. Disponible en: <https://doi.org/10.1016/j.jaci.2020.08.013>
 72. Mordovin VF, Zyubanova IV, Manukyan MA, Dorzhieva IK, Vtorushina AA, Khunkhinova SA, et al. The role of immune-inflammatory mechanisms in the pathogenesis of hypertension. *SJCEM* [Internet]. 8 de abril de 2023 [citado 9 de diciembre de 2024];38(1):21-7. Disponible en: <https://www.sibjcem.ru/jour/article/view/1573>
 73. Sandhu APS, Tanvir ., Singh K, Singh S, Antaal H, Luthra S, et al. Decoding Cancer Risk: Understanding Gene-Environment Interactions in Cancer Development. *Cureus* [Internet]. 19 de julio de 2024 [citado 9 de diciembre de 2024]; Disponible en: <https://doi.org/10.7759/cureus.64936>
 74. Veiga VC, Cavalcanti AB. Age, host response, and mortality in COVID-19. *Eur Respir J* [Internet]. julio de 2023 [citado 21 de marzo de 2024];62(1):2300796. Disponible en: <http://erj.ersjournals.com/lookup/doi/10.1183/13993003.00796-2023>
 75. Huntley J, Ostfeld AM, Taylor JO, Wallace RB, Blazer D, Berkman LF, et al. Established populations for epidemiologic studies of the elderly: Study design and methodology. *Aging Clin Exp Res* [Internet]. febrero de 1993 [citado 1 de abril de 2024];5(1):27-37. Disponible en: <http://link.springer.com/10.1007/BF03324123>
 76. Novelli L, Raimondi F, Ghirardi A, Galimberti C, Biza R, Trapasso R, et al. Frequency, characteristics, and outcome of patients with COVID-19 pneumonia and «silent hypoxemia» at admission: a severity-matched analysis. *Panminerva Med* [Internet]. diciembre de 2022 [citado 22 de marzo de 2024];64(4). Disponible en: <https://www.minervamedica.it/index2.php?show=R41Y2022N04A0442>
 77. Gogia B, Pujara D, Thottempudi N, Ghanayem T, Ajam Y, Singh A, et al. Impact of Acute Confusional State in Patients With COVID-19 and a Predictive Score. *Cureus* [Internet]. 28 de septiembre de 2021 [citado 20 de marzo de 2024]; Disponible en: <https://www.cureus.com/articles/72011-impact-of-acute-confusional-state-in-patients-with-covid-19-and-a-predictive-score>
 78. Chatterjee NA, Jensen PN, Harris AW, Nguyen DD, Huang HD, Cheng RK, et al. Admission respiratory status predicts mortality in COVID-19. *Influenza Resp Viruses* [Internet]. septiembre de 2021 [citado 21 de marzo de 2024];15(5):569-72. Disponible en: <https://onlinelibrary.wiley.com/doi/10.1111/irv.12869>

79. Battaglini D, Lopes-Pacheco M, Castro-Faria-Neto HC, Pelosi P, Rocco PRM. Laboratory Biomarkers for Diagnosis and Prognosis in COVID-19. *Front Immunol* [Internet]. 27 de abril de 2022 [citado 6 de noviembre de 2023];13:857573. Disponible en: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.857573/full>
80. Dymicka-Piekarska V, Dorf J, Milewska A, Łukaszyk M, Kosidło JW, Kamińska J, et al. Neutrophil/Lymphocyte Ratio (NLR) and Lymphocyte/Monocyte Ratio (LMR) – Risk of Death Inflammatory Biomarkers in Patients with COVID-19. *JIR* [Internet]. mayo de 2023 [citado 22 de marzo de 2024];Volume 16:2209-22. Disponible en: <https://www.dovepress.com/neutrophillymphocyte-ratio-nlr-and-lymphocytemonocyte-ratio-lmr--risk--peer-reviewed-fulltext-article-JIR>
81. Rajanna A, Ramanna B, Bellary V. Correlation of platelet counts with Coronavirus Disease 2019 outcome in the second wave at a tertiary care center. *APIK J Int Med* [Internet]. 2022 [citado 22 de marzo de 2024];10(4):250. Disponible en: https://journals.lww.com/10.4103/ajim.ajim_106_21
82. Biecek P, Burzykowski T. Explanatory Model Analysis. 1era ed. New York: CRC Press; 2021. (DATA SCIENCE SERIES).
83. Arora A, Basu N. Machine Learning in Modern Healthcare. *IJAMST* [Internet]. 30 de junio de 2023 [citado 1 de abril de 2024];3(4):12-8. Disponible en: <https://www.ijamst.latticescipub.com/portfolio-item/D3037063423/>
84. Silva Aycaguer LC. Una ceremonia estadística para identificar factores de riesgo. *Salud colectiva* [Internet]. 2005 [citado 18 de febrero de 2013];1(3):309-22. Disponible en: http://www.scielo.org.ar/scielo.php?pid=S1851-82652005000300004&script=sci_arttext
85. Silva Aycaguer LCS, Utra IMB. Selección algorítmica de modelos en las aplicaciones biomédicas de la regresión múltiple. *Med Clin (Barc)* [Internet]. 2001 [citado 18 de febrero de 2013];116:741-5. Disponible en: <http://www.elsevier.es/sites/default/files/elsevier/pdf/2/2v116n19a13014323pdf001.pdf>
86. Silva Aycaguer PS. Cultura estadística e investigación científica en el campo de la salud: una mirada crítica. Madrid: Diaz de Santos; 1998.
87. Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. *Biom J* [Internet]. mayo de 2018 [citado 9 de mayo de 2021];60(3):431-49. Disponible en: <http://doi.wiley.com/10.1002/bimj.201700067>
88. Keusińska E. Generalized Walker-Duncan Procedure for Obtaining Maximum Likelihood Estimates of Parameters of Logistic Discriminant Function. *Biometrical J* [Internet]. enero de 1988 [citado 31 de octubre de 2024];30(3):259-74. Disponible en: <https://onlinelibrary.wiley.com/doi/10.1002/bimj.4710300302>

89. Leung A. Maximum likelihood estimation. En: Actuarial Principles [Internet]. Elsevier; 2022 [citado 31 de octubre de 2024]. p. 117-22. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/B9780323901727000257>
90. Allison PD. Measures of fit for logistic regression. En: Proceedings of the SAS global forum 2014 conference. 2014. p. 1-13. Disponible en: <https://support.sas.com/resources/papers/proceedings14/1485-2014.pdf>
91. Lee S, Jun CH. Fast incremental learning of logistic model tree using least angle regression. Expert Systems with Applications [Internet]. mayo de 2018 [citado 18 de marzo de 2024];97:137-45. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/S0957417417308308>
92. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. BMC Med [Internet]. diciembre de 2019 [citado 13 de junio de 2021];17(1):230. Disponible en: <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-019-1466-7>
93. Aburass S. Quantifying Overfitting: Introducing the Overfitting Index [Internet]. arXiv; 2023 [citado 4 de noviembre de 2024]. Disponible en: <https://arxiv.org/abs/2308.08682>
94. Fernández-Díaz R, Hoang TL, Lopez V, Shields DC. A new framework for evaluating model out-of-distribution for the biochemical domain [Internet]. 2024 [citado 4 de noviembre de 2024]. Disponible en: <http://biorxiv.org/lookup/doi/10.1101/2024.03.14.584508>
95. Nevalainen R, Ruiz A, Varkoi T. Making Software Safety Assessable and Transparent. En: McCaffery F, O'Connor RV, Messnarz R, editores. Systems, Software and Services Process Improvement [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013 [citado 10 de diciembre de 2024]. p. 1-12. (Communications in Computer and Information Science; vol. 364). Disponible en: http://link.springer.com/10.1007/978-3-642-39179-8_1
96. WHO. Ethics and Governance of Artificial Intelligence for Health: WHO Guidance [Internet]. 1st ed. Geneva: World Health Organization; 2021. 1 p. Disponible en: <https://iris.who.int/bitstream/handle/10665/341996/9789240029200-eng.pdf?sequence=1>
97. Van Calster B, Steyerberg EW, Wynants L, Van Smeden M. There is no such thing as a validated prediction model. BMC Med [Internet]. 24 de febrero de 2023 [citado 22 de marzo de 2024];21(1):70. Disponible en: <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-023-02779-w>
98. Rentroia-Pacheco B, Bellomo D, Lakeman IMM, Wakkee M, Hollestein LM, Van Klaveren D. Weighted metrics are required when evaluating the performance of prediction models in nested case-control studies. BMC Med Res Methodol [Internet]. 17 de mayo de 2024 [citado 10 de diciembre de 2024];24(1):115. Disponible en:

- <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-024-02213-6>
99. Ensor J, Martin EC, Riley RD. Package 'pmsampsize'. 2022; Disponible en: <https://cran.r-project.org/web/packages/pmsampsize/pmsampsize.pdf>
 100. Sepandi M, Alimohamadi Y, Hailu Tola H, Abbasi-Ghahramanloo A, Janani M. Case Fatality Rate of COVID-19: A Systematic Review and Meta-Analysis. *Journal of Preventive Medicine and Hygiene* [Internet]. 30 de julio de 2021 [citado 11 de marzo de 2024];E311 Pages. Disponible en: <https://www.jpmmh.org/index.php/jpmmh/article/view/1627>
 101. Ma J, Dhiman P. Poor handling of continuous predictors in clinical prediction models using logistic regression: a systematic review. *Journal of Clinical Epidemiology* [Internet]. septiembre de 2023 [citado 30 de octubre de 2023];161:140-51. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/S0895435623001890>
 102. Hadley Wickham GG. R for Data Science [Internet]. 2.^a ed. Vol. 1. O'Reilly; 2023 [citado 7 de diciembre de 2024]. Disponible en: <https://es.r4ds.hadley.nz/>
 103. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* [Internet]. 1 de enero de 2012 [citado 16 de marzo de 2024];28(1):112-8. Disponible en: <https://academic.oup.com/bioinformatics/article/28/1/112/219101>
 104. Gómez-Méndez I, Joly E. On the consistency of a random forest algorithm in the presence of missing entries. *Journal of Nonparametric Statistics* [Internet]. 6 de junio de 2023 [citado 16 de marzo de 2024];1-35. Disponible en: <https://www.tandfonline.com/doi/full/10.1080/10485252.2023.2219783>
 105. Huang TM. A Knot Selection Algorithm for Splines in Logistic Regression. En: *Proceedings of the 2020 3rd International Conference on Mathematics and Statistics* [Internet]. Paris France: ACM; 2020 [citado 16 de marzo de 2024]. p. 29-33. Disponible en: <https://dl.acm.org/doi/10.1145/3409915.3409921>
 106. Jamil S, Agarwal M. Fusion of Feature Selection Techniques and Machine learning Algorithms for Attack Classification on 802.11 Wi-Fi AWID Dataset. En: *2023 IEEE Guwahati Subsection Conference (GCON)* [Internet]. Guwahati, India: IEEE; 2023 [citado 18 de marzo de 2024]. p. 01-6. Disponible en: <https://ieeexplore.ieee.org/document/10183427/>
 107. Frank E, Harrell J. rms: Regression Modeling Strategies. R package version 5.1-4. 2019 [citado 12 de diciembre de 2024]; Disponible en: <https://CRAN.R-project.org/package=rms>
 108. Cole TJ, Altman DG. Statistics Notes: Percentage differences, symmetry, and natural logarithms. *BMJ* [Internet]. 16 de agosto de 2017 [citado 10 de diciembre de 2024];j3683. Disponible en: <https://www.bmj.com/lookup/doi/10.1136/bmj.j3683>

109. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North. 1era ed. New York: Taylor & Francis; 2023.
110. Gupta RK, Marks M, Samuels THA, Luintel A, Rampling T, Chowdhury H, et al. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study. *Eur Respir J* [Internet]. diciembre de 2020 [citado 4 de octubre de 2021];56(6):2003498. Disponible en: <http://erj.ersjournals.com/lookup/doi/10.1183/13993003.03498-2020>
111. National Health Service [Internet]. [citado 26 de febrero de 2024]. Disponible en: <https://www.england.nhs.uk/>
112. Zhang K, Zhang X, Ding W, Xuan N, Tian B, Huang T, et al. The Prognostic Accuracy of National Early Warning Score 2 on Predicting Clinical Deterioration for Patients With COVID-19: A Systematic Review and Meta-Analysis. *Front Med* [Internet]. 9 de julio de 2021 [citado 26 de febrero de 2024];8:699880. Disponible en: <https://www.frontiersin.org/articles/10.3389/fmed.2021.699880/full>
113. Paraskevas T, Michailides C, Karalis I, Kantanis A, Chourpiliadi C, Matthaiakaki E, et al. External validation of the 4C Mortality Score and PRIEST COVID-19 Clinical Severity Score in patients hospitalized with COVID-19 pneumonia in Greece. *Romanian Journal of Internal Medicine* [Internet]. 1 de diciembre de 2022 [citado 1 de marzo de 2024];60(4):244-9. Disponible en: <https://www.sciendo.com/article/10.2478/rjim-2022-0015>
114. Wickstrøm KE, Vitelli V, Carr E, Holten AR, Bendayan R, Reiner AH, et al. Regional performance variation in external validation of four prediction models for severity of COVID-19 at hospital admission: An observational multi-centre cohort study. Lazzeri C, editor. *PLoS ONE* [Internet]. 25 de agosto de 2021 [citado 26 de febrero de 2024];16(8):e0255748. Disponible en: <https://dx.plos.org/10.1371/journal.pone.0255748>
115. Cruz JC, Martins CK, Katchborian Neto A, De Araújo CM, Dellalibera-Joviliano R, Barbosa F. Age, gender, and 11 comorbidities as risk factors associated with COVID-19 mortality: A retrospective cohort including 1.8 million individuals. *Journal of Toxicology and Environmental Health, Part A* [Internet]. 18 de julio de 2023 [citado 2 de diciembre de 2024];86(14):491-500. Disponible en: <https://www.tandfonline.com/doi/full/10.1080/15287394.2023.2223598>
116. Masudah AF, Purwanti LE, Ernawati H, Rohma NM, Muslimah VA, Muslimah VI. Age and Comorbidities as The Main Factors of Mortality in Covid-19 Patients : A Retrospective Cohort Study. *IJNP* [Internet]. 27 de junio de 2023 [citado 2 de diciembre de 2024];7(1). Disponible en: <https://journal.umy.ac.id/index.php/ijnp/article/view/17347>
117. Tredinnick-Rowe J, Symonds R. Rapid systematic review of respiratory rate as a vital sign of clinical deterioration in COVID-19. *Expert Review of Respiratory Medicine* [Internet]. 2 de noviembre de 2022 [citado 2 de diciembre

- de 2024];16(11-12):1227-36. Disponible en: <https://www.tandfonline.com/doi/full/10.1080/17476348.2023.2169138>
118. Yangchen T, Korashy FM, Xu C, Hou W, Rohatgi R. Initial mean arterial blood pressure (MABP) measurement is a risk factor for mortality in hypertensive COVID-19 positive hospitalized patients. Wang Y, editor. PLoS ONE [Internet]. 30 de marzo de 2023 [citado 2 de diciembre de 2024];18(3):e0283331. Disponible en: <https://dx.plos.org/10.1371/journal.pone.0283331>
 119. Callejas D, Mosquera-Sulbaran J, Silva L, Carbonell A, Rivera-Porras D, Bermudez V, et al. Neutrophil/Lymphocyte Ratio and C-Reactive Protein in Patients with COVID-19: Correlations in Deceased and Surviving Patients [Internet]. 2024 [citado 2 de diciembre de 2024]. Disponible en: <https://www.preprints.org/manuscript/202410.1481/v1>
 120. Bi X, Su Z, Yan H, Du J, Wang J, Chen L, et al. Prediction of severe illness due to COVID-19 based on an analysis of initial Fibrinogen to Albumin Ratio and Platelet count. Platelets. 3 de julio de 2020;31(5):674-9. Disponible en: <https://doi.org/10.1080/09537104.2020.1760230>
 121. Malik P, Patel U, Mehta D, Patel N, Kelkar R, Akrmah M, et al. Biomarkers and outcomes of COVID-19 hospitalisations: systematic review and meta-analysis. BMJ EBM [Internet]. junio de 2021 [citado 6 de noviembre de 2023];26(3):107-8. Disponible en: <https://ebm.bmj.com/lookup/doi/10.1136/bmjebm-2020-111536>
 122. Chung H, Ko H, Lee H, Yon DK, Lee WH, Kim T, et al. Development and validation of a deep learning model to diagnose COVID-19 using time-series heart rate values before the onset of symptoms. Journal of Medical Virology [Internet]. febrero de 2023 [citado 26 de noviembre de 2024];95(2):e28462. Disponible en: <https://onlinelibrary.wiley.com/doi/10.1002/jmv.28462>
 123. Subasi A. Hospital readmission forecasting using artificial intelligence. En: Applications of Artificial Intelligence in Healthcare and Biomedicine [Internet]. Elsevier; 2024 [citado 3 de diciembre de 2024]. p. 455-520. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/B9780443223082000068>
 124. Moyenuddin M. Future Pandemic Could Be Worse. Infect Dis Clin Pract [Internet]. noviembre de 2023 [citado 6 de diciembre de 2024];31(6). Disponible en: <https://journals.lww.com/10.1097/IPC.0000000000001330>
 125. Alinejad T, Zareh D, Hao Z, Zhou T, Chen C shui. SARS-CoV-2 Mutation Mechanism, Features, and Future Perspective. En: J. Rodriguez-Morales A, editor. Infectious Diseases [Internet]. IntechOpen; 2023 [citado 6 de diciembre de 2024]. Disponible en: <https://www.intechopen.com/chapters/83476>
 126. Adla K, Dejan K, Neira D, Dragana Š. Degradation of ecosystems and loss of ecosystem services. En: One Health [Internet]. Elsevier; 2022 [citado 6 de diciembre de 2024]. p. 281-327. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/B9780128227947000083>

ANEXOS

Anexo 1. Participación en eventos científicos y publicaciones a partir de los resultados de la tesis

	
Participación en evento 1	Participación en evento 2
	
Publicación 1	Publicación 2

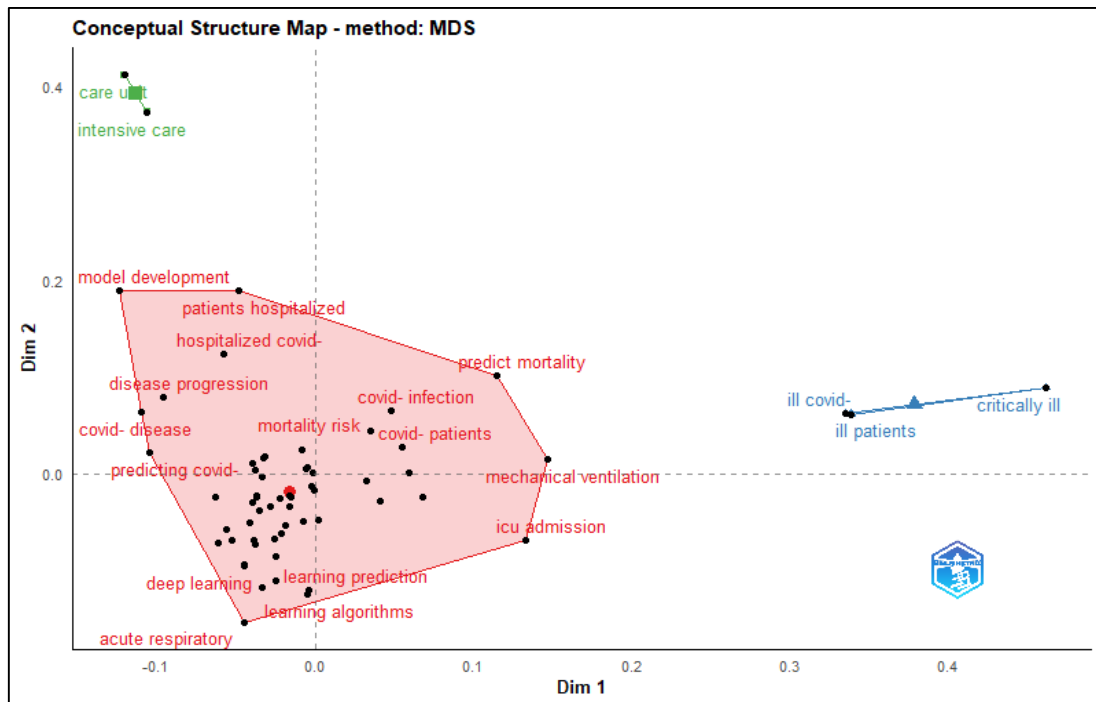
Anexo 2. Principales recursos de información, autores y artículos sobre el tema según la base de datos Dimension

Tabla 1. Principales recursos de información, autores y artículos sobre el tema según la base de datos Dimension

Recursos de información	Publicaciones	Citas	MC
Scientific Reports	22	389	17,68
PLOS ONE	17	766	45,06
Frontiers in Medicine	16	103	6,44
BMJ Open	14	119	8,50
Journal of Medical Internet Research	11	289	26,27
Nombre, organización, país	Publicaciones	Citas	MC
Maarten Van Smeden, University Medical Center Utrecht, Netherlands	6	2 231	371,83
Gary Stephen Collins, University of Oxford, United Kingdom	6	1 964	327,33
Ting Wang, Sichuan University, China	4	103	25,75
Aziz Sheikh, University of Edinburgh, United Kingdom	4	159	39,75
Zong-An Liang, Sichuan University, China	4	103	25,75
Artículos (DOI)	Citas	FCR	RCR
doi.org/10.1136/bmj.m1328	1 947	1084,01	177,66
doi.org/10.1007/s10654-020-00649-w	187	67	13
doi.org/10.1038/s41598-021-81844-x	48	-	-
doi.org/10.1016/s2213-2600(20)30559-2	123	-	-
doi.org/10.1021/acsinfecdis.0c00522	95	28	9,58

Leyenda: MC = Media de citas, FCR = Proporción de citas de campo, RCR = Proporción relativa de citas.
Fuente: Monzón-Pérez M, Sánchez-Valdés L, Lage-Dávila A. Estado y tendencias en la investigación sobre modelos de predicción clínica para la estratificación del riesgo de severidad en pacientes confirmados de Covid-19. 22 de marzo de 2023 [citado 22 de marzo de 2023];22(1). Disponible en: <https://revhabanera.sld.cu/index.php/rhab/article/view/5257>

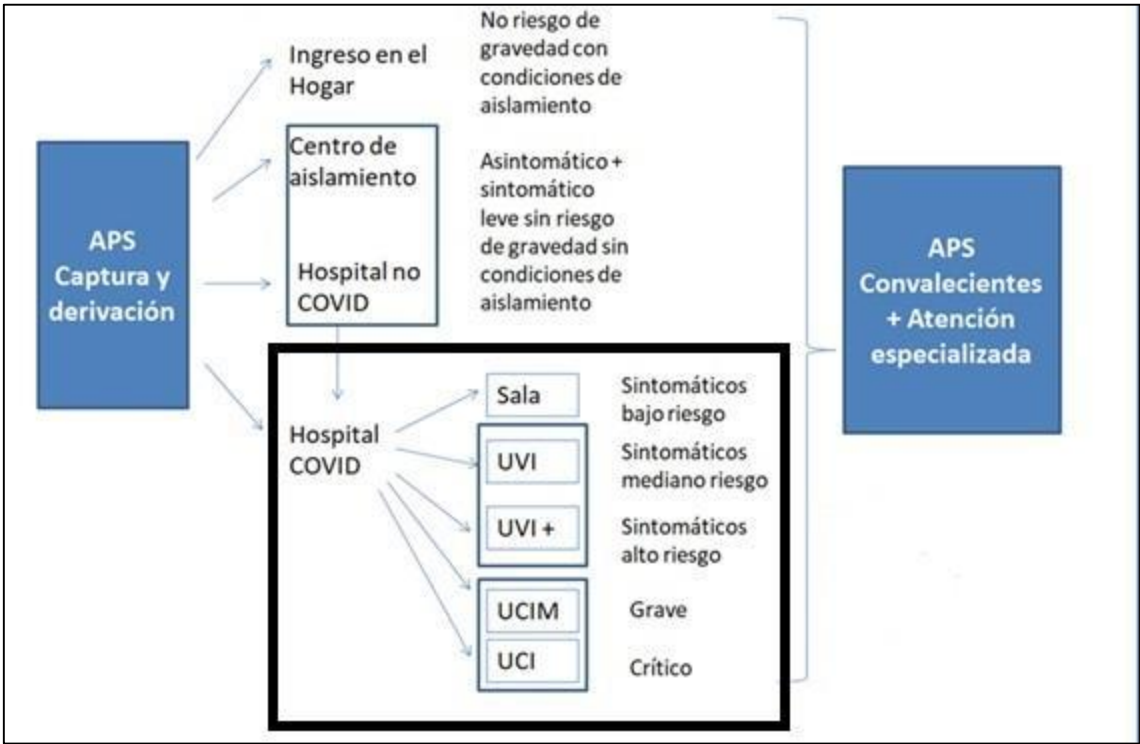
Anexo 3. Principales conceptos de los trabajos seleccionados



Fuente: Monzón-Pérez M, Sánchez-Valdés L, Lage-Dávila A. Estado y tendencias en la investigación sobre modelos de predicción clínica para la estratificación del riesgo de severidad en pacientes confirmados de COVID-19. 22 de marzo de 2023 [citado 22 de marzo de 2023];22(1). Disponible en: <https://revhabanera.sld.cu/index.php/rhab/article/view/5257>

Figura 1. Principales conceptos de los trabajos seleccionados.

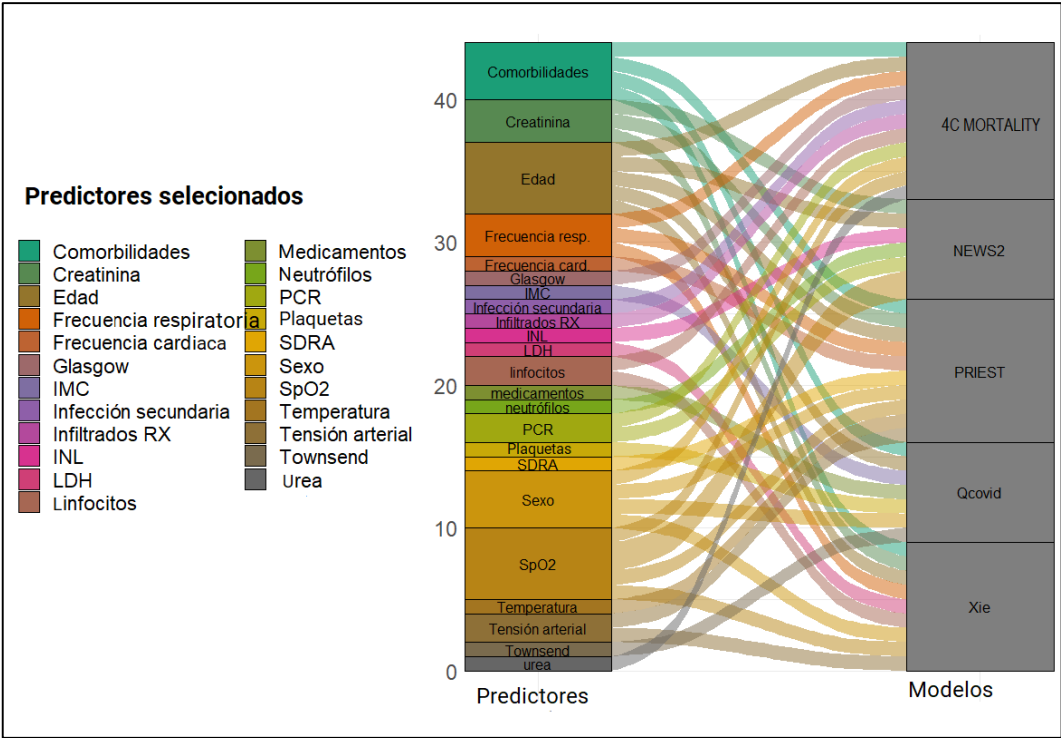
Anexo 4. Escenarios de atención médica en Cuba para el enfrentamiento a la COVID-19 y su relación con la estratificación del riesgo



Nota: Adaptado de 1. Minsap. Protocolo de Actuación Nacional para la COVID-19 (versión 1.5) [Internet]. 2020 [citado 6 de junio de 2023]. Disponible en: <https://covid19cubadata.github.io/protocolos/protocolo-version-5.pdf>

Figura 2. Escenarios de atención médica en Cuba para el enfrentamiento a la COVID-19.

Anexo 5. Relación entre modelo de mayor calidad y sus predictores



Fuente: Monzon-Perez M, Sanchez-Valdés L, Lage-Dávila A. Identificación de pacientes de muy bajo riesgo de transición a la severidad en confirmados de COVID-19. Cuba. Marzo 2020-Mayo 2021. 2 de febrero de 2022 [citado 3 de marzo de 2022];22(1). Disponible en: <http://www.revhabanera.sld.cu/index.php/rhab/article/view/4943.5>

Figura 3. Relación entre modelo de mayor calidad y sus predictores.

Anexo 6. Indicadores de estrato

Los modelos de árboles de clasificación aportan una estratificación del riesgo de muerte por COVID-19 de la cual se derivan tres indicadores. Para cada estrato, los indicadores son el Tamaño Relativo del Estrato (TRE), Riesgo de Muerte (RM) y el Riesgo de Muerte Explicada (RM), cuyas fórmulas de cálculos son las siguientes:

$$\begin{aligned} \textbf{Tamaño relativo del estrato (TR):} \\ TR &= \frac{\text{Pacientes del estrato}}{\text{Total de pacientes}} \\[10pt] \textbf{Riesgo de muerte (RM):} \\ RM &= \frac{\text{Pacientes fallecidos}}{\text{Total de pacientes del estrato}} \\[10pt] \textbf{Riesgo de muerte explicada (ME):} \\ ME &= \frac{\text{Pacientes fallecidos del estrato}}{\text{Total de pacientes fallecidos}} \end{aligned}$$

Nota: los indicadores se pueden utilizar en forma de proporción o en forma de porcentaje si se multiplican el resultado del cociente por 100.

Figura 4. Fórmulas para el cálculo de los indicadores de estrato.

El tamaño relativo del estrato (TRE) mide la frecuencia relativa de determinados pacientes (estratos) en comparación con el total de individuos expuestos al riesgo. Esta es una medida de la importancia relativa de determinadas características en la población estudiada.

- Proporciones bajas del TRE: Implica que hay menos pacientes en ese estrato, lo que permite redirigir recursos a otros grupos que lo necesiten. Sin embargo, es importante considerar el contexto general y el resto de los indicadores para determinar si se trata de pacientes de alto riesgo o bajo riesgo.

- Proporciones altas del TRE: Indica que hay un gran número de pacientes en esa categoría, lo que podría involucrar una alta demanda de recursos de salud para ese grupo y una especial atención al riesgo. Algunas decisiones que se podrían derivar de ese contexto son las siguientes:
- Proporciones altas del TRE con cifras bajas de RM y ME: Se trata de una oportunidad para atenuar la atención médica y ahorrar recursos. Por ejemplo, los pacientes de bajo riesgo en el estrato (menores de 65 años y sin comorbilidades) se pueden tratar en escenarios comunitarios y realizar acciones como ingreso en el hogar con seguimiento por la APS y tratamiento sintomático.
- Proporciones altas de TRE con RM y ME altos: La combinación de cifras elevadas en TRE, RM y ME indica una situación crítica con alto riesgo de estrés de los servicios médicos y saturación de los recursos de salud. Esto requiere una respuesta rápida y una planificación efectiva. Si el número de pacientes en el estrato supera la cantidad de recursos de salud específicos, como el número de camas, se puede considerar una situación de alarma, dado que se ha alcanzado el punto de saturación para ese recurso.

El riesgo de muerte (RM) es la tasa de muertes en función del número de sujetos expuestos al riesgo en el estrato. Este indicador que define el riesgo de muerte para grupos con características específicas. En los pacientes de bajo riesgo está justificada la atenuación de la conducta médica. Por el contrario, el grupo de peor pronóstico, es aquel al cual se debe orientar el tratamiento más intensivo.

El riesgo de muerte explicada (RM) se calcula como la proporción de muertes dentro del total de fallecidos. Este indicador también mide la importancia relativa

de grupos de fallecidos con características específicas. Un RM bajo sugiere que este grupo no está contribuyendo significativamente a la mortalidad total, lo que puede permitir una atención menos intensiva. Aunque en este contexto es probable una atenuación de la conducta médica, es importante monitorear a estos pacientes, asegurarse de que no se ignoren sus necesidades, y priorizar la prevención de salud, la educación y la atención de emergencias. Por el contrario, un RM alto en un estrato específico indica que este grupo está contribuyendo significativamente a la mortalidad total, lo que requiere atención prioritaria. La asignación de recursos significativos a este grupo se justifica, ya que su atención puede tener un impacto considerable en la reducción de la mortalidad.

Anexo 7. Variaciones temporales a lo largo de la pandemia en Cuba

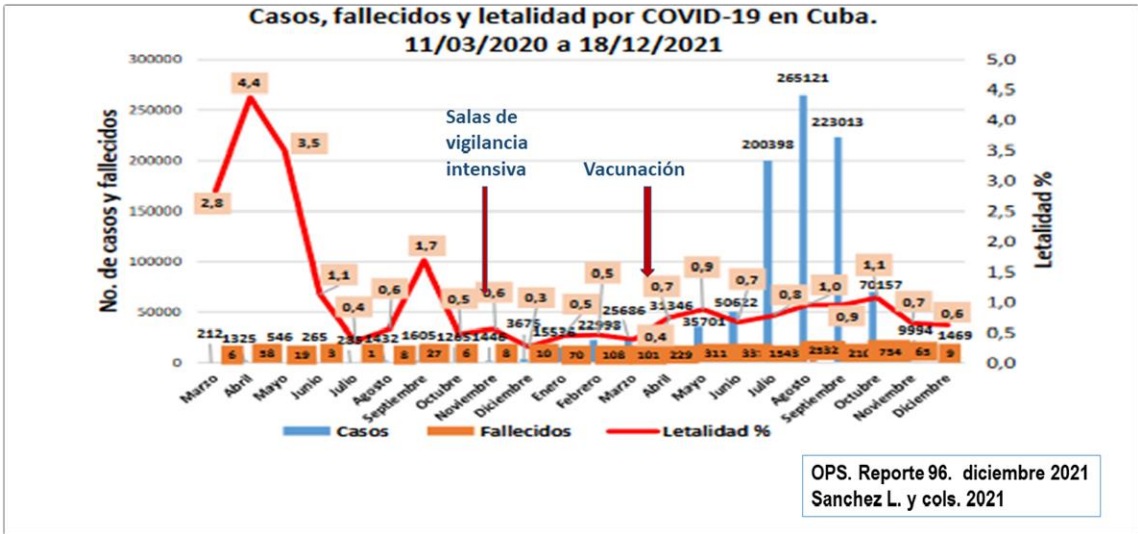


Figura 5. Variaciones temporales a lo largo de la pandemia en Cuba.

Anexo 8. Versiones del PAN y frecuencia de casos por versiones

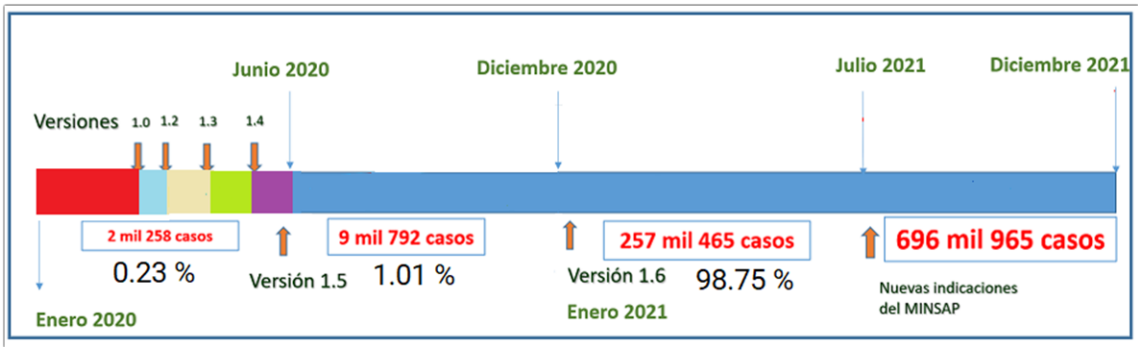
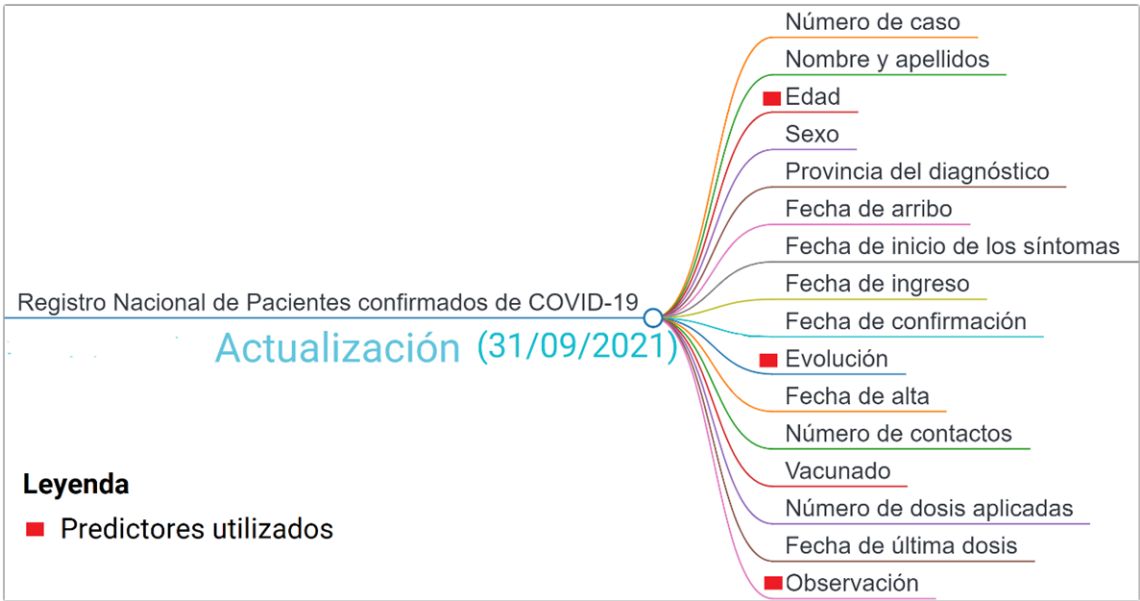


Figura 6. Versiones del PAN y frecuencia de casos por versiones.

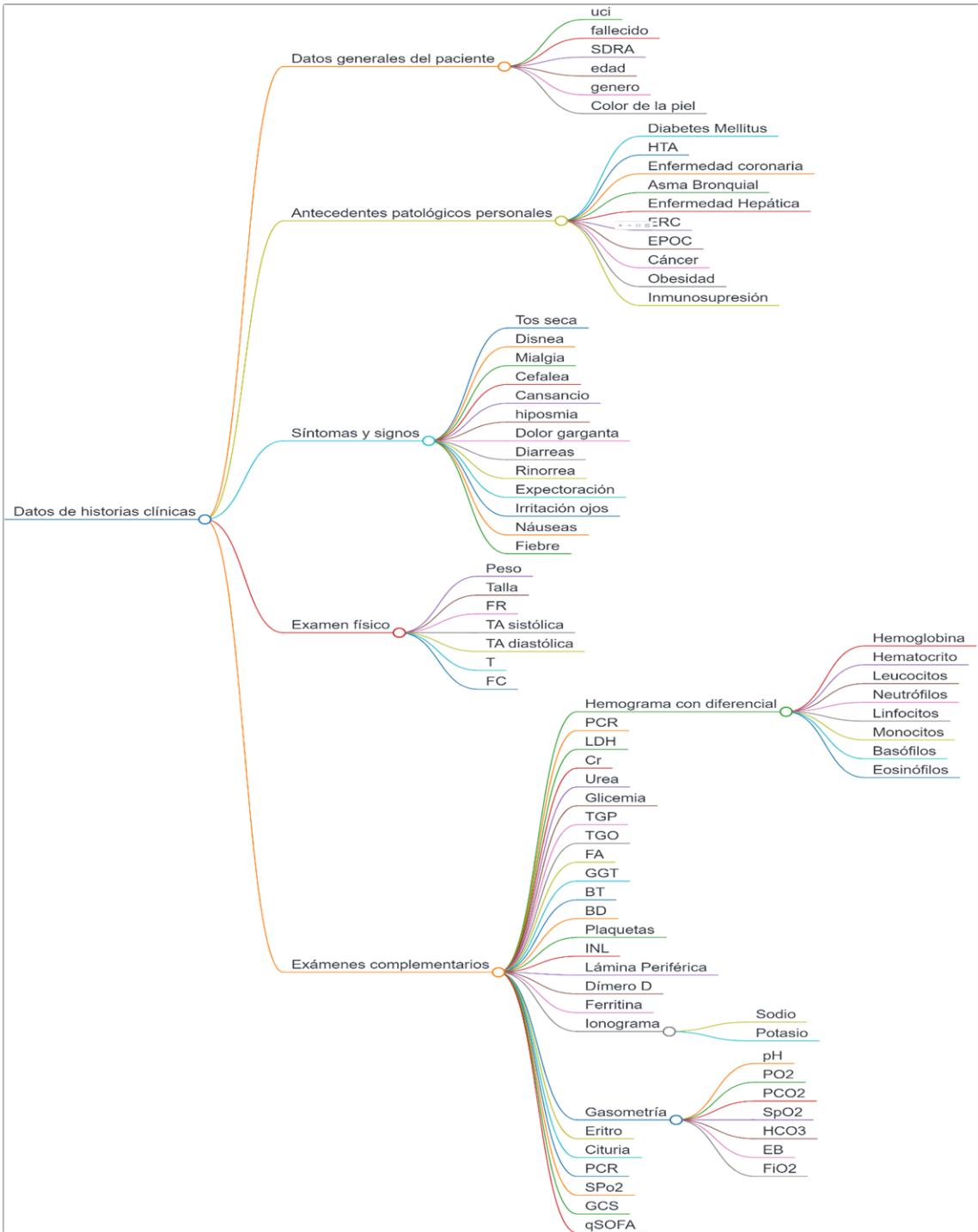
Anexo 9. Variables disponibles en RNPC



Nota: Observación es una variable cualitativa de respuesta abierta, que puede incluir múltiples situaciones o no contener información, lo que no significa que no exista algún hallazgo relevante, pues depende de lo que haya considerado el que registró el caso.

Figura 7. Variables disponibles en el Registro Nacional de Pacientes confirmados de COVID-19 (RNPC).

Anexo 10. Variables disponibles en RICPA



Fuente: Amaylid Arteaga García, Maicel Monzón. PROYECTO COVID ICPA. 2024 [citado 1 de diciembre de 2024]; Disponible en: <https://rgdoi.net/10.13140/RG.2.2.29975.18087>

Figura 8. Variables disponibles en el Registro del proyecto Impacto de la COVID-19 en pacientes adultos en Cuba. 2020-2023 (ICPA).

Anexo 11 Extrapolación y comparación de los modelos

Procedimiento para comparar modelos

1. Se aplica la fórmula que se muestra a continuación al intercepto del modelo original con el modelo de referencia según la prevalencia deseada para el ajuste.
2. Se utiliza los modelos ajustados para hacer predicciones en el conjunto de datos de la investigación.
3. Se identifican las diferencias en los desempeños de los modelos.

A continuación, se explica el procedimiento para la estandarización

Dada la función logit

$$P(Y = 1) = \frac{1}{1 + \exp(-a^* - b_1 X_1 - b_2 X_2 - \dots - b_k X_k)}$$

La fórmula para corregir el término independiente a^* fue la siguiente:

$$a^* = a - \ln \left[\left(\frac{f_1}{1-f_1} \right) \left(\frac{1-f_2}{f_2} \right) \right]$$

A continuación, se muestra para cada modelo la tasa de prevalencia reportada en el artículo y el factor de corrección para su comparación con COVIDCENCEC con una tasa de letalidad de 50 % por ser un estudio de casos y controles (tabla 3).

Tabla 3. Corrección del término independiente según modelos seleccionados

Modelos	Prevalencia reportada (%)	Término independiente corregido (a^*)
PRIEST	13.5	2.696
4C MORTALITY	32.2	1.750
NEWS 2	20.3	2.283
Modelo de Xie	11.0	2.591

Anexo 12 Estrategias para la identificación y mitigación del sesgo

Tabla 4. Estrategias para la identificación y mitigación del sesgo

Identificación	Tipo de sesgo	Estrategia de mitigación
Selección de Casos y Controles	Sesgo de Selección	Se definieron criterios estrictos de inclusión/exclusión para casos y controles basados en medidas objetivas (muerte hasta los 30 días).
Medición de Predictores	Sesgo de Información	Se utilizaron medidas estandarizadas y objetivas para la evaluación de predictores, aplicadas consistentemente a casos y controles (El PAN proporciona uniformidad de criterios)
		Considerar cuidadosamente el momento de la medición del predictor en relación con el resultado. Los predictores medidos después del resultado podrían estar influenciados por el mismo, introduciendo sesgo
Datos Faltantes de Predictores	Sesgo de Información	Emplear métodos para manejar datos faltantes, como la imputación múltiple
		Uso de múltiples fuentes de datos
		Estandarización y uniformidad de criterios por el empleo generalizado del PAN
Valores extremos	Sesgo de Información	Se utilizaron técnicas para el tratamiento de datos atípicos
Especificación y Selección del Modelo	Sesgo de Selección	Pre-especificar un número limitado de predictores candidatos basado en una revisión exhaustiva de la literatura existente y experiencia clínica.
Validación del Modelo	Sobreajuste	Diseño metacéntrico
		Se utilizaron técnicas de remuestreo como la validación cruzada para evaluar la validez interna del modelo y estimar el grado de optimismo en su rendimiento.
		Se realizó validación externa con criterio temporal y espacial un conjunto de datos independiente para evaluar su validez externa y generalización a otras poblaciones.

Clasificación errónea del resultado	Sesgo de Información	Se midió la variable resultado variable estado al egreso como la muerte hasta los 30 días con un criterio preciso.
Falta de representatividad por cambios de la conducta y tratamientos diferenciales	Sesgo de confusión	Diseño multicéntrico
		Aleatorización por hospitales

Anexo 13. Determinación del tamaño de muestra según criterios y parámetros seleccionados

Tabla 5. Determinación del tamaño de muestra según criterios y parámetros seleccionados

Criterio	Tamaño de muestra	AUC	R-cuadrado de Cox-Snell	R-cuadrado máximo	R-cuadrado de Nagelkerke	Eventos por parámetro predictivo (EPP)
1	2930	0,900	0,03	0,11	0,29	2,93
2	1854	0,851	0,03	0,11	0,29	1,85
3	1600	0,900	0,03	0,11	0,29	0,02
Final	2930	0,900	0,03	0,11	0,29	2,93

Nota: número de predictores = 10

Anexo 14. Operacionalización de las variables según etapas de la investigación

Tabla 6. Operacionalización de las variables según etapas de la investigación

Variable	Tipo	Operacionalización	
		Escala	Descripción
Etapa 1. Escenario comunitario			
Estado al egreso	Cualitativa nominal dicotómica	vivo/fallecido	Según se refiere el estado al egreso en la base de datos confirmados.
Edad	Cuantitativa continua	-	La edad del paciente en años referida en la base de datos confirmados
Sexo	Cualitativa nominal dicotómica	masculino/femenino	Sexo de nacimiento referido en la base de datos confirmados
Síntomas	Cualitativa nominal dicotómica	sintomático/asintomático	Según se refiere el antecedente o no de síntomas en la base de datos confirmados
Comorbilidades	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
HTA	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
Asma bronquial	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
Diabetes mellitus	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
EPOC	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la

			base de datos confirmados
Cardiopatía isquémica	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
Enfermedades cardiovasculares	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
Obesidad	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
Enfermedad renal crónica	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
Neoplasias	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
Enfermedad cerebrovascular	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
Inmunodeficiencias	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la base de datos confirmados
Etapla 2. Escenario hospitalario			
Estado al egreso	Cualitativa nominal dicotómica	vivo/fallecido	Según se refiere el estado al egreso en la historia clínica
Edad	Cuantitativa continua	-	Según edad en años cumplidos reflejados en la historia clínica
Sexo	Cualitativa nominal dicotómica	masculino/femenino	Según se refiere el sexo de nacimiento en la historia clínica
Suma de comorbilidades	Cuantitativa discreta	-	Número total de afecciones médicas preexistentes

			referidas en la historia clínica
HTA	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
Asma bronquial	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
Diabetes mellitus	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
EPOC	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
Enfermedad coronaria	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
Enfermedades cardiovasculares	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
Obesidad	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
ERC	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
Neoplasias	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
Enfermedades cerebrovasculares	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
Inmunodeficiencias	Cualitativa nominal dicotómica	sí/no	Según si se identifica o no este antecedente en la historia clínica
IRPS	Cualitativa ordinal	0,1,2,3,4,5	Medida agregada de la severidad de los síntomas

Frecuencia cardiaca	Cuantitativa discreta	-	Número de latidos del corazón por minuto referido en la historia clínica al ingreso
Frecuencia respiratoria	Cuantitativa discreta	-	Número de respiraciones por minuto referido en la historia clínica al ingreso
Presión arterial media	Cuantitativa continua	-	PAM (dos tercios de la Presión Diastólica más un tercio de la Presión Sistólica) referido en la historia clínica al ingreso
Velocidad de sedimentación globular	Cuantitativa continua	-	Según valores Velocidad de sedimentación globular (VSG) referidos en la historia clínica al ingreso
INL	Cuantitativa continua	-	Según valores de INL referidos en la historia clínica al ingreso
Conteo de plaquetas	Cuantitativa continua	-	Según cifras del PLT referidas en la historia clínica al ingreso
PCR	Cuantitativa continua	-	Según valores de proteína C reactiva (PCR) referidos en la historia clínica al ingreso
Creatinina sérica	Cuantitativa continua	-	Según concentración de creatinina en $\mu\text{mol/L}$ referidos en la historia clínica al ingreso
Urea	Cuantitativa continua	-	Según concentración de urea en $\text{mg}/100\text{ ml}$ referidos en la historia clínica al ingreso

LDH	Cuantitativa continua	-	Concentración de Lactato deshidrogenasa en mg/100 ml referidos en la historia clínica al ingreso
SpO ₂	Cuantitativa discreta	-	Porcentaje de hemoglobina en la sangre que está saturada con oxígeno referidos en la historia clínica al ingreso

Nota: A excepción del estado al egreso, el resto de las variables se obtienen al ingreso.

Anexo 15. Índice de riesgo ponderado de síntomas (IRPS)

Tabla 7. Ajuste del modelo de regresión logística a la variable estado al egreso, según síntomas para establecer la variable Índice de riesgo ponderado de síntomas (IRPS)

Síntomas	OR	Valor aproximado
Fiebre	0,10	-
Tos seca	1,28*	1
Disnea	2,30**	2
Alteración del estado mental	2,02**	2
Cansancio	0,10	-
Anosmia o hiposmia	0,41	-
Dolor de garganta	0,58	-
Diarreas	1,00	-
Rinorrea	0,42	-
Expectoración	0,68	-
Irritación en los ojos	0,42	-
Náuseas	0,36	-

Anexo 16. Frecuencia de datos perdidos en pacientes hospitalizados

Tabla 8. Frecuencia de datos perdidos en pacientes hospitalizados

Característica	Global N = 3 683	Entrenamiento N = 2 577	Prueba, N = 1 106
Estado al egreso, n (%)	292 (7,93)	200 (7,76)	92 (8,32)
Sexo, n (%)	212 (5,76)	143 (5,55)	69 (6,24)
FC Lat. x min, n (%)	198 (5,38)	138 (5,36)	60 (5,42)
FR Resp. x min, n (%)	139 (3,77)	96 (3,73)	43 (3,89)
PAM, mm/Hg, n (%)	156 (4,24)	112 (4,35)	44 (3,98)
VSG, n (%)	2 928 (79,5)	2 035 (79,0)	893 (80,7)
INL, n (%)	1 773 (48,1)	1,224 (47,5)	549 (49,6)
PLT, n (%)	1 582 (43,0)	1 119 (43,4)	463 (41,9)
PCR, n (%)	3 384 (91,9)	2 357 (91,5)	1 027 (92,9)
Cr, n (%)	2 128 (57,8)	1 488 (57,7)	640 (57,9)
Urea, n (%)	3 197 (86,8)	2 226 (86,4)	971 (87,8)
Ldh, n (%)	3 068 (83,3)	2 147 (83,3)	921 (83,3)
SpO2, n (%)	3 061 (83,1)	2 142 (83,1)	919 (83,1)

Nota: variables sin datos perdidos no se incluyen en la tabla.

Anexo 17. Factor de inflación de varianza del ajuste para el modelo COVIDCENCEC



Figura 11. Factor de inflación de varianza del ajuste para el modelo COVIDCENCEC.

Anexo 18. Estrategia para controlar las interacciones

Tabla 9. Efectos de las interacciones según variables seleccionadas

Características	OR [†]	95% CI [†]
Edad	1.06	1.01, 1.12
FR	1.21	1.02, 1.44
Suma comorbilidades	16.9	2.22, 130
PAM	1.01	0.98, 1.03
IRP síntomas	2.04	0.89, 4.65
INL	1.39	1.29, 1.49
Plaquetas	1.00	1.00, 1.00
Edad * FR	1.00	1.00, 1.00
Suma comorbilidades * PAM	0.97	0.95, 1.00
IRP síntomas * INL	0.97	0.93, 1.02
IRP síntomas * Plaquetas	1.00	1.00, 1.00
FR * IRP síntomas	1.01	0.97, 1.05
[†] OR = Odds Ratio, CI = Confidence Interval		

Anexo 19. Ecuación del modelo COVIDCENCEC

Se ajustó un modelo de regresión logística binaria múltiple para predecir la probabilidad de un evento binario, en este caso, el resultado estado al egreso (vivo = 0, fallecido = 1). La relación entre la variable resultado y los predictores se expresa como:

$$\text{Logit}(p) = \ln \left(\frac{p}{1-p} \right) = \beta_0 + \sum_{i=1}^n \beta_i X_i$$
$$p = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^n \beta_i X_i)}}$$

Donde los coeficientes son:

$$\begin{aligned}\beta_0 &= 0.00 \\ \beta_1 &= 1.11 \quad (\text{Edad}) \\ \beta_2 &= 0.96 \quad (\text{Edad}') \\ \beta_3 &= 1.13 \quad (\text{FR}) \\ \beta_4 &= 1.00 \quad (\text{FR}') \\ \beta_5 &= 0.96 \quad (\text{TAM}) \\ \beta_6 &= 1.05 \quad (\text{TAM}') \\ \beta_7 &= 1.86 \quad (\text{INL}) \\ \beta_8 &= 0.46 \quad (\text{INL}') \\ \beta_9 &= 1.00 \quad (\text{PLT}) \\ \beta_{10} &= 2.00 \quad (\text{PLT}') \\ \beta_{11} &= 1.92 \quad (\text{Suma de comorbilidades}) \\ \beta_{12} &= 1.94 \quad (\text{IRPS})\end{aligned}$$

Anexo 20. Métricas de rendimiento del modelo ABR según partición de datos

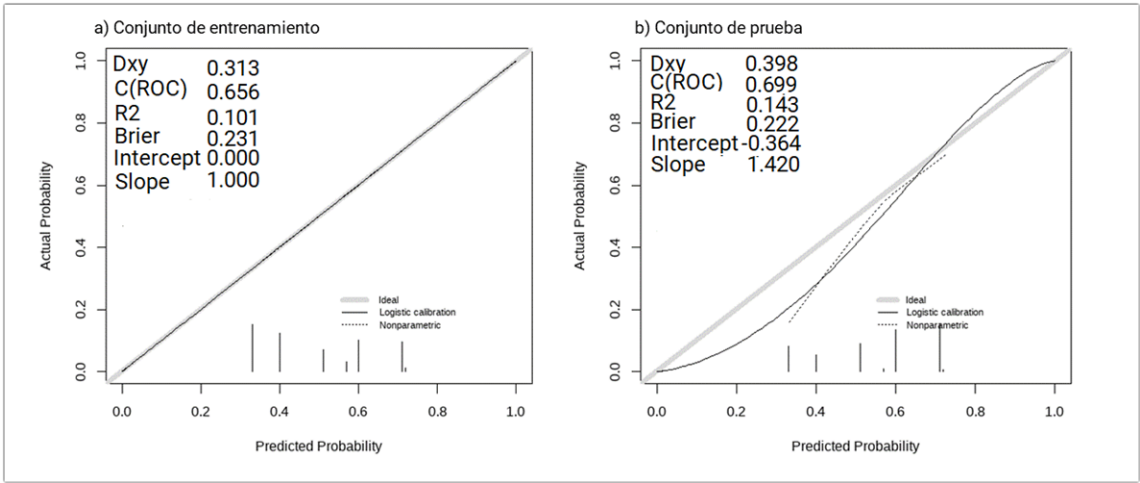


Figura 15. Métricas de rendimiento del modelo ABR según partición de datos.

Anexo 21. Métricas de calibración y discriminación del modelo AAR según partición de datos

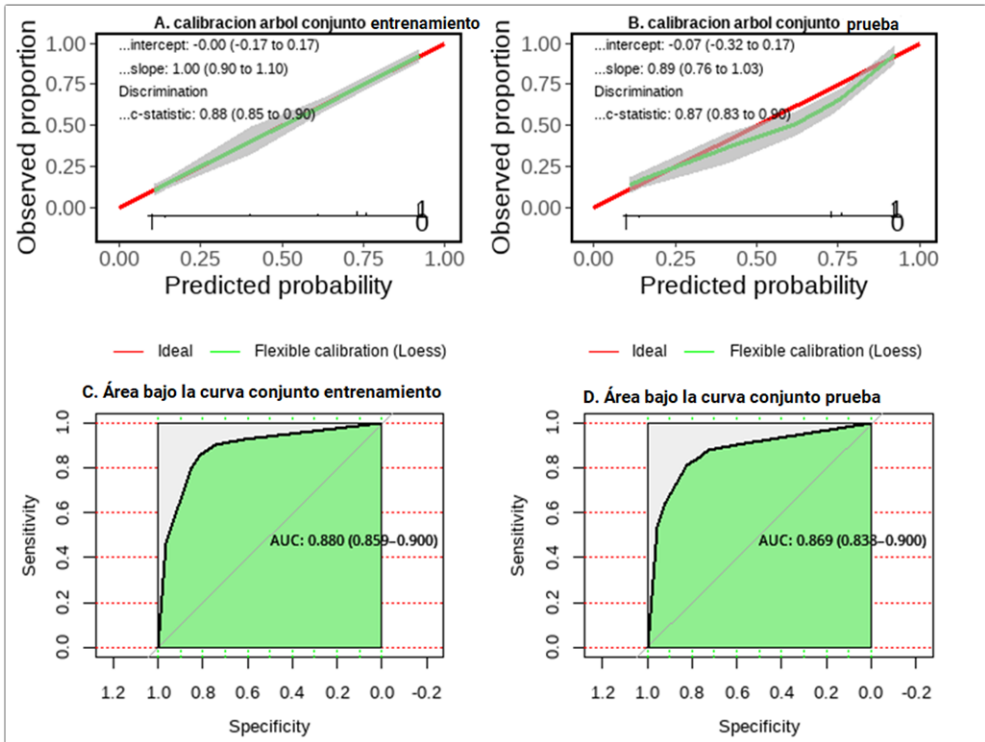


Figura 16. Métricas de calibración y discriminación del modelo AAR según partición de datos.

Anexo 22. Métricas de calibración y discriminación del modelo COVIDCENCEC según partición de datos

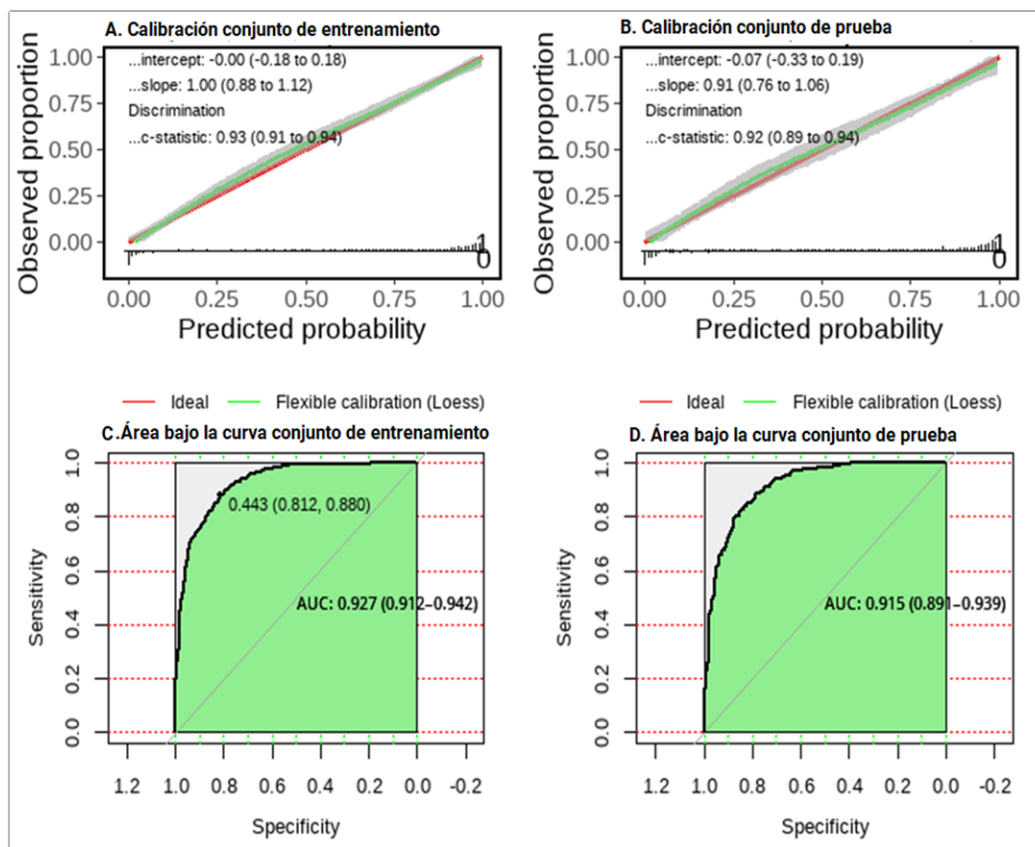
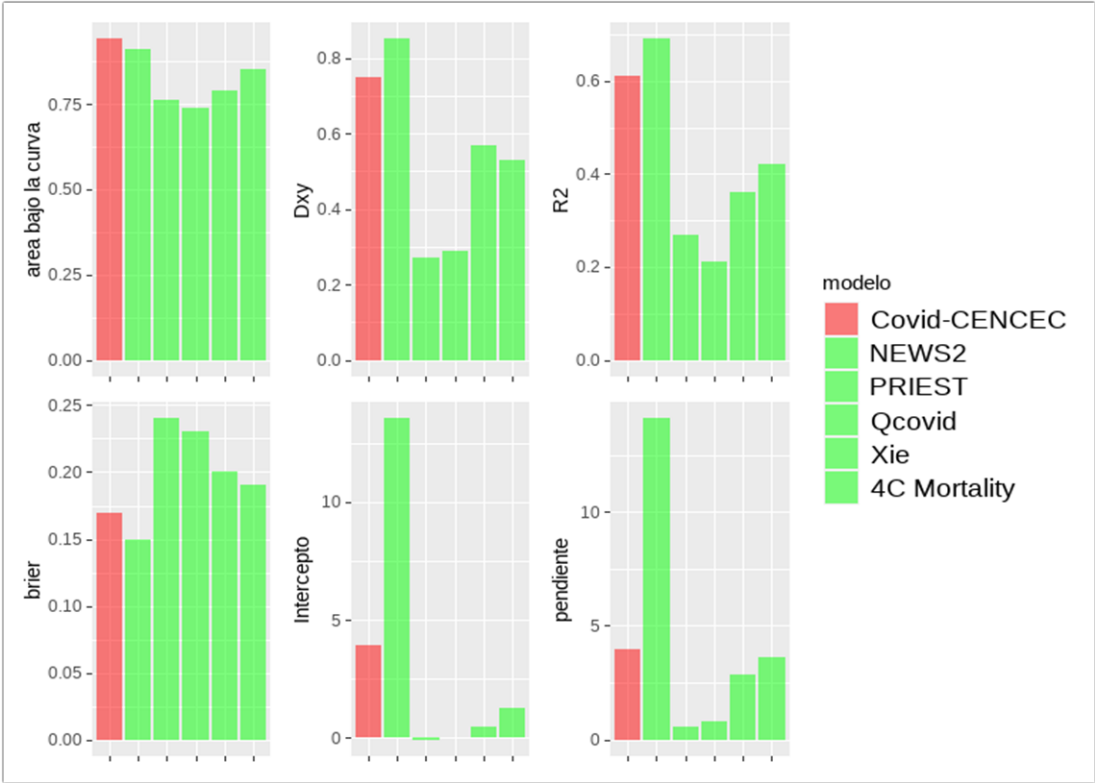


Figura 17. Métricas de calibración y discriminación del modelo COVIDCENCEC según partición de datos.

Anexo 23. Rendimiento de los modelos propuestos en el estudio e identificados como de mayor calidad



Fuente: Tabla 19

Figura 18. Rendimiento de los modelos propuestos en el estudio e identificados como de mayor calidad.

Anexo 24. Nomograma del modelo COVIDCENCEC

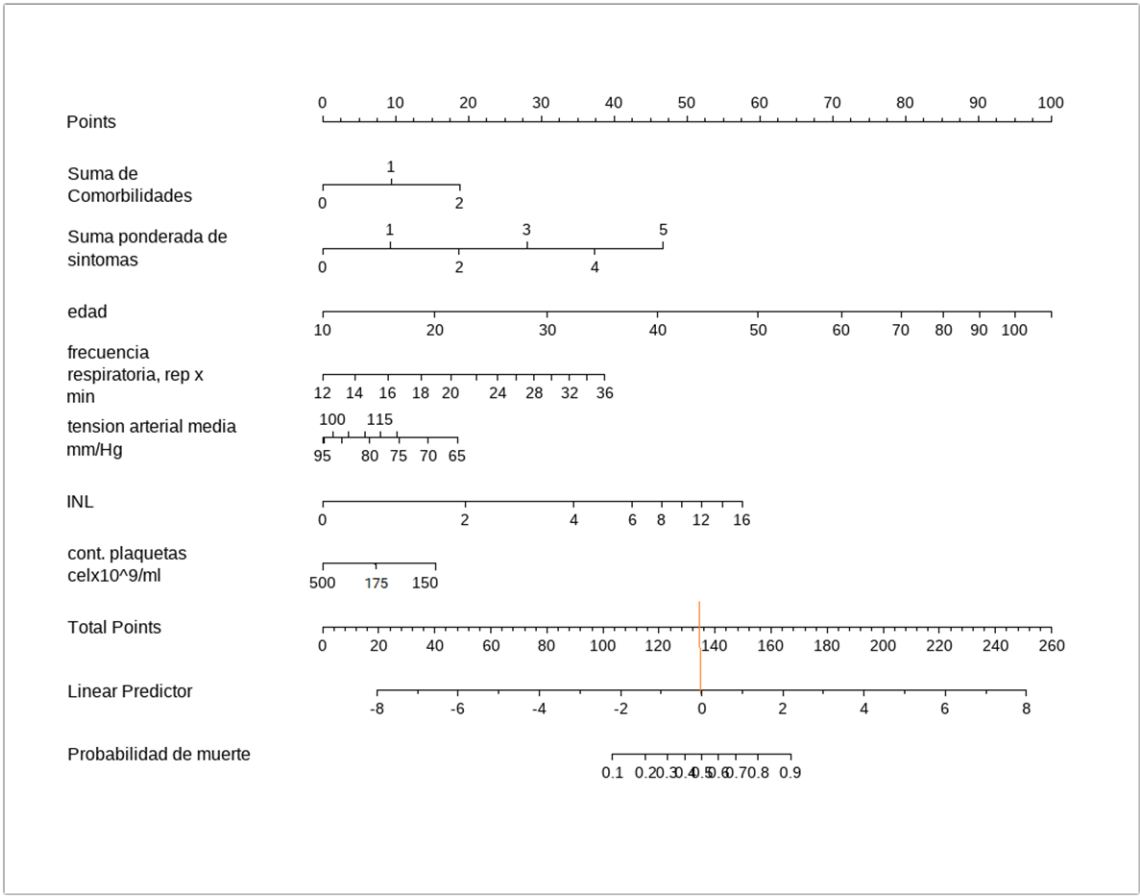


Figura 19. Nomograma del modelo COVIDCENCEC.

Anexo 25. Interfaz de la aplicación COVIDCENCEC-APK

The screenshot displays the COVID CENCEC application interface on a mobile device. At the top, the status bar shows the time as 9:17, signal strength, and battery level at 93%. The app title "COVID CENCEC" is prominently displayed. Below it, a section titled "Datos del paciente" contains several input fields for patient information: "Edad (años): 40", "Suma de comorbilidades (0-3): 0", "IRP síntomas (0-5): 3", "Frecuencia respiratoria (Resp x min): 18", "Presión arterial media (PAM): 80", "Índice neutrófilo-linfocito (INL): 4.8", and "Plaquetas (x10⁹/L): 175". A "Calculate" button is positioned below these fields. The result is shown as "Probabilidad:" followed by a horizontal bar chart. The bar is green and represents 43.06%, with the value "43.06%" displayed on the bar. The x-axis of the chart is labeled from 0 to 100.

Parameter	Value
Edad (años)	40
Suma de comorbilidades (0-3)	0
IRP síntomas (0-5)	3
Frecuencia respiratoria (Resp x min)	18
Presión arterial media (PAM)	80
Índice neutrófilo-linfocito (INL)	4.8
Plaquetas (x10 ⁹ /L)	175

Calculate

Probabilidad:

43.06%

0 100

Figura 20. Interfaz de la aplicación COVIDCENCEC-APK.

Anexo 26. Informe final de validación del software COVIDCENCEC-APK

Introducción

Este informe presenta un resumen del proceso de validación de la aplicación para teléfonos móviles COVIDCENCEC. La validación se llevó a cabo con el objetivo de asegurar que la aplicación cumpla con los requisitos funcionales, normativos y de documentación necesarios para su uso en un entorno clínico.

Metodología utilizada

El proceso de validación incluyó las etapas siguientes:

1. Pruebas funcionales

Se definieron casos de prueba que abarcan las funcionalidades clave del software.

Se realizaron pruebas con un grupo de colegas médicos, quienes proporcionaron retroalimentación sobre la usabilidad y el rendimiento.

2. Validación de requisitos

Se organizó una revisión colaborativa con médicos para asegurar que los requisitos fueran claros y relevantes.

Se ajustaron los requisitos según los comentarios recibidos.

3. Validación de cumplimiento normativo

Se investigaron las normativas aplicables para garantizar que el software cumpliera con los estándares requeridos.

Se creó una lista de verificación para evaluar el cumplimiento normativo.

4. Validación de la documentación

Se revisó toda la documentación relacionada con el software para asegurar su precisión y actualización.

Se solicitó retroalimentación a usuarios sobre la claridad y utilidad de la documentación.

Resultados obtenidos

Pruebas funcionales: Se identificaron y corrigieron errores menores en la interfaz y se mejoró la usabilidad del software.

Requisitos: La revisión colaborativa permitió ajustar varios requisitos, asegurando su alineación con las necesidades clínicas.

Cumplimiento normativo: El software cumplió con los estándares normativos relevantes tras la revisión y consulta a expertos.

Documentación: La documentación fue actualizada y mejorada sobre la base de los comentarios recibidos, garantizando su utilidad para los usuarios.

Recomendaciones para futuras mejoras

- Establecer un proceso continuo de retroalimentación con usuarios finales para identificar áreas de mejora.
- Realizar revisiones periódicas del cumplimiento normativo a medida que cambien las regulaciones.
- Mantener actualizada la documentación conforme se realicen modificaciones en el software.

Conclusión

La validación del software COVIDCENCEC-APK ha demostrado ser un proceso efectivo para asegurar su funcionalidad y utilidad en la práctica clínica. Los resultados obtenidos indican que la herramienta es adecuada para ayudar a médicos y pacientes en la evaluación del riesgo asociado a la COVID-19. Este informe servirá como referencia para futuras implementaciones y mejoras del software.