

Modelando casos COVID-19 para estimar la oferta de PCR y camas hospitalarias en servicios de salud

Modeling COVID-19 cases to estimate the supply of PCR and hospital beds in health services

Rodolfo Herrera Medina^{1,2}, Helmer Zapata Ossa^{3,4,5}, Luz Adriana Libreros Arana², John Housseman Bolaños Ramos⁵

1 Universidad Libre, Cali, Colombia. Grupo de Investigación en Epidemiología y Servicios GRIEPIS y grupo EPICRITICA.

2 Coomeva, Cali, Colombia

3 Universidad del Valle; Grupo de Investigación en Gestión y Estudios de Salud GIGES, Cali, Colombia

4 Fundación Universitaria San Martín, Grupo de Investigación de Salud Pública GISAP, Cali, Colombia

5 Secretaría Departamental de Salud del Valle, Cali, Colombia.

Correspondencia: Rodolfo Herrera. correo rodolfo-herreram@unilibre.edu.co

Recibido: 30 mayo 2020

Aceptado: 29 julio 2020

Publicado: 30 julio 2020

Palabras clave: COVID-19, predicción, predicción estadística, función Gompertz, PCR, camas hospitalarias, servicios de salud, estados de una epidemia

Keywords: COVID-19, prediction, statistical models, Gompertz function, PCR, hospital bed, health services, epidemic stages.

Citación: Herrera MR, Zapata OH, Libreros ALA, Bolaños RJH. Modeling COVID-19 cases to estimate the supply of CRP and hospital beds in health services. *ijEPH*. 2020; 3(1): e-6604. Doi: 10.18041/2665-427X/ijeph.1.6604

Resumen

Objetivo: Determinar la capacidad instalada de servicios necesaria para atender la pandemia del COVID-19 considerando como principal predictor los modelos estadísticos y matemáticos.

Métodos: Se describen cuatro fases: 1. Modelación de número de casos COVID-19, 2. Uso de información histórica del fenómeno, 3. Modelamiento estadístico no lineal, 4. Modelamiento matemático. Para este último se usó la función Gompertz con el cual se realizaron las proyecciones de casos confirmados. La fuente de datos principal fue la suministrada por el Instituto Nacional de Salud de Colombia. Los parámetros para estimar la oferta de servicios se basaron en la proyección de casos y los parámetros establecidos según los lineamientos del INS.

Resultados: Los modelos de mayor ajuste para las diferentes etapas de la pandemia fueron: la exponencial, el modelo de regresión no lineal cúbico con un coeficiente de determinación del 99.7%; para las etapas de crecimiento exponencial y de estabilización de la pandemia fue el modelo matemático Gompertz. La combinación de modelos permitió predecir para el 31 de agosto un acumulado de 899,708 casos confirmados para Colombia con Fecha de Inicio de Síntomas. Estas predicciones permitieron estimar servicios para una población objetivo de un millón seiscientos (1,600,000) de habitantes en 47,328 pruebas de PCR para el mes de agosto, 1,072 camas hospitalarias y 365 camas en Unidades de Cuidados Intensivos -UCI para ese mes.

Conclusiones: La combinación de técnicas estadísticas y matemáticas sencillas y robustas permitió estimaciones confiables para proyectar las necesidades de servicios relacionadas con PCR, camas hospitalarias y camas UCI.

Abstract

Objective: To determine the installed capacity of services necessary to attend the COVID-19 pandemic, considering statistical and mathematical models as the main predictor.

Methods: The article describes four phases for modeling the number of COVID-19 cases, use of historical information on the phenomenon, non-linear statistical modeling, and mathematical modeling. For the latter, the Gompertz function was used as a model with which the projections of confirmed cases were made. The main data source was the data provided by the National Institute of Health - INS of Colombia. The parameters to estimate the supply of services were based on the projection of cases and the parameters established according to the INS guidelines.

Results: The models with the best fit for the different stages of the pandemic were: in the exponential, the cubic non-linear regression model with a coefficient of determination of 99.7%; for the exponential growth and stabilization stages, the Gompertz mathematical model. The combination of models allowed predicting for August 31 a cumulative 899,708 confirmed cases for Colombia with Symptom Onset Date - FIS. These predictions allowed estimating services for a target population of one million six hundred (1,600,000) inhabitants in 47,328 PCR tests for the month of August, 1,072 hospital beds and 365 beds in Intensive Care Units-ICU for that month.

Conclusions: The combination of simple and robust statistical and mathematical techniques allowed reliable estimates to project the needs of services related to CRP, hospital beds and ICU beds.

Contribución clave del estudio

| | |
|------------------------------|---|
| Objetivo | Determinar la capacidad instalada de servicios necesaria para atender la pandemia del COVID-19 considerando como principal predictor los modelos estadísticos y matemáticos |
| Diseño del estudio | Estudio cuantitativo descriptivo |
| Fuente de información | La fuente de datos principal fue suministrada por el Instituto Nacional de Salud de Colombia. Los parámetros para estimar la oferta de servicios se basaron en la proyección de casos y los parámetros establecidos según los lineamientos del INS |
| Población / muestra | Población de Colombia para la predicción de casos Covid; la población objeto para la estimación de pruebas PCR y camas hospitalarias y en UCI para presentar resultados corresponde a la población afiliada a una Empresa Administradora de Planes de Beneficio de cobertura nacional en Colombia. Se entiende como muestra los casos observados desde inicio de la Pandemia en Colombia hasta julio de 2020 |
| Análisis estadísticos | Combinación de modelos estadísticos y matemáticos para la proyección de casos Covid. La estimación de pruebas PCR y camas hospitalarias se hizo con los resultados de los datos proyectados y algunos parámetros establecidos por el gobierno de Colombia. |
| Principales hallazgos | Los modelos estadísticos de regresión no lineal y el cubico presentaron buen ajuste para las predicciones en los dos primeros meses de la pandemia. En las fases de crecimiento exponencial el modelo que mejor ajustó fue Gompertz. Se espera que el punto de inflexión, máxima incidencia diaria, se dé en el mes de noviembre de 2020. Los casos esperados para prueba PCR se hicieron con el promedio histórico de la población objetivo, que puede ser una EAPB, con respecto a las encuestas de sospechosos al Covid-19. Al número de casos presentados se les aplicó fórmula |

INTRODUCCIÓN

El virus SARS-COV2 quien produce la enfermedad COVID-19 generó una crisis hospitalaria a nivel mundial. Causó una emergencia mundial que obligó a los gobiernos a declarar la cuarentena y el distanciamiento físico. El 4 de julio de 2020 se registraron 11,267,309 casos a nivel mundial con un total de muertes acumuladas de 530,754 (1), con un balance positivo entre los recuperados y los que fallecen (2). En Colombia, aunque se logró pasar de la fase de contención a la de mitigación (3), se continúa con un crecimiento exponencial, al día 4 de julio se reportaron 113,389 casos y 3,942 fallecidos por esta enfermedad. Estas cifras están condicionadas a la disponibilidad y realización de pruebas PCR (4).

Uno de los retos a enfrentar ante una pandemia es estimar a futuro, los servicios (oferta) que deriva una nueva enfermedad, con la dificultad que se conoce poco sobre la historia natural de la enfermedad y día a día se construye con la nueva evidencia que surge de la investigación, la clínica y experiencia. Para la estimación, lo más contraproducente es buscar modelos estadísticos que se ajusten con mayor precisión, con la poca información de datos que se dispone, por ej en COVID-19. Una alternativa es identificar la combinación de estrategias y métodos con modelos sencillos y estructurales que puedan apoyar la toma de decisiones para aumentar la oferta; existen modelos epidemiológicos, estadísticos y matemáticos que son usados para tal propósito, no obstante en este artículo se tienen en cuenta los dos últimos que son de fácil aplicación (5-8).

Los modelos epidemiológicos que describen un proceso epidémico básicamente consideran cuatro etapas o fases: 1) **Crecimiento lineal**, etapa en donde se incluyen los primeros casos, primordialmente importados, pero su crecimiento es muy lento, con una marcada irregularidad en la incidencia de casos diarios; las siguientes fases (9) comprende las otras tres etapas: 2) **crecimiento exponencial**, comportamiento de forma exponencial hasta llegar al punto de inflexión. Este punto de inflexión corresponde al tiempo “t” en el que se alcanzaría el máximo número de casos confirmados diarios o máxima incidencia diaria registrada. En esta fase el crecimiento de la población con el evento no se ve afectada por ningún factor externo a la misma dinámica que rodea el fenómeno; 3) **fase de interacción con el medio**, en la que el crecimiento poblacional del evento se afecta por los factores que intervienen en el medio; su comportamiento se ve afectado por las diferentes intervenciones de política pública y autocuidado de la población haciendo que se desacelere su crecimiento, esta fase, también es conocida como de decrecimiento exponencial; 4) **fase de equilibrio o de estabilización**, la interacción que mantiene el crecimiento del evento en la población con el medio, continúa su desaceleración hasta alcanzar un equilibrio en el mismo.

Estos modelos ajustan muy bien la aparición de los casos en la etapa de crecimiento pero, bajo algunos supuestos como son la continuidad de una tasa de crecimiento reproductiva o R_0 , el cual es un parámetro teórico que proporciona información relacionada con la velocidad con que una enfermedad puede propagarse en una población determinada (10,11), se determina con los datos que proporciona la región para el modelo, esto es, que no recoge el

efecto de las intervenciones gubernamentales que pueda tener hacia adelante como las cuarentenas o medidas de aislamiento físico para frenar la tasa de crecimiento o de contagio como la decretada por el gobierno colombiano (12). Además, este tipo de modelo es bueno bajo ciertos supuestos en la etapa de crecimiento, tiene la debilidad de que no tienen una asíntota. En otras palabras, es que la curva no presenta un tope en el que se presenta un punto máximo.

Entre los modelos matemáticos usados en el modelamiento de la pandemia del COVID 19, se ha considerado la función Gompertz, la cual pertenece a la familia de las curvas sigmoidales. Inicialmente en su comportamiento son cóncavas y tras pasar el punto de inflexión, su comportamiento es convexo, por lo que son también conocidas como curvas en forma de “S” (13,14). Esta función tiene la propiedad de tener una asíntota, es decir que en algún momento se obtiene el máximo de crecimiento y la curva empieza a estabilizarse; se pueden identificar en su recorrido, tres fases: a) el crecimiento exponencial, b) la interacción con el medio y c) el equilibrio. Todas las fases se encuadran dentro los procesos epidémicos de una pandemia (9).

El propósito de este artículo fue presentar una combinación de técnicas estadísticas y matemáticas para proyectar los casos confirmados de COVID 19 teniendo en cuenta el comportamiento en las diferentes fases de la pandemia y una propuesta de estimación de pruebas PCR y camas hospitalarias necesarias a una fecha determinada, usando los resultados de los datos proyectados y algunos parámetros establecidos por el gobierno de Colombia.

MATERIALES Y MÉTODOS

Tipo de estudio

Corresponde a un estudio cuantitativo descriptivo sobre la población de Colombia, para la predicción de casos Covid y para la estimación de pruebas PCR, camas hospitalarias y en UCI que se necesitan para cubrir la posible demanda. La población objeto para presentar las proyecciones correspondió a la población afiliada a una empresa administradora de planes de beneficio de cobertura nacional en Colombia. La empresa administradora de planes de beneficio tiene 1,600,000 usuarios.

Variables

La variable de interés correspondieron al registro de casos confirmados por COVID-19 durante el período marzo a julio de 2020. Para la proyección de los casos se consideró como variable de tiempo la fecha de inicio de los síntomas, las tasas de crecimiento y la clasificación de casos. Los datos se obtuvieron de la data suministrada en la pagina web del Instituto Nacional de Salud y del registro del Sistema de Vigilancia en Salud Pública de la entidad.

Administración de datos

La fuente primaria correspondieron a la data suministrada por el Instituto Nacional de Salud. La organización de la tendencia de casos COVID 19 y las respectivas proyecciones se realizaron con base en la fecha de inicio de síntomas. Para el caso de los eventos confirmados asintomáticos, su fecha de inicio de síntomas se imputó teniendo en cuenta la mediana de retraso en número de

días transcurridos entre la fecha de inicio de síntomas y la fecha de confirmación de resultado positivo para COVID-19.

Etapas del modelamiento

Para el modelamiento se contemplaron cuatro etapas 1) Modelamiento del crecimiento lineal, realizado con base en la información disponible del comportamiento de los eventos COVID de 19 países con características culturales y demográficas similares a Colombia 2) Modelamiento del crecimiento exponencial, que tuvieron en cuenta los modelos estadístico para cuando el número de casos confirmados fue mayor a 30, 3) fase de interacción con el medio, en el que usó un modelamiento mixto considerando varios métodos estadísticos y matemáticos; 4) Proyección de casos, con base en el resultado del modelamiento de la fase 3.

Primera etapa

En la fase inicial de la pandemia de COVID 19 para Colombia, una limitante fue no contar con casos suficientes para el modelamiento del crecimiento lineal. Por lo anterior, se consideró el comportamiento de los casos COVID 19 reportados por España durante los primeros 40 días. Las razones de seleccionar a España para modelar la fase de crecimiento lineal de la pandemia para Colombia, fueron la similitud en tamaño poblacional, sus características culturales y religiosas. También se evaluaron las características de su sistema de salud en términos de su estructura hospitalaria y su capacidad de respuesta ante la pandemia.

Segunda etapa

Correspondió a la **modelación estadística**: se evaluaron varios modelos estadísticos que se ajustaron a la tendencia de los registros de casos confirmados de COVI 19 en la fase de crecimiento exponencial del proceso epidémico; entre los que se encuentran los modelos de regresión curvilíneos como el de potencia, el cúbico, y el cuadrático.

Tercera etapa

Denominada **mixta con modelos estadísticos y modelos matemáticos**. Los modelos matemáticos y estadísticos son la base para estudiar la evolución de enfermedades, epidemias y pandemias a nivel poblacional; entenderlos y saber aplicarlos es de gran utilidad para profesionales, estudiantes e investigadores de Ciencias de la Salud que deseen realizar investigaciones en esta área.

Modelos estadísticos

Con relación al modelamiento estadístico se evaluaron diferentes modelos de regresión curvilíneos. Lo primero que se realizó fue un gráfico de dispersión (Scatter). Posteriormente se examinaron los modelos con mayor ajuste a los datos; de éstos, los de potencia, el cúbico y el cuadrático fueron los que presentaron mayor ajuste. El de Potencia, proporcionó un R^2 o coeficiente de determinación superior al 0.996. Se interpreta que la altura del día de la pandemia en el territorio explicó en más del 99.6% el número de casos confirmados de COVID-19 para Colombia. Para el día 156 de la pandemia en Colombia, es decir, hasta el 31 de julio de 2020, el modelo seleccionado presentó la siguiente ecuación:

$$Y = -16621.17 + 1532.02 * X - 34.15 * X^2 + 0.253 * X^3$$

En donde Y corresponde al número de casos confirmados acumulado que se espera en el día X para Colombia.

A partir de este modelo se estimaron las proyecciones de casos para COVID-19 en Colombia hasta el día 40 de la pandemia.

Modelos matemáticos

En el modelamiento matemático se tuvo en cuenta la función Gompertz, para lo cual se usó la siguiente ecuación:

$$P(t) = k \cdot e^{-\ln\left(\frac{k}{P_0}\right) \cdot e^{-rt}}$$

Donde, $P(t)$, indica el tamaño de la población existente en un tiempo establecido (t) expresado en años, meses o días. Su comportamiento se ve influenciado por la tasa de crecimiento intrínseca (r) que refleja la interacción de las ganancias y pérdidas en la dinámica de crecimiento del evento a estudiar; y la máxima población que el ecosistema puede soportar definida como capacidad de carga (K , capacidad carga) (10). A la expresión $(\ln(k/P_0))$ se le conoce también como la constante (b) que representa el desplazamiento horizontal que determina cuántos días tardará en llegar a la estabilización. El punto de inflexión se alcanza en k/e .

Estrategia a) El primer, estimando el K de acuerdo a comentarios recogidos de expertos en el que predicen un punto de inflexión. Se observa el comportamiento hasta el día que se tiene datos reales. Con este punto de inflexión se recurre al mejor modelo de regresión curvilínea al punto de corte. Con estos parámetros el modelo estima el K , capacidad máxima de carga, es igual al acumulado en donde ocurre el punto de inflexión multiplicado por e (2,718). El segundo dato necesario es la constante de crecimiento, que es la que determina la velocidad de crecimiento en el eje de las ordenadas (Y), o lo que es lo mismo la inclinación de la curva. Este valor se puede hallar mediante un ajuste a los datos. Para lo cual se sugiere utilizar una función de Excel “Buscar objetivo para encontrar el resultado deseado mediante el ajuste de un valor de entrada” (Datos->Análisis de hipótesis->Buscar objetivo) en el que se equipara el valor estimado de Gompertz con el valor de casos observados a la fecha y se recalcula el valor “ r ” de tal manera que el valor de “ r ” se ajusta para que la curva de la función coincida con la de los datos.

Estrategia b) para estimar los parámetros de Gompertz se utilizó el método de Gauss Newton para hacer interacciones hasta encontrar los parámetros que mejor se ajusten.

Se inicia dando una semilla de valores para los parámetros: K , de r y de b . En Stata el comando se denomina NL, nl = Non Linear Least Squares – Iteraciones Gauss.Newton.

Con este comando se estiman los valores de los parámetros referidos anteriormente (como ilustración se usan los mismos valores del método anterior, para el valor de K se puede recurrir para su estimación, a modelos de regresión no lineal):

$$nl(\text{Casos_FIS} = \{k\} * \exp(\{b\} * \exp(\{r\} * t));$$

en donde $\text{Casos_FIS} = Y = \text{Casos acumulados por Covid-19 según fecha de inicio de síntomas (FIS)}$ y $t = \text{día de la pandemia}$.

Cuarta etapa

En esta etapa se realizó nuevamente la proyección con los datos conseguidos hasta el punto de inflexión o unos quince días después. Esto depende de la velocidad con la que se viene propagando la epidemia y va acorde al proceso de retardamiento realizado a través de las medidas gubernamentales como la cuarentena y las restricciones que tiene para lograr el distanciamiento físico. Para Colombia, se espera que el punto de inflexión se dé en el mes de noviembre de 2020. El proceso de estimación de casos COVID-19 sería similar al descrito en la fase tres.

Estimaciones de las necesidades de los servicios de Salud por COVID-19

Pruebas PCR y utilización de camas hospitalarias (proyecciones).

Una vez se estimaron los casos esperados de enfermos para la entidad o población objetivo, se procedió a estimar algunos de los servicios a utilizar para atender los enfermos por COVID-19. Para lo anterior, se usaron fórmulas matemáticas en función de los casos proyectados por COVID-19, teniendo en cuenta los parámetros establecidos por el gobierno colombiano en el decreto 417 de 2020.

Estimaciones de las pruebas PCR

Los casos esperados para prueba PCR se hicieron con el promedio histórico de la población objetivo, que puede ser una EAPB, con respecto a las encuestas de sospechosos al Covid-19. Al número de casos presentados se les aplicó una fórmula interpretativa de los autores para el último lineamiento del INS para pruebas del Ministerio de mayo de 2020 encontrando la proporción de pruebas realizadas por casos confirmados a esa fecha.

El cálculo de las PCR se hizo con la siguiente concepto: sospechosos para prueba PCR (Los captados mediante alguna herramienta para captar solicitudes de usuarios) + los evento reportados en SIVIGILA (evento 346, para IRAG e IRA por virus nuevo) + fallecidos esperados + falsos negativos + segunda prueba de los casos + tercera prueba para los positivos que salieron positivos en la segunda.

La estimación de pruebas PCR para diagnosticar el COVID-19 se realizó considerando los lineamientos del Ministerio de salud del 26 de marzo de 2020 (16):

$\text{PCR} = \text{Sospechosos para COVID} + \text{Contactos sintomáticos} + \text{falsos negativos} + \text{segunda prueba de los casos} + \text{tercera prueba para los positivos que salieron positivos en la segunda prueba} + \text{fallecidos}$.

Con base en lo anterior y en la información disponible a la fecha, los autores proponen la siguiente fórmula:

$$\text{PCR} = 2.6 * \text{casos COVID} + 3 * \text{casos COVID} * 50.9\% + (\text{Sospechosos por encuestas y por SIVIGILA} - \text{casos COVID}) (1 - \text{sensibilidad del PCR}, 0.886) * 18.5\% + \text{segunda prueba de los casos (14 días después de la primera toma de muestra)} + \text{tercera prueba para los positivos que salieron positivos en la segunda prueba (5\% de los de la primera toma a los 21 días)} + \text{fallecidos}.$$

Donde, el coeficiente de 2.6 es una equivalencia de “sospechosos para la toma de PCR” de acuerdo al comportamiento histórico de la población objeto con respecto a los casos COVID; los sospechosos para la toma PCR son los reportados por encuesta y por síntoma, nexo epidemiológico y otras variables demográficas y de comorbilidades de base, son clasificados como “sospechosos para COVID-19” y se les ordena la prueba; el coeficiente de 3, para los contactos estrechos, se refiere al número promedio por familia que es contacto estrecho a un caso confirmado por COVID-19 de los cuales, por datos históricos, se estima que el 50.9% sean sintomáticos.

Estimación de camas hospitalarias

Se calculó la relación de servicios hospitalarios siguiendo los parámetros definidos por el decreto 417 de 2020 en el que se declaró el estado de emergencia económico, social y ecológico y cuyas proyecciones se basaron en cifras entregado por el INS (ambulatorios 81.5%, moderados el 13.8% y críticos el 4.7%) y comparativamente se presentaron las cifras de acuerdo a los parámetros suministrados por el INS el 3 de julio de 2020. Los parámetros definidos por el decreto 417 son muy similares a los parámetros establecidos por la Imperial College (3) con la diferencia que están clasificados por grupo de edad en hospitalización y UCI siendo para el grupo de 0 a 9 años 0.1% y 5% respectivamente, de 30 a 39 años 3.2% y 5.0%, 80+ 27.3% para hospitalización y 70.9% para UCI.

Estos parámetros se han ido reformulando en la medida que se ha tenido mayor información (15) para el país; para el 5 de julio el INS reportó en su página 90.4% (leves + asintomáticos) de los enfermos con cuidado en casa, el 7.3% en servicio hospitalario y el 2.5% en UCI (graves + la mitad de los fallecidos), sin embargo, para la estimación de camas se adoptó los parámetros establecidos en el decreto 417 de 2020.

RESULTADOS

Modelando casos COVID-19 con técnicas estadísticas y matemáticas

Etap 1. Modelamiento del crecimiento lineal de los casos confirmados. Los casos presentados para España en los primeros días de la Pandemia fueron utilizados para Colombia homologado para Colombia el día 1 (6/03/2020) con el día 1 de España (1/02/2020). En la Tabla 1 se observa que en los primeros treinta días, el número de casos confirmados de COVID-19 para España, utilizados para las estimaciones empíricas de Colombia, están muy por debajo de lo que realmente se observó en Colombia, no obstante, en el día 40 son equivalentes, pero a partir del día 45 las estimaciones de casos para España superan significativamente a las de Colombia.

Etap 2. Modelamiento del crecimiento exponencial de los casos confirmados. Los modelos de regresión no lineal evaluados presentaron buen ajuste para las predicciones en los dos primeros meses de la pandemia, periodo en el cual aún los casos confirmados presentan un comportamiento exponencial. Los resultados de los modelos regresión curvilíneos: cuadrático, cúbico y de potencia con mayor ajuste a los datos con corte al 25 de julio (día 142) fueron:

Tabla 1. Casos utilizados para Colombia en los primeros 40 días de la pandemia ajustados con los casos de España.

| Colombia | | España | | |
|----------|------------|------------------|------------|------------------|
| Día | Fecha | Casos acumulados | Fecha | Casos acumulados |
| 1 | 6/03/2020 | 1 | 1/02/2020 | 1 |
| 15 | 20/03/2020 | 128 | 15/02/2020 | 2 |
| 30 | 4/04/2020 | 1,406 | 1/03/2020 | 84 |
| 40 | 14/04/2020 | 2,979 | 11/03/2020 | 2,277 |
| 45 | 19/04/2020 | 3,792 | 16/03/2020 | 9,942 |

Fuente: Universidad John Hopkins, www.shorturl.at/afwyS

Cuadrático:

$$Y = 28,332.66 - 1,751.09 * X + 20,969 * X^2; \text{ con un } R^2 \text{ de } 0.964$$

Cúbico:

$$Y = -16,621.17 + 1,532.02 * X - 34.15 * X^2 + 0.25 * X^3; \text{ con un } R^2 \text{ de } 0.997$$

Potencia:

$$Y = \text{EXP}(\ln(0.022) + (3.196 * \ln(X))); \text{ con un } R^2 \text{ de } 0.993$$

En donde Y predice el número de casos a ocurrir en el día X.

En la Figura 1 se observan los casos confirmados de COVID-19 observados en el período referido identificados con líneas punteadas y la proyección de casos confirmados con los tres modelos curvilíneos de mayor ajuste al 31 de agosto. Se observa que los tres modelos tienen crecimiento exponencial sin una

asíntota, no obstante, el modelo que al parecer presenta un mejor ajuste es el cúbico.

Etapa 3. Modelamiento mixto de métodos estadísticos y matemáticos o de interacción con el medio para identificar los parámetros de la función Gompertz.

a) Estimación de parámetros para la función Gompertz usando modelos curvilíneos.

Valor k o capacidad de carga: El modelo de regresión curvilínea de mayor ajuste al punto de corte fue el cúbico con los siguientes parámetros $Y = -16,621.17 + 1,532.02 * X - 34.15 * X^2 + 0.25 * X^3$; con un R^2 de 0.997; este modelo predice al 20 de noviembre ($X = 268$) un acumulado de 2,811,413 casos positivos para Colombia con el que se estima el valor de $K = e * 2,811,413 = 7,642,213$.

Estimación de la constante de crecimiento r: es la que determina la velocidad de crecimiento en el eje de las ordenadas (Y), o lo que es lo mismo la inclinación de la curva. Este valor de $r = 0.0092814$ se halló mediante la herramienta de Excel, explicada en la metodología, en el que se equiparan los casos acumulados observados al 31 de julio (434,161) contra el dato a esa fecha que ofrece la función Gompertz, a la que previamente se le ha alimentado con parámetros K, Po y r ficticio, con el objetivo de obtener el r definitivo.

Estimación P_0 : P_0 en la serie de datos representó un valor de 27 casos acumulados después de haber estabilidad en la aparición de casos, ocurrió en el octavo día de FIS, en el que se observa un comportamiento regularizado de casos nuevos para Colombia en dicho momento.

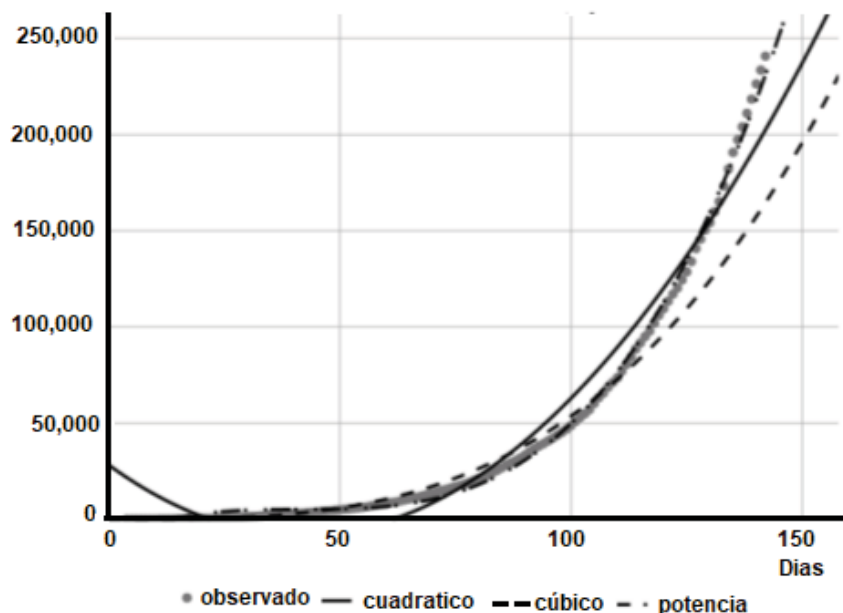


Figura 1. Comparativo de tres modelos de regresión no lineal para predecir hasta el día 25 de julio, día 142 de la pandemia para Colombia. Nota: Se muestran los datos observados hasta el 25 de julio para validar las proyecciones con los tres modelos propuestos en esta fase.

Esto da la siguiente ecuación:

$$P(t) = 7,610,961 * \text{EXP}((- \text{LN}(7,610,961/27)) * \text{EXP}((-0.01485) * t)).$$

Al simplificar la expresión $\text{LN}(7,610,961/27) = b$ (parámetro de desplazamiento horizontal). La ecuación se expresa de la siguiente manera:

$$P(t) = 7,610,961 * \text{EXP}((-12.57145)) * \text{EXP}((-0.0092814) * t),$$

b) Estimación de parámetros para la función Gompertz usando la función o método de Gauss Newton.

Valor de k: Se logró estimar estos valores (aquí se utiliza como ilustración los mismos valores del método anterior pero que no necesariamente deben ser estos, al menos para el K si se debería recurrir a hacer estimaciones según modelos de regresión no lineal):

$$\ln(\text{casos FIS} = \{k=6,000,000\} * \exp(\{b=-12,899\} * \exp(\{r=-0.0092814\} * \text{Dia})))$$

En donde casos FIS= casos COVID-19 según fecha de inicio de síntoma y día= t o día de altura de los casos COVID-19. Con la semilla inicial se re-calcularon los parámetros para K, b y r. Los gráficos de datos observados y los proyectados mediante la función de Gompertz se pueden visualizar en la Figura 2.

Los modelos utilizados en la etapa tres y cuatro tienen los parámetros resumen del modelo según Tabla 2.

Etapa 4. Proyección de casos, con base en el modelo seleccionado. En esta etapa se realiza la proyección con los datos obtenidos hasta el punto de inflexión o quince días después. Este proceso es similar al descrito en la fase tres.

Tabla 2. Parámetros de resumen de los diferentes modelos

| Resumen del modelo | | | |
|--------------------|----------------|-----------|-----------|
| Modelo | R ² | F | Res. dev. |
| Cuadrático | 0,9640 | 1806,257 | |
| Cúbico | 0,9970 | 16445,791 | |
| Potencia | 0,9930 | 20878,238 | |
| Gompertz | 0,9996 | | 2053,255 |

Con los datos proyectados según FIS observados hasta el punto de inflexión (línea roja, Figura 2), se procedió a predecir para Colombia (Figura 2), los siguientes 120 días (hasta el 30 noviembre), periodo en el cual se proyectó encontrar el punto de inflexión. Con los datos predichos para COVID-19 por el modelo Gompertz se ubica la fecha en la que se alcanzó la incidencia diaria máxima de casos o pico de la epidemia, en la Figura 2 se representa con la línea roja vertical a la altura del 22 de noviembre. Este punto se estimó bajo el supuesto que la epidemia sigue creciendo al ritmo del mes de julio e inicio de agosto, esto es con un número de reproducción efectivo, R_t de 1.15 según estimaciones emitidas por el Instituto Nacional de Salud de Colombia para el 25 de julio (17) lo que significa un distanciamiento físico al 50% o una persona contagia a 1.15 personas en promedio en 5 días. Para la función Gompertz se utilizó el r y se sugiere un “r” como semilla de 0.0092814. Con las anteriores consideraciones Gompertz proyectó para el 31 de agosto 899,708 confirmados.

Estimación de PCR y camas hospitalarias

Estimaciones de PCR

De acuerdo a la fórmula para PCR y los datos predichos al 31 de agosto= $2.6 * \text{casos COVID} + 3 * \text{casos COVID} * 50.9\% + (\text{Sospechosos por encuestas y por SIVIGILA} - \text{Casos COVID})$

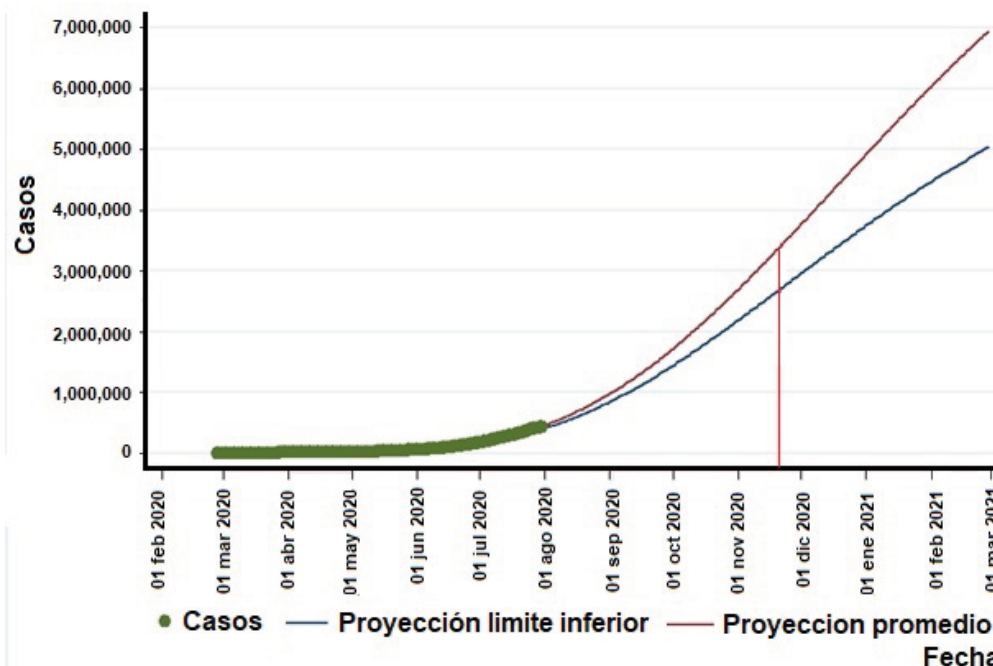


Figura 2. Número de casos COVID-19 para Colombia observados hasta el 31 de julio y proyectados hasta el 28 de febrero de 2021

Tabla 3. Estimaciones del número de PCR necesarios para cubrir la posible demanda.

| Periodo | EAPB | | INS | |
|---------------|-----------------|---------------------------|--------------------------|-------------------------------------|
| | PCR Proyectados | PCR anterior (68 x 1,000) | Lineamiento (68 x 1,000) | Lineamiento de agosto (115 x 1,000) |
| Junio | 10,061 | | 3,537 | 10,475 |
| Julio | 18,755 | | 6,592 | 19,526 |
| Agosto | 47,328 | | 16,636 | 49,273 |
| Septiembre | 52,863 | | 18,581 | 55,036 |
| Octubre | 61,298 | | 21,547 | 63,817 |
| Noviembre | 60,041 | | 21,104 | 62,509 |
| Diciembre | 58,098 | | 20,422 | 60,486 |
| Total 7 meses | 308,445 | | 108,419 | 321,122 |

EAPB: empresa administradora de planes de beneficio, INS: Instituto Nacional de Salud

* La Meta del INS era de 68 pruebas por cada mil afiliados

**La Meta del INS es de 200,000 pruebas por cada millón de afiliados

(1-sensibilidad del PCR= 0.886)*18.5% + segunda prueba de los casos (14 días después de la primera toma de muestra) + tercera prueba para los positivos que salieron positivos en la segunda prueba (5% de los de la primera toma a los 21 días) + fallecidos.

| | |
|--|-------------------|
| Casos COVID-19 acumulados= | 16,309 |
| Sospechosos por encuestas y por SIVIGILA = | 321,987 |
| Sensibilidad del PCR = | 0.886 |
| Segunda prueba de los casos = | 12,738 |
| Tercera prueba para los positivos = | 538 |
| Fallecidos acumulados= casos covid*letalidad para población objetivo al 31 de julio= | 16,309*3.1% = 506 |

Los PCR acumulados hasta el 31 agosto serían:

$$2.6*16,309 + 3*16,309*50.9\% + (321,987 - 16,309) * (1-0,886) * 18.5\% + 12,738 + 506 = 87,536;$$

a la anterior cifra acumulada de PCR se restan los PCR acumulados hasta Julio 31 de 40,208 lo que daría un neto de PCR para el mes de agosto de 47,328 pruebas. En la Tabla 3, se puede observar la proyección de PCR por mes según sea la fórmula anterior, aplicable a una empresa administradora de planes de beneficio con una población igual a la objetivo, o las pruebas estimadas por el INS como meta establecida en dos momentos; la primera en mayo para realizar 68 pruebas de PCR x 1,000 afiliados en un periodo de 7 meses y la segunda, ajustando la anterior, establecida en agosto para realizar 115 pruebas x 1,000 afiliados en un periodo de 5 meses.

Estimaciones de camas

En la Tabla 4 la proyección de camas de hospitalización y de UCI para el mes de agosto según el cálculo provenga del Decreto 417 o del INS del 3 de julio, cuya diferencia porcentual radica en que los parámetros del Decreto, tiene porcentajes hospitalarios mayores al de los parámetros del INS porque al inicio de la pandemia se establecen parámetros sin medidas de aislamiento. Se estima los siguientes valores de cama hospitalaria o de UCI multiplicando los casos de COVID-19 esperados para el mes por el parámetro, representado por el porcentaje, según provenga del Decreto o del INS así: por ejemplo, considerando que se espera, para la población objeto, 7,767 casos incidentes durante el mes de agosto las camas hospitalarias según el Decreto 417 serían de $7,767*13.8\% = 1,072$.

DISCUSIÓN

¿Porque preferiblemente la función Gompertz?. Al comparar la eficiencia de los diferentes grupos de modelos en las fases evaluadas (20), entre ellos los modelos estadísticos de máximo verosimilitud y las funciones matemáticas como Gompertz, se identificaron pocas diferencias entre los resultados del indicador de bondad de ajuste con un R^2 superiores a 0.9. El modelo con mejor ajuste terminó siendo el GOMPERTZ con un R^2 de 0.9996, resultados muy similares a los presentados por Villalobos en las proyecciones realizadas para Italia y España (5).

No obstante la función Gompertz tiene la característica de poder representar el comportamiento de la pandemia en las cuatro fases de la epidemia. Es prudente preciar que dicho modelo

Tabla 4. Estimaciones de servicios hospitalarios necesarios para cubrir la posible demanda

| Clasificación del caso | % según decreto 417 de 2020 | Proyectado EAPB para el mes de agosto (criterio Decreto 417) | % según parámetros INS del 6 de junio | Proyectado EAPB para el mes de agosto (criterio INS 3 de julio) |
|-----------------------------|-----------------------------|--|---------------------------------------|---|
| Ambulatorio | 81.5 | 6,330 | 87.4 | 6,785 |
| Urgencias - Hospitalización | 13.8 | 1,072 | 11.2 | 872 |
| Críticos-UCI | 4.7 | 365 | 1.4 | 111 |
| Casos esperados en agosto | | 7,767 | | |

puede presentar limitaciones al no explicar necesariamente la heterogeneidad de varios factores que puedan influir en la predicción como la estructura poblacional (habrá zonas que tienen un mayor índice de envejecimiento que otras), o regiones con mayor riesgo en su población con enfermedades de base para infectarse por el COVID-19, ni tampoco, las variaciones en las tendencias relacionadas con las intervenciones relacionadas con el aislamiento social, el confinamiento y la movilidad; de ahí la necesidad de combinar varios modelos y técnicas para la comprensión del fenómeno (6, 8, 10, 13).

Hay otros modelos de comportamiento, como los SIR, que son muy eficientes en la predicción en la fase de crecimiento exponencial pero después de esta fase pierden la precisión.

Hay muchos factores que pueden explicar la precisión que puedan tener los pronósticos para los casos de COVID-19, cuando se parte desde la proyección de un territorio mayor, acogiendo a la proporcionalidad de su población:

1. La densidad poblacional y la participación que se tiene en la sub-zona frente a la población del territorio mayor de donde originalmente se hacen las proyecciones. El comportamiento del Covid-19 para Colombia está siendo “jalonado” por pocas ciudades como Bogotá quien es la ciudad con mayor densidad poblacional por ser la capital de Colombia. Para el 24 de junio se registraron 23,367 casos que representaron el 30.3% del total de los casos para Colombia y una tasa de contagio de 279 casos x 100,00 muy por encima a la de Colombia que fue de 156 x 100,000 habitantes. Hay otras ciudades como Barranquilla con una tasa de contagio de 789 casos x 100,000 habitantes y Cali, con una tasa de contagio de 246 casos x 100,000 habitantes.

2. El comportamiento social de la ciudadanía juega un papel importante en el número de casos que se podrían presentar en una Región o una ciudad.

3. La estructura poblacional se conjuga con las dos anteriores; el COVID-19 que es una enfermedad con alto riesgo para los mayores de 65 años debería tener una mayor incidencia de casos graves en regiones en las que se tiene una población con mayor envejecimiento, como es el caso del Eje Cafetero con un índice de vejez alto, pero afortunadamente con un buen comportamiento social hasta la fecha. No es el caso para la ciudad de Cali, con un índice de envejecimiento superior al 60% donde se han presentado eventos sociales no permitidos.

En las estimaciones realizadas para una empresa administradora de planes de beneficio con el algoritmo interpretativo estimado por los autores para la proyección de las pruebas PCR necesarias, basada en las proyecciones de caso y en últimos lineamientos sobre muestras PCR a tomar por el INS (21), se llegó a un estimativo de una tasa de 95 x 1,000 afiliados. Las cifras son más altas a las metas fijadas por el INS de 68 por mil.

¿Para qué realizar estimaciones en Servicios de Salud?

Las estimaciones de frecuencias de casos COVID-19 y servicios en salud son fundamentales para la toma de decisiones y la

planeación para incrementar el control y la posibilidad para responder eficientemente a lo que no se pueda controlar (18), teniendo herramientas estadísticas y epidemiológicas para la toma de decisiones en tiempo real. Lo anterior debe basarse en hechos y datos obtenidos de la realidad de un mundo cambiante y complejo, cuyo fin principal es el mejoramiento de las condiciones de vida de la población, con procesos costo-efectivos que permitan obtener los mejores resultados en salud.

El sistema nacional de salud requiere herramientas esenciales que la epidemiología y la estadística, ofrecen, a fin de tomar decisiones que resulten apropiadas para un buen diseño de políticas de salud basadas en evidencias epidemiológicas. Para todos los actores del Sistema, es vital el conocimiento del comportamiento histórico de los eventos o los comparadores que permitan acercar al contexto actual la situación en salud de la población, para con ello planear eficientemente los recursos finitos y proyectar estabilizaciones o desviaciones que impacten no solo en los costos de la atención, sino en elementos como la calidad en los servicios, la oportunidad y la accesibilidad en épocas tan difíciles como la actual, donde el mundo debe prepararse para enfrentar retos cada vez más complejos y cambiantes.

Las pautas para considerarse en la toma de decisiones basadas en proyecciones de servicios y costos parten de:

- Conocimiento del comportamiento histórico del Sistema de Salud, la Organización y la población.
- Caracterización de la población con el fin de determinar las prioridades, el impacto en resultados en salud y el mejoramiento de estrategias poco costo-efectivas usadas en el pasado.
- Los recursos disponibles en términos financieros, personal, capacidad instalada, condiciones del entorno y control a riesgos detectados.
- Comparadores nacionales e internacionales cuya finalidad no es más que aprender de los aciertos y desaciertos del entorno, así como adaptar las propias estrategias para el logro de los objetivos.
- Finalmente, las acciones individuales y colectivas que permitan a todos los actores del sistema aportar a los mejores resultados en salud costo-efectivos y sostenibilidad de los mismos en el tiempo.

Con respecto a la Pandemia por COVID-19, las estimaciones son útiles para determinar las medidas que se deben implementar para controlar o mitigar el impacto del evento. Las medidas de control como la cuarentena, las restricciones de viaje y la inspección en el aeropuerto para los viajeros han sido ampliamente implementadas para contener la propagación de infecciones. Sin embargo, la efectividad de estas contenciones para controlar el brote no es concluyente en la evidencia actual de modelado. Los estudios sobre COVID-19 sugieren que las restricciones de viaje que conducen a una transmisibilidad reducida pueden ser altamente eficaces en contener la extensión, mientras que el cierre de la escuela es menos efectivo que el distanciamiento del lugar de trabajo o cuarentena de individuos expuestos; una estrategia

combinada que implementa las tres medidas era más efectiva para reducir la propagación (19).

REFERENCIAS

1. CSSEGISandData / COVID-19. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. [citado 4 de julio de 2020]. Disponible en: <https://github.com/CSSEGISandData/COVID-19/>
2. González-Jaramillo V, González-Jaramillo N, Gómez-Restrepo C, Palacio-Acosta CA, Gómez-López A, Franco OH. Impact of the COVID-19 pandemic on the Colombian population according to mitigation measures. Preliminary data from epidemiological models for the period March 18 to April 18, 2020. *Rev Salud Publica*. 2020; 22(1): e185789. Doi: 10.15446/rsap.v22.85789
3. MRC Centre for Global Infectious Disease Analysis. COVID-19 reports. Faculty of Medicine, Imperial College: London; 2020. Available from: <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/>
4. Plasencia TA. COVID-19: por qué los países deben informar sobre el número de pruebas realizadas. *The Conversation*; 2020; Available from: <https://theconversation.com/covid-19-por-que-los-paises-deben-informar-sobre-el-numero-de-pruebas-realizadas-133307>
5. Villalobos-Arias M. Estimation of population infected by Covid-19 using regression Generalized logistics and optimization heuristics. Cornell University; 2020; Available from: <http://arxiv.org/abs/2004.01207>
6. Damian VM. Simulaciones con el modelo SEIR del comportamiento epidémico del COVID-19 en el departamento de Guairá-Paraguay. Universidad Católica Nuestra Señora de la Asunción. Paraguay; 2020. Available from: https://www.researchgate.net/publication/340136048_Simulaciones_con_el_modelo_SEIR_del_comportamiento_epidemico_del_COVID-19_en_el_departamento_de_Guaira-Paraguay.
7. Gómez EA, Rosendo MJA, Gonzalez CMA. Modelado y análisis de la evolución de una epidemia vírica mediante filtros de Kalman: el caso del COVID-19 en España. Universidad de Sevilla: España; 2020. Disponible en: <https://idus.us.es/handle/11441/94508>
8. Sanz I. Modelos epidemiológicos basados en ecuaciones diferenciales (Trabajo de Grado). Univ Rioja. 2016; 67. Available from: https://biblioteca.unirioja.es/tfe_e/TFE002211.pdf
9. Rodríguez J, Ulloa J. Alternativa didáctica para el estudio del modelo Gompertz. *Investig Innov Matemática Educ*. 2017; 2: 98-114.
10. Ridenhour B, Kowalik JM, Shay DK. El número reproductivo básico (R0): consideraciones para su aplicación en la salud pública. *Am J Public Health*. 2018; 108(Suppl 6): S455-65.
11. Navarro-Robles E, Martínez-Matsushita L, López-Molina R, Fritz-Hernández J, Flores-Aldana BA, Mendoza-Pérez JC. Modelo para estimación del comportamiento epidémico de la influenza A (H1N1) en México. *Rev Panam Salud Publica*. 2012; 31(4): 269-74
12. González-Jaramillo V, González-Jaramillo N, Gómez-Restrepo C, Franco OH, Palacio-Acosta CA, Gómez-López A. Proyecciones de impacto de la pandemia COVID-19 en la población colombiana, según medidas de mitigación. Datos preliminares de modelos epidemiológicos para el periodo del 18 de marzo al 18 de abril de 2020. *Rev Salud Pública*. 2020; 22(2): 1-6.
13. Mínguez FS. Procesos de difusión Logístico y Gompertz. Métodos numéricos clásicos en la estimación paramétrica. Master en Estadística Aplicada Departamento de Estadística e Investigación Operativa, Universidad de Granada: Granada, España. 2016. Disponible en: [https://masteres.ugr.es/moea/pages/curso201516/tfm1516/simon_minguez_tfm/!](https://masteres.ugr.es/moea/pages/curso201516/tfm1516/simon_minguez_tfm/)
14. Martínez CA, Rodríguez AP, Jiménez A, Manrique C. Descripción Matemática De La Función Gompertz Aplicada Al Crecimiento De Animales. *Rev la Fac Med Vet y Zootec*. 2010; 57(1): 76-80.
15. Osejo C. Pronóstico de utilización total de la capacidad instalada combinada en los hospitales de la ciudad de Pasto, Colombia, con la epidemia del COVID-19, a marzo 17 de 2020. *Rev Sigma*. 2020; 14:14-24.
16. Presidencia de la República. Decreto 417. Por el cual se declara un Estado de Emergencia Económica, Social y Ecológica en todo el Territorio Nacional. Bogotá, Colombia; 2020. Disponible en: <https://dapre.presidencia.gov.co/normativa/normativa/DECRETO%20417%20DEL%2017%20DE%20MARZO%20DE%202020.pdf>
17. Instituto Nacional de Salud. Número reproductivo efectivo Rt Nacional. Modelos de Estimación. Bogotá Colombia; 2020. Disponible en: <https://www.ins.gov.co/Direcciones/ONS/modelos-de-estimacion>.
18. Porter ME. What Is Strategy? *Harvard Business Review*. 1996; 74(6): 61-78. Disponible en: <https://www.hbs.edu/faculty/Pages/item.aspx?num=10698>
19. Park M, Cook AR, Lim JT, Sun Y, Dickens BL. A Systematic Review of COVID-19 Epidemiology Based on Current Evidence. *J Clin Med*. 2020; 9(4): 967. Doi: 10.3390/jcm9040967.
20. Yang CY, Wang J. A mathematical model for the novel coronavirus epidemic in Wuhan, China. *Math Biosci Eng MBE*. 2020; 17(3): 2708-24.
21. Ministerio de Salud y Protección Social. Lineamientos, orientaciones y protocolos para enfrentar la COVID-19 en Colombia. Ministerio de Salud y Protección Social: Bogotá; 2020. Disponible en: <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/lineamientos-orientaciones-protocolos-covid19-compressed.pdf>.

