

Uso de los modelos epidemiológicos para mejorar el pronóstico de los contagios durante una pandemia

Use of epidemiological models to improve the forecasting of infections during a pandemic

Carmen González-Velasco, Francisco-José Sáez-Trujillo y Simón-Pablo Contreras-González

Recibido 3 noviembre 2022 / Enviado para modificación 23 diciembre 2024 / Aceptado 28 diciembre 2024

RESUMEN

Objetivo Esta investigación propone una estrategia basada en el uso de los modelos epidemiológicos tipo SIR (susceptibles, infectados y recuperados) para mejorar los pronósticos de la tasa de contagios durante las primeras etapas de un brote pandémico.

Métodos Se utilizan datos de la pandemia COVID-19 en Chile durante el 2020, se comparan cinco métodos de proyección de uso frecuente en la literatura: (i) la hipótesis del paseo aleatorio, que sirve de benchmark o estimación de referencia, (ii) un modelo AutoRegressive Integrated Moving Average, (iii) la función de probabilidad Weibull, (iv) la estimación mínimo-cuadrática y (v) un modelo de vectores autorregresivo. Los dos últimos enfoques incorporan como variables explicativas las propias estimaciones de la tasa de contagio extraídas de modelos tipo SIR (en cinco de sus versiones más simples).

Resultados La evaluación de la calidad de pronósticos muestra que el uso de la información del modelo SIR reduce el error cuadrático de proyección en las estrategias MCO y VAR con respecto a las metodologías alternativas que no aprovechan la información que proviene de los modelos teóricos. Este resultado es robusto a diferentes horizontes temporales.

Conclusiones El estudio concluye que los modelos epidemiológicos, estilo SIR, contienen información relevante para pronosticar las tendencias subyacentes de la tasa de contagios en el corto plazo. Esto implica que las autoridades sanitarias, que hacen seguimiento a la evaluación de una pandemia, además de basar sus previsiones en los datos oficiales, pueden utilizar información indirecta proveniente de modelos teóricos estructurales que han sido previamente calibrados o estimados.

Palabras Clave: COVID-19; Modelos SIR; modelos epidemiológicos (*fuente: DeCS, BIREME*).

ABSTRACT

Objective This research proposes a strategy based on the use of SIR-type epidemiological models (susceptible, infected and recovered) to improve the forecasts of the infection rate during the early stages of a pandemic outbreak.

Methods Data from the COVID-19 pandemic in Chile during 2020 are used, five projection methods frequently used in the literature are compared: (i) the random walk hypothesis, which serves as a Benchmark or reference estimate, (ii)) an AutoRegressive Integrated Moving Average model, (iii) the Weibull probability function, (iv) the least-squares estimation and (v) an autoregressive vector model. The last two approaches incorporate as explanatory variables the contagion rate estimates themselves extracted from SIR-type models (in five of their simplest versions).

Results The evaluation of the quality of forecasts shows that the use of the information from the SIR model reduces the quadratic error of projection in the OLS and VAR strate-

CG: Econ. Ph. D. Ciencias Empresariales.
Departamento de Dirección y Economía de la Empresa, Facultad de Ciencias Económicas y Empresariales, Universidad de León.

León, España.
mgonzv@unileon.es

FS: Econ. Ph. D. Economía. Departamento de Dirección y Economía de la Empresa, Facultad de Ciencias Económicas y Empresariales, Universidad de León. León, España.

fsaet@unileon.es

SC: Ing. Civ. Ind. M. Sc. Ciencias Actuariales y Financieras. Universidad de León. Caja Los Andes. Santiago de Chile, Chile. simon.contreras.g@gmail.com



DOI: <https://doi.org/10.15446/rsap.V27n1.105629>

gies with respect to alternative methodologies that do not take advantage of the information that comes from the theoretical models. This result is robust to different time horizons.

Conclusions The study concludes that the epidemiological models, SIR style, contain relevant information to forecast the underlying trends in the infection rate in the short term. This implies that health authorities, which monitor the evaluation of a pandemic, in addition to basing their forecasts on official data, can use indirect information from structural theoretical models that have been previously calibrated or estimated.

Key Words: COVID-19; SIR models; epidemiological models (*source: MeSH, NLM*).

La exploración de nuevos métodos estadísticos para mejorar las previsiones durante una pandemia es una tarea permanente debido, sobre todo, al carácter endémico de algunas enfermedades, la aparición de nuevas cepas y la necesidad de definir, de manera oportuna, políticas sanitarias que reduzcan el impacto de los contagios. En este sentido, durante los últimos cien años, y desde el trabajo pionero de Kermack y McKendrick (1), el pilar fundamental para la construcción de estimaciones sobre la evolución de los brotes pandémicos ha sido el modelo SIR (2).

Diferentes versiones de estos modelos han sido adaptadas para evaluar los éxitos relativos de estrategias sanitarias y determinar el alcance de las pandemias en diversas dimensiones socioeconómicas o incluso políticas (3,4). Aunque estas extensiones teóricas del modelo SIR han mejorado significativamente la compresión sobre las variables que determinan la propagación y la evolución de las enfermedades epidémicas, sus predicciones son, en general, muy pobres (5,6).

En este marco conceptual, el número reproductivo básico (R_0), que determina el número medio de personas infectadas por un individuo infectado, resulta clave para comprender la evolución de una pandemia y el comportamiento probable de los contagios futuros. Debido a su importancia, es natural comprender por qué, desde el inicio del brote pandémico del COVID-19, la estimación de este parámetro ha desempeñado un rol central en la literatura. Estimar y predecir su valor resulta clave para que las autoridades sanitarias puedan implementar políticas públicas de prevención.

El principal objetivo de este documento es proponer una estrategia para estimar la tasa de contagio (R_0) mediante la utilización de la información que proveen los modelos SIR. La utilización de modelos teóricos (calibrados o estimados) como herramienta de medida se justifica, en parte, por los problemas de confiabilidad de los datos primarios (7,8) y por errores observacionales o de naturaleza muestral.

Así, en una primera fase, se recurre a la estimación de cinco versiones del modelo SIR con datos de la experiencia chilena durante el brote de COVID-19. En una segunda fase de la metodología, se proyecta la tasa de contagio con cuatro métodos de proyección: dos de estos son univariantes (modelo AutoRegressive Integrated Moving Average (ARIMA) y función logística Weibull) y se es-

timan usando exclusivamente datos oficiales obtenidos del Ministerio de Salud. Las otras dos aproximaciones son multivariantes: mínimo-cuadrática (MCO) y vectores autorregresivos (VAR). En este último caso se utiliza información sobre la tasa de contagio, proveniente tanto de datos oficiales como de los modelos SIR calibrados o estimados en la primera etapa.

Los modelos multivariantes generan un pronóstico de la tasa de contagios que es “corregido” o “afectado” por la especificación teórica del modelo SIR. Las proyecciones se realizan fuera del periodo de estimación y a diferentes horizontes temporales. A partir de la evaluación de la calidad de pronóstico se argumenta que, aun cuando analizados individualmente, las previsiones de los modelos SIR puedan ser relativamente pobres, de forma conjunta pueden aportar información relevante para mejorar los pronósticos de la tasa de contagio cuando se combinan con las técnicas estadísticas o econométricas estándar.

MÉTODOS

El modelo SIR

El modelo SIR se presenta como un sistema de ecuaciones en las siguientes variables relacionadas con las clases epidemiológicas de la población objeto de estudio: susceptibles (S), expuestos (E), infectados (I) y recuperados (R). Inicialmente, el modelo divide a la población de tamaño N en tres clases o estados. De acuerdo con estas variables, se observa que todas las personas que se contagian deberían inicialmente estar sanas, después se enferman y posteriormente se recuperan, limitándose así la posibilidad de que una persona perteneciente al grupo R(t) se vuelva a contagiar.

Inicialmente, el modelo SIR tradicional divide una población N fija, suficientemente grande, en tres clases diferentes: Susceptibles S(t), Infectados I(t) y Recuperados R(t), donde la evolución de cada una de ellas es interpretada por una ecuación diferencial específica. De esta forma, cuando la población es variable se debe cumplir: $N(t)=S(t)+I(t)+R(t)$.

El modelo considera que existe solo una enfermedad capaz de transmitir la infección, desestima la posibilidad de reinfección cuando una persona ya esté recuperada y

asume que toda persona de la población N tiene la misma probabilidad de contagiarse por otro individuo, y, por tanto, cada individuo susceptible tiene la misma probabilidad de contraer la enfermedad.

Estos conceptos básicos se aplican al modelo SIR de población variable, es decir, en un escenario cuando la población se modifica en el tiempo, y se definen los siguientes parámetros: la probabilidad de contagio β , interpretada como la posibilidad de que una persona de la clase $S(t)$ sea realmente contagiada y se traslade a la clase $I(t)$ en un intervalo de tiempo Δt ; el número de contactos k , que representa la cantidad de interacciones que el individuo haya tenido, y de las cuales no se tiene certeza a qué clase corresponden, y por tanto tienen la probabilidad $i(t)$ de estar infectados, que representa la fracción de personas infectadas sobre la población variable de tamaño $N(t)$; la tasa de contagio para una persona $c(t)$, que es la tasa a la que personas pertenecientes a la clase $S(t)$ se mueven a la clase $I(t)$, definida para toda la clase $S(t)$ como $c(t) = S(t) \beta i(t) k \Delta t$; y la tasa de recuperación $v(t)$, que corresponde al ritmo de integración a la clase $R(t)$ de todos aquellos individuos de la clase $I(t)$ que logren recuperarse, expresado como el producto de la cantidad de infectados $I(t)$ y la tasa de retiro γ , donde γ se interpreta como la razón en que se trasladan pacientes enfermos al grupo de recuperados. De esta manera, se obtiene que $v(t) = I(t) \gamma$. Además, se consideran las tasas de mortalidad ϕ y de natalidad η derivadas del ciclo vital del ser humano y la tasa de mortalidad λ como consecuencia del brote pandémico.

Modelo SIR: calibración

La calibración requiere estimar los valores de ϕ , η y λ mediante los datos de mortalidad y natalidad obtenidos desde el Departamento de Estadísticas e Información de Salud:

- **Tasa de mortalidad (ϕ):** en el año 2019, en Chile se produjeron 109 369 defunciones, con una tasa anual de mortalidad de 0,57% y una mortalidad diaria de $\phi=0,0016\%$.
- **Tasa de natalidad (η):** en el mismo año se registraron 210 863 nacimientos en el país, con una tasa anual de natalidad de 1,10% y una tasa diaria de natalidad de $\eta=0,00302\%$.
- **Tasa de mortalidad derivada del COVID-19 (λ):** este parámetro se estima con la información de individuos que se han contagiado, pero que ya no son infecciosos, ya sea por recuperación o por muerte. En efecto, a 10 de agosto de 2020 este número de personas era 347 342, de los cuales 10 139 habían fallecido debido al virus. Por tanto, λ se corresponde con la proporción de individuos que superaron la infección, pero que habían fallecido, es decir, $\lambda = 2,92\%$.
- Sea $N_0 = 19\,134\,898$; $S_0 = 19\,134\,897$; $I_0 = 1$; $R_0=0$; que representan los valores iniciales de la población número de susceptibles, infectados y número reproductivo básico.

Sin embargo, también es posible aproximar el número reproductivo básico teóricamente a través de la calibración del modelo SIR, tanto (i) en su versión de población fija (R_o^{PF}) como (ii) con población variable, (R_o^{PV}) (iii) la aproximación exponencial (R_o^{CE}) (9), (iv) la estimación por mínimos cuadrados ordinarios (R_o^{MC}) (véase, por ejemplo, 10,11) y (v) el uso de la ecuación de susceptibles (R_o^{ES}) (12). A continuación, se presentan brevemente las cinco versiones del modelo SIR utilizadas para aproximar R_o . Posteriormente, estos resultados se combinan con los valores de publicaciones oficiales (11) para realizar proyecciones de la tasa de contagio.

RESULTADOS

Análisis de la evolución de la tasa de contagios

La Figura 1 muestra el comportamiento aproximado de R_o para cada periodo t obtenido a partir de los modelos SIR expuestos anteriormente. Todas las estimaciones tienden hacia un valor de R_o menos volátil y próximo a uno, lo cual es coherente con la información publicada por fuentes oficiales de la situación de la pandemia en Chile, específicamente con el número efectivo de reproducción (R_e) (13,14). Este parámetro mide el número promedio de individuos que una persona infectada puede contagiar cuando una fracción de la población es susceptible, de modo que, para contrastarlo con las estimaciones obtenidas, la evolución de este parámetro debe transformarse (15) a $R_0 = R_e(t) \cdot \left(\frac{s(t-1)}{N(t)} \right)$. El número reproductivo básico que se utiliza en este trabajo se relaciona con el efectivo, según esta última formulación.

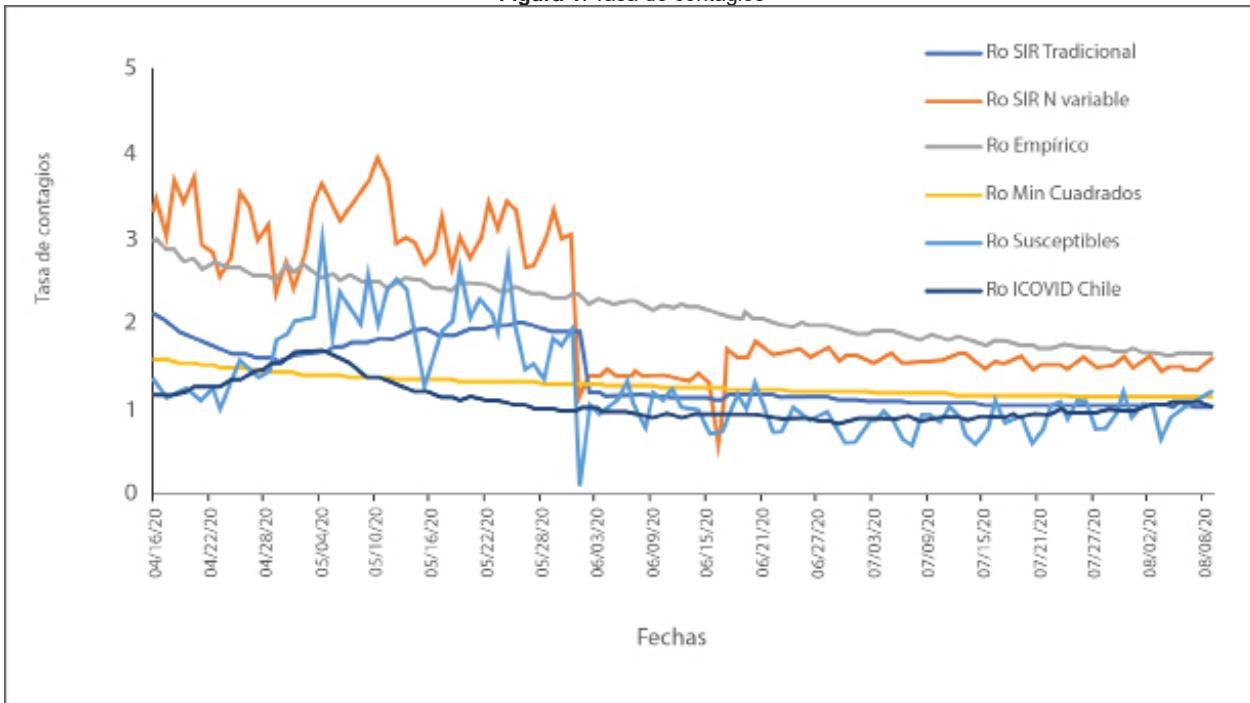
En el caso de Chile, las estimaciones oficiales del número reproductivo efectivo no están condicionadas por las especificaciones de ningún modelo teórico, sino más bien por un método puramente estadístico. La metodología oficial para el cálculo de esta variable sigue a (16) que, a su vez, se basa en las propuestas de (17-19), entre otros (Figura 1).

Modelos de pronósticos

En esta sección se presentan los resultados de los cinco modelos de pronóstico que pudieron ser utilizados al comienzo del brote pandémico, a fin de evaluar la dinámica probable del número de contagiados. El periodo de estimación se extiende desde el 16 de abril hasta el 8 de agosto de 2020 (el 3 de marzo de 2020 en Chile se diagnosticó el primer caso de COVID-19). La calidad de los pronósticos de estos enfoques se evalúa en el siguiente apartado.

Paseo aleatorio

La estimación más sencilla posible se basa en la hipótesis del paseo aleatorio. Implica que la tasa de contagios es igual a la observada, más una perturbación aleatoria que tiene

Figura 1. Tasa de contagios

Fuente: elaboración propia a partir de datos extraídos de (11).

valor esperado de cero: $R_0(t) = y_t = y_{t-1} + \varepsilon_t$ con $\varepsilon_t \sim N(0, \sigma_t)$. Es decir, el componente secular de la serie es un paseo aleatorio sin deriva, el elemento cíclico es estacionario y ambos componentes no están correlacionados. En este caso, el valor esperado de este proceso viene dado por el último dato observado, de forma tal que: $E[y_t|I(t)] = y_{t-1}$.

Modelo ARIMA

El modelo AutoRegressive Integrated Moving Average ARIMA (q, d, k) viene dado por:

$$\nabla^d y_t = \sum_{i=1}^q a_i \cdot \nabla^d y_{t-i} + \sum_{i=1}^k \beta_i \cdot \varepsilon_{t-i} + \varepsilon_t,$$

donde las diferencias de orden d para lograr un proceso estacionario vienen dadas por $\nabla^d y_t = \nabla^{d-1} y_t - \nabla^{d-1} y_{t-1}$. Una vez estimado el modelo ARIMA para y_t , es posible predecir la observación diferencial de orden d -énesimo en $t+1$ como $\nabla^d \hat{y}_{t+1}$ y recuperar posteriormente el nivel de la estimación con: $\nabla^d \hat{y}_{t+1} = \nabla^d \hat{y}_{t+1} + \sum_{i=0}^{d-1} \nabla^d \hat{y}_{t-i+1}$. Teniendo en cuenta el número reproductivo básico obtenido de las fuentes oficiales, se obtiene $d=1$. Los residuos de la regresión están normalmente distribuidos y no presentan evidencia de estructura residual. Ni la incorporación de una tendencia determinística, ni la heterocedasticidad condicional contribuyeron a la capacidad predictiva del modelo.

Aproximación por Weibull

Entre las funciones matemáticas de comportamiento sinusoidal que pueden ser utilizadas para modelar la pandemia se distinguen la función de Gompertz, la función

logística, la función log-logística y la función Weibull. Según el criterio de información de Akaike (AIC), la que mejor se ajusta a los datos de Chile es la última. Esta función de distribución fue estimada de forma no lineal mediante métodos numéricos (véase, por ejemplo, 20).

Estimación MCO

En este caso el sistema se define como: $y = X\beta + \varepsilon$, donde y es un vector $n \times 1$ de observaciones de la variable dependiente R_0 , β es un vector $k \times 1$ de parámetros desconocidos y X una matriz de n observaciones y k filas, donde $n=115$ observaciones, con cinco variables explicativas [$R_o^{ST}, R_o^{SN}, R_o^{CE}, R_o^{MC}, R_o^{ES}$], aunque la primera y la última variable (SIR tradicional de población fija y estimación de susceptibles), no se incluyen en la especificación final de la ecuación de estimación por no ser significativas, por lo que finalmente $k=3$. Es importante destacar que, al igual que en el modelo ARIMA, los retardos más significativos están entre el día 5 y el 7, lo cual es consistente con la aparición promedio de los síntomas de la enfermedad, después de estar expuesto al contagio.

Modelo VAR

En términos generales, el VAR puede ser representado como un sistema de ecuaciones: $y_t = A(\iota)y_{t-1} + e_t$, donde y_t, e_t son vectores $m \times 1$; $e_t \sim (0, \Sigma_e)$ y ι es el operador de retardos, de forma que: $A(\iota) = a_0y_t + a_1y_{t-1} + a_2y_{t-2} + \dots + a_qy_{t-q}$. El vector de variables viene dado por R_0 observado u oficial

y por el conjunto de estimaciones extraídas de los modelos SIR, es decir: $y_t = [R_o, R_o^{SN}, R_o^{CE}, R_o^{MC}]$. Al igual que en el modelo MCO, se excluyen de este enfoque el R_0 estimado por el modelo SIR tradicional (R_o^{ST}) y el R_0 estimado por la ecuación de susceptibles (R_o^{ES}) por no mostrar ninguna relación con el R_0 oficial. El número óptimo de retardos se obtuvo a partir del criterio de AIC y el orden de integración se mantuvo igual para todas las variables ($d=1$).

Errores de pronóstico

La Tabla 1 recoge los errores absolutos, cuadráticos y relativos de las diferentes propuestas de estimación. Se resaltan aquellas metodologías que, en promedio, muestran la menor discrepancia con los datos en diferentes horizontes temporales de proyección. La evaluación de la calidad del pronóstico fuera de la muestra indica que la combinación de pronósticos supera a la hipótesis del paseo aleatorio en todos los horizontes temporales considerados. Asimismo, la propuesta empírica permite mejorar medidas alternativas basadas en modelos ARIMA y la función de distribución de contagios Weibull.

Tabla 1. Errores de proyección de la tasa de contagio a diferentes horizontes temporales y usando diferentes metodologías

Paseo aleatorio	Absoluto	Cuadrático	Relativo (%)
Días			
7	0,0104	0,0001	0,0109
20	0,0319	0,0016	0,0316
30	0,0403	0,0022	0,0397
45	0,1908	0,3351	0,7021
ARIMA			
Días			
7	0,0114	0,0001	0,0120
20	0,0478	0,0034	0,0474
30	0,0601	0,0047	0,0593
45	0,0483	0,0035	0,0481
Weibull			
Días			
7	0,0230	0,0005	0,0242
20	0,0580	0,0044	0,0578
30	0,0688	0,0057	0,0681
45	0,0546	0,0042	0,0544
MCO			
Días			
7	0,0348	0,0012	0,0366
20	0,0282	0,0009	0,0287
30	0,0280	0,0009	0,0282
45	0,1885	0,3347	0,7007
VAR			
Días			
7	0,0040	0,0000	0,0042
20	0,0351	0,0021	0,0347
30	0,0467	0,0030	0,0459
45	0,1936	0,3356	0,7045

Esta tabla incluye los valores del error de proyección de la tasa de contagios en Chile, según cuatro métodos de proyección: ARIMA, Weibull, mínimos cuadrados ordinarios y vectores autorregresivos. Se añade, además, un caso de referencia, o base, donde se asume que la tasa de contagio sigue un paseo aleatorio, de forma tal que su valor esperado es igual al último valor observado. Para cada metodología, y a partir del 9 de agosto del 2020, se reporta el valor promedio del error de estimación en el horizonte temporal considerado (+7, +20, +30, +45, días). Se han resaltado en gris los modelos de proyección que muestran los menores errores de previsión. Se observa que el VAR y la estimación por mínimos cuadrados son los únicos modelos que utilizan información de los modelos SIR y muestran el mejor desempeño a corto plazo (menos de 30 días).

DISCUSIÓN

A muy corto plazo (siete días o menos) la proyección de la tasa de contagio obtenida mediante un modelo VAR supera a todas las alternativas analizadas, debido a que sus discrepancias con los datos observados son notablemente inferiores al resto de las aproximaciones. Estos resultados son coherentes con los obtenidos, por ejemplo, en (21), en el sentido de que los modelos VAR pueden mejorar la predicción del número de contagios y fallecimientos en relación con otras aproximaciones empíricas. El modelo mínimo cuadrático se mantiene como el más apropiado para horizontes de proyección ligeramente superiores (7-20 días), aunque apenas se diferencia del paseo aleatorio cuando se extiende el periodo de pronóstico a 30 o 45 días.

Como es bien sabido, los modelos ARIMA muestran su mayor potencial de pronóstico en las proyecciones de corto plazo (22). El menor error de predicción de estos modelos a horizontes temporales más elevados ($t=45$), que se muestra en la Tabla 1, debe ser interpretado con cautela. Dado que los modelos univariantes están mejor preparados para extraer la información de las series temporales, en su comportamiento más inmediato no es apropiado atribuir este menor error de pronóstico a su capacidad para incluir elementos de más largo plazo. El conjunto de estimaciones pierde calidad y esta relación resulta más casual que causal.

La propuesta metodológica que incorpora, en la estimación econométrica, información extraída o construida a partir de los modelos SIR, proporciona una aproximación empírica que supera a otras alternativas de estimación, incluyendo la hipótesis del paseo aleatorio, modelos ARIMA y las estimaciones de la función de distribución de contagios Weibull. Estos resultados validan la relevancia de las técnicas de calibración y estimación de la tasa de contagios basadas en el uso de modelos estructurales.

Este resultado contrasta con la visión, ampliamente difundida, sobre la pobre capacidad predictiva de los modelos SIR (16). Se confirma que, analizados de forma individual, los modelos SIR muestran severas limitaciones para proyectar la tasa de contagio. No obstante, si se combinan con la información oficial mediante el uso de técnicas econométricas estándar, su uso puede contribuir a mejorar notablemente las previsiones de corto plazo (30 días o menos). De esta forma, los resultados sugieren que los diferentes modelos teóricos tipo SIR (aunque no todos ellos) contienen información relevante para el pronóstico de los contagios.

Este tipo de pronósticos puede ayudar a las autoridades sanitarias a controlar los impactos de contingencias de una pandemia de forma más oportuna y efectiva. Esto es de suma importancia porque las diferentes políticas sanitarias tienen costes económicos y sociales muy diferen-

tes. Por tanto, la intensidad, el alcance, e incluso el tipo de medidas adoptadas, dependen de la evolución prevista de los contagios, incluso en el muy corto plazo ♠

Conflictos de interés: Ninguno.

REFERENCIAS

1. Ogilvy Kermack W, McKendrick AG. A contribution to the mathematical theory of epidemics. *Proc R Soc Lond A.* 1927; 115(772):700-21. <https://doi.org/10.1098/rspa.1927.0118>.
2. Getz WM, Salter R, Mgbara W. Adequacy of SEIR models when epidemics have spatial structure: Ebola in Sierra Leone. *Philos Trans R Soc Lond B Biol Sci.* 2019; 374(1775):20180282. <https://doi.org/10.1098/rstb.2018.0282>.
3. Korennoy FI, Gulenkin VM, Gogin AE, Vergne T, Karaulov AK. Estimating the basic reproductive number for African swine fever using the Ukrainian historical epidemic of 1977. *Transbound Emerg Dis.* 2017; 64(6):1858-66. <https://doi.org/10.1111/tbed.12583>.
4. World Health Organization. Preguntas y respuestas sobre la enfermedad por coronavirus (COVID-19). 2020. Disponible en: <https://bit.ly/3SC5pxV>.
5. Roda WC, Varughese MB, Han D, Li MY. Why is it difficult to accurately predict the COVID-19 epidemic? *Infect Dis Model.* 2020; 5:271-81. <https://doi.org/10.1016/j.idm.2020.03.001>.
6. Moein S, Nickaeen N, Rooointan A, Borhani N, Heidary Z, Javanmard SH, Gheisari Y, et al. Inefficiency of SIR models in forecasting COVID-19 epidemic: a case study of Isfahan. *Sci Rep.* 2021; 11(1):1-9. <https://doi.org/10.1038/s41598-021-84055-6>.
7. Sáez C, Romero N, Conejero JA, García-Gómez JM. Potential limitations in COVID-19 machine learning due to data source variability: a case study in the nCov2019 dataset. *J Am Med Inform Assoc.* 2021; 28(2):360-4. <https://doi.org/10.1093/jamia/ocaa258>.
8. Ashofteh A, Bravo JM. A study on the quality of novel coronavirus (COVID-19) official datasets. *Stat J IAOS.* 2020;36(2):291-301. <https://doi.org/10.3233/SJI-200674>.
9. Iglesias I, Pérez AM, Sánchez-Vizcaíno JM, Muñoz MJ, Martínez M, De La Torre A. Reproductive ratio for the local spread of highly pathogenic avian influenza in wild bird populations of Europe, 2005-2008. *Epidemiol Infect.* 2011; 139(1):99-104. <https://doi.org/10.1017/S0950268810001330>.
10. Chang HJ. Estimation of basic reproduction number of the Middle East respiratory syndrome coronavirus (MERS-CoV) during the outbreak in South Korea. *Biomed Eng Online.* 2017; 16(1):1-11. <https://doi.org/10.1186/s12938-017-0370-7>.
11. Zhang L, Huang J, Yu H, Liu X, Wei Y, Lian X, et al. Optimal parameterization of COVID-19 epidemic models. *Atmos Ocean Sci Lett.* 2021; 14(4):100024. <https://doi.org/10.1016/j.aosl.2020.100024>.
12. Tang L, Zhou Y, Wang L, Purkayastha S, Zhang L, He J, Wang F, Peter X, Song A review of multi-compartment infectious disease models. *Int Stat Rev.* 2020; 88(2):462-513. <https://doi.org/10.1111/insr.12402>.
13. Our World in Data (OIID). Estimated cumulative excess deaths per 100,000 people during COVID-19, Chile. Ministerio de Salud (Chile). 2025..Disponible en: <https://bit.ly/44GbwZA>.
14. Leipzig Health Atlas. COVID-19 Trajectory Viewer. 2020. Disponible en: <https://bit.ly/3GQfLb3>.
15. Arroyo-Marioli F, Bullano F, Kuinskas S, Rondón-Moreno C. Dynamics of transmission and control of COVID-19: a real-time estimation using the Kalman filter. 2020. <https://doi.org/10.1101/2020.04.19.20071886>.
16. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am J Epidemiol.* 2013; 178(9):1505-12. <https://doi.org/10.1093/aje/kwt133>.
17. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am J Epidemiol.* 2004; 160(6):509-16. <https://doi.org/10.1093/aje/kwh255>.
18. Cauchemez S, Boëlle P, Donnelly CA, Ferguson N, Thomas G, Leung GM, et al. Real-time Estimates in Early Detection of SARS. *Emerg Infect Dis.* 2006; 12(1):110-113. <https://doi.org/10.3201/eid1201.050593>.
19. Hens N, Van Ranst M, Aerts M, Robesyn E, Van Damme P, Beutels P. Estimating the effective reproduction number for pandemic influenza from notification data made publicly available in real time: a multi-country analysis for influenza A/H1N1v 2009. *Vaccine.* 2011; 29(5):896-904. <https://doi.org/10.1016/j.vaccine.2010.05.010>.
20. Seber GA, Wild CJ. Nonlinear regression. Hoboken (NJ): John Wiley & Sons; WILEY-INTERSCIENCE. 2003.
21. Rajab K, Kamalov F, Cherukuri AK. Forecasting COVID-19: Vector autoregression-based model. *Arab J Sci Eng.* 2022; 47:1-10. <https://doi.org/10.1007/s13369-021-06526-2>.
22. Canova F. Detrending and business cycle facts. *J Monet Econ [Internet].* 1998; 41(3):475-512. Disponible en: <https://bit.ly/44zPK9D>.