

THE CONVERSATION

Rigor académico, oficio periodístico



martin sanchez ezY rc mUE unsplash.

COVID-19: pandemia de modelos matemáticos

24 abril 2020 19:27 CEST

Los modelos matemáticos para predecir la evolución de la epidemia y el impacto de las medidas para su control son una *bola de cristal* de moda en todo el planeta, con intención más o menos académica y ejecutiva. Ya hay incluso una revisión Cochrane de los modelos y su utilidad.

La mayoría parte de un modelo simple y elegante que se formuló hace casi un siglo, en 1927, y fue popularizado en 1991 por Kermack-McKendrick. Es el llamado modelo SIR (Susceptible, Infectado, Recuperado), o SEIR, que introduce a los Expuestos.

Este modelo se basa en ecuaciones diferenciales para describir la dinámica de los contagios en una población cerrada con N individuos que inicialmente son susceptibles (S) al patógeno y que, a partir de un infectado inicial, van contagiándose a una determinada velocidad y pasando a ser infectados (I). Tras un período de enfermedad activa, los que no fallecen pasan al estado de inmunes: se han recuperado (R) y ya no contagiarán más. Por tanto, la población susceptible va disminuyendo hasta que ya no se produzcan más contagios. La popularidad actual de estos modelos se plasma en el uso de

Autores



Beatriz Gonzalez Lopez-Valcarcel

Catedrática de Métodos Cuantitativos en Economía y Gestión, Universidad de Las Palmas de Gran Canaria



Laura Tomaino

University of Milan



Lluís Serra Majem

la calculadora epidemiológica basada en ellos y disponible *online* para jugar a simular los efectos de la epidemia.

El problema es que este modelo ha de aplicarse en tiempo real, con pocos datos y gran incertidumbre sobre los parámetros. Los datos son poco fiables y se revisan y cambian retrospectivamente, por lo que los modelos han de cambiar también. Pero, por muy inciertos que sean los modelos, puesto que hay que tomar decisiones rápidamente, mejor es hacerlo con la tenue luz de esa linterna que a oscuras.

La gran pregunta es cómo y cuánto podremos aplanar la curva de contagios (que es la curva de nuevos casos diarios) y dilatarla en el tiempo, para no colapsar los recursos disponibles, en particular las camas de UCI y los respiradores. Los modelos pueden ayudar a prever las necesidades de ampliación de esos recursos. Un ejemplo encomiable en este sentido es el modelo disponible desde el principio de la pandemia, diseñado para los hospitales de Pennsylvania, que se puede ejecutar *online* para prever el número de camas necesarias en un hospital específico.

En su versión más simple, el modelo únicamente se alimenta de los datos del número diario de nuevos casos positivos por COVID-19 en el territorio que se analiza. Si llevamos un mes desde el inicio, tendremos 30 datos. Ese es nuestro tamaño muestral. El modelo debe hacer asunciones sobre la dinámica del contagio, con y sin medidas de intervención, y sobre el comportamiento del virus (distribución del tiempo de incubación y contagio, duración de la fase activa de la enfermedad, tasa de letalidad por edad, distribución de la gravedad y por tanto de la necesidad de recursos sanitarios como camas de UCI y en planta de hospitalización). Mucha demanda de información para pocos datos, faltan grados de libertad.

Ajuste y predicción

Los modelos tienen dos fases: una de ajuste y otra de predicción. En la de ajuste, se determinan qué funciones matemáticas y valores de los parámetros son compatibles con los datos que observamos de evolución diaria de la epidemia en el pasado, incorporando también el conocimiento biomédico *a priori* sobre la historia natural de la enfermedad, en un enfoque bayesiano. Una vez elegido el modelo que mejor se ajusta a los datos del pasado –el que mejor *predice* el pasado– se utiliza para simular o predecir la evolución de los casos en el futuro, bajo distintos escenarios de intervención. El gran problema de esta fase es que no podemos estar seguros de que los parámetros seguirán en el futuro el patrón que se les supone.

Concretemos. Un parámetro fundamental es el número básico de reproducción (R_0), que mide el número medio de personas a las que cada infectado contagia. Los modelos parten de un valor para cuando el virus circula libremente, antes de haber tomado medidas. Ese depende de factores

Catedrático de Medicina Preventiva y Salud Pública y Director de Instituto de Investigación, Universidad de Las Palmas de Gran Canaria



Patricia Barber

Universidad de Las Palmas de Gran Canaria



Silvia Rodríguez-Mireles

Médica Especialista en Medicina Preventiva y Salud Pública, Hospital Universitario de Gran Canaria Dr. Negrín, Servicio Canario de la Salud, Universidad de Las Palmas de Gran Canaria

culturales y, por tanto, será muy distinto en Suecia, donde las personas viven aisladas y apenas se tocan, que en las grandes urbes donde el metro cataliza los contagios o en los países mediterráneos del sur, donde nos apiñamos en bares, besamos para saludar a desconocidos y convivimos varias generaciones en el mismo hogar.

Hasta hace poco, la mayor parte de los modelos asumían que el R_0 sin intervención podría estar en el intervalo entre 2,25 y 2,5. También el influyente modelo del 16 de marzo del Imperial College, define escenarios con R_0 entre 2 y 2,6 para evaluar el impacto probable de las intervenciones.

Esos valores de R_0 son consistentes con el modelo publicado en *Science* el 16 de marzo con datos de China. Más recientemente, otro modelo que saldrá publicado en el número de julio de la revista del CDC sugiere que en China el R_0 podría haber sido mucho mayor, hasta 5, antes del confinamiento. Otros proponen escenarios alternativos basados en distintos valores de R_0 , por ejemplo entre 1,5 y 3.

El problema es que los resultados son extremadamente sensibles a los supuestos de partida.

Faltan datos

Esta gran incertidumbre y variabilidad en las estimaciones está causada en último término porque desconocemos el dato fundamental: cuántos pacientes hay realmente infectados. Solo sabemos cuántos han dado positivo, pero este dato depende del número de pruebas de diagnóstico microbiológico realizadas. Así pues, la variable endógena del modelo se mide con error. Y a destiempo: a lo largo de la pandemia, en nuestro país, se ha efectuado en los hospitales y servicios de urgencia, y en la fase actual como cribado de grupos de riesgo asintomáticos.

El artículo de *Science* estimaba que el 86 % de los casos habían pasado inadvertidos en Wuhan antes del 23 de enero, cuando empezó el confinamiento. Otros modelos como el del Imperial College del 30 de marzo sugieren que el número de infectados puede ser de millones de personas (en España, un 15 %, o sea, unos siete millones).

El hecho de que no sepamos el número de casos implica que la tasa de letalidad del virus sea desconocida y se estime de forma imprecisa. Este es, sin embargo, un dato fundamental para estimar los costes sociales de la enfermedad. Hasta tal punto es una incógnita el numerador (número de casos) que hay modelos diseñados para estimarlo a partir de los datos (más ciertos) de hospitalizados y graves (en UCI), como este.

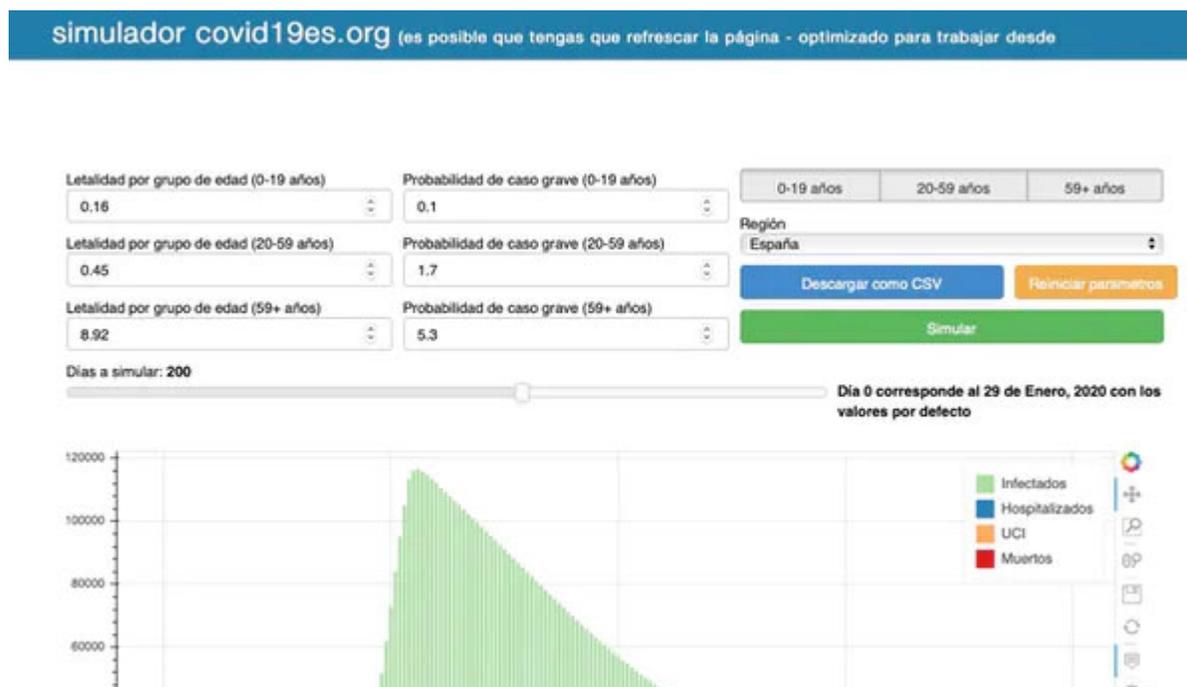
De hecho, si pudiéramos saber qué porcentaje de los casos son graves y requieren UCI, por edad, seríamos capaces de estimar el número de casos a partir del número de pacientes ingresados en UCI. Pero sin conocer la incidencia de la enfermedad nos faltan grados de libertad. Esa es la triste verdad. Por eso es tan útil y esencial la evidencia que aportará la gran encuesta de seroprevalencia en España, cuyo trabajo de campo se iniciará próximamente.

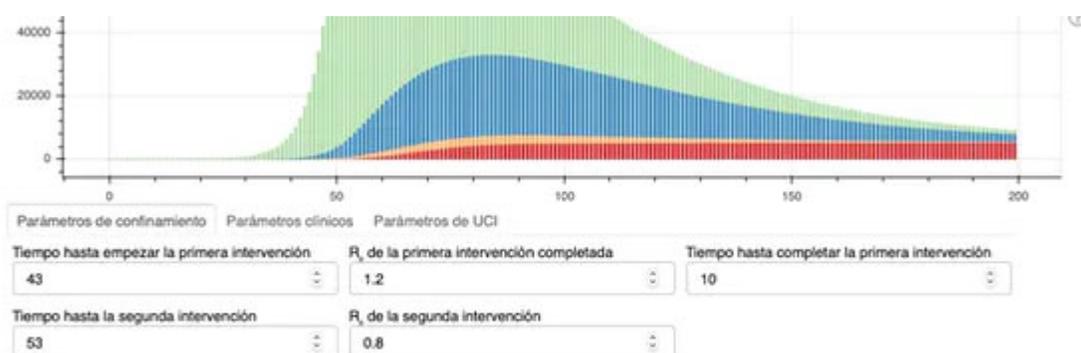
Los modelos estiman la efectividad de las intervenciones para bajar el R_0 , pero en último término la evolución de este parámetro depende del comportamiento humano, de la adherencia al confinamiento. Aunque estemos ya en fase de desaceleración o descenso en la curva de nuevos contagios, solo manteniendo la vigilancia epidemiológica y evitando activamente los contagios se conseguirá alcanzar una cota casi nula de nuevos contagios.

Algunos analistas intentan resolver el problema de la falta de grados de libertad alimentando a los modelos con más información (covariables y territorios), incorporando la movilidad de las personas entre áreas y la intensidad de las interacciones sociales. Un modelo para prever camas de UCI necesarias en Francia, por ejemplo, utiliza la matriz de contactos publicada en 2017 para 152 países, para 17 grupos de edad, que a su vez se basa en millones de datos de movilidad y en técnicas de inteligencia artificial: con cuántas personas de cada grupo de edad interactúa a lo largo de un día cada individuo.

También incorporan datos masivos de interacción social otros modelos como el de ZZData labs, una empresa de inteligencia artificial para los negocios. Esto puede ayudar en la primera fase (estimación o ajuste antes de la intervención), pero no resuelve la fase de predicción porque necesita asumir hipótesis sobre la reducción de las interacciones sociales, y con ellas, del R_0 , bajo confinamiento. A medida que pasan los días, algunos modelos han de rectificar estas asunciones a la luz de la evidencia. Otros, como el de la UPV, han quedado suspendidos en cuarentena durante días para ajustarse, sin reportar predicciones diarias.

Una ventaja de los modelos multiterritoriales es que algunos parámetros de la historia natural de la enfermedad asociados a la biología (tiempos, gravedad por grupos de edad) puede suponerse que son homogéneos o universales, y esto no es supuesto heroico típico de economista.





covid19es.org

Covid19es

Modelos locales y globales

Unos modelos son locales; otros, globales. El IHME lanzó el suyo recientemente, paramétrico y de estructura simple para prácticamente todos los países del mundo, y estima el número de camas UCI y de hospital necesarias y disponibles. Es muy osado, no solo por los problemas que antes mencionamos, sino porque asume, en el modelo para España, que el conjunto del país es una única unidad en la que las camas UCI y los pacientes tienen movilidad en el territorio. Da la impresión de que, como no se puede controlar al virus, al menos podemos alimentar la ilusión de controlarlo sobre el papel.

Una alternativa científicamente honesta, sensata y creativa es la de Jeff Harris del 6 de abril en NBER, para mostrar que la tasa de contagios en Nueva York ya ha entrado en fase de descenso. Con argumentos empíricos a partir del modelo SIR más simple, y con razonamientos económicos, desgana una secuencia de argumentos fundados en pro y en contra de esa conclusión.

Otro modelo de gran interés es el liderado por la UPC, con predicciones por países y por CCAA en España hasta cinco días hacia delante. Estima dinámicamente el R_t , lo que permite monitorizar de forma consistente en el tiempo la efectividad de las medidas públicas contra la expansión del virus. Su primer informe se publicó el 17 de marzo y desde entonces publican un informe diario de seguimiento. Asimismo, el trabajo del grupo CeMat con la iniciativa “Acción matemática contra el coronavirus” contribuye al conocimiento de la epidemia desde las matemáticas.

Estamos ya en el descenso de la curva de contagios, pero seguimos en la fase ascendente de modelos predictivos. Cada día aparecen nuevos modelos, cuyas predicciones parecen también estar sujetas a contagio, pues resulta chocante que con tal grado de incertidumbre haya tanto consenso en los resultados entre modelos. El tiempo dirá. De momento, es tiempo de agradecer los esfuerzos de la comunidad científica de modelizadores para sacar brillo a la bola de cristal. Algunos de esos modelos, como el de la figura que acompaña a esta entrada (accesible en <https://www.covid19es.org/simulador>), se pueden ejecutar *online*. Seamos optimistas, la colaboración científica, compartiendo

códigos abiertos de modelos como EpiNow, en R, y librerías de recursos disponibles para modelizar el COVID-19 aportarán información útil para tomar decisiones acertadas.

[datos](#)[big data](#)[pandemia](#)[coronavirus](#)[COVID-19](#)[SARS-CoV-2](#)