

Análisis del impacto de la vacunación sobre la incidencia acumulada de COVID-19 en España

Alba Bueses Suárez
Facultad de informática
Universidad de A Coruña
A Coruña, España
alba.bsuares@udc.es

Resumen—En este trabajo se estudia el impacto de la vacunación en la incidencia de COVID-19 en España, a partir de datos procedentes del Centro Nacional de Epidemiología y el Instituto de Salud Carlos III. Con este fin se han formulado modelos estadísticos para explicar el número de pacientes hospitalizados en función de la incidencia acumulada (IA) de COVID-19 y el porcentaje de población vacunada. El efecto de la vacunación se ha valorado desde dos perspectivas: la primera, introduciendo el porcentaje de vacunación como variable explicativa en el modelo, y la segunda, formulando dos modelos distintos, uno con los datos correspondientes a la etapa previa al comienzo de la vacunación, y otro con los datos correspondientes a la etapa posterior, donde el porcentaje de vacunación ya casi no varía. Los resultados de los análisis permiten concluir que existe una diferencia significativa entre, la proporción de pacientes que han sido hospitalizados con respecto al número de ellos que ha sido diagnosticado como positivo para COVID-19, antes y después de la vacunación. Llegando a obtener una relación 7 veces mayor de esta proporción, hospitalizados frente a casos positivos en el modelo antes de la vacunación, que en el de después.

1. Introducción

1.1. Inicio de la pandemia

El SARS-CoV-2 o Síndrome Respiratorio Agudo Severo Coronavirus-2, denominado así por el Comité Internacional de Taxonomía de Virus, es un virus de la familia Coronaviridae, dentro del orden Nidovirales, orden caracterizado por presentar un genoma de ARN monocatenario positivo e infectar una amplia cantidad de especies animales, entre las que se encuentra los humanos [1], [2]. El 11 de febrero de 2020 la Organización Mundial de la Salud (OMS) anunció el nombre de CoV o COVID-19, para la enfermedad causada por este virus, cuya gravedad varía desde síntomas leves hasta potencialmente mortales, declarándola un mes después como una pandemia [3].

Los primeros casos se sitúan en diciembre de 2019, cuando varios pacientes con neumonía de etiología desconocida fueron ingresados en hospitales de Wuhan [4], [5]. Según un documento publicado por la OMS, hasta

el 3 de enero de 2020 se habían notificado 44 pacientes afectados por neumonía de etiología desconocida, 11 de los cuales estaban gravemente enfermos, mientras que los 33 restantes se encontraban estables. El 8 de enero de 2020, la secuenciación genética demostró un nuevo CoV como posible organismo causal [4]. Desde entonces, el virus se ha propagado muy rápidamente, infectando hasta la fecha a más de 572 millones de personas y se ha cobrado más de 6 millones de vidas [3]. Aunque la tasa de mortalidad del SARS-CoV-2 es más baja que la del SARS-CoV y el MERS-CoV, su transmisión es mucho más rápida, habiéndose extendido a más de 160 países desde Wuhan en tan solo tres meses [6], [7].

1.2. Medidas de prevención

En el 2020 se demostró que la edad avanzada y las afecciones subyacentes aumentan sustancialmente la tasa de letalidad causada por este virus [4], [5] y los gobiernos pusieron en marcha medidas no farmacéuticas como el distanciamiento social, las mascarillas y el rastreo de contactos, que resultaron ser medidas eficaces para reducir la propagación viral y limitar las demandas de atención médica [8], [9]. Sin embargo, estas medidas también han tenido un impacto negativo en la economía y salud mental de gran parte de la población [10]), haciendo necesario la búsqueda y aplicación de otros métodos que permitan la vuelta gradual a la normalidad [11]. Por ello las investigaciones se centraron en desarrollar rápidamente una vacuna, al considerarse una medida rutinaria y eficaz contra las enfermedades infecciosas [12], [13].

Aunque el desarrollo de las vacunas ha tenido lugar en un tiempo récord, algunos países comenzaron las vacunaciones masivas antes de contar con los resultados de los ensayos clínicos [14], [15]. El 2 de diciembre de 2020, se aprobó en el Reino Unido la primera vacuna probada en un gran ensayo clínico aleatorizado [16] y se publicaron los resultados preliminares de la eficacia de otras vacunas en diferentes poblaciones, como estudios a nivel nacional en Israel [17] y Escocia [18], y estudios en los que participan trabajadores esenciales de primera línea en lugares específicos en los Estados Unidos [19], [20], [21]. Una vez que el desarrollo y producción de múltiples vacunas se demostró exitoso,

los gobiernos tomaron la vacunación como la solución a la pandemia, haciéndose necesario un seguimiento de las vacunaciones para determinar la efectividad de las mismas una vez suministradas a gran escala [22].

1.3. Efectividad de la vacunación

Los primeros análisis de la efectividad de las vacunas (VE), se centraron en aprobar su uso a gran escala [23], realizándose sobre todo estudios de cohortes sobre diferentes grupos de población [24], [25], [26]. Estos análisis de VE siguen siendo relevantes por diferentes motivos, por ejemplo, la determinación de la durabilidad del efecto de las vacunas [27], [28], pero también se ha hecho necesario ampliar el espectro de población a analizar, poniendo de manifiesto las posibles ventajas de realizar estudios sobre datos de salud pública [11], [29]. En primer lugar, tales conjuntos de datos cubren una amplia gama de personas, lo que resuelve las dudas sobre que las diferencias en el estado serológico entre las poblaciones puedan limitar la utilidad de las estimaciones que se basan solo en poblaciones específicas [30]. Así, un diseño de prueba negativa (TND, por sus siglas en inglés) o un estudio de cohortes con trabajadores de la salud, que es más probable que tengan inmunidad previa inducida por infecciones [31], podría no proporcionar estimaciones válidas de VE para la población general [32], [33]. En segundo lugar, los conjuntos de datos de salud pública suelen ser grandes, lo que permite a las autoridades sanitarias estimar la VE frente a resultados raros pero graves [ingresos o muertes en la unidad de cuidados intensivos (UCI)]. Por último, las estimaciones de VE que utilizan datos de salud pública de toda la población también pueden ayudar a combatir las dudas con respecto a la seguridad y eficacia de las vacunas [23].

Sin embargo, los conjuntos de datos de salud pública en la mayoría de los países no contienen información adecuada para calcular la VE utilizando métodos estándar como TND [34]. En particular, la información sobre el estado de vacunación a menudo no se registra para las personas que se hicieron las pruebas de COVID. Como resultado, casi toda la evidencia a gran escala hasta la fecha sobre la efectividad real de las vacunas proviene de ciertos países donde se han registrado estos datos. Por ejemplo, view-hub.org, una base de datos integral de 181 estudios de VE de COVID-19 (a agosto de 2022), la mayoría de estos son de cohortes o TND (por lo tanto, el estado de vacunación de todos los participantes se conoce en cada caso) y solo unos pocos se basan en datos de salud pública a gran escala [11], [35], [36].

En este trabajo se analizan los datos correspondientes a España por grupos de edad facilitados por el Sistema para la Vigilancia en España, Centro Nacional de Epidemiología (SIVIES) del Instituto de Salud Carlos III para tratar de estimar el efecto de la vacunación en la incidencia de covid.

2. Material

La información necesaria para la realización de este trabajo ha sido obtenida a través del proyecto CEDCOVID, donde se han llevado a cabo dos consultas sobre los datos facilitados por el Centro Nacional de Epidemiología y el Instituto de Salud Carlos III:

- **Datos de casos de COVID-19** positivos, hospitalizados e ingresados en UCI: a partir de las actualizaciones diarias de la página oficial del Centro Nacional de Epidemiología para datos sobre COVID [37]. Se tiene en cuenta la provincia, la edad (decanales) y la fecha (desde el 1 de diciembre del 2019 hasta el 28 de julio de 2022) de cada entrada.
- **Datos de vacunación** obtenidos a través del Sistema para la Vigilancia en España [38]. Anonimizados y agrupados por edad, ubicación (provincia, comunidad autónoma o y a nivel nacional) y fecha (del 27 de diciembre de 2020 al 8 de agosto de 2022). Se diferencia número de dosis.

Adicionalmente se han tomado datos de la web del Instituto Nacional de Estadística para disponer de información demográfica de cada ubicación por grupos de edad quinquenales [39].

3. Métodos

3.1. Preprocesamiento

3.1.1. Estandarización.

Los datos disponibles de casos (número de casos positivos) y hospitalizados, así como, el número de vacunados hasta ese momento, se proporcionan de manera diaria, de modo que se dispone de cada uno de estos valores por día, y también por grupo de edad. Pero como se corresponden a fuentes distintas, y cada consulta se ha realizado por separado, ha sido necesario estandarizar los atributos de los datos, de forma que los atributos fecha y edad tuviesen la misma estructura en todos los casos. Una vez hecho esto, para que la estructura de los datos fuese más fácil de manejar para su análisis, se ha generado una función que selecciona los datos correspondientes a una ubicación concreta (provincia, comunidad autónoma o a nivel nacional).

3.1.2. Cálculo de la IA.

El Manual Docente de la Escuela Nacional de Sanidad publicado por el Instituto de Salud Carlos III en Madrid en octubre de 2009, “Método Epidemiológico” [40], define la incidencia acumulada (o proporción de incidencia) como un parámetro adecuado en el análisis de datos epidemiológicos. La incidencia acumulada (IA) relaciona el número de casos con el tamaño poblacional al comienzo de un periodo de tiempo (Δ_t):

$$IA = \frac{n^\circ \text{ de casos nuevos en } \Delta_t}{n^\circ \text{ de personas al inicio del periodo } (t_0)} \quad (1)$$

Para realizar este cálculo se codificó una función básica y luego se creó una más específica para la obtención de la IA en función del grupo de edad y tomando 14 días como valor por defecto para el intervalo en el que calcularla. De este modo disponemos de la proporción de personas que han enfermado, sido hospitalizadas o ingresadas en la UCI en este periodo, y el valor del atributo fecha se corresponde al primer día de los 14 que se valoran. Este dato lo expresamos por cada 100 000 habitantes multiplicándolo por 100 000, ya que se trata de datos epidemiológicos, y en estos se suele usar estas proporciones que facilitan la visualización de los datos. Para diferenciar estos valores de los valores de origen se les añade el sufijo *-ia*: *ia + (casos/hosp/uci) + grupo de edad*.

3.1.3. Cálculo del porcentaje de vacunación.

En el caso de los datos de vacunación se optó por realizar el cálculo del porcentaje de personas que ya habrían recibido un cierto número de dosis en cada fecha. Se seleccionó la cantidad de dos dosis para realizar este cálculo por ser considerada la pauta completa recomendada para la población, en general, con la mayoría de las vacunas que se inocularon en España [41]. En este caso se sigue la misma nomenclatura que en el apartado anterior, a pesar de que no se trata del mismo tipo de datos, por simplificar el diseño de la función.

3.1.4. Análisis exploratorio.

Una vez que los datos están preparados se realiza un análisis exploratorio para determinar el tipo de distribución de los mismos y si presentan correlación entre los diferentes grupos de edad de la misma variable.

3.2. Modelos de regresión

Debido a la aplicación de medidas preventivas y a que este análisis se realiza sobre los datos de población y no sobre una cohorte a la que se realiza un seguimiento, en lugar de analizar los datos de contagios, hospitalizados o ingresos en UCI directamente, vamos a analizar cómo varía la relación entre el número de pacientes diagnosticados de COVID-19 y el número de ellos que han sido hospitalizados. Primero se consideró realizar divisiones temporales en función de la llegada de cada nueva cepa, pero la falta de datos genéticos y la coexistencia de las diferentes cepas hizo que se descartase esta aproximación. Se optó por analizar los datos de forma global. Se analizaron 3 aproximaciones:

- Añadir el porcentaje de vacunación como variable explicativa.
- Generar dos modelos, uno antes de comenzar las vacunaciones y otro una vez que el porcentaje de vacunados se estabiliza en valores altos (estrategia Antes y Después).

3.2.1. Modelos lineales generalizados.

Los modelos lineales generalizados (GLM) son una extensión de los modelos lineales para casos en los que la distribución condicional de la variable respuesta no sea

necesariamente normal, pero sí perteneciente a la familia exponencial. Para ello se introduce lo que se conoce como una función de enlace (o link) g de forma que:

$$g(E(Y|X)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (2)$$

Bajo esta definición en el modelo lineal clásico la función de enlace sería la función identidad:

$$E(Y|X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (3)$$

3.2.2. Modelos aditivos generalizados.

Aunque los modelos GLM son más flexibles, siguen suponiendo linealidad, lo que no sucede en los modelos aditivos generalizados (GAM). Los GAM relajan las restricciones que presentan los GLM, pasando de asumir que el resultado debe ser una suma ponderada a modelarlo como una suma de funciones:

$$g(E(Y|X)) = \beta_0 + f_1(X_1) + f_2(X_2) + \dots + f_p(X_p) \quad (4)$$

Para generar funciones no lineales se usan funciones spline que se pueden combinar para aproximar de la forma más realista la relación entre las variables explicativas y la respuesta. Los GAM también introducen un término de penalización para los pesos, para que estos se mantengan cerca de cero y evitar el sobreajuste [42].

3.2.3. Modelos con autocorrelación.

Para determinar si los datos tienen dependencia temporal se analiza la correlación de los residuos, aplicando la función *ar()* del paquete *stats*, que ajusta modelos autorregresivos. El resultado obtenido de esta función determina si efectivamente hay dependencia temporal y cuántos días, en el caso que ocupa a este trabajo, debemos retroceder en los datos. Para poder aplicar este nuevo parámetro se usa el método de mínimos cuadrados generalizados (GLS), adecuado para incorporar la dependencia, y los modelos aditivos mixtos generalizados (GAMM), que permiten incorporar la dependencia en los GAM [43] [44].

3.2.4. Estructura de datos.

Para aplicar los modelos se selecciona una ubicación y se prueban diferentes subgrupos de datos:

- **Totales:** Se usan solo las variables generadas en base a la población completa de la región.
- **Grupo de edad:** Tras aplicar un análisis de correlación sobre las variables correspondientes al número de diagnosticados de COVID-19, el número de ellos que están hospitalizados y el porcentaje de vacunación entre los diferentes grupos de edad, se selecciona un único grupo de edad. El objetivo es seleccionar los datos correspondientes a un solo grupo de edad que puedan ser representativos de los demás grupos en relación a esa correlación observada, y que haya recibido las dos dosis de la vacuna pronto. De este modo se obtiene un modelo que únicamente representa la relación entre los hospitalizados y los casos diagnosticados como positivos correspondientes al grupo de edad seleccionado.

- **Totales + grupo de edad:** Se usan conjuntamente, las variables correspondientes al total de la población, y las correspondientes al grupo de edad seleccionado descrito en el apartado anterior. Como variable respuesta se considera el total de la población.
- **Población vulnerable:** Dado que la población de edad más avanzada fue vacunada primero, y presenta una mayor sensibilidad a los efectos más graves de la COVID-19, se seleccionó, en lugar de un único grupo de edad, un subgrupo de edades. En la región seleccionada para estos análisis, España, los cuatro grupos correspondientes a la población de mayor edad alcanzan valores altos de porcentaje de vacunación mucho antes que el resto de grupos, lo que supone que se dispone de más datos correspondientes a los días en los que el resto de la población todavía esta siendo vacunada. Por ello se realizó una reagrupación de los datos correspondientes a la población de 70 años o mayor, de modo que se recalculan los valores de IA y porcentajes, y se analizan los datos reagrupados de esta población como si se tratase de un nuevo grupo de edad. Esto supone que, como en casos anteriores, tanto las variables explicativas como la respuesta se corresponden a datos de este subgrupo, que se corresponde a la población mayor de 70 años, no a la población total.

Tras realizar las primeras pruebas con los métodos mencionados, se realizó una revisión de los datos y la hemeroteca con la intención de determinar posibles fuentes de ruido. Esta revisión determinó que era mejor realizar estos análisis solo sobre una sección de los datos, en lugar de sobre todo el conjunto. Se propuso el 20 de mayo de 2020 como fecha de inicio, por ser una fecha aproximada a aquella en la que se empezó a disponer de material y test suficientes en los hospitales para determinar los posibles positivos, y el 4 de marzo de 2022 como fecha final, por anunciarse que tras esa fecha solo se reportarían los casos de la población mayor de 60 años.

4. Resultados

Como ubicación de referencia, para mostrar los resultados, se selecciona España.

4.1. Correlación

En la gráfica de la figura 1 se puede observar cómo los datos correspondientes a la IA de los casos de las personas de entre 60 y 69 son los que presentan mayor correlación con el resto de grupos, por lo que se seleccionan para aplicar los modelos, tanto de forma individual, como conjuntamente con los datos correspondientes a la población total.

El análisis de autocorrelación de los residuos de los primeros modelos indica que existe dependencia temporal, presentando un valor de ϕ cercano al 1. Esto quiere decir que hay una dependencia fuerte de los residuos de cada día con

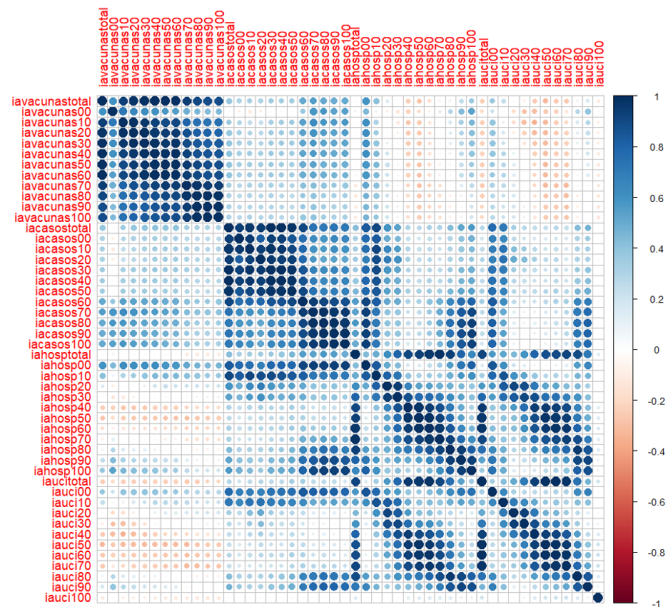


Figura 1. Gráfica de la matriz de correlación de los datos.

los residuos del día inmediatamente anterior. Es por ello que a la hora de ajustar los modelos GLS y GMM seleccionamos los valores $p = 1$ y $q = 0$ ($corARMA(p=1, q=0)$), para indicar que valore los datos correspondientes a la variable respuesta el día anterior, como parte del modelo.

4.2. Modelos

Como se puede observar, en la tabla 1 los valores del coeficiente de determinación (R^2) alcanzan valores de ajuste muy altos en la mayoría de los casos, sobre todo en el caso de los modelos para Antes y Después. Los valores bajos en los modelos GLM y GLS para Totales pueden deberse al ruido producido al valorar todos los casos en conjunto. Se debe tener en cuenta que cada grupo de edad ha recibido las vacunas en periodos distintos y que los efectos de la COVID-19 no han sido los mismos en cada uno de ellos. La población de edad avanzada ha sido la que se ha visto más afectada por la enfermedad y la que ha recibido la vacuna en una fase más temprana, mientras que los grupos de edad más jóvenes, que padecieron en menor medida los efectos de esta enfermedad fueron vacunados posteriormente.

Por todo lo anterior, se seleccionan los modelos de Antes y Después, de los cuales se calcula el intervalo de confianza para su variable explicativa. En concreto, se seleccionan los modelos lineales (GLM y GLS) con los subgrupos de datos correspondientes a **Totales, 60 años y Mayores de 70**. Esta selección se corresponde con los modelos más sencillos, que facilitan la interpretación posterior de los resultados, ya que no se observan grandes diferencias de ajuste con los demás.

No se descartan los modelos GLS debido a que, el análisis realizado sobre los residuos de los modelos, reflejó una fuerte dependencia temporal con los valores del día anterior.

Tabla 1. VALORES DEL COEFICIENTE DE DETERMINACIÓN (R^2) OBTENIDOS DE LOS DIFERENTES MODELOS.

R^2	Totales			60 años			Totales + 60 años			Mayores de 70		
	Único	Antes	Después	Único	Antes	Después	Único	Antes	Después	Único	Antes	Después
GLM	0.513	0.993	0.974	0.637	0.987	0.835	0.846	0.993	0.976	0.690	0.993	0.945
GAM	0.941	0.996	0.987	0.910	0.885	0.993	0.993	0.998	0.998	0.931	0.998	0.986
GLS	0.513	0.993	0.974	0.497	0.987	0.835	0.808	0.992	0.975	0.672	0.993	0.948
GAMM	0.719	0.994	0.981	0.604	0.991	0.805	0.959	0.995	0.988	0.747	0.997	0.948

Tabla 2. INTERVALO DE CONFIANZA (IC) PARA LOS MODELOS GLM Y GLS DE ANTES Y DESPUÉS.

			Totales		60 años		Mayores de 70	
			Antes	Despues	Antes	Despues	Antes	Despues
GLM	Coef.		0.0755244	0.0107114	0.1423906	0.0260657	0.295576	0.075092
	IC al 95 %	Extremo inferior	0.0746889	0.0104591	0.1423906	0.0246543	0.2923783	0.072941
		Extremo superior	0.0763599	0.0109638	0.1440559	0.0274772	0.2987744	0.077242
GLS	Coef.		0.0627676	0.008988	0.1201641	0.0230756	0.258614	0.072095
	IC al 95 %	Extremo inferior	0.0603548	0.008503	0.1145757	0.0209996	0.2497356	0.068955
		Extremo superior	0.0651762	0.009472	0.1257526	0.0251516	0.2674918	0.075234

En la tabla 2 se muestra que en ninguno de los modelos los intervalos de confianza para las variables explicativas se solapan entre los obtenidos en el modelo antes de la vacunación y el obtenido después de la vacunación.

En la figura 2 se puede observar la estimación realizada por los GLM, un modelo lineal para Antes y otro para Después de la vacunación con los datos correspondientes a la población total. Las fórmulas correspondientes a los modelos son:

- Antes (rojo):

$$f(x) = 0,0032273 + 0,0755244x + \epsilon \quad (5)$$

- Después (verde):

$$f(x) = 4,6108222 + 0,0107114x + \epsilon \quad (6)$$

5. Conclusión

Se puede concluir que para los datos utilizados los modelos aditivos no generan un ajuste mucho mejor que los modelos lineales y que los modelos de Antes y Después presentan los mejores valores de R^2 .

Los intervalos de confianza de los modelos, antes y después de la vacunación, no muestran solapamiento en ninguno de los casos analizados. Para interpretar la magnitud de estos resultados se valoran los modelos como si no tuviesen intercepto (término independiente), y podemos observar como en los casos de Totales y 60 años la relación es entre 6 y 7 veces mayor antes que después de la vacunación, mientras que, en el caso de los Mayores de 70, es de 3 a 4 veces mayor. Teniendo estas aproximaciones en cuenta, en un escenario en el que no se produjese la vacunación el número de hospitalizados sería entre 3 y 7 veces mayor de lo que es.

Se puede concluir que existe una diferencia en la relación entre los casos y los hospitalizados antes y después de la vacunación, y que el efecto de esta ha supuesto un decremento de hospitalizados con respecto al número de casos. Estos resultados reflejan el efecto tan determinante

que ha tenido la vacunación para frenar la pandemia, y lo graves que podrían ser las consecuencias de no haberse aplicado esta medida en la mayor parte de la población, sobre todo aquella más susceptible a padecer los efectos más graves del COVID-19.

5.1. Perspectivas futuras

Podría resultar de interés aplicar diferentes valores del parámetro $corARMA$ a los modelos GLS y GAMM para observar su efecto sobre el ajuste de los modelos y los coeficientes de las demás variables explicativas. Por ejemplo, tomar el valor de 7 días antes, como dato autorregresivo, dado que la actualización de datos en algunas regiones era semanal, puede darnos información nueva sobre la relación entre las variables.

También se plantea analizar en el futuro el impacto de la vacunación tomando como variable respuesta el número de fallecidos. En este caso podrían considerarse como variables explicativas tanto el número de casos, como el número de hospitalizados o, incluso, ambas.

A su vez, se espera aplicar estos análisis a nivel de provincia y comunidad autónoma. El objetivo sería determinar si la vacunación presenta el mismo efecto independientemente de la región seleccionada y, por tanto, si tiene un efecto homogéneo en todo el país.

Agradecimientos

Agradecer al Dr. Ricardo Cao Abad y al Dr. Manuel Oviedo de la Fuente su tiempo y asesoramiento como tutores de este trabajo.

Investigación parcialmente financiada por el proyecto CED-COVID (Ciencia e Ingeniería de datos para la evaluación, predicción poblacional y personalizada de la evolución de la enfermedad COVID-1: referencia IN845D2020/26), subvencionado por la Axencia Galega de Innovación de la Xunta de Galicia. Se ha tenido acceso a los datos de SIVIES y REGVACU en el marco del acuerdo entre el Centro de Coordinación de Alertas y Emergencias Sanitarias

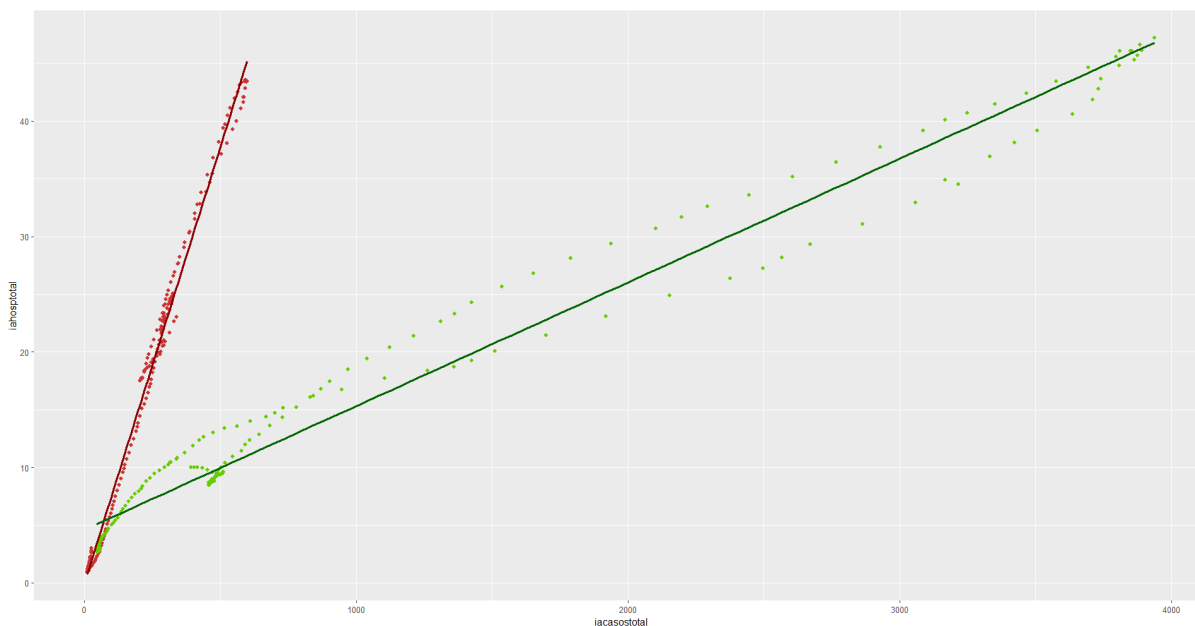


Figura 2. Representación grafica de los datos y los modelos GLM de Antes (rojo) y Después (verde) para los datos Totales.

(CCAES) del Ministerio de Sanidad y el Comité Español de Matemáticas (CEMat), dentro de la Acción Matemática contra el Coronavirus, con el fin de realizar análisis para facilitar la predicción sobre la evolución de la pandemia de Covid-19 en distintos ámbitos territoriales, así como la evaluación de distintos escenarios y su impacto en el control de la pandemia.

Y una mención especial al Dr. Óscar Crego por su ayuda y paciencia a lo largo de los últimos meses.

Referencias

- [1] Chih-Cheng Lai, Tzu-Ping Shih, Wen-Chien Ko, Hung-Jen Tang, and Po-Ren Hsueh. Severe acute respiratory syndrome coronavirus 2 (sars-cov-2) and coronavirus disease-2019 (covid-19): The epidemic and the challenges. *International journal of antimicrobial agents*, 55(3):105924, 2020.
- [2] Pavan Kumar Samudrala, Pramod Kumar, Kamlesh Choudhary, N-gender Thakur, Gaurav Suresh Wadekar, Richa Dayaramani, Mukta Agrawal, and Amit Alexander. Virology, pathogenesis, diagnosis and in-line treatment of covid-19. *European journal of pharmacology*, 883:173375, 2020.
- [3] World Health Organization. World health organization. covid-19 global situation. *31 de julio*, 2022.
- [4] Isaac I Bogoch, Alexander Watts, Andrea Thomas-Bachli, Carmen Huber, Moritz UG Kraemer, and Kamran Khan. Pneumonia of unknown aetiology in wuhan, china: potential for international spread via commercial air travel. *Journal of travel medicine*, 27(2):taaa008, 2020.
- [5] Hussin A Rothan and Siddappa N Byrareddy. The epidemiology and pathogenesis of coronavirus disease (covid-19) outbreak. *Journal of autoimmunity*, 109:102433, 2020.
- [6] SA Meo, AM Alhowikan, T Al-Khlaiwi, IM Meo, DM Halepoto, M Iqbal, AM Usmani, W Hajjar, and N Ahmed. Novel coronavirus 2019-ncov: prevalence, biological and clinical characteristics comparison with sars-cov and mers-cov. *Eur Rev Med Pharmacol Sci*, 24(4):2012–2019, 2020.
- [7] Mujeeb Khan, Syed F Adil, Hamad Z Alkhatlan, Muhammad N Tahir, Sadia Saif, Merajuddin Khan, and Shams T Khan. Covid-19: a global challenge with old history, epidemiology and progress so far. *Molecules*, 26(1):39, 2020.
- [8] Yixuan Yang, Fei Hao, Doo-Soon Park, Sony Peng, Hyejung Lee, and Makara Mao. Modelling prevention and control strategies for covid-19 propagation with patient contact networks. *Hum.-Cent. Comput. Inf. Sci*, 11:45, 2021.
- [9] Inés Barbeito, Daniel Precioso, María José Sierra, Susana Vegas-Azcárate, Sonia Fernández Balbuena, Begoña Vitoriano, David Gómez-Ullate, Ricardo Cao, Susana Monge, et al. Effectiveness of non-pharmaceutical interventions in nine fields of activity to decrease sars-cov-2 transmission (spain, september 2020-may 2021). 2022.
- [10] Brett Marroquín, Vera Vine, and Reed Morgan. Mental health during the covid-19 pandemic: Effects of stay-at-home policies, social distancing behavior, and social resources. *Psychiatry research*, 293:113419, 2020.
- [11] Alejandro Jara, Eduardo A Undurraga, Cecilia González, Fabio Paredes, Tomás Fontecilla, Gonzalo Jara, Alejandra Pizarro, Johanna Acevedo, Katherine Leo, Francisco Leon, et al. Effectiveness of an inactivated sars-cov-2 vaccine in chile. *New England Journal of Medicine*, 385(10):875–884, 2021.
- [12] Inaya Hajj Hussein, Nour Chams, Sana Chams, Skye El Sayegh, Reina Badran, Mohamad Raad, Alice Gerges-Geagea, Angelo Leone, and Abdo Jurjus. Vaccines through centuries: major cornerstones of global health. *Frontiers in public health*, 3:269, 2015.
- [13] Qiang Wang, Liuqing Yang, Hui Jin, and Leesa Lin. Vaccination against covid-19: A systematic review and meta-analysis of acceptability and its predictors. *Preventive medicine*, 150:106694, 2021.
- [14] Shabir A Madhi, Anthonet L Koen, Alane Izu, Lee Fairlie, Clare L Cutland, Vicky Baillie, Sherman D Padayachee, Keertan Dheda, Shaun L Barnabas, Qasim Ebrahim Bhorat, et al. Safety and immunogenicity of the chadox1 ncov-19 (azd1222) vaccine against sars-cov-2 in people living with and without hiv in south africa: an interim analysis of a randomised, double-blind, placebo-controlled, phase 1b/2a trial. *The lancet HIV*, 8(9):e568–e580, 2021.

- [15] Shcheblyakov DV et al Logunov DY, Dolzhikova IV. Safety and efficacy of an rad26 and rad5 vector-based heterologous prime-boost covid-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in russia. *Lancet*, 397:671–81, 2021.
- [16] Heidi Ledford, David Cyranoski, Richard Van Noorden, et al. The uk has approved a covid vaccine—here’s what scientists now want to know. *Nature*, 588(7837):205–206, 2020.
- [17] Kepten E Dagan N, Barda N et al. Bnt162b2 mrna covid-19 vaccine in a nationwide mass vaccination setting. *Nature*, 384:1412–23, 2021.
- [18] Eleftheria Vasileiou, Colin R Simpson, Chris Robertson, Ting Shi, Steven Kerr, Utkarsh Agrawal, Ashley Akbari, Stuart Bedston, Jillian Beggs, Declan Bradley, et al. Effectiveness of first dose of covid-19 vaccines against hospital admissions in scotland: national prospective cohort study of 5.4 million people. 2021.
- [19] Mark G Thompson, Jefferey L Burgess, Allison L Naleway, Harmony L Tyner, Sarang K Yoon, Jennifer Meece, Lauren EW Olsho, Alberto J Caban-Martinez, Ashley Fowlkes, Karen Lutrick, et al. Interim estimates of vaccine effectiveness of bnt162b2 and mrna-1273 covid-19 vaccines in preventing sars-cov-2 infection among health care personnel, first responders, and other essential and frontline workers—eight us locations, december 2020–march 2021. *Morbidity and Mortality Weekly Report*, 70(13):495, 2021.
- [20] William Daniel, Marc Nivet, John Warner, and Daniel K Podolsky. Early evidence of the effect of sars-cov-2 vaccine at one medical center. *New England Journal of Medicine*, 384(20):1962–1963, 2021.
- [21] Amadea Britton, Kara M Jacobs Slifka, Chris Edens, Srinivas Acharya Nanduri, Stephen M Bart, Nong Shang, Adora Harizaj, Jillian Armstrong, Kerui Xu, Hanna Y Ehrlich, et al. Effectiveness of the pfizer-biontech covid-19 vaccine among residents of two skilled nursing facilities experiencing covid-19 outbreaks—connecticut, december 2020–february 2021. *Morbidity and Mortality Weekly Report*, 70(11):396, 2021.
- [22] Edouard Mathieu, Hannah Ritchie, Esteban Ortiz-Ospina, Max Roser, Joe Hasell, Cameron Appel, Charlie Giattino, and Lucas Rodés-Guirao. A global database of covid-19 vaccinations. *Nature human behaviour*, 5(7):947–953, 2021.
- [23] Pier Luigi Lopalco and Frank DeStefano. The complementary roles of phase 3 trials and post-licensure surveillance in the evaluation of new vaccines. *Vaccine*, 33(13):1541–1548, 2015.
- [24] Malta Luxembourg. Interim analysis of covid-19 vaccine effectiveness against severe acute respiratory infection due to laboratory-confirmed sars-cov-2 among individuals aged 50 years and older, ecdc multi-country study—first update. 2022.
- [25] Ruth Link-Gelles, Matthew E Levy, Manjusha Gaglani, Stephanie A Irving, Melissa Stockwell, Kristin Dascomb, Malini B DeSilva, Sarah E Reese, I-Chia Liao, Toan C Ong, et al. Effectiveness of 2, 3, and 4 covid-19 mrna vaccine doses among immunocompetent adults during periods when sars-cov-2 omicron ba. 1 and ba. 2/ba. 2.12. 1 sublineages predominated—vision network, 10 states, december 2021–june 2022. *Morbidity and Mortality Weekly Report*, 71(29):931, 2022.
- [26] Karen Lutrick, Patrick Rivers, Young M Yoo, Lauren Grant, James Hollister, Krystal Jovel, Sana Khan, Ashley Lowe, Zoe Baccam, Hanna Hanson, et al. Interim estimate of vaccine effectiveness of bnt162b2 (pfizer-biontech) vaccine in preventing sars-cov-2 infection among adolescents aged 12–17 years—arizona, july–december 2021. *Morbidity and Mortality Weekly Report*, 70(51-52):1761, 2021.
- [27] Daniel R Feikin, Melissa M Higdon, Laith J Abu-Raddad, Nick Andrews, Rafael Araos, Yair Goldberg, Michelle J Groome, Amit Huppert, Katherine L O’Brien, Peter G Smith, et al. Duration of effectiveness of vaccines against sars-cov-2 infection and covid-19 disease: results of a systematic review and meta-regression. *The Lancet*, 2022.
- [28] Yair Goldberg, Micha Mandel, Yinon M Bar-On, Omri Bodenheimer, Laurence Freedman, Eric J Haas, Ron Milo, Sharon Alroy-Preis, Nachman Ash, and Amit Huppert. Waning immunity after the bnt162b2 vaccine in israel. *New England Journal of Medicine*, 385(24):e85, 2021.
- [29] Abhiroop Mukherjee, George Panayotov, Rik Sen, Harsha Dutta, and Pulak Ghosh. Measuring vaccine effectiveness from limited public health datasets: Framework and estimates from india’s second covid wave. *Science advances*, 8(18):eabn4274, 2022.
- [30] Erkan Kalafat, Laura A Magee, Peter von Dadelszen, Paul Heath, and Asma Khalil. Covid-19 booster doses in pregnancy and global vaccine equity. *Lancet (London, England)*, 399(10328):907, 2022.
- [31] Loris Wauthier, Quentin Delefortrie, Nicolas Eppe, Patrick Van-kerkhoven, Eva Wolff, Mélanie Dekeyser, Jean-Michel Cirriez, and Alexandre Grimmelprez. Sars-cov-2 seroprevalence in high-risk health care workers in a belgian general hospital: Evolution from the first wave to the second. *Acta Clinica Belgica*, pages 1–9, 2021.
- [32] Shmuel Benenson, Yonatan Oster, Matan J Cohen, and Ran Nir-Paz. Bnt162b2 mrna covid-19 vaccine effectiveness among health care workers. *New England Journal of Medicine*, 384(18):1775–1777, 2021.
- [33] Victoria Jane Hall, Sarah Foulkes, Ayoub Saei, Nick Andrews, Blanche Oguti, Andre Charlett, Edgar Wellington, Julia Stowe, Natalie Gillson, Ana Atti, et al. Covid-19 vaccine coverage in health-care workers in england and effectiveness of bnt162b2 mrna vaccine against infection (siren): a prospective, multicentre, cohort study. *The Lancet*, 397(10286):1725–1735, 2021.
- [34] Heather M Scobie, Michael Edelstein, Edward Nicol, Ana Morice, Nargis Rahimi, Noni E MacDonald, M Carolina Danovaro-Holliday, Jaleela Jawad, et al. Improving the quality and use of immunization and surveillance data: Summary report of the working group of the strategic advisory group of experts on immunization. *Vaccine*, 38(46):7183–7197, 2020.
- [35] Thiago Cerqueira-Silva, Srinivasa Vittal Katikireddi, Vinicius de Araujo Oliveira, Renzo Flores-Ortiz, Juracy Bertoldo Júnior, Enny S Paixão, Chris Robertson, Gerson O Penna, Guilherme L Werneck, Maurício L Barreto, et al. Vaccine effectiveness of heterologous coronavac plus bnt162b2 in brazil. *Nature Medicine*, 28(4):838–843, 2022.
- [36] Leonardo Arregocés-Castillo, Julián Fernández-Niño, Maylen Rojas-Botero, Andrés Palacios-Clavijo, Maryory Galvis-Pedraza, Luz Rincón-Medrano, Mariana Pinto-Álvarez, Fernando Ruiz-Gómez, and Belem Trejo-Valdivia. Effectiveness of covid-19 vaccines in older adults in colombia: a retrospective, population-based study of the esperanza cohort. *The Lancet Healthy Longevity*, 3(4):e242–e252, 2022.
- [37] Gobierno de España Ministerio de Ciencia e Innovación. Centro nacional de epidemiología. <https://cnecovid.isciii.es/>, Última visita: 1 de Agosto de 2022.
- [38] Instituto de Salud Carlos III Centro nacional de epidemiología. Sistema para la vigilancia en españa. <https://sivies.isciii.es/>, Última visita: 9 de Agosto de 2022.
- [39] Adscrito al Ministerio de Asuntos Económicos y Transformación Digital. Instituto nacional de estadística. <https://www.ine.es/>, Última visita: 15 de Agosto de 2022.
- [40] Instituto de Salud Carlos III. Método epidemiológico. <https://repisalud.isciii.es/>, 2009.
- [41] Gobierno de España. Estrategia de vacunación covid-19. <https://www.vacunacovid.gob.es/>, 1 de Agosto de 2022.
- [42] Michael Clark. Statistics, data, science. generalized additive models. <https://m-clark.github.io/generalized-additive-models/>, Agosto de 2022.
- [43] Luca Basile, Manuel Oviedo de la Fuente, Nuria Torner, Ana Martínez, and Mireia Jané. Real-time predictive seasonal influenza model in catalonia, spain. *PloS one*, 13(3):e0193651, 2018.
- [44] María L. Durbán Reguera. Chapter 2, generalized least squares. *Departamento de Estadística*, Universidad Carlos III de Madrid.