

Lessons from the pandemic. Usefulness of datathons and open bases in the analysis of Covid-19

Jelissa Reynoso-García, PhD¹, Ariel Leonardo Fernández MgSc², Patricia Ordoñez PhD³, and Cleva Villanueva, MD PhD⁴

¹University of California-Davis, USA, jreynoso@ucdavis.edu, orcid: 0000-003-5285-6295, ²Sociedad Argentina de Terapia Intensiva, Argentina, ariel.fernandez@hardineros.com, orcid: 0000-0002-4438-8743, ³University of Maryland Baltimore County, USA, patti.ordonez@umbc.edu, orcid: 0000-0002-5707-857X, ⁴Instituto Politecnico Nacional, México, gvillanueval@ipn.mx, orcid: 0000-0003-2694-5902

Abstract. Covid-19 Latam Datathon was the origin of the present study. Analyzing open databases, it was detected that in Chile the deaths of confirmed Covid-19 cases were concentrated in regions with greatest access to hospitals, which was the opposite of what was detected in other countries of Latin America. The objective was to determine the relationship between the distribution of hospitals and mortality from COVID -19 in Chile. The percentage of confirmed cases from March 2020 to August 2021 was analyzed by regions of the country. The analysis was based on the open database of the Chilean Ministry of Health. Geolocation was used to analyze the distribution of health establishments, confirmed cases and deaths. A positive correlation was observed between mortality and the number of hospitals ($R = 0.54$, $p=0.02$). It is known that in large cities there are factors such as the segregation of vulnerable groups and environmental contamination that determine the higher mortality from infectious-contagious diseases. However, in the analyzed database it was not possible to find regional information that would help in the interpretation of the results. The datathon allowed us to glimpse the global need to have open, complete, updated databases, which would make it possible to analyze the variables that determine the causes and prevent the consequences of natural disasters or pandemics such as COVID-19. This will help in the future to reduce analysis biases and to help governments to act promptly in the most vulnerable sites to reduce mortality and economic consequences.

Keywords: Geolocation, bias, Covid-19, open databases, datathon

Enseñanzas de la pandemia. Utilidad de datathones y bases de datos abiertas en el análisis de COVID-19

Jelissa Reynoso-García, PhD¹, Ariel Leonardo Fernández MgSc², Patricia Ordoñez PhD³, and Cleve Villanueva, MD PhD⁴

¹University of California-Davis, USA, jreynoso@ucdavis.edu, orcid: 0000-003-5285-6295, ²Sociedad Argentina de Terapia Intensiva, Argentina, ariel.fernandez@hardineros.com, orcid: 0000-0002-4438-8743, ³University of Maryland Baltimore, County, USA, patti.ordonez@umbc.edu, orcid: 0000-0002-5707-857X, ⁴Instituto Politécnico Nacional, México, gvillanueva@ipn.mx, orcid: 0000-0003-2694-5902

Resumen– El Latam COVID Datathon 2021 dio lugar al presente estudio. Analizando bases de datos abiertos, se detectó que en Chile las muertes confirmadas por COVID-19 se concentraban en las regiones con mayor acceso a hospitales, relación inversa a lo que sucedía en otros países de la región. El objetivo fue determinar la relación entre la distribución de hospitales y la mortalidad por COVID-19 en Chile. Se hizo un análisis por región del porcentaje de mortalidad en casos confirmados entre marzo de 2020 y agosto de 2021, utilizando la base de datos abierta del Ministerio de Salud de Chile. A través de geolocalización se analizó la distribución de establecimientos de salud, los casos confirmados y las muertes. Se observó correlación positiva ($R = 0.54$, $p = 0.02$) entre la mortalidad y el número de hospitales. Se sabe que en las grandes ciudades hay factores como la segregación de grupos vulnerables y la contaminación ambiental que condicionan la mayor mortalidad de enfermedades infecto-contagiosas. En la base de datos abierta no fue posible encontrar información regional que ayudara a la interpretación de los resultados. El datathon permitió vislumbrar la necesidad global de bases de datos abiertas, completas, actualizadas, las cuales posibiliten analizar las variables que determinan las consecuencias de desastres naturales. Esto ayudará en el futuro a disminuir los sesgos de análisis y actuar con prontitud en los sitios más vulnerables para así disminuir la mortalidad y las consecuencias económicas.

Palabras clave: Geolocalización, sesgos, COVID-19, bases de datos abiertas, datathon.

I. INTRODUCCIÓN

El Latam COVID Datathon 2021 [1], organizado por la Universidad de Puerto Rico Río Piedras y desarrollado online, permitió que los autores del presente estudio, provenientes de diferentes países, evaluarán la situación de COVID-19 en Latinoamérica. En el caso de Chile, una primera aproximación denota que los porcentajes más altos de mortalidad en los casos confirmados se concentraban en las regiones con mayor acceso a hospitales de alta complejidad [2], relación inversa a lo que sucedía en otros países de la región. Ese fue el origen del estudio.

La mortalidad por COVID-19 se ha asociado a diferentes factores incluyendo la edad, el sexo, comorbilidades, nivel socioeconómico, urbanización y contaminación ambiental [3-7].

En los Estados Unidos de América (EUA) la mortalidad por COVID-19 en las áreas rurales y urbanas ha cambiado a lo largo de la pandemia [8]. Durante la primera ola de la pandemia en EUA (febrero a marzo de 2020) la mortalidad fue mayor en las ciudades que en las áreas rurales. Esto se atribuyó a la movilidad de las personas infectadas, la llegada de portadores del virus y al retraso en la implementación de medidas de confinamiento y restricción de la movilidad [9]. Posteriormente, de abril a diciembre de 2020, la mortalidad fue mayor en las áreas rurales, a pesar de que el número de casos confirmados era menor. La mayor mortalidad fue consecuencia de la baja capacidad e infraestructura hospitalaria, particularmente la disponibilidad de ventiladores, las grandes distancias entre el área rural y los hospitales de la zona urbana, los niveles bajos de educación en el área rural, y la cobertura inadecuada de los seguros de gastos médicos [10].

En diferentes regiones de Latinoamérica se observó un panorama semejante. En Colombia la mortalidad fue mayor en áreas urbanas de marzo a octubre del 2020, lo cual se explicó porque en las grandes ciudades hay grupos con alta densidad poblacional, bajo nivel socioeconómico y acceso pobre a los servicios de salud [11]. Perú, el país que llegó a tener la más alta mortalidad internacional tuvo el número más alto de porcentaje de defunciones de casos confirmados en la costa, el área más poblada, que en las montañas, el área menos poblada [12]. En México se observó que la más alta mortalidad por COVID-19 se presentó en personas indígenas (febrero a agosto de 2020) y fuera de los hospitales. La explicación que se dio fue que esa población tiene menos acceso a los servicios de salud [13].

Otro factor que contribuye a la mayor mortalidad es la falta de agua en las zonas más pobres de las ciudades y en las áreas rurales lo que afecta la higiene [14].

El objetivo del estudio fue determinar la relación entre el número de hospitales por región y el porcentaje de mortalidad en casos confirmados por COVID-19 en Chile.

II. MATERIAL Y MÉTODOS

En septiembre de 2021 se analizó la base de datos abierta reportada en la página web del Ministerio de Salud de Chile [15], en la que se mostraban por región y por cada diez mil habitantes, los casos confirmados de COVID-19 y la mortalidad

Digital Object Identifier: (only for full papers, inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).
DO NOT REMOVE

de casos confirmados, de marzo de 2020 a agosto de 2021. Se calculó el porcentaje de mortalidad por región.

Se utilizaron técnicas de análisis espacial para analizar la distribución de los casos, la mortalidad y la localización de los establecimientos sanitarios con capacidad de hospitalización. Los límites administrativos se obtuvieron de la página web “Chile Subnational Administrative Boundaries” [16]. Se incluyeron las 16 regiones que mostraba datos de casos confirmados y defunciones por COVID-19, las islas del oeste se omitieron. Los datos geoespaciales se transformaron en una tabla incluyendo los datos de COVID-19 por región. Posteriormente, se creó una base de datos a partir de la red abierta del Ministerio de Salud y del Gobierno de Chile [17]. Sólo se incluyeron los hospitales, los centros de atención primaria, los centros de vacunación y las clínicas de consulta externa, debido a que se consideraron sitios a los que pudieron haber acudido los pacientes con COVID-19. Se excluyeron las clínicas dentales y los laboratorios.

Las coordenadas (latitud y longitud) de cada centro de salud incluido en el estudio se transformaron en objetos de características simples (SF, por sus siglas en inglés). Subsecuentemente, se crearon mapas utilizando los paquetes tidyverse, ggplot2, readxl del programa R [18]. La correlación entre los establecimientos de salud y el porcentaje de muerte se hizo con el programa Prism 9 [19].

Por último, se utilizaron los datos de establecimientos de salud en Chile con infraestructura hospitalaria adecuada para el tratamiento de pacientes con COVID-19 [20] para realizar un segundo análisis de correlación entre los porcentajes de mortalidad de pacientes confirmados por el Ministerio de Salud de Chile y los centros de salud optimizados para atender pacientes en estado crítico, utilizando los datos acumulados más recientes.

III. RESULTADOS

Se incluyeron un total de 3714 establecimientos de salud con capacidad de atender pacientes con COVID-19. En la tabla I se muestra la población, casos confirmados y muertes por COVID-19 en cada región, de marzo de 2020 a agosto de 2021. En la figura 1 se muestran las 16 regiones de Chile. En la figura 2 se muestran la densidad de población y distribución de establecimientos de salud. La figura 3 muestra el mapa coroplético de Chile con los casos confirmados, la distribución del porcentaje de muerte de esos casos y distribución de hospitales. En la figura 4 se muestra la correlación significativa entre el porcentaje de mortalidad y el número de establecimientos de salud.

Se consultaron los datos más recientes reportados por el Ministerio de Salud [12, 15] y Global Health Intelligence [17]. Se incluyeron 360 hospitales acondicionados con 1819 ventiladores para recibir pacientes con COVID-19. En la figura 5 se muestra la correlación entre los hospitales y el porcentaje de mortalidad (A) y entre los ventiladores y el porcentaje de mortalidad (B) por región.

Ambas fueron correlaciones positivas, es decir, más porcentaje de mortalidad en las regiones donde hay más hospitales y más ventiladores. Se corrobora lo que se encontró en el análisis inicial (figura 4), donde se asociaron los lugares con más hospitales con el mayor porcentaje de mortalidad por COVID-19 de casos confirmados.

TABLA I
POBLACIÓN, CASOS Y MUERTES CONFIRMADAS POR COVID-19 EN CADA REGIÓN DE CHILE

Regiones de Chile	Casos confirmados por cada 10,000 habitantes	Porcentaje de mortalidad de casos confirmados	Población
Antofagasta	8504.1	1.99	700074
Arica y Parinacota	10429.3	1.71	280642
Atacama	846.0	1.18	314738
Aysén Región del Gral. Ibáñez del Campo	7558.9	0.87	106484
Coquimbo	5997.0	1.93	847829
La Araucanía	9370.0	1.42	10288887
Los Lagos	9773.2	1.51	892830
Los Ríos	11854.2	1.26	407070
Magallanes y Chile Antártico	16155.0	1.65	180219
Ñuble	6875.7	1.75	511773
Valparaíso	6063.0	2.99	1979713
Tarapacá	10384.4	1.93	387042
Bio-Bio	8847.3	1.49	1899348
Libertador Bernardo O'Higgins	6756	2.16	996018
Maulé	8751	1.67	1135482
Santiago Región Metropolitana	7914.2	2.70	8090017

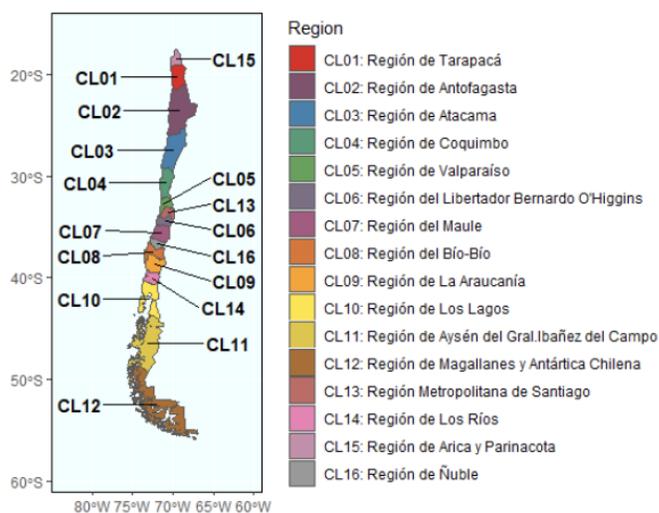


Figura 1. Las 16 regiones de Chile

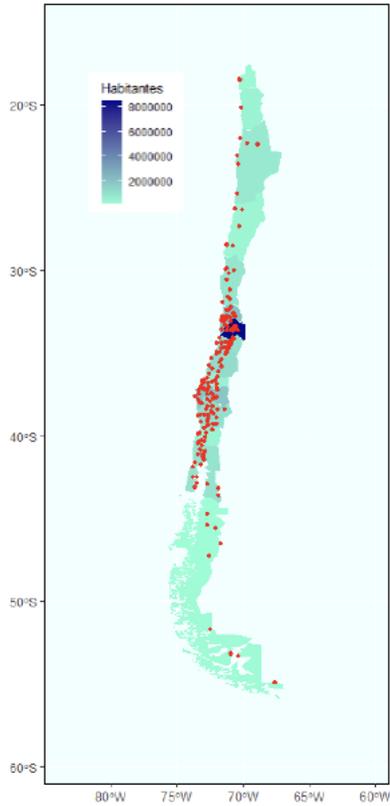


Figura 2. Densidad de población en colores y distribución de los establecimientos de salud

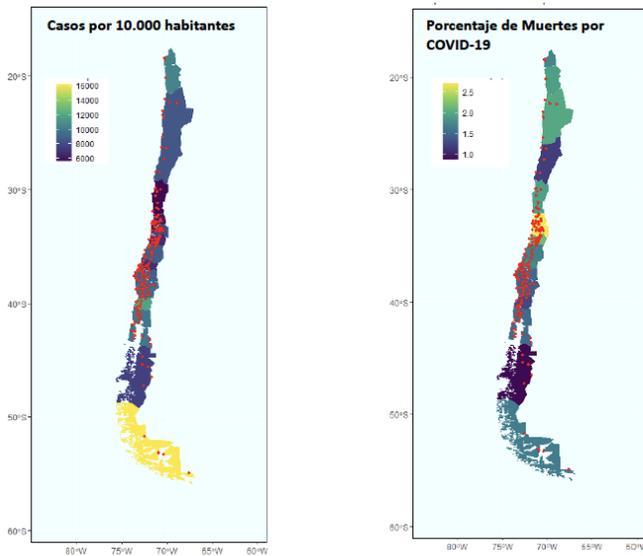


Figura 3. Mapa coroplético de Chile con casos confirmados por 10,000 habitantes y porcentaje de muertes confirmadas por COVID-19. Los puntos rojos representan los hospitales

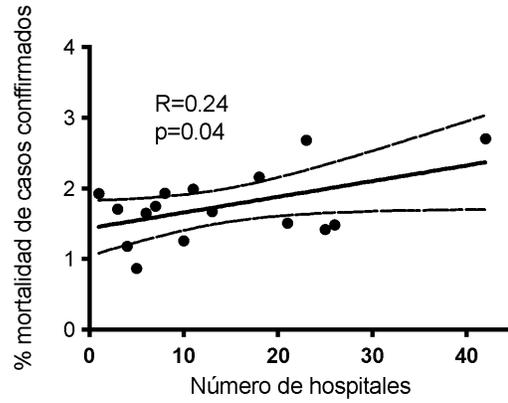


Figura 4. Correlación significativa ($p=0.04$) entre el porcentaje de mortalidad de casos confirmados y el número de hospitales de marzo de 2020 a agosto de 2021. Se presentan el coeficiente de correlación R y los intervalos de confianza del 95%.

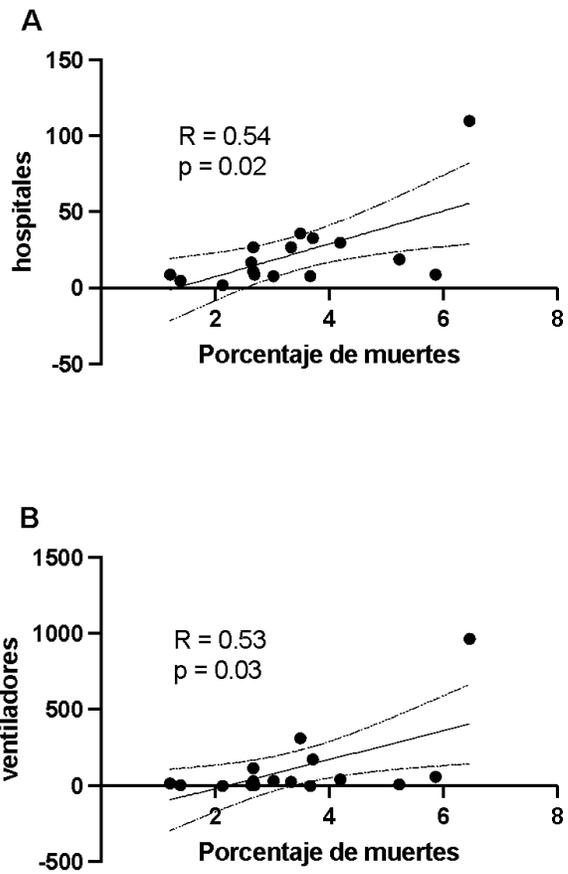


Figura 5. Correlación significativa entre el porcentaje de muertes por COVID-19 y: 1. El número de hospitales (A) y 2. El número de ventiladores (B). Se presentan el coeficiente de correlación R y los intervalos de confianza del 95%.

Finalmente, en la figura 6 se muestra la correlación positiva entre el número de habitantes y el número de hospitales (A) y

entre el número de habitantes y el porcentaje de muertes (B). Lo que sugiere que el porcentaje de muertes fue mayor en los sitios con más hospitales porque la población es mayor.

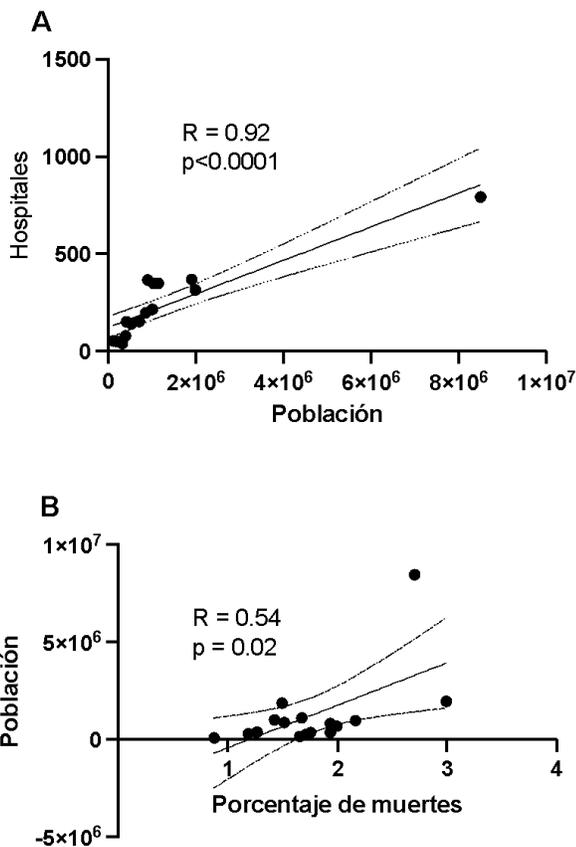


Figura 6. Correlación significativa entre: 1. La población de cada región y el número de hospitales (A) y 2. La población y el porcentaje de muertes por COVID-19 en cada región (B). Se presentan el coeficiente de correlación R y los intervalos de confianza del 95%

IV. DISCUSIÓN Y CONCLUSIONES

La pandemia de COVID-19, que aún estamos viviendo, probablemente es una de las más prolongadas de la historia de la humanidad. No sólo ha afectado la salud de la población, ha incidido en todas las actividades humanas y todavía no tenemos una visión clara de las consecuencias a largo plazo.

El análisis de datos realizado durante el Datathon permitió describir los factores relevantes para la mortalidad confirmada por COVID-19 en América Latina y observar un patrón diferente en Chile.

En general, en las grandes ciudades se encuentran áreas de población segregada cuya mayor vulnerabilidad probablemente se debe, entre otros factores, al bajo poder socioeconómico. Con menor acceso a infraestructura sanitaria y menor posibilidad de recibir atención médica de calidad. Además, la mayor parte de la población en las grandes ciudades está expuesta a mayor contaminación ambiental, lo que se ha

relacionado con mayor mortalidad por COVID-19 no sólo en las grandes ciudades de EUA [22], sino también en ciudades de países Latinoamericanos, incluyendo Chile [23].

El principal hallazgo del presente estudio es que el porcentaje de mortalidad por COVID-19 en Chile se dio en los sitios con más hospitales, los cuales se localizan en las regiones más pobladas. Se decidió utilizar el porcentaje en lugar del número total de defunciones de casos confirmados porque es más representativo cuando se trata de asociar factores que influyen en la mortalidad.

Asimismo y hasta donde sabemos, este es el primer estudio en el que se utiliza georreferenciación y mapa coroplético para mostrar la distribución de mortalidad en pacientes confirmados con COVID-19 y su correlación con los establecimientos de salud en Chile.

Es importante destacar que los datos fueron obtenidos de las bases abiertas del Ministerio de Salud de Chile [15], en las que no fue posible obtener más información para explicar los hallazgos. En esa base de datos hay información de comorbilidad nacional que constituye un factor de riesgo de mortalidad, pero no hay información de la comorbilidad de los pacientes con COVID-19, a nivel regional. Esto es común en las bases de datos abiertas, donde se encuentra información general pero no regional. Por ejemplo, en las bases de datos abiertas de Perú [24] y México [25], los países de América con mayor porcentaje de mortalidad de casos confirmados con 4.90% y 4.5%, respectivamente y cuyas tasas fueron calculadas a partir de los datos obtenidos de la Organización Mundial de la Salud [26], se muestra información regional de mortalidad pero no se vierte la información de la morbilidad de los casos regionales, lo cual ayudaría a detectar las zonas más vulnerables. En Brasil se hizo una aproximación importante: se comparó la región norte con la centro-sur. En el norte la mortalidad fue de 82%, mientras que en centro-sur fue de 18%. En ambas regiones la comorbilidad de los pacientes con COVID-19 fue semejante, la diferencia se dio a nivel racial. La mayor mortalidad fue en la población negra y mestiza que es la que habita en el norte del país, mientras que fue menor en la blanca, la cual habita predominantemente la región centro-sur. La etnicidad se asoció al poder socioeconómico, que es mayor en el norte del país. Los autores indicaron que es necesario que las autoridades gubernamentales enfoquen el apoyo en la región más vulnerable (el norte) [27].

Otro factor que influye en la elevada mortalidad en los sitios con más hospitales (grandes ciudades) es el aislamiento entre regiones. Los pacientes atendidos en un establecimiento sanitario de primer nivel pueden evolucionar sin disponibilidad de unidades de terapia intensiva o apoyo logístico de tercer nivel. Cuando son trasladados se encuentran en condiciones más severas y muerte en los hospitales. En las bases de datos abiertas no se muestra la procedencia de los pacientes que mueren a nivel regional, por lo que no es posible identificar la circunstancia que se menciona.

En un estudio realizado en Tijuana, ciudad localizada en el borde del norte de México, y que hace frontera con San Diego,

EUA, se detectó un aumento de mortalidad por COVID-19 fuera de los hospitales o a la llegada a las salas de urgencias. La explicación fue que, en las regiones alejadas de los hospitales, los pacientes no dan importancia a la hipoxemia cuando acuden al establecimiento de salud más cercano. Evolucionan a condición severa que se agrava durante el traslado [8].

Más allá de los resultados científicos que se pudieron obtener, es fundamental destacar el valor que representó el Datathon para esta investigación. Toda la información analizada fue producto de consultas de bases de datos abiertas y publicadas con acceso libre en Internet. Además, nos permitió generar nuestras hipótesis de manera preliminar y nos dio las primeras herramientas para explorar más a fondo el conjunto de datos. Finalmente y más importante aún, el Datathon permitió que los miembros de este equipo interdisciplinario de investigadores provenientes de distintas partes de América Latina (Puerto Rico, Argentina, Colombia y México), puedan intercambiar diferentes enfoques y ópticas para abordar el problema desde una óptica y perspectiva diferentes. En nuestro caso, la colaboración entre diversos investigadores con conocimientos complementarios e intereses comunes en beneficio de la salud de América Latina junto con la posibilidad de acceder a grandes bases de datos públicas, permitió generar nuevos conocimientos, como los resultados presentados en este trabajo. Otra más de tantas enseñanzas que nos dejó la pandemia.

La pandemia ha mostrado la importancia de la adecuada gestión de los datos para evitar sesgos, duplicidad, demora de registros, falta de precisión y calidad de los datos. La publicación de datos abierta falla en términos de apertura, accesibilidad y transparencia. Generalmente se presentan desorganizados, con poca granularidad y trazabilidad temporal. A su vez refuerza la necesidad de disponer de investigadores de ciencia de datos que tengan formación óptima no solo académica, sino ética y con conocimiento del terreno, para gestionar la información de manera adecuada. Y principalmente ha colocado en el foco del panorama político internacional la importancia de la información pública. Los datos utilizados por los gobiernos para informar sobre el avance del COVID-10 y justificar medidas como el confinamiento social preventivo y obligatorio han resaltado las carencias de algunos sistemas de información relacionados con la gestión sanitaria. Durante el transcurso de la pandemia, se han presenciado retrasos en la actualización de datos, duplicación de registros e incluso faltas de precisión en la interpretación de los mismos [29]. Los datos publicados reflejan más la oferta de datos disponibles que la demanda por los temas específicos.

Los sesgos en la información y su manejo, en la mayor parte de los casos resultado de la ignorancia o inoperancia en menor parte producidos intencionalmente, han afectado diferentes ámbitos incluyendo la economía. Un análisis realizado recientemente, evidenció la labilidad del mercado de valores, sobre todo cuando la diseminación de la enfermedad fue rápida y en los momentos que se perdió control de la pandemia. Los modelos económicos se nutrieron de fuentes inciertas, lo que afectó prácticamente el mercado global [30].

Otro ejemplo es la indecisión para recibir vacunación, considerada por la Organización de la Salud como una de las diez principales amenazas a la salud global [31]. En China por ejemplo, se realizó una encuesta a más de 3000 personas y se llegó a la conclusión de que la comunicación mediática tenía una influencia fuerte en la decisión de recibir los refuerzos de la vacuna. Los medios de comunicación utilizados por la población desempleada, con bajos nivel de educación y nivel socioeconómico, influyeron para que esa población rechace a recibir el refuerzo de la vacuna contra COVID-19, lo que demuestra las consecuencias que genera la desinformación mediática de sociedad en la pandemia [32].

Estas situaciones demuestran que simplemente tener acceso a la información pública no es suficiente. El acceso a la información pública es uno de los derechos esenciales para ejercer una ciudadanía activa en un estado democrático. Se han puesto en marcha agencias, mecanismos de coordinación y se han desarrollado capacidades para gestionar los datos abiertos en varios países de Latinoamérica [33], sin embargo, aún queda mucho por hacer. Si la información pública es un componente vital, la gestión de datos de alta calidad es todavía una deuda esencial para ampliar el derecho y mejorar la toma de decisiones.

El Latam COVID Datathon 2021 [1] permitió concluir, a través del análisis de la información obtenida de las bases de datos abiertas de Chile, que el mayor porcentaje de mortalidad de casos confirmados de COVID-19 se dio en las regiones con más hospitales y que a su vez son las más pobladas de ese país. Es importante destacar que el elemento de análisis fue de porcentaje, no del número de defunciones. Y que los fallecidos se registraron en el lugar donde ocurrieron, sin posibilidad de conocer la región de origen de esos pacientes. Con los datos obtenidos no se puede llegar a una conclusión tácita, solo se puede postular, con base en lo que se ha reportado en la literatura internacional, que los posibles factores involucrados fueron: 1) traslado de pacientes con COVID-19, provenientes de lugares con menor infraestructura y atención médica, que llegaron a los hospitales en condición avanzada y crítica con poca probabilidad de sobrevivir, 2) la falta de registro del lugar de residencia de los pacientes que murieron en hospitales, 3) retraso en la actualización de los datos debida a la abundante información generada durante la pandemia.

Otra conclusión importante del Datathon es la que conduce a la necesidad global de contar con bases de datos abiertas, completas, actualizadas, curadas y con información a nivel regional, que puedan ser consultadas a nivel global de forma libre. Los datos abiertos que producen los Estados frecuentemente son complicados de interpretar, descargar y manipular. La conformidad (cumplimiento con normas aceptadas), la credibilidad (seguridad en las fuentes de los datos), la accesibilidad (facilidad de lectura), y la relevancia (valor, importancia y cantidad adecuada de los datos) son los aspectos clave de la gestión de datos [34] que deben ser considerados al momento de crear repositorios públicos de información, y sobre todo con datos de salud.

Entrevistas con expertos en áreas sociales, académicas y económicas de Chile confirmaron nuestras conclusiones. También mencionaron que los datos abiertos no siempre muestran la “situación real” porque los datos no se actualizan en “tiempo real”. Sugieren que el análisis de los datos disponibles requiere recopilar información en el momento en que ocurre, de manera auditable y confiable, para validar y explicar los resultados dentro del contexto [35].

Es necesario crear una agenda digital compartida de todos los países de Latinoamérica que incluya reformas para mejorar la calidad de los datos abiertos publicados por los Gobiernos. Se requieren cambios en tecnología, instituciones y participación ciudadana. Pero sobre todo Gobernanza de datos: planificación, supervisión y control en la gestión y el uso de datos [36]. Esto ayudará a que en un futuro sea posible ampliar el análisis de nuestra investigación, incluyendo otras fuentes de datos confiables y abiertos, permitiendo que otros investigadores puedan seguir explorando nuevas posibles causas y consecuencias de las observaciones descritas en este estudio. Pero principalmente permitirá que los Gobiernos puedan analizar las variables que determinen las causas y prevengan las consecuencias de desastres naturales o pandemias como el COVID-19, para diseñar estrategias efectivas dirigidas al cuidado de la salud en las poblaciones más vulnerables.

REFERENCIAS

- [1] Latam COVID Datathon [Internet], [citado 5 sep 2022]. Disponible en: <https://www.latanmdatathon.hpcf.upr.edu/>
- [2] <https://www.misal.cl/nuevo-coronavirus-2019-ncov/casos-confirmados-en-chile-covid-19>
- [3] O. Dadrás, S. Seyed Alinaghi, A. Karimi, A. Shamsabadi, K. Qaderi, M. Remezani, et al. “COVID-19 mortality and its predictors in the elderly: A systematic review.” *Health Sci. Rep.*, vol. 5, pp. 1-15, Junio 2022. Doi: 10.1002/hrs2.657.
- [4] A.C. Danielsen, K.M.N. Lee, M. Boilicault, T. Rushovich, A. Gompers, A. Tarrant, et al. “Sex disparities in COVID-19 outcomes in the United States: Quantifying and contextualizing variation.” *Social Sci. Med.* vol. 294 pp. 1-13, Enero 2022. Doi: 10.1016/j.soscimed.2022.11476.
- [5] V.J. McGowan, C. Bamba. “COVID-19 mortality and deprivation: Pandemic, syndemic, and endemic health inequalities.” *Lancet Pub. Health* vol. 7 pp e966-3975. Noviembre 2022. Doi: 10.1016/S2468-2667(22)00223-7.
- [6] Y. Fan, M. Fang, X. Zhang, Y. Yu. “Will the economic growth benefit public health? Health vulnerability, urbanization and COVID-19 in the USA.” *Ann. Reg. Sci.* pp. 1-19. Enero 2022. Doi: 10.1007/s00168-021-01103-9.
- [7] A. Belconi, P. Vounatsou. “Long-term air pollution exposure and COVID-19 case severity: An analysis of individual-level data from Switzerland.” *Environ. Res.* Vol. 216 pp. 2023. Doi: 10.1016/j.envres.2022.114481.
- [8] D.F. Cuadros, A.J. Branscum, W. Mukadavire, F.D. Miller, N. MacKinnon. “Dynamics of the COVID-19 epidemic in urban and rural areas in the United States.” *Ann. Epidemiol.* vol. 59 pp. 16-20. Julio 2021. Doi: 10.1016/j.anaepidem.2021.04.007.
- [9] F. Natale, s.M. Lacus, A. Conte, S. Spyrtos, F. Semi. “Territorial differences in the spread of COVID-19 in European regions and US counties. arXiv: 210308321 [physics, q-fin, stat] [internet]. Marzo 2021. [citado Mayo 2, 2022]: disponible en: <http://arxiv.org/abs/2103.08321>.
- [10] R. Ahmed, H. Williams, M.A. Hamid, N. Ashraf. “United States county-level COVID-19 death rates and case fatality rates vary by region and urban status.” *Healthcare* vol. 8 no. 3, pp. E330. Septiembre 2020. Doi: 10.3390/healthcare8030330.
- [11] M.P. Cifuentes, I.A. Rodríguez-Villamiz, M.L. Rojas-Botero, C.A. Álvarez Moreno, J.A. Fernández-Niño. “Socioeconomic inequalities associated with mortality for COVID-19 in Colombia: a cohort nationwide study. *J. Epidemiol. Comm. Health* vol. 75. no. 7 pp.610-615. Doi: 10.1136/jech-2020-216275.
- [12] M.G.F. López, A.S. Tarazona, J.A.D.I. Cruz-Vargas. “Distribución regional de mortalidad COVID-19 en Perú.” *Revista de la Facultad de Medicina Humana* vol. 22, no. 2 [internet]. Febvreo 2021. [citado el 2 de mayo de 2022]. Disponible en: <https://revistas.urp.edu.pe/index.php/FRMH/article/view/3721>
- [13] I. Ibarra-Nava, K.G. Flores-Rodríguez, V. Ruiz-Herrera, H.C. Ochoa-Bayona, A. Salinas-Zertuche, M. Padilla-Orozco, et al. “Ethnic disparities in COVID-19 mortality in Mexico: A cross-sectional study based on national data.” *PLoS One* vol. 6 no. 3 pp. e0239168, 2021.
- [14] Caribbean EC for LA and the COVID-19 and social protection for poor vulnerable groups in Latin America: a conceptual framework [internet]. CEPAL 2020 [citado septiembre 2022]. Disponible en: <https://www.cepal.org/en/publications/46939-covid-19-and-social-protection-poor-and-vulnerable-groups-latin-america>.
- [15] Departamento de Estadísticas e Información en Salud [internet]. [citado 5 sep 2022]. Disponible en: <https://deis.minsal.cl/>.
- [16] Chile – Subnational Administrative Boundaries – Humanitarian Data Exchange [internet]. [citado 5 sep 2022]. Disponible en: <https://data.humdata.org/dataset/cod-ab-chl>.
- [17] Base de Datos COVID-19 [internet]. Miniciencia.gob.cl. 2021 [citado 3 ago 2022]. Disponible en: <https://miniciencia.gob.cl/covid19>.
- [18] R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- [19] Graphpad Software. LLC. San Diego California. www.graphpad.com.
- [20] Interferencia. 5 regiones de Chile tienen infraestructura hospitalaria crítica contra el COVID-19 comparable con Sierra Leona o Guatemala [internet]. [citado 5 sep 2022]. Disponible en: <https://interferencia.cl/articulos/5-regiones-de-chile-tienen-infraestructura-hospitalaria-critica-contra-el-covid-19>.
- [21] www.covid19.who.int.
- [22] M. Jerret, C.L. Nau, D.R. Young, R.K. Butler, C.M. Batterate, J. Su, et al. “Air pollution and meteorology as risk factors for COVID-19 death in a cohort from South California.”. *Environ. Int.* vol. 171 pp 1-11. 2023. Doi: 10.1016/j.envint.2022.107675.
- [23] J.A. Bonilla, A. López-Feldmann, P.C. Pereda, N.M. Rivera, and J.C. Ruiz-Tagle. “Association between long-term air pollution exposure and COVID-19 mortality in Latin America.” *PLoS One* vol. 18, no.1 pp. 1-21. Enero 2023. Doi: 10.1371/journal.pone.0280355.
- [24] Ministerio de Perú. Mortalidad COVID-19 [internet] [citado 5 sep 2022]. Disponible en: https://dge.gob.pe/portalnuevo/informacion_publica/ortalidad-covid-19.
- [25] Sistema Nacional de Información Estadística y Geografía de México. Perspectiva COVID-19 [internet]. [citado 5 sep 2022]. Disponible en: <https://www.inegi.org.mx/investigacion/covid/ligas.html>.
- [26] OMS. Dashboard COVID-19 [internet]. [citado 5 sep 2022]. Disponible en: <https://covid19.who.int>.
- [27] P. Baqui, I. Bica, V. Marra, A. Ercole, M. van der Shaar. “Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study.” *Lancet Global Health* vol. 8, no. 8 pp. e1018-1026. Agosto 2020. Doi: 10.1016/S2214-109X(20)30285-0.
- [28] J. Friedmann, A. Calderón-Villareal, I. Bojorquez, C. Vera-Hernández, D.I. Schrier, E. Tovar Hirashima. “Excess out-of-hospital mortality and declining oxygen saturation: The sentinel role of emergency medical services data in the COVID-19 crisis in Tijuana. Mexico. *Ann. Emerg. Med.* vol. 6, no. 4, pp. 413-426. Octubre 2020. Doi: 10.1016/j.annemergmed.2020.07.035.
- [29] Comisión Económica para América Latina y el Caribe (CEPAL). Los impactos sociodemográficos de la pandemia de COVID-19 en América Latina y el Caribe (LC/CRPD.4/3), Santiago, 2022.
- [30] L. Xu, X. Zhang, J. Zhao. “Limited investor attention and biased reaction to information. Evidence from the COVID-19 pandemic.” *J. Financial Markets* vol. 62, pp. 1-16, 2023. Doi: 10.1016/j.finmar.2022.100757.
- [31] D. Bussnik-Voored, J.L.A. Hautvast, L. Vanderbert, O. Visser, M.E.J. Hulsher. “A systematic literature review to clarify the concept of vaccine

hesitancy.” *Nature Human Behavior* vol. 6, pp. 1634-1648. Diciembre 2022. Doi: 10.1038/s41562-022-01431-6.

[32] R. Wang, C. Qin, M. Du, Q. Liu, L. Tao, J. Liu. “The association between social media use and hesitancy toward COVID-19 vaccine booster shots in China: A web-based cross-sectional survey.” *Human vaccines & Immunother.* vol. 18, no. 5, pp.1-10. Junio 2022. Doi: 10.1080/21645515.2022.2065167.

[33] Barómetro Global de Datos (2022). Primera Edición Informe – Doi: 10.5821/zenodo.6488349.

[34] Comisión Económica para América Latina y el Caribe (CEPAL). Metodología del marco lógico para la planificación, el seguimiento y la evaluación de proyectos y programas. Hugo Navarro (LC/L.22188-P;LC/IP/L.254), Santiago, 2005.

[35] C. Hullin, A. García, S. Hartel (comunicación personal, 178 enero, 2022).

[36] A. Naser (coord.), “Gobernanza digital e interoperabilidad gubernamental: una guía para su implementación.” Documentos de Proyectos (LC/TS.2021/80), Santiago, Comisión Económica para América Latina y el Caribe (CEPAL), 2021.

AGRADECIMIENTOS

Carol Hullin, Alejandra García y Steffen Hartel por compartir sus opiniones de este artículo a través de sus importantes visiones particulares y experiencias de trabajo en el terreno de Chile.

Grant NIH. Increasing diversity in interdisciplinary BD2K (IDI-BD2K).

Grant CONACYT-México 311866.