Rev Med Chile 2023; 151: 823-829

Incidence moments: Short term forecast of the COVID-19 incidence rate in Chile

MAURICIO CANALS L.[1], ANDREA CANALS C.[2,3]

[1]Programa de Salud Ambiental, Escuela de Salud Pública, Facultad de Medicina, Universidad de Chile. Santiago, Chile.
[2]Dirección Académica, Clínica Santa María. Santiago, Chile.
[3]Programa de Bioestadística, Escuela de Salud Pública, Facultad de Medicina, Universidad de Chile. Santiago, Chile.

Interest Conflict None declared.

Recibido el 24 de marzo de 2022, aceptado el 12 de junio de 2023.

Corresponding author:
Mauricio Canals L.
Programa de Salud Ambiental,
Escuela de Salud Pública,
Facultad de Medicina,
Universidad de Chile.
Santiago, Chile.
mcanals@uchile.cl

Introduction: *The predictability of any epidemic is highly uncertain, especially regarding a new emerging pathogen such as SARS-CoV-2.* ***Objectives:*** *We studied the predictability of the incidence series of COVID-19 in Chile (whole country) and three regions with different population sizes. The analysis included the period intervened by vaccination campaigns and when new variants of the SARS-CoV-2 virus arrived. This study also is focused on possible differences in predictability between epidemic and "inter-epidemic" periods.* ***Methods:*** *We studied the daily incidence of COVID-19 in Chile for the Metropolitan, Biobío, Arica, and Parinacota regions from March 2020 to February 2022, with the recently proposed method of the third moment of incidence. We assessed the predictive capacity with the corrected mean arctangent absolute percent error.* ***Results:*** *The predictability of the daily incidence of COVID-19 was on the limit between good and reasonable for the entire epidemic process. The third moment of incidence produced reasonable predictions for regions with large population sizes and insufficient predictions for smaller regions. We found lower prediction capacity during the start of the pandemic and the epidemic caused by the Omicron variant.* ***Conclusion:*** *The third incidence moment method is suitable for short-term forecasting of COVID-19 with an error of around 30%. This forecast represents a short time of predictability in mainly chaotic dynamics. The predictability decreased only slightly due to pharmacological interventions and the income of new virus variants. We found low predictability in the initial periods of the epidemic and during the Omicron epidemic outbreak.*

(Rev Med Chile 2023; 151: 823-829)

Key words: *COVID-19; Forecasting; Incidence.*

Momentos de Incidencia: Pronóstico a Corto Plazo de la Tasa de Incidencia de COVID-19 en Chile

Introducción: *La predictibilidad de cualquier epidemia es muy incierta, y más aún cuando se trata de un nuevo patógeno emergente como el SARS CoV-2.* ***Objetivos:*** *Estudiamos la predictibilidad de series de incidencia de COVID-19 en Chile y en tres regiones de distinto tamaño poblacional, incluyendo el período de intervención de las campañas de vacunación y llegada de variantes del virus SARS CoV-2. Estudiamos también las posibles diferencias en la predictibilidad entre los períodos epidémicos e "inter-epidémicos".* ***Métodos:*** *Se estudió la incidencia diaria de COVID-19 en Chile, para las regiones Metropolitana, Biobío y Arica y Parinacota de marzo de 2020 a febrero de 2022 con el método del tercer momento de incidencia propuesto recientemente. La capacidad de predicción se evaluó con el arcotangente del error porcentual absoluto medio* *corregido.* ***Resultados***: *La predictibilidad de la incidencia diaria de COVID-19 estuvo en el límite entre buena y razonable durante todo el proceso epidémico. El tercer momento de incidencia produce predicciones razonables para regiones con gran tamaño de población y predicciones insuficientes para regiones más pequeñas. Se encontró menor capacidad de predicción durante el inicio de la pandemia y durante la epidemia provocada por la variante Omicron.* ***Conclusión:*** *El método del tercer momento de incidencia es un buen método para el pronóstico a corto plazo de COVID-19 que tiene un error de alrededor del 30%. Esto representa un tiempo corto de predictibilidad en una dinámica principalmente caótica. La predictibilidad disminuyó solo levemente debido a las intervenciones farmacológicas y al ingreso de nuevas variantes del virus. Se encontró baja predictibilidad en los períodos iniciales de la epidemia y en el brote epidémico de Omicron.*

Palabras clave: *Incidencia; Predicción; COVID-19.*

In confronting the COVID-19 pandemic, countries have tried to base their political response using prediction tools to determine the magnitude of the epidemic and the capacity of their health systems to respond to the demands for healthcare[1]. The quality of the response to the epidemic will depend on the ability to monitor and predict the behavior of the epidemic and to extract useful information to implement timely effective public health interventions[2]. However, the predictability of any epidemic is highly uncertain, and even more so with a new emerging pathogen[3,4]. Even with complex models, the predictability has proved to be limited in the long term[5,6]. This is due in part to multiple environmental and social factors and demographic differences that affect the parameter estimation, and also to endogenous factors of the case-susceptible interaction that are essentially non-linear, reducing the predictability of the disease dynamics. The dynamics of COVID-19 has been very difficult to predict, with sudden and unexpected changes in the incidence in different countries. Recent studies have proposed that the dynamics of COVID-19 shows chaotic behavior that would complicate the prediction of its dynamics[5-10].

Canals et al (2022b)[10], recently proposed a method to study the temporal memory of the time series of the daily incidence of the world and of 44 countries during 2020 and the beginning of 2021, proposing the concept of incidence moments. This method is based on the successive products between the mean incidence (I) and the effective reproductive number (R). The best predictor was determined to be the third moment of incidence (IR^3) in a short temporal prediction window of 15 days; the prediction capacity dropped significantly after 15 days. These authors proposed that this is due to an intrinsic chaotic behavior in the dynamics of COVID-19. However, they also proposed that the third moment of incidence could be used for short-term prediction of COVID dynamics, with rapid implementation in different settings, even with limited epidemiological technical capability and without requiring a large amount of computational data.

The pandemic has strongly affected the population in Chile; the first year (March 2020-February 2021) saw a strong epidemic in the Metropolitan region, a period of relative tranquility with incidences of less than 10 cases/one hundred thousand inhabitants from August to December 2020 and a rise from December, 2020 to February, 2021. During this first year the interventions were non-pharmacological, mainly quarantines, suspension of educational activities, national curfew, social distancing and basic sanitary measures[11]. During this period no systematic genotyping of viral variants was performed. The second year of the pandemic was different with a very successful vaccination campaign, with high coverage that was quickly achieved. Vaccination began in February, 2021, the second dose (full vaccination) began to be administered in March, adding the first third booster dose since August and a fourth dose since December, 2021. The vaccines used in the campaign were CoronaVac from the Sinovac

Life Sciences laboratory (60.9%), BNT162b2 from Pfizer-BioNTech (29.6%), ChAdOx1-S -recombinant from AstraZeneca (7.9%) and Ad5.nCoV recombinant from CanSino Biologics Inc. (1.5%). This period included a large outbreak that encompassed all the territory during the first semester where the dominant variant was γ (P.1, VOC) and outbreaks associated with the emergence of the delta variants (δ; B.1.617.2, VOC) and the omicron variant (O, VOC)[12].

This study has two objectives. First, to study the predictive capacity of the method of the third moment of incidence (IR3) in Chile and in three regions of different population size, including the period strongly intervened by vaccination campaigns and in which the γ, δ and O variants of SARS CoV-2 virus entered successively in our country. Second, to study possible differences in predictability between epidemic and "inter-epidemic" periods.

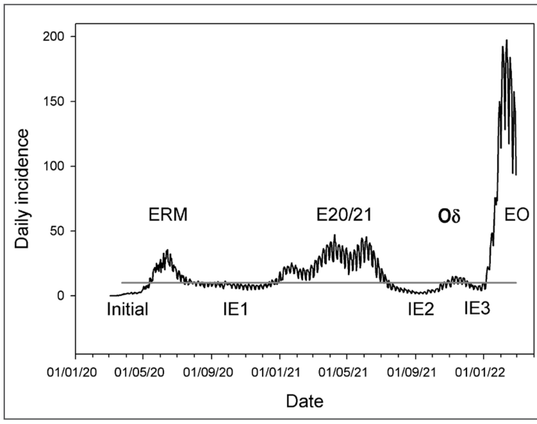

Figure 1. Epidemic and inter-epidemic periods in the evolution of the COVID-19 pandemic in Chile. Initial; initial period previous to the first epidemic outbreak, epidemic outbreak in the Metropolitan region (ERM), epidemic at the end of 2020 and first half of 2021 (E20/21), a small outbreak associated with the arrival of §-variant (O§) and a great epidemic associated with the arrival of O-variant (EO). There were three inter-epidemic periods (IE1, IE2 and IE3).

Methods

We conducted a study based on official daily public reports from the Ministry of Health of Chile, including daily new confirmed cases of COVID-19 at the national and sub-national levels (administrative regions)[13]. A period between March 24, 2020 and February 28, 2022 was analyzed for Chile, and for the Metropolitan, Bío Bío and Arica & Parinacota regions, representing large, medium and small population sizes. The incidence series of Chile was also studied, considering epidemic and inter-epidemic periods. The following criterion was used: An inter-epidemic period was defined as a period with a persistent daily incidence rate < 10 cases/100,000 inhabitants; otherwise it was considered an epidemic period. With this criterion we recognize an initial period previous to the first epidemic outbreak, four epidemic periods: the first epidemic outbreak in the Metropolitan region (ERM), a great epidemic at the end of 2020 and first half of 2021 (E20/21), a small outbreak associated with the arrival of the δ-variant (Oδ) and a great epidemic peak associated with the arrival of the O-variant (EO). Finally between the outbreaks we recognize three interepidemic periods (IE1, IE2 and IE3) between the outbreaks (Figure 1).

Model: We used the incidence moments method[10]. This is based on the following assumption: If on a day t there is I_t incidence and with an effective reproductive number R_t, then under the scenario of a population of constant size, it is expected that after a serial interval (τ), on $t + \tau$, there is an incidence I_tR_t. If also $R_t = R$ is also constant, at $t + 2\tau$ an incidence I_tR^2 would be expected, and at $t + 3\tau$ an incidence I_tR^3 etc. We used IR^3 , the 3rd moment of incidence, a good estimation of the incidence that will occur fifteen days later under the assumption of incidence and that an effective reproductive number will be constant *(R)*. Here, I is the moving average of the incidence of the last 7 days *(I = (It+It-1+… It-6)/7)*[10].

Statistical Methods: We estimate the effective reproductive number (R_t) with the RKI method[14] with a serial interval of 5 days[15-17] and we calculate IR^3, representing predictions for 15 days. We compare the observed values with those expected by the third moment using the mean arctangent absolute percentage error (MAAPE). MAAPE = 100[∑N1 arctan(| *(It-Ft)/It* |), where N is the number of data points, It and *Ft* are the current and forecasted incidence rates. The unit of measurement of this metric is in radians, so in deviations values lower than 50%, the values of MAPE and MAAPE are very similar and then this metric has been proposed as an alternative for the mean absolute percentage error (MAPE), avoiding the disadvantage that MAPE produces infinite or undefined values for zero or close to zero values of *It*[18]. We consider MAAPE values ≤ 30 as good, and values between 30-50 reasonable forecasting[10]. This analysis was carried out for all Chile and for each of the three studied administrative studied regions of Chile, and for epidemic and inter-epidemic periods for the whole incidence series of Chile.

Results

The predictability of daily incidence of COVID-19 with the third moment of incidence method was at the limit between good and reasonable for the entire epidemic process from March 2020 to February 2022; MAAPE = 32.28. We find that the third moment of incidence produces reasonable predictions for regions with large population sizes and insufficient predictions for smaller regions such as Arica (Figures 2-3). The prediction capacity decreased slightly with respect to that reported by Canals et al (2022b)[10] prior to the arrival of the γ, δ and O variants of SARS CoV-2 during the 2021/2022 period (Table 1).

Considering all the data for both periods, less than one percentage point (0.7 in MAAPE) of differences in predictability were obtained between

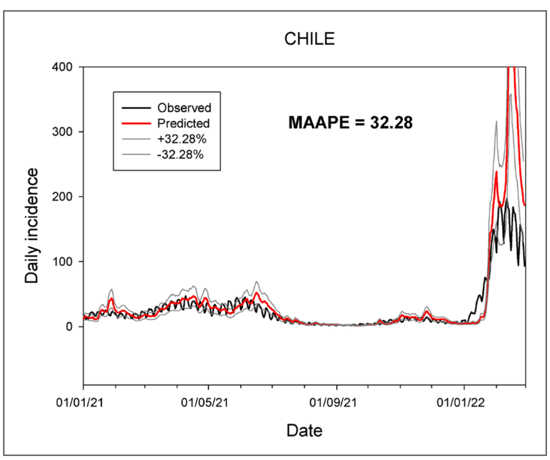

Figure 2. Evolution of the reported incidence (black) and the expected values of Chile, using the third incidence moment method (IR3) (red), from March 1, 2021 to February 28, 2022. Mean absolute percentaje error bands (with arctangent correction; i.e. MAAPE) in grey.

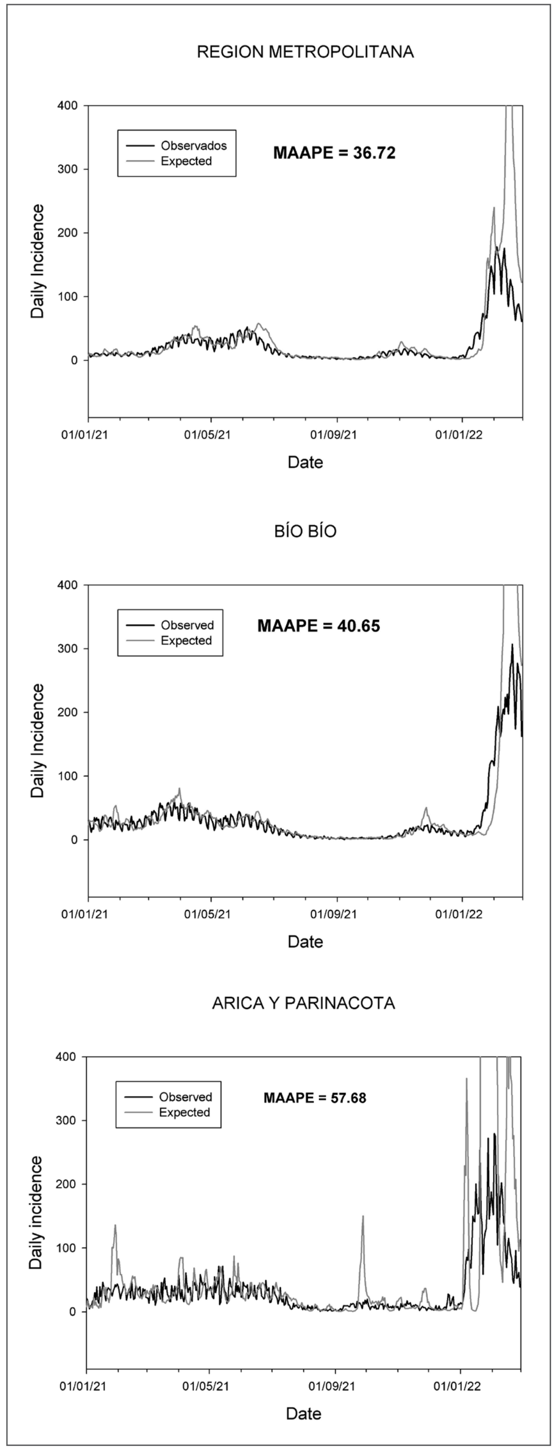

Figure 3. Evolution of the reported incidence (black) and the expected values for the Metropolitan, Biobío and Arica and Parinacota regions, by the third incidence moment method (IR3) (grey), from March 1, 2021 to February 28, 2022.

Table 1. Comparison between the mean arctangent absolute percentage error (MAAPE) between the reported and the expected incidence of COVID-19 in Chile before (IR320/21) and after the emergence of γ, δ and O variants (IR320/22), in three administrative regions of different population size (INE 2022)

	IR^3 20/21	IR^3 20/22	Population size
Arica & Parinacota	51.52	57.58	252,110
R. Metropolitana	33.06	36.72	8,125,072
Biobío	42.22	40.65	1,663,696
Chile	30.54	32.52	19,458,310

Table 2. Comparison between the mean arctangent absolute percentage error (MAAPE) between the reported and the expected incidence of COVID-19 in Chile during epidemic and inter-epidemic periods. Initial; initial period previous to the first epidemic outbreak, three interepidemic periods (IE1, IE2 and IE3). Epidemic outbreak in the Metropolitan region (ERM), epidemic at the end of 2020 and first half of 2021 (E20/21), a small outbreak associated with the arrival of the δ-variant (Oδ) and a great epidemic associated with the arrival of the O-variant (EO)

Inter-Epidemic periods	MAAPE	Epidemic periods	MAAPE
Initial	68.28	ERM	31.66
IE1	19.51	E20/21	27.10
IE2	29.41	Oδ	31.69
IE3	30.09	EO	56.72
All Inter-epidemic period	31.95	All Epidemic period	32.65

the epidemic and inter-epidemic periods. However, in a finer analysis studying the predictability of each small period, lower prediction capacity can be observed during the start of the pandemic and during the epidemic caused by the Omicron variant (Table 2).

Discussion

The incidence moments model proposes that the cases at each moment of an infectious disease do not depend on the past history (i.e. years ago) of the disease but on the number of recent cases that constitute the source of infection of the new cases. Thus, if at a given moment there are I cases, then IR cases would be expected after a serial interval τ, where R is the effective reproductive number, since this represents the number of new cases originating from the initial I. In two serial intervals IR^2 would occur, in three, IR^3 and so on[10]. It has been reported that for the COVID-19 epidemic this would be fulfilled with relatively good predictive capacity until the third serial interval, that is, this pandemic would have a memory of 3τ, which represents approximately 15 days if we consider that the serial interval of COVID-19 has been proposed to be 5 days[15-17]. This represent a short period of predictability in a chaotic dynamics. In this sense, epidemics are always restarting from potentially infective cases at every moment with a very irregular dynamic trajectory. When used for short term forecasting, this method assumes that the effective reproductive number R remains constant over a certain interval. And it is precisely because R is changing that the memory of the incidence series is being lost.

The method depends on the serial interval, which is interpreted as the average time between infection and infection[19]. Thus, the validity of the estimates will depend on an adequate choice of the serial interval. In the case of COVID-19 there are estimates of the serial interval between 1.9 and 7.5 days[15-17], but in the meta-analysis of Izadi et al (2020)[15] an average value of 4.45 days was estimated, in that of Zhang et al (2020)[16] a value of 5.35 days and 5.2 days in that of Alene et al. (2021)[17], so the value of 5 days used in this study seems to be reasonable. However, it has been reported that the serial interval can be shortened over time by non-pharmaceutical interventions[20]. Also, if emerging variants modify the infectious dynamic of the agent (e.g. shorter incubation periods or increased infectivity in pre-symptomatic period) this can have an impact on serial interval, therefore affecting the predictive capacity of the method if this parameter is specified incorrectly. For example, shorter serial intervals have been reported for the δ and O variants[21,22].

We found a decrease of approximately 2 percentage points in the predictive capacity of the method (i.e. two points increase in the MAAPE) between the period prior to the vaccination campaigns and the entry of the γ δ and O variants between March 2020 and February, 2021 and the subsequent period (March, 2021-February, 2022). During this last period in Chile we had a large epidemic (E20/21) dominated by the γ variant of SARS CoV-2 virus (P.1 VOC) that between May and June 2021 constituted more than 66% of the positive samples for SARS CoV-212, a small outbreak dominated by the δ variant (B1.617.2 VOC) that between October and November, 2021 constituted 100% of the samples, and in recent times a very important epidemic associated with the omicron variant (O VOC), which already constituted 55% of the samples in December and 94% in January, 2022. The transmission dynamics of SARS CoV-2 were intervened during this period with strong vaccination campaigns that included two initial doses which start in February and March, 2021, mainly with the CoronaVac vaccine (Sinovac), a third dose begun in August, 2021 and currently a booster dose since December, 2021. A variety of schemes have been used, combinations of vaccines of Sinovac, AstraZeneca and Pfizer-BioNTech and to a lesser extent, CanSino labs. Also during this period, since October 2021, epidemiological interventions, national curfew and restrictions on travel and population mobility were relaxed. All these factors affect the predictive capacity of the model, and yet the drop was only two percentage points, which reveals a certain average resilience of the system to changes in the exogenous factors that affect its dynamics. Population size, however, is a factor that contributes strong uncertainty in the short-term predictive capacity of this model, as previously reported[10].

When we compare the epidemic and inter-epidemic periods, there is generally only a drop of one percentage point in the predictive capacity of the method, however, in the finer analysis two periods of high unpredictability are detected. First, the initial stage, which is explained by the low number of cases and their disaggregated distribution, which implies a high uncertainty in contagion. A probably chaotic behavior with high unpredictability (MAAPE = 56.72) was found during the last epidemic outbreak associated with the O variant. This can be explained by several factors. On one hand, reasons intrinsic to the virus variants, which in addition to high transmission apparently have a shorter serial interval[22] and, on the other hand, factors associated with the reporting of daily cases. This outbreak of the Omicron variant has been very impressive, with a large number of official cases reported, with a peak of 38,446 (11/2), while at the end of 2021 the highest incidence in Chile was 2,895 (6/11). That is, the incidence was multiplied by 13 times in approximately 1 to 2 months. There were problems in the report, which can be seen in the differences between the new daily cases reported and the successive differences between total cases. Also during this period, the existence of a saturation in diagnostic tests due to the unusual increase in cases is highly probable because while the daily incidence increased 13 times, the number of test performed only increased 25%. This is a clear indication of saturation of diagnostic tests. Also, the responsibility for traceability was transferred to the patients, which has the consequence that there were many family outbreaks that were not reported. Some people did not have access to diagnostic tests and others did not consider it necessary since the source of infection and the symptoms were clear.

In conclusion, our results show that the third incidence moment method (IR^3) is a good method for short-term forecasting of COVID-19 that has a mean absolute percentage error of around 30%. This represent a short time of predictability in a mainly chaotic dynamics. That the predictability only decreased on average by one percentage point, affected by pharmacological interventions and the arrival of variants of the SARS CoV-2 virus. It also showed periods of low predictability in the initial periods of the epidemic and in periods of high incidence where there is significant underreporting of cases.

Funding The authors acknowledge the partial funding of OPS/OMS CON21-00013967 grant.

Ethics Our study is based on official secondary data reported by the Ministry of Health of Chile; approved by Faculty of Medicine ethics committee (Acta 129, 23/11/2021).

References

1. WHO-Europe. 2020. Strengthening the health system response to COVID-19. Recommendations for the WHO European Region. Policy brief (1 April 2020). Available:http://www.euro.who.int/__data/assets/pdf_file/0003/436350/strengthening-health-system-response-COVID-19.pdf?ua=1[Accesed 1 April 2020].
2. Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand". Imperial College London 2020; https://www.imperial.ac.uk/media/imperialcollege/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf.
3. Mangiarotti S, Peyre M, Zhang Y, Huc M, Roger F, Kerr Y. Chaos theory applied to the outbreak of COVID-19: an ancillary approach to decision making in pandemic context. Epidemiology & Infection 2020; 148: e95, 1–9. https://doi.org/10.1017/ S0950268820000990.
4. Canals M, Labra F. Análisis no lineal de la dinámica de enfermedades infecciosas en Chile. Rev Med Chile 1999; 127: 1086-92.
5. Jones A, Strigul N. Is spread of COVID-19 a chaotic epidemic. Chaos, Solitons and Fractals 2021; 142: 110376. https://doi.org/10.1016/j.chaos.2020.110376.
6. Chen X, Yu B. First two months of the 2019 Coronavirus disease (COVID-19) epidemic in China: real time surveillance and evaluation with a second derivative model. Global Health Research Policy 2020; 5: 7 https://doi.org/10.1186/s41256-020-00137-4.
7. Bant C. Entropy ratio and entropy concentration coefficient, with application to COVID-19 pandemic. Entropy 2020; 22, 1315; doi:10.3390/e22111315.
8. Sivakumar B, Deepthi B. Complexity of COVID-19 dynamics. Entropy 2022; 24, 50. https://doi.org/10.3390/e24010050.
9. Canals M, Maass A, Monreal J. Non-linear analysis of incidence and mortality time series of COVID-19 reveals high unpredictability. Science reports 2022; (submitted).
10. Canals M, Canals A, Cuadrado C. Incidence moments: A simple method for study the memory and short term forecast of the COVID-19 incidence time-series. Epidemiol Methods 2022; https://doi.org/10.1515/em-2021-0029.
11. Canals M, Cuadrado C, Canals A, Johannessen K, Lefio LA, Bertoglia MP, et al. Epidemic trends, public health response and health system capacity: The Chilean experience in COVID-19 epidemic. Rev Panam Salud Pública 2020; 44: e99. https://doi.org/10.26633/RPSP.2020.99.
12. Angulo J, Martínez-Valdebenito C, Pardo-Roa C, Almonacid LI, Fuentes-Luppichini E, Contreras AM, et al. Assesment of mutations associated with genomic variants of SARS CoV-2: RT-qPCR as a rapid and affordable tool to monitoring known circulating variants in Chile, 2021. Frontiers in Medicine 2022; 9: 841073 doi: 10.3389/fmed.2022.841073.
13. MINSAL, Chilean Government. 2020.. COVID-19 Official Reports. Available at https:// www.gob.cl/coronavirus/cifrasoficiales/ [Accesed 1 April 2021].
14. An der Heiden M, Hamouda O. "Schätzung der aktuellen Entwicklung der SARSCoV-2-Epidemie in Deutschland – Nowcasting". Epididemiol Bull 2020; 17:S.10–15. DOI 10.25646/6692.
15. Izadi N, Taherpour Y, Mokhayeri Y, Ghorbani SS, Rahmani K, Hashemi SS. The epidemiologic parameters for COVID-19: A Systematic Review and Meta-Analysis; Available at Medrix https://doi.org/10.1101/2020.05.02.20088385doi [Accesed at 1/2/2022].
16. Zhang P, Wang T, Xie SX. Meta-analysis of several epidemic characteristics of COVID-19. Journal of Data Science 2020; 18(3): 536-49 doi: 10.6339/JDS.202007_18(3).0019.
17. Alenne M, Yismaw L, Assemie MA, Ketema DB, Gietaneh W, Birhan TY. Serial interval and incubation period of COVID-19:a systematic review and meta-analysis. Infectious Diseases 2021; 21: 257 https://doi.org/10.1186/s12879-021-05950-x.
18. KimS, Kim H. A new metric of absolute percentage error for intermittent demand forecasts. Int. Jour. Forecast. 2016; 32: 669-79.
19. Canals M. Revisión del concepto de inmunidad poblacional, a propósito de la epidemia COVID-19 y del desarrollo de una vacuna. Rev Chilena Infectol 2021; 38(4): 495-9.
20. Alí ST, Wang L, Lau EHY, Xu RK, Du Z, Wu Y, et al. Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. Science 2022; 369: 1106–9.
21. Pung R, Mak TM, CMMID COVID Group, Kucharski AJ, Lee VJ. Serial interval in SARS CoV-2 B.1.617.2 variant cases. The Lancet. 2021; 398: 837-8.
22. Backer JA, Eggink D, Andeweg SP, Veldhuijzen IK, van Maarseveen N, Vermaas K, et al. Shorter serial intervals in SARS-CoV-2 cases with Omicron variant compared to Delta variant in the Netherlands, 13 – 19 December 2021. 2022; Available in medRxiv https://doi.org/10.1101/2022.01.18.22269217doi (Accesed at 9/3/2022).