Apéndice Metodológico a: Ciclos de trabajo-confinamiento para reducir la transmisión de COVID-19: evidencia y recomendaciones en el contexto de México

Carolina Pérez Ferrer
Nancy Lopez Olmedo
Sergio Bautista
Arantxa Colchero
Dalia Stern
Rodrigo Zepeda-Tello
Rossana Torres-Álvarez
Francisco Reyes-Sánchez
Tonatiuh Barrientos-Gutiérrez

El modelo es un proceso de ramificación para los enfermos sintomáticos de COVID-19 asociado a otro proceso estocástico para el tiempo hasta el aislamiento. La descripción actual está organizada de la siguiente forma:

- Se explica el funcionamiento donde un individuo siempre genera un único caso nuevo.
- Se explica el funcionamiento general donde un individuo infectado genera una cantidad aleatoria de nuevos casos.

1. Mecanismo con un único descendiente

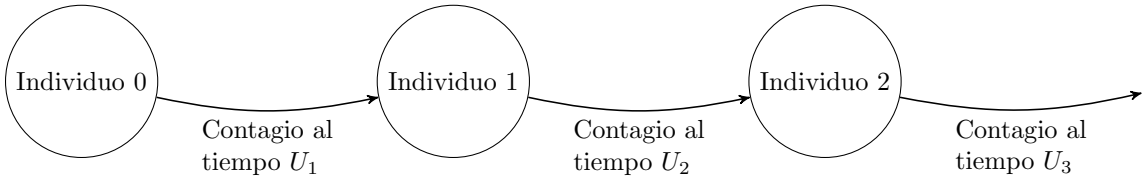


Figura 1: Mecanismo con un único descendiente

En esta sección explicamos una versión simplificada del modelo. Aquí, consideraremos inicialmente un infectado y cada infectado genera a su vez sólo un nuevo caso. Para ello, tomemos S_0 como el tiempo que un individuo tarda en desarrollar síntomas a partir de ser contagiado en el tiempo $t_0 = 0$. Suponemos que la distribución de S_0 es una distribución paramétrica $F(\Theta)$ donde Θ es un vector de parámetros. En la literatura F suele ser una distribución Gamma, Lognormal, Weibull o Erlang. Para el análisis principal utilizaremos $F(\Theta)$ Lognormal con parámetros $\mu = 1.621$ y $\sigma = 0.418$ estimados por Lauer *et.al* [1].

Supondremos además que este individuo sólo puede contagiar hasta antes de desarrollar síntomas pues en cuanto desarrolla los síntomas se aísla. Un nuevo individuo contagiado por el primero será llamado su descendiente (epidémico) y el individuo que lo contagió es su padre (epidémico). El momento de contagio del descendiente ocurre en un intervalo de tiempo (t_0, S_0) y está dado por la variable aleatoria U_1 definida en

dicho intervalo. En nuestro caso tomaremos $U_1 \sim \text{Uniforme}(0, S_0)$. El tiempo hasta desarrollo de síntomas (desde el inicio de la epidemia) para este nuevo individuo está dado por $T_1 = U_1 + S_1$ donde S_1 tiene la distribución $F(\Theta)$ (es el tiempo hasta desarrollo de síntomas desde su contagio). En general para el k-ésimo descendiente de un enfermo, U_k es una variable aleatoria definida en $(U_{k-1}, U_{k-1} + S_{k-1})$ donde U_{k-1} es el tiempo de contagio de su padre (el $k-1$-ésimo descendiente del enfermo) y S_{k-1} es el tiempo hasta el desarrollo de síntomas del mismo antecesor. Esto significa que el k-ésimo descendiente sólo se pudo haber contagiado en el intervalo en el que su padre ya había sido contagiado pero aún no presentaba suficientes síntomas para aislarse. Este modelo considera $U_k \sim \text{Uniforme}(U_{k-1}, U_{k-1} + S_{k-1})$ con $U_0 = 0$ y la colección $\{S_k\}_{k\geq 0}$ forma un conjunto de variables aleatorias independientes idénticamente distribuidas con distribución $F(\Theta)$.

2. Mecanismo con múltiples descendientes

En esta sección explicamos el modelo general (diagramado en la figura 2). Para ello, consideraremos distintas generaciones de la enfermedad donde cada caso (padre) genera nuevos infectados (descendientes) de acuerdo a un proceso de Galton-Watson. Recordamos que dentro de esta clase de modelos, la cantidad de individuos por "generación" a partir del primer infectado se denota Z^k donde k representa el número de generación. Sin pérdida de generalidad consideramos inicialmente un solo contagiado $Z^0 = 1$. Para cada generación $k > 0$ el número de enfermos en dicha generación está dado por la siguiente igualdad:

$$Z^k = \begin{cases} \sum_{j=1}^{Z^{k-1}} \eta_j^{k-1} & \text{si } Z^{k-1} > 0, \\ 0 & \text{si } Z^{k-1} = 0, \end{cases} \tag{1}$$

donde η_j^{k-1} representa el número de descendientes del j-ésimo padre en la generación $k-1$. Enumeramos a los individuos de la generación k mediante el índice $i = 1, 2, \ldots, Z^k$. Para un individuo i denotamos su padre en la generación anterior como $p(i, 1)$ de tal forma que $p(i, 1) = j$ si i es uno de los η_j^{k-1} descendientes de j. Como notación, si i es descendiente de j y j es descendiente de ℓ escribimos $\ell = p(i, 2) = p(j, 1) = p\big(p(i, 1), 1\big)$. De manera general el m-ésimo ancestro de i se denota como:

$$p(i, m) = p\Big(p(i, m-1), 1\Big) \tag{2}$$

donde $p(i, 0) = i$.

Asociado a cada individuo i de la generación k existe el momento en el que desarrolla síntomas a partir de su contagio, S_i^k. El momento de contagio del individuo i de la generación k está dado por la variable aleatoria U_i^k definida en el intervalo donde su padre fue infeccioso $(U_{p(i,1)}^{k-1}, U_{p(i,1)}^{k-1} + S_{p(i,1)}^{k-1})$. De manera general, sea Q un intervalo o bien la unión (posiblemente infinita pero numerable) de intervalos denotando el periodo de cuarentena: aquellos momentos en los que un individuo está forzosamente aislado y no puede contagiar. Denotamos $Q_i^{k-1} = (U_i^k, U_i^k + S_i^k) \setminus Q$ al intervalo de tiempo en el que el individuo i de la generación k no está en cuarentena y es contagioso.

Por otro lado, en (1) la variable η_j^{k-1} es una variable aleatoria sobre los enteros positivos con distribución dada por:

$$\eta_j^{k-1} | Q^k \sim \text{Poisson}(\lambda(t) \cdot m(Q^k)) \tag{3}$$

donde $m(Q^k)$ es la medida de Lebesgue ("longitud") del conjunto Q^k y $\lambda(t)$ es el número esperado de descendientes por día de un individuo infectado al tiempo t.

Denotamos los tiempos asociados al desarrollo de síntomas de cada individuo i (con $0 \leq i \leq Z^k$) de la generación k como S_i^k. Una vez el padre de i, $p(i, 1)$ es contagiado en el momento $U_{p(i,1)}^{k-1}$ el momento de contagio de i ocurre al tiempo U_i^k en el intervalo $(U_{p(i,1)}^{k-1}, U_{p(i,1)}^{k-1} + S_{p(i,1)}^{k-1})$ donde $U_0 = 0$. Las variables aleatorias $\{S_i^k\}_{(i,k)}$ son independientes idénticamente distribuidas con distribución $F(\Theta)$ y denotan el tiempo desde contagio hasta inicio de síntomas.

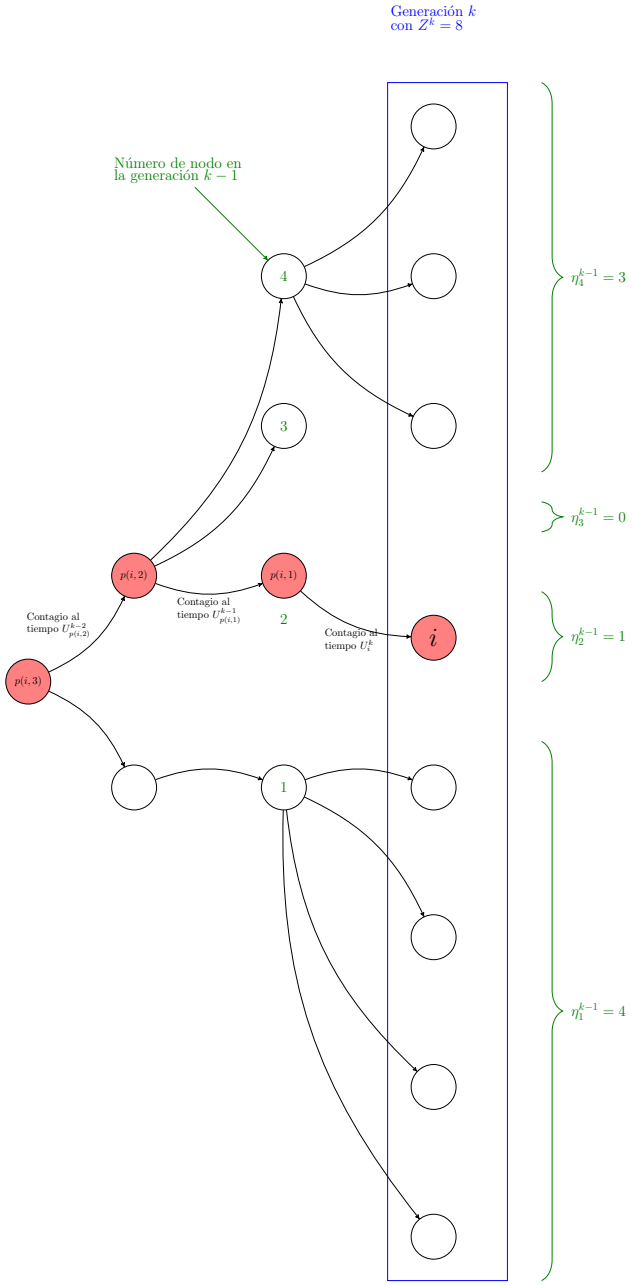


Figura 2: Mecanismo con múltiples descendientes

Finalmente notamos que la cantidad de enfermos infectados al tiempo s bajo el esquema de cuarentena Q está dada por la variable aleatoria:

$$N_s(Q) = \sum_{k=0}^{\infty} \sum_{i=1}^{Z^k} \mathbb{I}\{U_i^k \in [0, s)\} \tag{4}$$

La comparación de interés es la variable aleatoria dada por la diferencia entre dos esquemas de cuarentenas Q_1 y Q_2:

$$N_s(Q_1) - N_s(Q_2) \tag{5}$$

donde Q_1, Q_2 puede incluir el conjunto vacío.

3. Implementación y algoritmo

El modelo fue implementado en R usando el paquete Rcpp [2, 3, 4, 5] y puede ser encontrado en GitHub [6]. El algoritmo 1 contiene el pseudo-código de la implementación del modelo.

Algoritmo 1 Proceso de ramificación para COVID-19

Proceso

Entrada:

$nsim$ ▷ Número de simulaciones

$maxk$ ▷ Cota superior para generaciones: cuando $k \geq maxk$, el modelo se detiene.

$maxt$ ▷ Número de días a simular

$prob_autoquarantine$ ▷ Probabilidad de aislamiento al desarrollar síntomas

$quarantine_flag$ ▷ Si $quarantine_flag$ = TRUE, el esquema de $\ell \times m$ se implementa.

l ▷ Número de días laborables en el esquema $\ell \times m$

m ▷ Número de días de confinamiento en el esquema $\ell \times m$

$init_proc$ ▷ Día desde el inicio de la epidemia cuando se implementa el esquema de $\ell \times m$

$dias_decaimiento$ ▷ El periodo infeccioso decae aproximadamente 8 días después de presentar síntomas [7]

Proceso:

$k \leftarrow 0$

$t \leftarrow 0$

$currentSim \leftarrow 0$

$Z_1^0 \leftarrow 1$

$U_1^0 \leftarrow 0$

mientras $k < maxk$ & $t < maxt$ & $currentSim < maxsim$ hacer:

  para i in $1 : Z_k^i(t)$ hacer:

    $\mathtt{S}_i^k \leftarrow \text{LogNormal}(\mu = 1.621, \sigma = 0.418)$

    si $t < init_proc$ entonces: ▷ Esquema 6×1 durante las primeras 8 semanas

      $l_1 \leftarrow 6$

      $m_1 \leftarrow 1$

    en otro caso: ▷ Esquema l×m a partir de la semana 9

      $l1 \leftarrow l$

      $m1 \leftarrow m$

    finalizar si

    si Uniforme$(0, 1) > prob_autoquarantine$ entonces:

      $\mathtt{S}_i^k(t) \leftarrow \mathtt{S}_i^k(t) + dias_decaimiento$

    finalizar si

    $m(Q_i^k(t)) \leftarrow$ Cantidad de días que el individuo no se aisló

    $\eta_i^k(t) \leftarrow$ Poisson$(\lambda(t) * m(Q_i^k(t)))$

    $U_i^k(t) \leftarrow$ Uniforme$(Q_i^k(t))$

  finalizar para

  $Z_i^k(t) \leftarrow \sum_{i=1}^{Z_i^{k-1}(t)} \eta_i^k(t)$

$k \leftarrow k + 1$

$t \leftarrow t + 1$

$currentSim \leftarrow sim + 1$

finalizar mientras

$N_s(Q) \leftarrow \sum_{k=0}^{\infty} \sum_{i=1}^{Z^k} \mathbb{I}\{U_i^k \in [0, s)\}$

finalizar Proceso

4. Ajuste del modelo

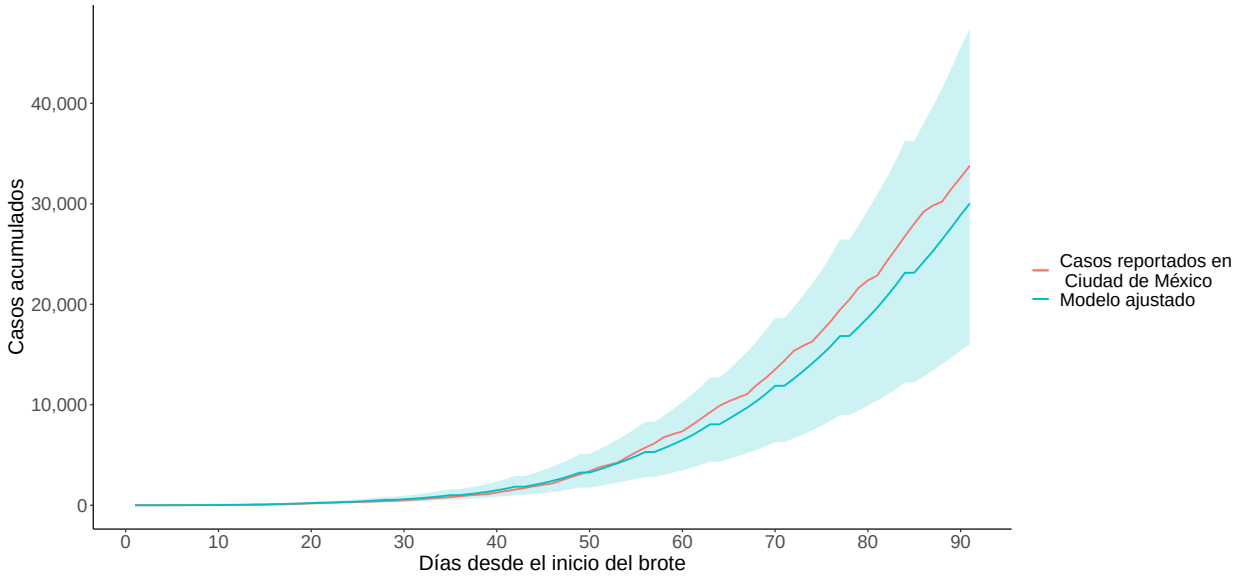
Para describir la evolución de la pandemia de COVID-19 en la Ciudad de México, se ajustó el modelo presentado de acuerdo a los casos reportados por la Secretaría de Salud, a partir del anuncio del primer caso en México (27 de Febrero de 2020) [8]. Los valores que toma el parámetro $\lambda(t)$, que corresponde al número de contactos por persona por unidad de tiempo, se presentan en la Tabla 1. La figura 3 muestra los resultados del ajuste del modelo. Asumimos que el esquema 6 días laborales y 1 día de aislamiento (6×1) es el usual en Ciudad de México y que los individuos infectados se aíslan al presentar síntomas ($prob_autoquarantine = 1$). Realizamos 5000 simulaciones para cada día, y presentamos las medianas para comparar con los casos acumulados diarios reportados en la Ciudad de México.

Tabla 1: Estimación diaria para el número esperado de nuevos contagios por individuo.

t	$\lambda(t)$
1	0.458054494
2	0.369597153
3	0.35966752
4	0.391829255
5	0.446481383
6	0.5389833
7	0.448411164
8	0.39633321
9	0.339127311
10	0.30298841
11	0.295770414
12	0.291741356
13	0.279323836
14	0.26842082
15	0.247184279
16	0.233583211
17	0.230153487
18	0.232625557
19	0.234004172
20	0.234010362
21	0.237982174
22	0.238316066
23	0.23981587
24	0.238305326
25	0.240491204
26	0.236959407
27	0.23658983
28	0.23674071
29	0.235121721
30	0.232173613
31	0.233496526
32	0.239419152
33	0.240433084
34	0.238946145
35	0.237154091
36	0.236242674
37	0.234305634
38	0.235879486
39	0.242161104

t	$\lambda(t)$
40	0.241845951
41	0.238907838
42	0.237327704
43	0.235423114
44	0.231124933
45	0.230748335
46	0.235363745
47	0.234628716
48	0.231704481
49	0.229125936
50	0.22775851
51	0.222482779
52	0.22128124
53	0.224193759
54	0.224745986
55	0.223563633
56	0.22110222
57	0.217578256
58	0.21572104
59	0.216155906
60	0.219771048
61	0.219743966
62	0.219982584
63	0.219513168
64	0.218023275
65	0.214512037
66	0.213788914
67	0.216296757
68	0.217071556
69	0.217237011
70	0.215915433
71	0.21479934
72	0.21223609
73	0.211236065
74	0.213712565
75	0.213823381
76	0.213219339
77	0.211852318
78	0.209878068
79	0.207058941
80	0.206098041
81	0.207753404
82	0.207860091
83	0.207490424
84	0.206574386
85	0.205540988
86	0.203961479
87	0.203446447
88	0.205018655
89	0.205423183
90	0.205749485

Figura 3: Comparación entre la mediana de simulaciones del modelo ajustado con los casos de COVID-19 reportados en la Ciudad de México. El sombrado considera el 20 % de los escenarios más probables.



5. Simulaciones con distintos esquemas de cuarentena y resultados complementarios

Para cada uno de los esquemas $l \times m$ considerados (l días laborales y m días de aislamiento), el esquema 6×1 se simula durante las primeras 8 semanas tras el inicio del brote. El esquema $l \times m$ es implementado a partir de la semana 9. Lo anterior se realizó para permitir el crecimiento de casos en los primeros días. Los resultados presentados en el manuscrito (figura 4) representan los casos acumulados por día obtenidos por las simulaciones del modelo ajustado (mediana de las simulaciones). La Tabla 2 contiene los casos acumulados hasta el día 90 para cada esquema. Para replicar resultados, utilizamos una semilla igual a 374590.

Figura 4: Comparación entre distintos esquemas de cuarentena simulados con el modelo ajustado (medianas).

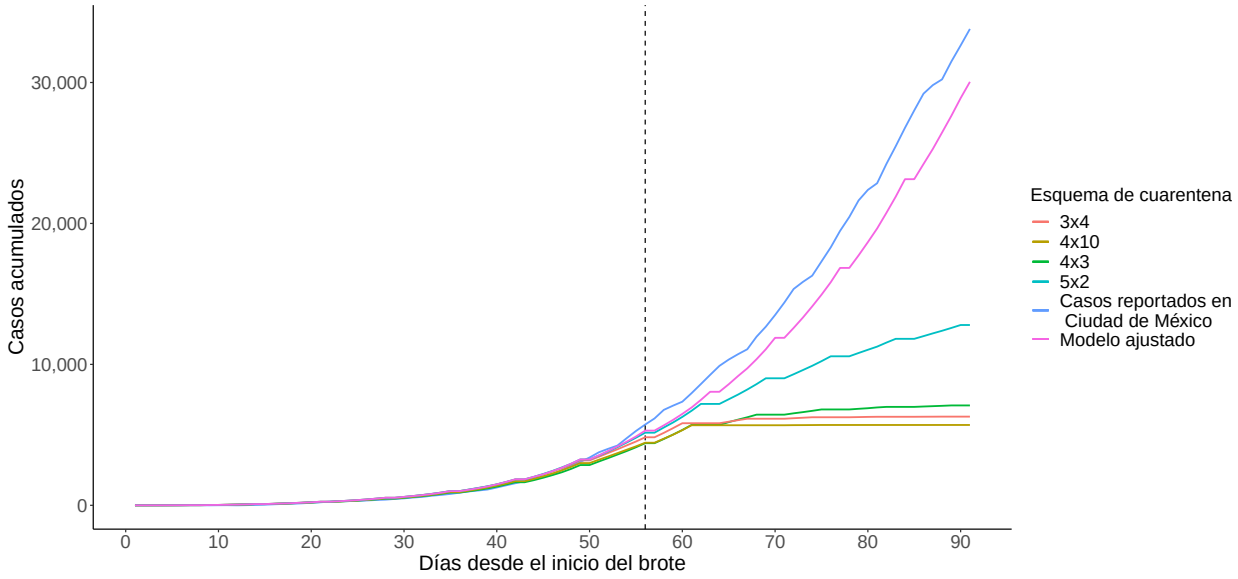


Tabla 2: Simulaciones de casos acumulados al día 90 (medianas) con el modelo ajustado, utilizando distintos esquemas.

	Casos acumulados hasta el día 90
Casos reportados en CDMX	33,800.0
Estimación con esquema 6×1 *(Modelo ajustado)*	30,046.5
Estimación con esquema 5×2	12,793.5
Estimación con esquema 4×3	7,087.5
Estimación con esquema 3×4	6,293.0
Estimación con esquema 4×10	5,701.0

6. Notación relativa a la descripción del modelo

Variable	Descripción	Valor/Rango
k	Número de generación	$k \in \{0, 1, 2, \ldots\}$
U_1^0	Día de inicio de infección del primer infectado.	$U_1^0 = 0$
U^k	Día de inicio de infección para cada individuo infectado en la generación k ($k \neq 0$).	$U^k \sim \text{Uniforme}(U^{k-1}, U^{k-1} + S^{k-1})$
S^k	Número de días transcurridos entre el día de contagio y el día de síntomas para cada individuo en la generación k	$S^k \sim \text{Lognormal}(1.621, 0.418)$.
T_k	Día de inicio de síntomas para cada individuo infectado en la generación k.	$T_k = U^k + S^k$
Q	Periodo de cuarentena: conjunto de días de confinamiento (de acuerdo al esquema) ocurridos durante el periodo infeccioso de cada individuo infectado en la generación k. Este conjunto es obtenido para cada individuo de acuerdo a su correspondiente periodo infeccioso.	
Q^k	Conjunto de días en los que un individuo infectado en la generación k no estuvo en cuarentena (de acuerdo al esquema) y fue contagioso.	$Q^k = (U_k, U_k + S_k) \setminus Q$
$m(Q^k)$	Número de días en los que un individuo infectado en la generación k no estuvo en cuarentena (de acuerdo al esquema) y fue contagioso.	$m(Q^k) \in \{0, 1, 2, \ldots\}$
$\lambda(t)$	Número esperado de nuevos casos por día por individuo infectado en el tiempo t.	Tabla 1
$proba_autoquarantine$	Probabilidad de auto-aislamiento al presentar síntomas.	1.0
η_j^k	Número de sujetos contagiados por el individuo j perteneciente a la generación $k \neq 0$. El individuo j fue contagiado en la generación $k - 1$.	$\eta_j^k \vert Q^k \sim \text{Poisson}(\lambda(t) \cdot m(Q^k))$
Z^k	Número de individuos infectados en la generación k ($Z^0 = 1$).	$Z^k \in \{0, 1, 2, \ldots\}$
$N_s(Q)$	Número de infectados al tiempo s bajo el esquema Q.	$N_s(Q) \in \{0, 1, 2, \ldots\}$

Referencias

[1] Lauer SA, Grantz KH, Samarin A, Bi Q, Jones FK, Zheng Q, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Annals of Internal Medicine. 2020;172(9):577–582. Available from: https://doi.org/10.7326/M20-0504.

[2] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2019. Available from: https://www.R-project.org/.

[3] Eddelbuettel D, François R. Rcpp: Seamless R and C++ Integration. Journal of Statistical Software. 2011;40(8):1–18. Available from: http://www.jstatsoft.org/v40/i08/.

[4] Eddelbuettel D. Seamless R and C++ Integration with Rcpp. New York: Springer; 2013. ISBN 978-1-4614-6867-7.

[5] Eddelbuettel D, Balamuta JJ. Extending extitR with extitC++: A Brief Introduction to extitRcpp. PeerJ Preprints. 2017 aug;5:e3188v1. Available from: https://doi.org/10.7287/peerj.preprints.3188v1.

[6] Zepeda-Tello R, Torres-Álvarez R, Reyes-Sánchez F. Modelo lxm: Modelo estocástico para COVID-19. GitHub; 2020. https://github.com/INSP-RH/Modelo-lxm.

[7] He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature Medicine. 2020;26:672—-675. Available from: https://doi.org/10.1038/s41591-020-0869-5.

[8] Secretaría de Salud. Datos abiertos - Dirección General de Epidemiología. Gobierno de México; 2020. [Datos al 06-Septiembre-2020]. https://www.gob.mx/salud/documentos/datos-abiertos-152127.