Multi-level multi-state modelling applied to hospital admission in mexican patients with COVID-19

Diaz Martinez Juan [*], Fuentes-García Ruth [*]

Department 1, Street, City, State, Zip
Mathematics Department 2, Mexico City, 04510

* Corresponding author: juan.diaz.martinez@mail.utoronto.ca
* Corresponding author: rfuentes@ciencias.unam.mx

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eget porta erat. Morbi consectetur est vel gravida pretium. Suspendisse ut dui eu ante cursus gravida non sed sem. Nullam sapien tellus, commodo id velit id, eleifend volutpat quam. Phasellus mauris velit, dapibus finibus elementum vel, pulvinar non tellus. Nunc pellentesque pretium diam, quis maximus dolor faucibus id. Nunc convallis sodales ante, ut ullamcorper est egestas vitae. Nam sit amet enim ultrices, ultrices elit pulvinar, volutpat risus.

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic was declared a Public Health Emergency of International Concern on January 30, 2020 by the World Health Organization.The Mexican Health Authorithies declared the first lockdown on March 26 with 585 cases and 8 deaths reported for COVID-19 [2]; by the end of the lockdown (June 5th, 2020) total number of cases and deaths were 110,026 and 13,170, respectively. By November 1, Mexico became the fourth country in number of deaths of SARS-CoV-19 (106,765 deaths), with 1,122,362 incident cases [3]; by April 15th 2021 the number of deaths had raised to 214,372 with 2,309,099 incident cases.

Over time it has become clear that comorbidy factors such as hypertension, diabetes, obesity and smoking increase the seriousness of the disease, leading to a higher rate of hospitalizations with an additional 25% of the cases requiring intensive care unit (ICU) admission and ultimately, intubation [4]. Mexico ranks second in obesity among OECD countries, with an obesity rate of 72.5% among the adult population, which is associated with the high prevalence of type 2 diabetes, estimated at 13% of the adult population in 2017, the highest rate among OECD countries; the rate of hypertension is also one of the highest chronic diseases among adult population with 30% [5]. The high prevalence of these comorbidities besides the precarious health care system could be among the main main reasons of the elevated severity of the number of cases and deaths rates in the country.

In Mexico the health institutions are divided in public and private services. There are different public institutions which provide care to different sets of the population: the state employees, the army and naval members, the oil state company (PEMEX) employees and private companies employees. There are also public hospitals for population with no health service coverage. In general, the care withing different

healthcare providers cannot be considered homogeneous, therefore it should be relevant for the final outcome of a COVID-19 patient.

There have been efforts to understand how patients with comorbilities are affected by COVID-19; the work by [8] proposed a clinical score to predict COVID-19 lethality, including different factors like diabetes and obesity among mexican population. This work lead to believe that obesity mediates 49.5% of the effect of diabetes on COVID-19 lethality. Early-onset diabetes conferred an increased risk of hospitalization while obesity increased the risk ICU for admission and intubation.

After onset of infection there is a period of time between symptom detection and hospitalization. The time elapsed before patients approach hospitals could be excessively long. Once patients are admitted to hospital, there is also a period of time between the admission and death. Estimation of these lengths of time through a multilevel model could enable a better information system to estimate incidence and transmission rates, particularly at regional level where differences can be apparent.

This work considers a multi-state model under a Bayesian framework to estimate times between symptom detection and hospitalization and between hospitalization and death. Data used in the modeling comes from the official database by the Mexican Ministry of Health; the analysis provides of general overview of hospitalizations in each state of the country and the different health institutions within. Variables affecting the patient's final outcome such as the aforementioned comorbidites are included in the model. Additionally regional heterogeneity is accounted for trough nested models that consider the regional contribution and also the health service provider. Other efforts in recent literature [12] ME PARECE QUE ESTA CITA NO APARECE have considered more states (hospitalization-ICU, ICU-death, ICU-discharged), which allows researchers to asses whether improvements in patient outcomes have been sustained, finding evidence that median hospital stays have lengthened. Unfortunately the data available for Mexico lacks the necessary granularity to determine such states. Nevertheless, we believe this model could better inform the estimation of the incidence and transmission rates, which is particularly important while new variants and increased transmission rates are present.

Methods and materials

Data Source and Study Population

We conducted a prevalence study of the official database from the Mexican Ministry of Health, these data provide an overview of hospital admissions, deaths and the period of time between hospitalizations and first COVID-19 symptoms between March and December 2020. The data analyzed included mexican adult population diagnosed with COVID-19 in the whole country. Exclusion criteria were the observations with incomplete data about hospital admission, symptoms or comorbidities. Additionally, patients whose time of initial symptoms was captured as the day they were admitted to hospital were removed, since this time was likely to be unknown. After applying exclusion criteria a total sample of 1200 registers of adult patients belonging to any healthcare provider, either private or public in the 32 states of Mexico was selected, preserving the population characteristics. JUAN PABLO ESTE NUMERO DE MUESTRA ES CORRECTO?

Comorbidities that could worsen the patient outcome, affecting the times we aim to model; such as diabetes, hypertension, obesity, chronic obstructive pulmonary disease (COPD), asthma, immunosuppression and chronic kidney disease were included as linear predictors for each state in the model and so two different relevant groups of variables

were statistically significant for each state:

$$\begin{aligned} x_{death} &= (\sim Diabetes + COPD + Obesity \\ &\quad + Hypertension + Diabetes * Obesity * Hypertension \\ &\quad + Kidney\_Disease) \\ x_{hospitalization} &= (\sim COPD + Obesity + Kidney\_Disease \\ &\quad + Asthma + Immunosuppression) \end{aligned}$$

About 87% of the population in Mexico is affiliated to some healthcare provider, but during this pandemic the mexican government has established a list of hospitals designated to treat COVID-19 patients without any affiliation distinction. In this study we identified 6 different healthcare providers which were classified according to their sectors IMSS, ISSSTE, SEDENA/SEMAR/PEMEX, SSA, ESTATALES (healthcare provider within each state) these 5 are public care providers while the sixth sector is private hospitals. It is worth mentioning that following the national hospital transformation plan [6], IMSS alone has transformed about 260 medical units to treat COVID-19 patients [9].

Modelling

We developed four different Bayesian models for trajectories of interest namely, *Symptoms-Hospitalization* and *Hospitalization-Death*, in which non-informative initial distributions were used, located near 0, to improve convergence.

Figure 1. Graphical representation of the model: directed acyclic grph (DAG).

Additionally a QR reparameterization for the covariate matrix was used, that is, if X is an $n \times m$ covariate matrix, corresponding to the aforementioned comorbidities, $X = QR$, where Q is an ortogonal matrix and R is an upper triangular matrix. In practice, considering $X = Q'R'$ where $Q' = Q\sqrt{n-1}$ and $R' = \frac{1}{\sqrt{n-1}}R$ is convenient. Hence if ζ is the N linear predictor vector such that $\zeta = X\beta$, with β an M coefficient vector, then $\zeta = X\beta = QR\beta = Q'R'\beta$. We used $\zeta = Q'R'\beta$ for numerical stability.

Each model captures different levels of information, as more levels were included it was possible to differentiate the results according to the added information. The four levels were:

Model I: One level

Vectors M and H correspond to survival times for deaths and hospitalizations, respectively. We considered a different set of covariables in each case and assumed deaths and hospitalizations are independent, we then define the model as

$$\begin{aligned} M &\sim Weibull(\alpha, \eta) \\ H &\sim Weibull(\alpha, \upsilon) \\ \eta &= \exp\left(-\frac{\mu_m + \mathbf{Q}^*\vartheta}{\alpha}\right) \\ \upsilon &= \exp\left(-\frac{\mu_h + \mathbf{Q}^{**}\theta}{\alpha}\right) \\ \alpha &= \exp(\alpha_r * \tau_\alpha) \\ \alpha_r &\sim N(0,1) \\ \mu_m, \mu_h &\sim N(0, \tau_\mu) \\ \vartheta, \theta &\sim U(-\infty, \infty) \end{aligned}$$

where Q^* and Q^{**} are matrices of standarized covariables for deaths and hospitalizations respectively and τ_α and τ_μ are given positive values QUIZA AQUI PONER LOS VALORES. This model is described in grey in Figure 1.

Model II: two levels

The second model is based on the first one, an additional level is added to account for each state of Mexico to model deaths. The hospitalization H remains unchanged and for each state $l = 1, \ldots, 32$, deaths are summarized in matrix M_l the model is defined as

$$\begin{aligned}
M_l &\sim Weibull(\alpha, \eta) \\
H &\sim Weibull(\alpha, \upsilon) \\
\eta &= \exp\left(-\frac{\mu_m + \mu_l^r + \mathbf{Q}^*\vartheta}{\alpha}\right), l = 1, ..., 32 \\
\upsilon &= \exp(-\frac{\mu_h + \mathbf{Q}^{**}\theta}{\alpha}) \\
\mu_l &= \sigma * \mu_l^r \\
\alpha &= \exp(\alpha^r * \tau_\alpha) \\
\alpha^r &\sim N(0, 1) \\
\mu_l^r &\sim N(0, 1) \\
\sigma &\sim t_3^+(0, 1) \\
\mu_m, \mu_h &\sim N(0, \tau_\mu) \\
\vartheta, \theta &\sim U(-\infty, \infty)
\end{aligned}$$

where Q^* and Q^{**} are matrices of standarized covariables for deaths and hospitalizations respectively and τ_α and τ_μ are given positive values. This model is described in green in Figure 1.

Model III: Three levels Based on Model II, we consider a third level to include the healthcare provider where patients are hospitalized, k, for patient i in state l we have $M_{l,k}$ for the corresponding death matrix and the model is given as

$$\begin{aligned}
M_{l,k} &\sim Weibull(\alpha, \eta) \\
H &\sim Weibull(\alpha, \upsilon) \\
\eta &= \exp\left(-\frac{\mu_m + \mu_l^r + \mu_k^r + \mathbf{Q}^*\vartheta}{\alpha}\right), l = 1, ..., 32, k = 1, ..., 5 \\
\upsilon &= \exp\left(-\frac{\mu_h + \mathbf{Q}^{**}\theta}{\alpha}\right) \\
\mu_l &= \sigma_l * \mu_l^r, l = 1, 2 \\
\mu_k &= \sigma_l * \mu_k^r \\
\alpha &= \exp(\alpha^r * \tau_\alpha) \\
\alpha^r &\sim N(0, 1) \\
\mu_l^r, \mu_k^r &\sim N(0, 1) \\
\sigma_l &\sim t_3^+(0, 1) \\
\mu_m, \mu_h &\sim N(0, \tau_\mu) \\
\vartheta, \theta &\sim U(-\infty, \infty)
\end{aligned}$$

where Q^* and Q^{**} are matrices of standarized covariables for deaths and hospitalizations respectively and τ_α and τ_μ are given positive values. This model is

described in yellow in Figure 1.

Model IV:

Based as well on Model II, we consider index $j = (l, k)$ where l accounts for the l-th state and k for the healthcare provider, $j \in \{1, ..., 153\}$ where the distribution for deaths is given by

$$
\begin{aligned}
M_j &\sim Weibull\,(\alpha, \eta) \\
H &\sim Weibull(\alpha, \upsilon) \\
\eta &= \exp\left(-\frac{\mu_m + \mu_j^r + \mathbf{Q}^*\vartheta}{\alpha}\right), \\
\upsilon &= \exp\left(-\frac{\mu_h + \mathbf{Q}^{**}\theta}{\alpha}\right) \\
\mu_j &= \sigma_j * \mu_j^r \\
\alpha &= \exp(\alpha^r * \tau_\alpha) \\
\alpha^r &\sim N(0, 1) \\
\mu_l^r, \mu_k^r &\sim N(0, 1) \\
\sigma_j &\sim t_3^+(0, 1) \\
\mu_j &\sim N(0, \tau_\mu) \\
\vartheta, \theta &\sim U(-\infty, \infty)
\end{aligned}
$$

where Q^* and Q^{**} are matrices of standarized covariables for deaths and hospitalizations respectively and τ_α and τ_μ are given positive values.

JUAN PABLO, EL ULTIMO MODELO HAY QUE CHECARLO PARA VER QUE SEA CORRECTO.

To choose the model that best fits the data we considered the leave-one-out cross-validation (LOO) proposed by [1], which estimates pointwise out-of-sample prediction accuracy, using the log-likelihood evaluated at the posterior simulations of the parameter values, Table 1.

Model	elpd leave one out	p leave one out
Model 1	-75473.3 (187.8)	26.9 (2.2)
Model II	-75352.2 (186.3)	84.1 (4.7)
Model III	-75284.9 (186.6)	92.5 (4.9)

Table 1. Expected log-pointwise predictive density (ELPD) for a new data set and effective number of parameters (standard deviation).

Results

The parameters were estimated using Stan DAR QUIZA MÁS DETALLES. We show results for model III which performed better in terms of the likelihood and showed good convergence of all parameteres. The posterior 0.95 credibility intervals for parameters of interest at different levels of the model are shown in Figures 2 and 3. It is worth pointing out that we are displaying the log hazard ratio, hence positive values for parameters will point to increasing risks for the corresponding transition and level.

Figure 2. Hazard rate in logarithmic scale by state, credible interval of 95%.

Figure 3. Hazard rate in logarithmic scale by health institution, credible interval of 95%.

Increased risk for hospitalization was observed at the global population level for chronic renal disease, whereas for death such was the case for COPD and the interaction of diabetes:hypertension:obesity.

Figure 4. Predictive distribution for deaths.

Our results show that there are differences in mortality between the states without accounting for institution and it is related to the prompt time of death or viceversa. Figure 2 shows the states in which the overall rate of mortality is higher such as Campeche, Colima, Guanajuato, Hidalgo, Jalisco, Morelos, Nayarit, Oaxaca, Puebla, Tabasco and Veracruz. The difference might be linked to the late hospitalization of patients.

Figure 4 displays evidence that 5 days after hospitalization there is a peak on mortality rate, which could be related due the late hospitalization of patients with mild symptoms who developed "happy hypoxemia," that is extremely low blood oxygenation, but without sensation of dyspnea. In Wuhan, within a cohort of patients infected with (SARS-COV-2) who id not present dyspnea 62% showed severe disease and 46% ended up intubated, ventilated or dead [7].

Regarding the 6 healthcare providers included in the analysis differences were also found. While state-managed hospitals and private sector showed lower risks, in contrast the IMSS seems to be the one with the highest risk (Figure 3). Although it is worth mentioning that following the national hospital transformation plan [6], IMSS alone has transformed about 260 medical units to treat Covid-19 patients [9]. Additionaly it has the largest number of affiliations and they are liklely to have higher risk exposure.

Disscusion

Multiple sources have shownthat the presence of comorbidities such as diabetes, hypertension, obesity, chronic obstructive pulmonary disease (COPD), asthma, inmmuno-suppression and chronic kidney disease are associated to a worse outcome for the patients diagnosed with COVID-19. Particularly for those who are hospitalized in the ICU due intubation. To our knowledge this is the first study in Mexico which analyzed the time elapsed between the patient´s first symptoms, hospitalization and death; these analyses were further broken down to the different states of the republic and the healthcare providers within them. Thus it is possible to identify providers and states with an increased risk of hospitalization and death. Mexico is the third place in obesity among the OCDE countries, which could be the main reason of the high number of severe cases of COVID-19.

One problem that aggravates the situation is the precarious state of the public healthcare system which universal coverage is estimated about 87% of the mexican population. It is clear that mexican healthcare has overrun during this pandemic and has appealed to private health providers to cope with the treatment of COVID-19 patients. Among all health care providers the IMSS is the one with the highest risk, however it is the largest healthcare provider across Mexico with hospitals from level 2-level 4 of which "Siglo XXI" is a country-leader in research and innovative treatments and procedures. In March, 2020 38 hospitals were "converted" [6] to exclusively treat COVID-19 among which 18 were IMSS hospitals only; after one year 960 hospitals across the country were converted to treat patients with COVID-19 of these 289 (30.10%) belong to IMSS [9].

The proposed modelling can be helpful to a regional level to improve healthcare assistance, it could additionally be useful to inform statistical estimation of parameters for an epidemiological model.

This study has shown that there are differences in mortality between the different states of the republic; there are states in which the overall rate of mortality is higher

due the late hospitalization of patients such as Veracruz, Nuevo Leon, San Luis Potosí, Guanajuato, Chiapas and Mexico City. Breaking down this analysis to state level we found a higher risk of hospitalization, specifically in Veracruz which has been historically unsteady regarding the public healthcare system in sectors like IMSS, ISSSTE, SEDENA/SEMAR/PEMEX. The fact that different final outcomes could be related to patient's late hospitalization, hence suggesting that the average patient waits until the symptoms are severe to seek professional healthcare, needs to be further investigated.

One of the limitations of the study is the reduced number of states we were able to include in the modls due to the lack of information regarding dates of disccharged of recovered patient´s after hospitalization.

References

1.
Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing. 2017;27: 1413–1432. doi:10.1007/s11222-016-9696-4

2.
Carrillo-Vega MF, Salinas-Escudero G, García-Peña C, Gutiérrez-Robledo LM, Parra-Rodríguez L. Early estimation of the risk factors for hospitalization and mortality by COVID-19 in Mexico. PLOS ONE. Public Library of Science; 2020;15: e0238905. Available: https://doi.org/10.1371/journal.pone.0238905

3.
Organización Panamericana de la Salud O. Documentos técnicos de la OPS - Enfermedad por el Coronavirus (COVID-19) [Internet]. 2021. Available: https://www.paho.org/es/documentos-tecnicos-ops-enfermedad-por-coronavirus-covid-19

4.
SEDESA. Base del Sistema Nacional de Vigilancia Epidemiologica para el seguimiento a posibles casos de COVID-19 en la Ciudad de México [Internet]. 2021. Available: http://sinave.gob.mx/

5.
OCDE. Health at a Glance 2019: OECD Indicators. [Internet]. OCDE; 2019 p. 4. Available: https://www.oecd.org/mexico/health-at-a-glance-mexico-ES.pdf

6.
Mendoza-Popoca CÚ, Suárez-Morales M. Hospital reconversion in response to the COVID-19 pandemic. Revista Mexicana de Anestesiologia. 2020;43: 151–156. doi:10.35366/92875

7.
Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. New England Journal of Medicine. 2020;382: 1708–1720. doi:10.1056/nejmoa2002032

8.
Bello-Chavolla OY, Bahena-López JP, Antonio-Villa NE, Vargas-Vázquez A, González-Díaz A, Márquez-Salinas A, et al. Predicting Mortality Due to SARS-CoV-2: A Mechanistic Score Relating Obesity and Diabetes to COVID-19 Outcomes in Mexico. Journal of Clinical Endocrinology and Metabolism. 2020;105: 2752–2761. doi:10.1210/clinem/dgaa346

9.
UNAM LI de TeIE(iSTARL del I de G de la. Sistema de Información de la Red IRAG [Internet]. 2021. Available: https://www.gits.igg.unam.mx/red-irag-dashboard/reviewHome