**SUPPLEMENTARY METHODS**

*Definition of suspected COVID-19 cases in Mexico*

The Mexican Ministry of Health defines a suspected COVID-19 case as an individual of any age whom in the last 7 days has presented cough, fever or headache (≥two), accompanied by either dyspnea, arthralgias, myalgias, sore throat, rhinorrhea, conjunctivitis or chest pain. Amongst these, the Ministry of Health establishes two protocols for case confirmation: 1) SARS-CoV-2 testing is done widespread for suspected COVID-19 cases with severe acute respiratory infection with signs of breathing difficulty or deaths with suspected COVID-19, 2) for all other suspected cases, a sentinel surveillance model is used, whereby 475 nationally representative health facilities sample ~10% of mild outpatient cases and all suspected severe acute respiratory infections. Demographic and health data are collected and uploaded to the epidemiologic surveillance database by personnel from each corresponding facility.

*Available variables for analysis*

Available variables within the dataset include age, sex, occupation, nationality, and facility of case detection. Health information includes status of diabetes, obesity, chronic obstructive pulmonary disease (COPD), immunosuppression, pregnancy, arterial hypertension, cardiovascular disease (CVD), chronic kidney disease (CKD), HIV/AIDS, and asthma. Type and date of symptom onset, hospital admission, and death are available for all outpatient or hospitalized patients, as well as information regarding clinical diagnosis of pneumonia, ICU admission, and mechanic ventilatory support (MVS). Healthcare workers were defined as a composite of physicians, nurses, dentists, laboratory personnel and associated healthcare workers. Previous exposure to suspected viral cases was self-reported and was non-specific for known exposure to SARS-CoV-2 positive cases.

*Follow-up of non-hospitalized SARS-CoV-2 cases*

After a positive real time RT-PCR test for SARS-CoV-2 in Mexico City in subjects who are deemed initially not to require hospital admission, undergo domiciliary surveillance whereby they are instructed to self-isolate during at least 14 days and are monitored remotely via telephone and are instructed to call or contact the Health Ministry via SMS should any symptoms present after testing. If cases are not reached via telephone, domiciliary visits to assess clinical evolution are carried out by government personnel. Prior to self-isolation, positive cases under domiciliary surveillance also receive a medical kit, economic and food support during self-isolation as well as a pulse oximeter for cases with known comorbidity [2].

*Correspondence analysis and symptom clustering*

To investigate how the studied symptoms clustered in patients with confirmed SARS-CoV-2, we performed correspondence analysis on the evaluated columns for dimensionality reduction and identified the ideal number of components using the elbow method to maximize explained variance. Using these components, we performed hyerarchical cluster analysis on column categories using coordinates from a high-dimensional space corresponding to a dimensionality equal to the selected components from correspondence analysis (Dim=4). Distance within the feature space was calculated using Euclidean distance and agglomeration was conducted using Ward’s minimum variance criterion. The optimal number of symptom clusters was selected using the maximum mean silhouette method using the *CAinterprTools*, *factoextra* and *FactoMineR R packages*.

*Definition of respiratory symptoms*

For symptom clustering, we identified a 4-dimension solution for the correspondence analysis, which comprised 35.7% of the explained variance by symptoms. After applying hierarchical clustering, we identified three main cluster symptoms as shown in **Supplementary Figure 1**.



**Supplementary Figure 1.** Dendrogram for symptom clusters based on hierarchical clustering from correspondence analysis

The two symptom clusters were labeled as respiratory symptoms (RS) and Non-respiratory symptoms (NRS) which were latter grouped in gastrointestinal symptoms and constitutional symptoms [3,4]. Based on these definitions and on WHO criteria for suspected COVID-19 cases, we included cough and fever within the RS cluster [5].

*Statistical analyses*

To evaluate predictors of RS and NRS we fitted a multivariable mixed effects logistic regression models, including facility of treatment as a random effect to account for the variability in case distribution and treatment across healthcare facilities. We calculated time-to-event for hospital admission and lethality for each group and modeled these outcomes using mixed-effects Cox Proportional Risk regression, including as a frailty penalty treatment facility. All statistical analyses were performed using R version 3.6.4.

**Supplementary Figure 2.** Seven-day rolling average positivity rates for SARS-CoV-2 positivity in Mexico City, comparing RS, NRS and asymptomatic cases.



**SUPPLEMENTARY TABLES**

**Supplementary Table 1.** Descriptivecharacteristics of symptomatic and asymptomatic SARS-CoV-2 cases in Mexico City. *Abbreviations*: COPD: Chronic obstructive pulmonary disease; HIV/AIDS: Human immunodeficiency virus and/or acquired immunodeficiency syndrome; CKD: Chronic kidney disease, CDV: Cardiovascular disease; HCW: Healthcare worker

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **RS****(n=93,423)** | **NRS****(n=8,545)** | **Asymptomatic****(n=6,112)** | **P-value** |
| Male sex (%) | 49334 (52.8) | 3670 (42.9) | 3190 (52.2) | <0.001 |
| Age (years) | 45.34±16.54 | 40.41±15.89 | 40.40±17.93 | <0.001 |
| Mortality (%) | 10119 (10.8) | 112 (1.3) | 32 (0.5) | <0.001 |
| Hospitalizacion (%) | 23156 (24.8) | 365 (4.3) | 185 (3.0) | <0.001 |
| ICU admission (%) | 2019 (2.2) | 32 (0.4) | 29 (0.5) | <0.001 |
| Mechanical-Ventilation (%) | 3543 (3.8) | 32 (0.4) | 12 (0.2) | <0.001 |
| Severe COVID-19 (%) | 11350 (12.1) | 130 (1.5) | 56 (0.9) | <0.001 |
| Previous viral exposure (%) | 39384 (42.2) | 4980 (58.3) | 4980 (57.0) | <0.001 |
| Diabetes (%) | 13878 (14.9) | 740 (8.7) | 362 (5.9) | <0.001 |
| COPD (%) | 1400 (1.5) | 47 (0.6) | 25 (0.4) | <0.001 |
| Asthma (%) | 2198 (2.4) | 148 (1.7) | 72 (1.1) | <0.001 |
| HIV/AIDS (%) | 424 (0.5) | 32 (0.4) | 13 (0.1) | 0.014 |
| Obesity (%) | 18026 (19.3) | 1135 (13.3) | 449 (7.3) | <0.001 |
| Smoking status (%) | 10070 (10.8) | 945 (11.1) | 405 (7.3) | <0.001 |
| CVD (%) | 2084 (2.2) | 118 (1.4) | 57 (0.9) | <0.001 |
| Hypertension (%) | 16688 (17.9) | 950 (11.1) | 467 (7.6) | <0.001 |
| Immunosupression (%) | 1369 (1.5) | 72 (0.8) | 56 (0.9) | <0.001 |
| CKD (%) | 1637 (1.8) | 79 (0.9) | 39 (0.6) | <0.001 |
| Healthcare worker (%) | 133058 (14.2) | 1375 (16.1) | 541 (0.9) | <0.001 |

**Supplementary Table 2.** Cox Proportional Regression model with frailty penalty for prediction of hospitalization and mortality in initially asymptomatic SARS-CoV-2 cases.

*Abbreviations*: COPD= Chronic obstructive pulmonary disease; CKD= Chronic kidney disease; CVD: Cardiovascular disease

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Parameters** | **β-coefficient** | **HR (95%CI)** | **p-value** |
| Severe COVID-19c-stat=0.567 | Age >60 years | 1.036 | 2.82 (1.60-4.97) | <0.001 |
| Previous viral exposure | -0.850 | 0.43 (0.24-0.76) | 0.004 |
| CKD | 2.354 | 10.53 (4.32-25.65) | <0.001 |
| Mortalityc-stat=0.702 | Age >60 years | 1.947 | 7.01 (3.40-14.44) | <0.001 |
| CKD | 2.812 | 16.66 (6.23-44.62) | <0.001 |

**Supplementary Table 3.** Cox Proportional Regression model with frailty penalty for prediction of hospitalization and mortality in subjects with NRS only.

*Abbreviations*: COPD= Chronic obstructive pulmonary disease; CKD= Chronic kidney disease; NRS: Non-respiratory symptoms

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Parameters** | **β-coefficient** | **HR (95%CI)** | **p-value** |
| Severe COVID-19c-stat=0.832 | Age >60 years | 1.204 | 3.33 (2.13-5.02) | <0.001 |
| Previous viral exposure | -1.381 | 0.25 (0.16-0.414) | <0.001 |
| COPD | 0.964 | 2.62 (1.13-6.09) | 0.025 |
| Rhinorrhea | -0.985 | 0.37 (0.19-0.72) | 0.003 |
| Healthcare workers | -1.330 | 0.26 (0.08-0.85) | 0.025 |
| Diabetes | 0.905 | 2.47 (1.64-3.72) | <0.001 |
| CKD | 1.083 | 2.95 (1.59-5.48) | <0.001 |
| Mortalityc-stat=0.867 | Previous viral exposure | -1.544 | 0.21 (0.12-0.37) | <0.001 |
| Age >60 years | 1.491 | 4.44 (2.86-6.91) | <0.001 |
| Diabetes | 1.438 | 2.38 (1.52-3.70) | <0.001 |
| COPD | 0.926 | 2.52 (1.05-6.08) | 0.039 |
| CKD | 1.438 | 4.21 (2.26-7.84) | <0.001 |

**SUPPLEMENTARY REFERENCES**

1. Secretaría de Salud Subsecretaría de Prevención y Promoción de la Salud Dirección General de Epidemiología. LINEAMIENTOESTANDARIZADO PARA LA VIGILANCIA EPIDEMIOLÓGICA Y POR LABORATORIO DE LAENFERMEDAD RESPIRATORIA VIRAL. Apr 1, 2020.
2. Secretaría de Salud, Mexico City Government. “Programa De Detección De Casos COVID-19 y Sus Contactos.” CDMX, 2020, cdmx.gob.mx/portal/articulo/programa-de-deteccion-de-casos-covid-19-y-sus-contactos.
3. Han C et al. Digestive Symptoms in COVID-19 Patients With Mild Disease Severity: Clinical Presentation, Stool Viral RNA Testing, and Outcomes. Am J Gastroenterol. 2020;115(6):916-923.
4. Jin et al. Epidemiological, Clinical and Virological Characteristics of 74 Cases of Coronavirus-Infected Disease 2019 (COVID-19) With Gastrointestinal Symptoms. Gut. 2020 Jun;69(6):1002-1009.
5. World Health Organization. Q&A On Coronaviruses (COVID-19). World Health Organization, 2020, www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/q-a-coronaviruses#:~:text=symptoms.