Epidemiology of COVID-19 in Mexico: Symptomatic profiles and presymptomatic people

Miguel A. Fernández-Rojas[a], Marco A. Luna-Ruiz Esparza[a], Abraham Campos-Romero[a],
Diana Y. Calva-Espinosa[b], José L. Moreno-Camacho[b,c], Ariadna P. Langle-Martínez[b,d],
Abraham García-Gil[b,c], Claudia J. Solís-González[b,d], Adrián Canizalez-Román[e,f],
Nidia León-Sicairos[e,g], Jonathan Alcántar-Fernández[a,*]

[a] Innovation and Research Department, Salud Digna, Culiacan, 80000, Sinaloa, Mexico
[b] Clinical Laboratory Division, Salud Digna, Culiacan, 80000, Sinaloa, Mexico
[c] Molecular Biology Laboratory, National Reference Center, Salud Digna, La Primavera, Culiacan, 80199, Sinaloa, Mexico
[d] Molecular Biology Laboratory, National Reference Center "Mexico's Valley", Salud Digna, Los Reyes, Tlalnepantla de Baz, 54075, Estado de Mexico, Mexico
[e] CIASaP, School of Medicine, Autonomous University of Sinaloa, 80246 Culiacan, Sinaloa, Mexico
[f] The Women's Hospital, Secretariat of Health, 80127 Culiacan, Mexico
[g] Pediatric Hospital of Sinaloa, 80200 Culiacan, Mexico

A R T I C L E   I N F O

A B S T R A C T

Objectives: The COVID-19 diagnosis is difficult and ambiguous due to nonspecific symptoms. Further, data from Mexico arehospitable population-based without signs and symptoms information. Thus, this work aims to provide epidemiology information about the burden of COVID-19 in Mexican outpatients and to identify symptomatic COVID-19 profiles that could help in the early diagnosis of the disease.
Methods: From June to September, epidemiological, clinical, and demographic data of 482,413 individuals diagnosed by RT-PCR test for SARS-CoV-2 in Salud Digna clinics were collected.
Results: We observed a 41% incidence of SARS-CoV-2 infections with a mean age of 36 years and with young adults (20–40 years) being the most affected. Among occupations, delivery persons (OR 1.38) or informal traders (OR 1.33) had a higher risk of COVID-19. Moreover, 13% of SARS-CoV-2 infections were in presymptomatic patients. Finally, we identified three different symptomatic profiles (common, respiratory, and gastrointestinal) associated with COVID-19.
Conclusion: The incidence of SARS-CoV-2 was high among outpatients with a significant proportion of presymptomatic carriers, and thus it is necessary to increase testing and continue SARS-CoV-2 surveillance with a better description of signs and symptoms; in this regard, we identified three symptomatic profiles that could help in the diagnosis of COVID-19.

Introduction

At the start of the COVID-19 pandemic, the initial reports showed that SARS-CoV-2 infection promotes pneumonia in mainly older adults, which were the group with a major risk of getting hospitalized and getting mechanical ventilation due to severe COVID-19 disease (Berumen et al., 2020; Giannouchos et al., 2020; Zheng et al., 2020). However, other age-groups had similar risks due to different factors. Moreover, with regard to control of the pandemic, asymptomatic people have attracted attention due to their ability to transmit the virus without manifesting infection symptoms, which makes them identifiable only by difficult massive tests on the general population.

'Asymptomatic' and 'presymptomatic' carriers are terms commonly used indistinctly; however, they are not equivalent. By definition, asymptomatic carriers refer to individuals without symptoms but confirmed as SARS-CoV-2 positive by a laboratory test. On the other hand, presymptomatic carriers can be defined as cases without symptoms at the moment of test, but who could develop them later. Thus, presymptomatic carriers can only be detected by retrospective studies (Nikolai et al., 2020; Savvides and Siegel, 2020).

Several studies about this issue have shown that asymptomatic carrier's proportions ranged from 15 to 45% (Nikolai et al., 2020). These variations are due to age, presymptomatic status, virus

---

incubation time, and diagnosis mistakes (Buitrago-Garcia et al., 2020; Byambasuren et al., 2020; He et al., 2020; Oran and Topol, 2020). Likewise, similar virus shedding was demonstrated in symptomatic and asymptomatic patients (Cevik et al., 2020), possibly translating in the high ability to spread the virus by presymptomatic and asymptomatic patients which highlights the necessity of studying them in detail (Buitrago-Garcia et al., 2020; Byambasuren et al., 2020).

Moreover, the COVID-19 pandemic showed different risks according to the population structure and the association of different factors, including chronic diseases. In this regard, the situation in Mexico could be riskier due to the high prevalence of chronic diseases in the adult population such as hypertension (26%), obesity (36%), and diabetes (14%), which has been related to severe COVID-19 disease and mortality (Basto-Abreu et al., 2020; Campos-Nonato et al., 2018; OECD, 2020; SSA, 2020).

On February 27th, 2020, the first case of COVID-19 in Mexico was detected. (Forster et al., 2020). From this date, 1,255,974 confirmed cases of COVID-19 had been detected, with an overall mortality of 9%. Hypertension, obesity, and diabetes are the main comorbidities that increase risk of death by COVID-19 (SSA, 2020). Mexico's information comes from the Health Ministry's primarily hospital-based registries and shows a high incidence in men and people around 40 years; meanwhile, people with any comorbidity are more at risk of death by COVID-19 (Suárez et al., 2020).

Detailed epidemiology profile of COVID-19 in the country is not readily available; further, the current COVID-19 diagnosis is ambiguous due to the lack of specificity of the symptoms (Struyf et al., 2020). Moreover, the incidence of COVID-19 in outpatients, including those who are presymptomatic and symptomatic, and their epidemiological profile remains unknown. For the above, we analyzed epidemiological, clinical, and demographic data from 482,413 outpatients tested by RT-PCR for SARS-CoV-2 from 26 states of Mexico. Thus, we investigated the relevance of different factors and symptoms for SARS-CoV-2 infection to contribute the epidemiology of COVID-19 in the country.

Methods

Study design and population selection

The present work was a retrospective and transversal study. The collected data correspond to 482,413 outpatients of any sex and age evaluated for SARS-CoV-2 infection by PCR test in Salud Digna Clinics of 26 Mexican states from June 1st to September 30th. Only patients having clinical and demographic information and confirmed positive or negative status from SARS-CoV-2 test and accepting the use of their data by informed consent signing were considered. For geographical COVID-19 incidence analysis, we included states with 382 or more subjects. This threshold was considered from the sample size calculations for cross-sectional studies (Pourhoseingholi et al., 2013); in this case, we use the general COVID-19 incidence (47%) reported by the Ministry of Health of Mexico in the same period of this study (Direccion General de Epidemiologia, 2020). Finally standardization of incidence rate by state was made by age and sex and calculated with the direct method, using the standard world population as reference. The confidence interval was obtained considering a Poisson distribution for ASR according to (Boyle and Parkin, 1991).

Procedures

Nasopharyngeal swabs were collected in a viral transport medium to preserve samples until processing at the molecular biology laboratories at the National Reference Center located in Mexico State in Culiacan, Sinaloa. Molecular biologists processed samples in an automated and semiautomated workflow. Viral RNA extraction was automated with the MagNA Pure^{TM} 24 System (Roche Diagnostics, US) and the KingFisher TM Flex instrument (Thermo Fisher Scientific, US). PCR tests were run on Roche's fully automated cobas® 6800 and QuantStudio 7 Flex (Thermo Fisher). For SARS-CoV-2 detection, we used the cobas® SARS-CoV-2 Test (Roche Diagnostics), TaqMan 2019-nCoV Assay Kit v1 (Thermo Scientific), and VIASURE SARS-CoV-2 Real-Time PCR Detection Kit (CerTest Biotec). The TaqMan assay detects viral genes orf1ab, S, and N; cobas® test targets orf1ab and E viral genes; and VIASURE targets orf1ab and N genes. Human RNase P is used as an internal control.

For cobas® SARS-CoV-2 Test (Roche Diagnostics) and VIASURE SARS-CoV-2 Real-Time PCR Detection Kit (CerTest Biotec), a positive result corresponds with the amplification of target genes under Ct<38, meanwhile Ct<37 was used in the TaqMan 2019-nCoV Assay Kit v1 (Thermo Scientific) according to the manufacturers' recommendations.

Data collection

We applied a standardized survey prior to the SARS-CoV-2 PCR test to collect information regarding symptoms, clinical history, occupation, and lifestyle risk factors. PCR test results were according to the manufacturer's recommendations.

Consent for the use of information and privacy protection

Information analyzed in this work belongs to the patients who signed the informed consent for this study provided at the tests. A unique ID code was assigned to anonymize records for data privacy protection and prevent data duplication. Further, we aggregated information to enhance data protection. All procedures were carried out in adherence to the Mexican Federal Law on Personal Data Protection (LFPDPPP) and the privacy policy of Salud Digna.

Ethical statement

This study followed the approved guidelines for clinical information management, Helsinki's declaration, and Mexico's national regulations. This study was approved by the Ethical Review and Research Board of Salud Digna (SDI-2020-2).

Statistical analysis

Descriptive statistics and incidence rate

According to the data type, we performed a χ^2 test for categorical variables expressed as frequencies or percentages. Continuous variables are shown as mean or median and interquartile ranges (IQR). Standardized rates were calculated with the direct method using the standard world population (Ahmad et al., 2000); standardization was made by age groups every ten years and sex of every 100 persons.

Multivariate analysis

Age and sex-adjusted stepwise multivariate logistic regression was performed to evaluate the association between SARS-CoV-2 infection and potential risk factors (e.g., lifestyle).

Tandem clustering

To date, 11 symptoms are considered to be predictors of COVID-19 positiveness; however, some look little related between them or confuse clinicians. To determine the relevance of symptoms to SARS-CoV-2 infection, tandem clustering by Multiple Correspondence

M.A. Fernández-Rojas, M.A. Luna-Ruiz Esparza, A. Campos-Romero et al.

International Journal of Infectious Diseases 104 (2021) 572–579

Analysis (MCA) and K-mean method were run in R 4.0.2 (CRAN project) with FactoMineR and factoextra libraries. The cluster number was defined through a hierarchical cluster for Principal Components developed with 45,000 cases by Elbow Method. Subsequently, we calculate the agreement between these groups by kappa index (Ki = 0.97). We tested Hartigan-Wong, Lloyd, Forgy, and MacQueen methods with 1000 repetitions and 100,000 iterations for each one. Finally, we choose the MacQueen method for k-means due to its better interclass inertia.

Finally, a $p \leq 0.05$ value was considered to be a statistically significant threshold in all tests. Analyses were performed with SPSS 23 (SPSS Inc., Chicago, Ill., USA) or R 4.0.2 (CRAN project), and graphics were designed with GraphPad Prism 8 (GraphPad Software Inc., San Diego, CA.).

Results

Geographical distribution of SARS-CoV-2 infections

This work includes data from 26 of 32 states (81%) from Mexico, founding heterogeneous incidence of SARS-CoV-2 with at least one state per region with significantly higher incidence versus national incidence (Figure 1 A, Supplementary Table 1). Even inside Mexico Valley, there exists a clear difference between Mexico City and Mexico State (32 per 100 (95% CI = 30.99–32.73) vs 41 per 100 (95% CI = 39.62–41.42), respectively; Figure 2). Further, we observe that 14 states (54%) show a statistically significant different incidence versus the national rate (38 per 100) (Supplementary Table 1). Thus, States like Veracruz (49 per 100), Quintana Roo (44 per 100), Nuevo León (43 per 100), and Tabasco (43 per 100) have higher incidence, while Morelos (30 per 100), Guerrero (30 per 100), and Ciudad de Mexico (32 per 100) had lower incidences (among others) compared to the national incidence (Figure 2, Supplementary Table 1). Moreover, the temporal course of the SARS-CoV-2 pandemic shows a maximum incidence in the second week of July (50%), whereas it decreases after that (36% in September 2020) (Figure 1B).

Demographic characteristics of the population studied

From June 1st to September 30th, 482,413 nasopharyngeal swabs were collected in Salud Digna clinics; 244,171 (51%) are males, and 238,242 (49%) are females. The median age of all people studied is 36 years (IQR: 16–56 years) (Table 1). The child population is under-represented due to government confinement measures (including closing schools) implemented since March 27th (Figure 1C).

The COVID-19 incidence is 41% (n = 196,738) with a median age of 37 years; the incidence is higher in males (51%) than females (49%) (p < 0.0001); however, the male to female ratio is 1.02:1. (Table 1). Little more than half of the people (53%) were in contact with SARS-CoV-2 positive cases, with family members (56%) and co-workers (29%) being the primary source of infection (Table 1). Moreover, COVID-19 incidence in occupations like office workers (23%) and customer service workers (6%) was higher; however, occupations like delivery persons (OR 1.38, 95% CI 1.29–1.48), informal traders (OR 1.33, 95% CI 1.29–1.37), and transport personal (OR 1.22 95% CI 1.18–1.26) showed a higher risk for SARS-CoV-2 infections than others (Supplementary Table 2). Among health and lifestyle factors, we observe that 10% of people were active smokers, 24% received the influenza vaccine, and the main comorbidities were Hypertension (6%), Diabetes (5%), and Obesity (2%) (Table 1).

Symptoms related to COVID-19 in outpatients

To date, around 11 symptoms have been related to COVID-19, making its diagnosis very difficult because they are not specific to the disease. The main symptoms among people tested were headache (50%), arthralgia or myalgia (38%), and sore throat (36%) (Table 2). No differences were found in symptom incidence by sex or age. For the above, we performed a tandem clustering analysis to identify more related symptoms in confirmed COVID-19 cases to help healthcare workers to diagnose COVID-19 early. For this, we took the remaining SARS-CoV-2-positive cases (159,598) after
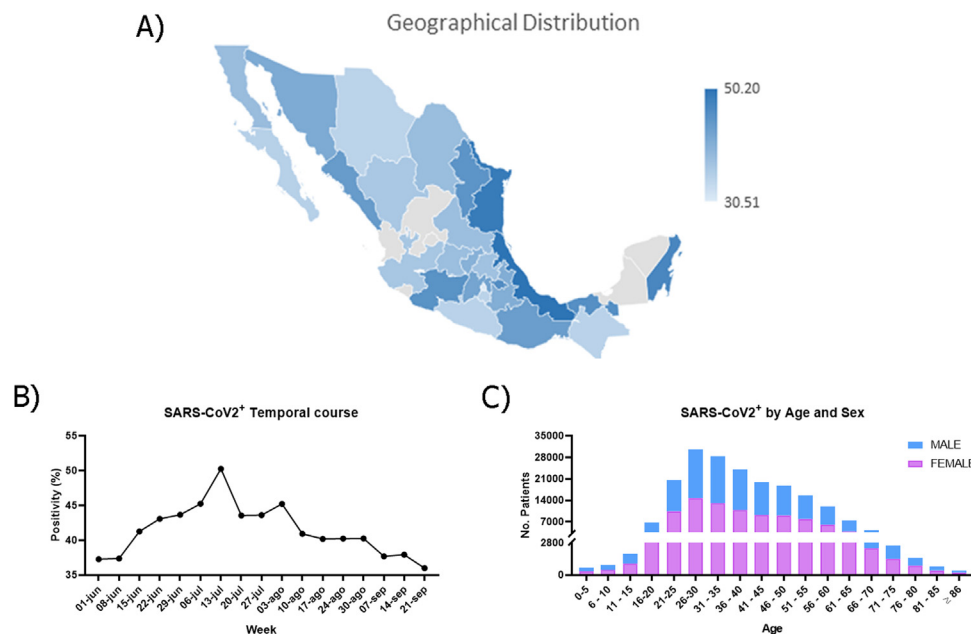
Figure 1. Demographic and Clinical data summary. (A) Coropletic map showing the incidence rate of SARS-CoV-2 infections in 26 states of Mexico; the dark-blue color indicates the states with a high incidence of infections, while light blue shows the states with a low incidence. The gray color in the map shows the states not included in this work; (B) Temporal trend of SARS-CoV 2 positive cases from June 1st to September 30th; (C) Sex and Age distribution of SARS-CoV 2 cases (n = 196,738).
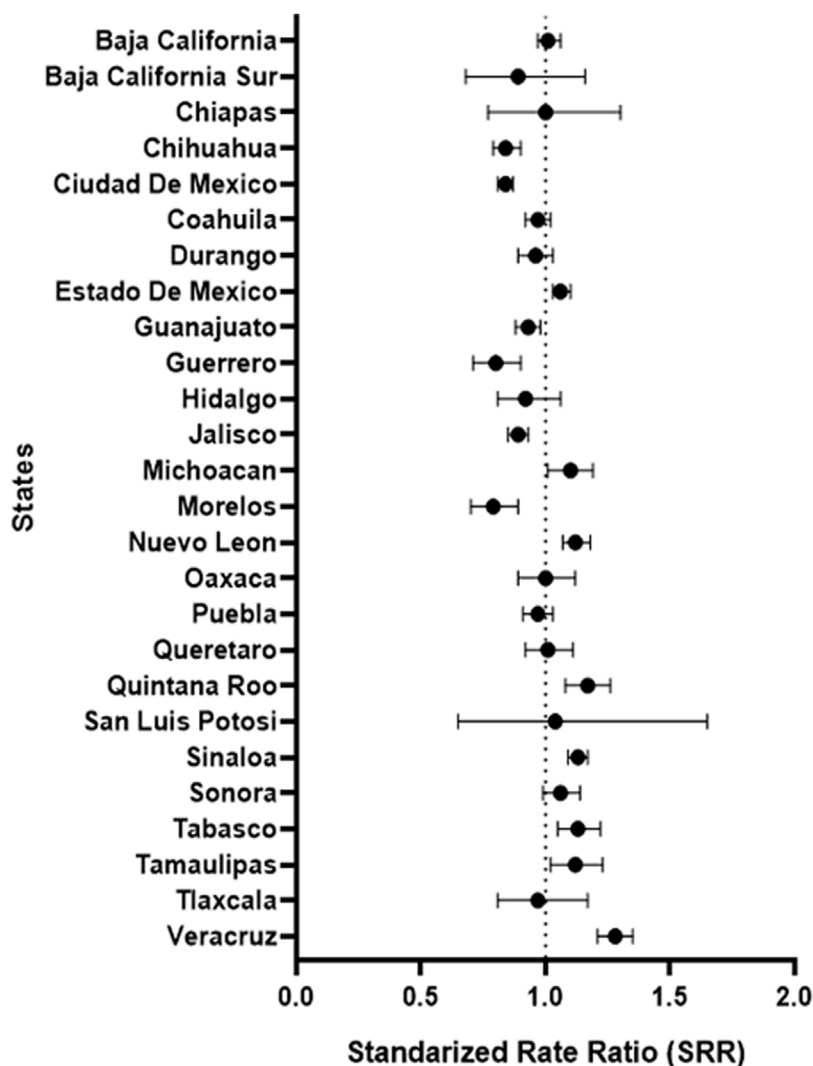
M.A. Fernández-Rojas, M.A. Luna-Ruiz Esparza, A. Campos-Romero et al.

International Journal of Infectious Diseases 104 (2021) 572–579

Figure 2. Standardized rate ratio (SRR) of SARS-CoV-2 infection in 26 states of Mexico, considering the general incidence as reference (1.0).

discounting presymptomatic people. This allowed us to identify three symptomatic groups associated with the incidence of COVID-19, who share headache, sore throat, and myalgia/arthralgia as frequent symptoms but differ between themselves due to other symptoms. We found one common cluster which was the most frequent (87,253 = 56%) in which none of the symptoms were currently prominent but had part of the core symptoms documented by the WHO (2020a): the respiratory cluster (52,827 = 34%) in which symptoms like fever (60%), cough (56%), chills (47%), runny nose (38%), and breathing difficulty (32%) are prominent; and the gastrointestinal cluster (16,518 = 10%) with a high frequency of diarrhea (74%), abdominal pain (63%), and vomiting (49%) symptoms in respect to the other two profiles (Figure 3 A and B).

Presymptomatic and symptomatic patients

Around 13% (25,520 cases) of SARS-CoV-2 positive patients had no symptoms on the test day and on the previous days. They were considered presymptomatic because they were not followed-up to determine whether they were asymptomatic. The sex distribution of presymptomatic and symptomatic carriers showed significant differences (presymptomatic males vs. females $p < 0.0001$; symptomatic males vs. females $p < 0.0001$); on the other hand, presymptomatic carriers were more common among older adults than children,

young people, and young adults (age <10 = 26%, age 10−20 = 27%, age 21−40 = 29%, age >65 = 31%, respectively), and these differences were statistically significant ($p < 0.0001$). Further, a similar trend was observed in symptomatic cases. This age-specific trend of presymptomatic and symptomatic cases was statistically significant ($p = 0.0273$, $p < 0.0001$, respectively) (Table 3).

Moreover, symptomatic cases were more in contact with COVID-19 cases (20%) than presymptomatic (9%); additionally, family members with COVID-19 were more frequently symptomatic (13%) than presymptomatic (6%). Moreover, non-communicable diseases (e.g., hypertension, diabetes, and obesity) were more prevalent in symptomatic than in presymptomatic cases. Finally, smokers and influenza vaccination were more frequent in symptomatic than in presymptomatic (Table 3).

Discussion

At present, the COVID-19 pandemic constitutes the most significant health challenge since the H1N1 and MERS-CoV epidemics. After 11 months of the Wuhan COVID-19 onset, over 71 million confirmed cases have been reported with a general mortality rate of 2% (WHO, 2020b). To date, America leads confirmed cases and death rates, with the US at the top (WHO, 2020b). Mexico is one of the countries with high SARS-CoV-2 incidences (43%), only behind Argentina (65%); meanwhile, other

M.A. Fernández-Rojas, M.A. Luna-Ruiz Esparza, A. Campos-Romero et al.

International Journal of Infectious Diseases 104 (2021) 572–579

Table 1

Demographic characteristics of the study population.

Characteristic	n[a]	% (95% CI)
SARS-CoV-2 Status		
Negative	285,675	59.22 (59.08–59.36)
Positive	196,738	40.78 (40.64–40.92)
Sex		
Female	238,242	49.39 (49.24–49.53)
Male	244,171	50.61 (50.47–50.76)
Age		
Median ± IQR	36 ± 20	−
<20	22,798	4.72 (4.67–4.79)
20–29	122,624	25.42 (25.30–25.54)
30–39	132,812	27.53 (27.40–27.66)
40–49	93,726	19.43 (19.32–19.54)
50–59	67,108	13.91 (13.81–14.01)
>60	43,345	8.98 (8.90–9.07)
Pregnant women[b]		
No	227,397	97.00 (96.93–97.07)
Yes	7038	3.00 (2.93–3.07)
Contact with COVID-19 patient[c]		
No	204,602	46.65 (46.50–46.80)
Yes	234,004	53.35 (53.20–53.50)
Relationship COVID-19 patient[d]		
Family	131,965	56.39 (56.19–56.60)
Co-worker	68,782	29.39 (29.21–29.58)
Friend	23,983	10.25 (10.13–10.37)
Other	20,112	8.59 (8.48–8.71)
Social security[e]		
No	87,009	18.83 (17.93–18.14)
Yes	375,106	81.17 (81.06–81.28)
Non-communicable diseases		
Hypertension	27,803	5.76 (5.70–5.83)
Diabetes	23,873	4.95 (4.89–5.01)
Obesity	7853	1.63 (1.59–1.66)
Smoking		
Non-Smokers	400,998	83.12 (83.02–83.23)
Passive smokers	32,272	6.69 (6.62–6.76)
Smokers	49,143	10.19 (10.10–10.27)
Influenza Vaccination[f]		
No	349,094	75.78 (75.66–75.91)
Yes	111,541	24.21 (24.09–24.34)

Note: Age groups was defined by decades, with the exception of <20 and >60.
[a] Total n = 482,413.
[b] Missing values = 3,807.
[c] Missing values = 43,807.
[d] It should be noted that some categories can be counted more than once.
[e] Missing values = 20,298.
[f] Missing values = 21,778.

Table 2

COVID-19-related symptoms in 482,413 outpatients.

Symptoms	Number of people (%)[a,b]	Positive SARS-CoV-2 (%)[*]
Abdominal pain	39,092 (8.10)	17,515 (44.80)
Breathing difficulty	69,513 (14.41)	34,588 (49.76)
Chills	69,515 (14.41)	39,003 (56.11)
Conjunctivitis	29,482 (6.11)	13,738 (46.60)
Constant dry cough	109,597 (22.72)	60,720 (55.40)
Diarrhea	71,829 (14.89)	32,843 (45.72)
Fever	106,126 (22.00)	60,209 (56.73)
Headache	238,987 (49.54)	106,103 (44.40)
Myalgia/Arthralgia	184,523 (38.25)	93,926 (50.90)
Runny nose	93,624 (19.41)	48,082 (51.36)
Sore throat	175,868 (36.46)	82,648 (46.99)
Vomit	31,812 (6.59)	15,841 (49.80)

Note: Chi-square calculated p-value.
[*] Statistically significant ($p < 0.0001$).
[a] Total n = 482,413.
[b] It should be noted that some categories could be counted more than once.

American countries with high confirmed cases like the US (~16 million cases) have lower incidence (7%) (WHO, 2020b). In this regard, Mexico's test rate for SARS-CoV-2 is 0.08% per thousand people versus 3% in the US or 0.43% in Argentina.

As a consequence, Mexico has a high reported mortality (9%) per confirmed cases of 114,298 confirmed deaths by COVID-19, which is higher than other American countries (the US–2%, Argentina–3% or Brazil–3%) (WHO, 2020b). In addition, the temporal analysis between July 1st to September 30th showed that after reaching a higher incidence in the second week of July (50%), this trend was reversed, showing a lower incidence (36%) by the end of September; however, the country-level incidence remains higher (43%).

At the state level, Suarez et al. reported that after the first month of the COVID-19 pandemic in Mexico, the states with higher incidence were Mexico City, Mexico State, Baja California, Sinaloa, and Tabasco. This work shows that main metropolises like Mexico City have lower COVID-19 incidence, while other states have ~18% more. Thus, as we can see, these data on the outpatients are similar to the hospitable database from the Mexican Health Ministry and indicate that we need to make some refinements to actual pandemic management to reduce the incidence and mortality of COVID-19.

Differences in the pandemic's onset in each state, population density, the obedience of mitigation actions to reduce virus spread, and disparities between rural and urban populations could help understand the regional incidence of COVID-19 in the country (Suárez et al., 2020). Similarly, in the US, people who live in rural areas have less access to the SARS-CoV-2 test than people who live in big cities like New York, which shows how the social, economic, and health discrepancies can impact the burden of COVID-19 (Souch and Cossman, 2020). Likewise, in Mexico, the South and Occident regions have a higher vulnerability index than other regions, limiting their access to health services (Santana-Casta-ñeda, 2020).

The first reports of COVID-19 proposed different risks of develop severe COVID-19 disease and consequently death between sexes (Englmeier, 2020; de Groot and Bontrop, 2020). Our results show a slight difference in the incidence of COVID-19 in females and males; however, the male to female incidence ratio was 1.02:1. The sex difference is more significant in mortality rate than in the incidence of COVID-19 (Peckham et al., 2020). Thus, even with the evidence of higher amounts of ACE2 receptors in men and the lower number of receptors involved in innate immunity (Eng-lmeier, 2020; de Groot and Bontrop, 2020), more epidemiological information is needed to clarify its relevance to COVID 19 etiology.

Similarly, an increase in COVID-19 severity has been associated with older people (Lusignan et al., 2020), while the risk of SARS-CoV-2 infection in this study was high in young adults. In this study, around 77% were people under 50 years (Figure 1B); this can be due to the population-age distribution in Mexico compared to other countries. De Souza et al. showed that other Latin-American countries like Brazil have a similar age-group distribution and found similar conclusions (Souza et al., 2020). Additionally, the infection in young adults could be because population groups around their thirties and forties are economically active, which means a higher exposure rate due to mobility (Amorim et al., 2020).

Thus, we found that logistics people, informal traders, or delivery persons have a higher risk of SARS-CoV-2 infection that could be explained by a high rate of contact due to their professions. In occupations like office workers, customer service, and public servants, the risk is lower than other jobs might be due to the reduction in mobility's; however, we cannot assure this statement due to a lack of information about working modalities in the last group. Furthermore, an increase in mobility is related to a higher risk of being infected, as seen in other countries, due to the relaxation of confinement measures.

Moreover, it is difficult to infer the viral infection's presence through the symptoms because they are not specific to COVID-19,
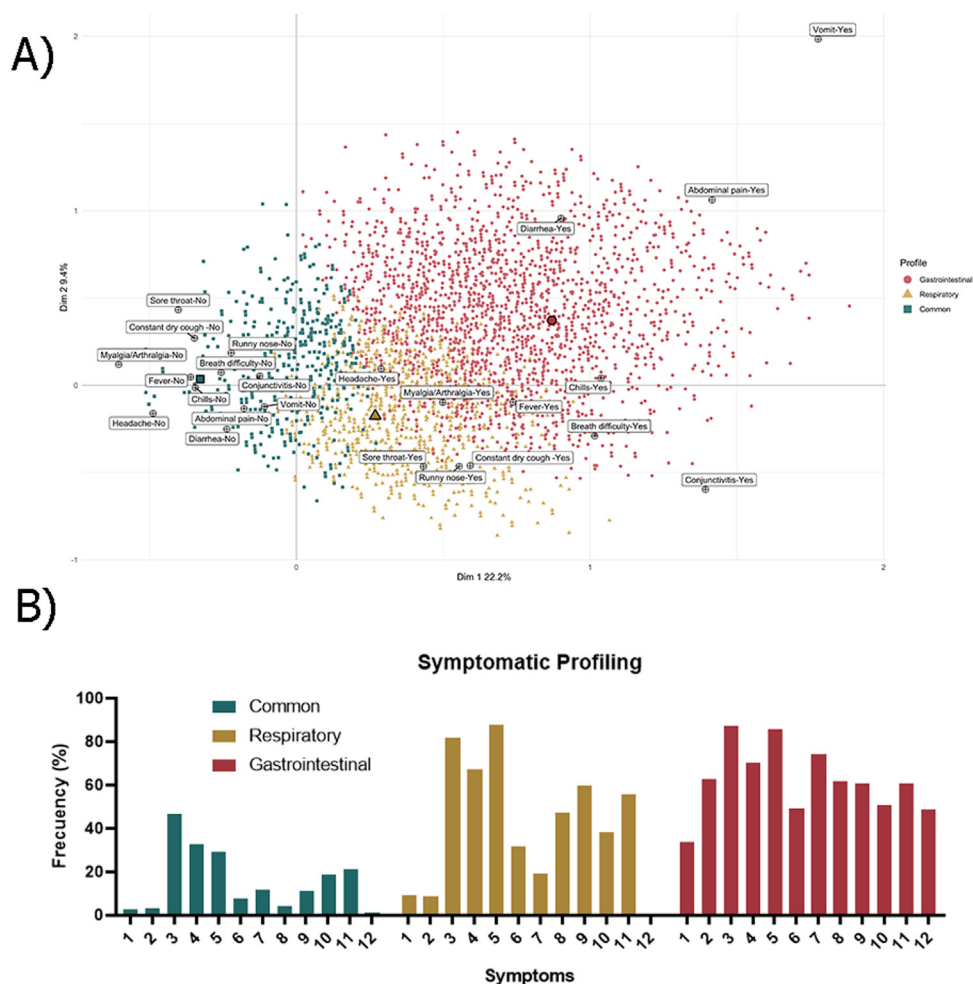
Figure 3. Tandem clustering analysis for COVID-19 related symptoms. Panel A shows the symptoms' dispersion plot through main component analysis (MCA); the red color indicates the gastrointestinal cluster, gold shows the respiratory cluster, and blue shows the common cluster (no symptoms stand out). Dimension 1 (Dim 1) and Dimension 2 (Dim 2) shows the variance explained by each one (22.2 and 9.4%, respectively). Panel B shows the frequency and distribution of COVID-19 related symptoms in each cluster identified. Symptoms are indicated as 1 = Conjunctivitis, 2 = Abdominal pain, 3 = Headache, 4 = Sore throat, 5 = Myalgia/Arthralgia, 6 = Breath difficulty, 7 = Diarrhea, 8 = Chills, 9 = Fever, 10 = Runny nose, 11 = Constant dry cough, and 12 = Vomiting).

which could confuse clinicians delaying the diagnosis. Therefore, we perform a tandem clustering analysis to identify possible relationships among symptoms reported. We found three symptomatic profiles: patients with non-specificity for any symptoms, others with mainly respiratory symptoms (Fever, Cough, Chills, Runny nose, and Breath difficulty), and those with gastrointestinal features (Diarrhea, Abdominal pain, and Vomiting); this classification could help healthcare workers to improve the diagnosis and treatment of COVID-19. Unfortunately, because of our study's nature, we could not determine the relationship between symptomatic profiling and clinical outcomes. Therefore, more studies are needed to assess if these clusters could predict clinical outcomes of COVID-19.

Finally, one issue that is still controversial is the presymptomatic/asymptomatic patient's contribution to the SARS-CoV-2 spread. In this regard, the Centers for Disease Control and Prevention (CDC) determined that one of the major SARS-CoV-2 infected people groups is asymptomatic. In the best-predicted scenario, these people have 75% efficiency to contagious other people (CDC, 2020). A meta-analysis reported a proportion of asymptomatic people from 4% to 50% (Oran and Topol, 2020). In this study, patients were not followed-up for any arising symptoms later, and people with positive SARS-CoV-2 tests who showed no symptoms on the test day or on previous days were considered to

be presymptomatic, which was about 13% of cases. Recent evidence in Korea showed that 36% of people analyzed were asymptomatic at the time of the PCR test, but 19% of these patients developed symptoms soon after the test and were virus transmitters even after three weeks (Lee et al., 2020), which emphasizes that the lack of symptoms does not mean less transmission capacity. In addition, we found that the proportion of SARS-CoV-2 positive and presymptomatic people is higher in older people (31%) versus children (26%). Current evidence is controversial; some studies indicated higher presymptomatic proportions in children versus older people arguing that lower ACE2 receptor amounts in young people and less aggressive immune system response explain the high rates of presymptomatic children (He et al., 2020; Ma et al., 2020); however, other studies showed no significant differences between age groups and presymptomatic proportions (Jung et al., 2020; Yu et al., 2020). Thus, more studies are necessary to characterize presymptomatic and asymptomatic carriers better, and increased testing could help identify them to reduce the virus spread.

Strengths and limitations

This work provides a general insight into the burden of COVID-19 in Mexican outpatients. It is important to note that we studied a

Table 3

Demographic characteristics of presymptomatic and symptomatic patients.

Characteristic	Presymptomatic (n = 86,196)			Symptomatic (n = 396,217)			Presymptomatic vs Symptomatic[£] p-value
	Positive SARS-CoV-2 (%)	95% CI	p-value	Positive SARS-CoV-2 (%)	95% CI	p-value	
Total	25,520 (29.61)	(29.30–29.91)		171,218 (43.21)	(43.06–43.37)		—
Sex							
Female	11,313 (13.12)	(12.90–13.35)	<0.0001	84,285 (21.27)	(21.15–21.40)	<0.0001	<0.0001
Male	14,207 (16.48)	(16.23–16.73)		86,933 (21.94)	(21.81–22.07)		
Age[€]							
<10	288 (25.62)	(23.07–28.18)	0.0003	955 (34.64)	(32.90–36.39)	<0.0001	<0.0001
10–20	1413 (27.41)	(26.19–28.63)	<0.0001	7493 (37.86)	(37.18–38.54)	<0.0001	
21–40	13,042 (28.80)	(28.39–29.23)	0.0008	90,248 (42.08)	(41.88–42.29)	<0.0001	
41–65	9418 (31.09)	(30.57–31.61)	ns	64,279 (45.31)	(45.06–45.58)	<0.0001	
>65	1359 (31.21)	(29.84–32.59)	ref	8203 (47.48)	(46.74–48.23)	ref	
Contact with COVID-19 cases[a]							
No	7401 (8.59)	(8.40–8.77)	<0.0001	81,591 (20.59)	(20.47–20.72)	<0.0001	<0.0001
Yes	8152 (9.46)	(9.26–9.55)		81,975 (20.69)	(20.56–20.82)		
Family member with COVID-19[b]							
No	3067 (3.56)	(3.43–3.68)	<0.0001	121,600 (30.69)	(30.55–30.83)	<0.0001	<0.0001
Yes	5085 (5.90)	(5.74–6.06)		49,618 (12.52)	(12.42–12.63)		
Non-communicable diseases							
Hypertension							
No	24,449 (28.36)	(28.06–28.67)	ns	159,917 (40.36)	(40.21–40.51)	<0.0001	<0.0001
Yes	1071 (1.24)	(1.17–1.32)		11,301 (2.85)	(2.80–2.90)		
Diabetes							
No	24,615 (28.56)	(28.26–28.86)	ns	160,904 (40.61)	(40.46–40.76)	ns	<0.0001
Yes	905 (1.05)	(0.98–1.12)		10,314 (2.60)	(2.55–2.65)		
Obesity							
No	25,306 (29.36)	(29.05–29.66)	ns	168,006 (42.40)	(42.25–42.56)	0.0013	<0.0001
Yes	214 (0.25)	(0.22–0.28)		3212 (0.81)	(0.78–0.84)		
Smoking							
Non-smokers	22,481 (26.08)	(25.79–26.37)	ref	144,547 (36.48)	(36.33–36.63)	ref	<0.0001
Passive smokers	1236 (1.43)	(1.35–1.51)	<0.0001	11,602 (2.93)	(2.88–2.98)	<0.0001	
Smokers	1803 (2.09)	(2.00–2.19)	<0.0001	15,069 (3.80)	(3.74–3.86)	<0.0001	
Influenza Vaccination[c]							
No	18,630 (21.61)	(21.34–21.89)	<0.0001	123,991 (31.29)	(31.15–31.44)	ns	<0.0001
Yes	4736 (5.49)	(5.34–5.65)		40,808 (10.30)	(10.20–10.39)		

Note:

[£] Presymptomatic vs. Symptomatic p-value, calculate by chi–square test for every feature. Age groups were defined as children (<10), young people (10–20), young adults (21–40), adults (41–65), and older adults (>65).

[€] The trend of age-specific incidence was evaluated with the chi-squared ($\chi 2$) test for trend and was statistically significant in presymptomatic (p < 0.0273) and symptomatic (p < 0.0001).

[a] Missing values: Presymptomatic = 9,967, Symptomatic = 7,652.

[b] It should be noted that some categories can be counted more than once.

[c] Missing values = Presymptomatic = 2,154, Symptomatic = 6,419.

self-selected population tested for SARS-CoV-2 infection, which involves a potential sub-representation of subjects with specific features such as age (e.g., children) and jobs (e.g., healthcare workers).

Therefore, because of patients' lack of follow-up in this study, we could not accurately identify presymptomatic and asymptomatic people. Further, this study's information does not represent the Mexican population as a whole; however, this work and the hospital-based information could help us understand the pandemic's behavior.

Moreover, in the case of symptoms, they were self-reported by patients, allowing the possibility that they could be misidentified in some cases, which could impact our tandem clustering analysis. Thus, more studies in other populations are needed to enhance the knowledge of clinical manifestations of COVID-19.

In conclusion, we found a high SARS-CoV-2 incidence (41%), with young adults (26–40 years) being the most affected. Further, we found a presymptomatic carrier's proportion of 13%, which increases with age; this could help implement triage in target populations to identify them and reduce virus spread. Furthermore, we identified three symptomatic COVID-19 profiles: common (56%), respiratory (34%), and gastrointestinal (10%) that we expected could help clinicians diagnose COVID-19 early. Further studies are needed to determine whether these profiles could predict a patient's clinical outcomes.

Conflict of interests

None declare.

Founding source

None declare.

Ethical approval

This study was approved by the Ethical Review and Research Board of Salud Digna (SDI-2020-2).

M.A. Fernández-Rojas, M.A. Luna-Ruiz Esparza, A. Campos-Romero et al.

International Journal of Infectious Diseases 104 (2021) 572–579

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ijid.2020.12.086.

References

Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJL, Lozano R, Inoue M. Age Standardization of Rates: a New WHO Standard. GPE Discussion Paper Series; 2000. p. 10–24.

Amorim L, Szwarcwald CL, Mateos Sde OG, de Leon ACMP, Medronho Rde A, Veloso VG, et al. Seroprevalence of anti-SARS-CoV-2 among blood donors in Rio de Janeiro, Brazil. Rev Saude Publica 2020;54, doi:http://dx.doi.org/10.11606/s1518-8787.2020054002643.

Basto-Abreu A, Barrientos-Gutiérrez T, Rojas-Martínez R, Aguilar-Salinas CA, López-Olmedo N, Cruz-Góngora VD la, et al. Prevalencia de diabetes y descontrol glucémico en México: resultados de la Ensanut 2016. Salud Publica Mex 2020;62:50–9.

Berumen J, Schmulson M, Alegre J, Guerrero G, Olaiz G, Wong-Chew RM, et al. Risk of infection and hospitalization by Covid-19 in Mexico: a case-control study. MedRxiv 2020;, doi:http://dx.doi.org/10.1101/2020.05.24.20104414 Submitted for publication.

Boyle P, Parkin DM. Cancer registration: principles and methods. Statistical methods for registries. IARC Sci Publ 1991;126–58.

Buitrago-Garcia D, Egli-Gany D, Counotte MJ, Hossmann S, Imeri H, Ipekci AM, et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: a living systematic review and meta-analysis. PLoS Med 2020;17(9):1–25, doi:http://dx.doi.org/10.1371/journal.pmed.1003346.

Byambasuren O, Cardona M, Bell K, Clark J, McLaws M-L, Glasziou P. Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis. MedRxiv 2020;5(4):223–34, doi:http://dx.doi.org/10.1101/2020.05.10.20097543.

Campos-Nonato I, Hernández-Barrera L, Pedroza-Tobías A, Medina C, Barquera S. Hipertensión arterial en adultos mexicanos: prevalencia, diagnóstico y tipo de tratamiento. Ensanut MC 2016. Salud Publica Mex 2018;60:233–43.

CDC. Coronavirus Disease 2019 (COVID-19). Centers for Disease Control and Prevention; 2020. https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html.

Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. Lancet Microbe 2020;2(1)13–22, doi:http://dx.doi.org/10.1016/S2666-5247(20)30172-5 0.

de Groot NG, Bontrop RE. COVID-19 pandemic: is a gender-defined dosage effect responsible for the high mortality rate among males?. Immunogenetics 2020;1–3, doi:http://dx.doi.org/10.1007/s00251-020-01165-7.

Direccion General de Epidemiologia. Sitio Oficial COVID-19 México Dirección General de Epidemiología-Gobierno Federal. COVID-19. Mexico: Datos Epidemiologicos; 2020. https://covid19.sinave.gob.mx/.

Englmeier L. A theory on SARS-COV-2 susceptibility: reduced TLR7-activity as a mechanistic link between men, obese and elderly. J Biol Regul Homeost Agents 2020;34(3):1125–9, doi:http://dx.doi.org/10.23812/20-221-L-49.

Forster P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-2 genomes. Proc Natl Acad Sci USA 2020;117:9241–3, doi:http://dx.doi.org/10.1073/pnas.2004999117.

Giannouchos TV, Sussman RA, Mier JM, Poulas K, Farsalinos K. Characteristics and risk factors for COVID-19 diagnosis and adverse outcomes in Mexico: an analysis of 89,756 laboratory–confirmed COVID-19 cases. Eur Respir J 2020;57(2):1–30, doi:http://dx.doi.org/10.1183/13993003.02144-2020.

He J, Guo Y, Mao R, Zhang J. Proportion of asymptomatic coronavirus disease 2019: a systematic review and meta-analysis. J Med Virol 2020;93(2):820–30, doi:http://dx.doi.org/10.1002/jmv.26326.

Jung C-Y, Park H, Kim DW, Choi YJ, Kim SW, Chang TI. Clinical characteristics of asymptomatic patients with COVID-19: a nationwide cohort study in South Korea. Int J Infect Dis 2020;99:266–8, doi:http://dx.doi.org/10.1016/j.ijid.2020.08.001.

Lee LYW, Cazier J-B, Starkey T, Briggs SEW, Arnold R, Bisht V, et al. COVID-19 prevalence and mortality in patients with cancer and the effect of primary tumour subtype and patient demographics: a prospective cohort study. Lancet Oncol 2020;21(10):1309–16, doi:http://dx.doi.org/10.1016/S1470-2045(20)30442-3.

Lusignan S de, Dorward J, Correa A, Jones N, Akinyemi O, Amirthalingam G, et al. Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of General Practitioners Research and Surveillance Centre primary care network: a cross-sectional study. Lancet Infect Dis 2020;20:1034–42, doi:http://dx.doi.org/10.1016/S1473-3099(20)30371-6.

Ma Y, Xu Q, Wang F, Ma X, Wang X, Zhang X, et al. Characteristics of asymptomatic patients with SARS-CoV-2 infection in Jinan, China. Microbes Infect 2020;22:212–7, doi:http://dx.doi.org/10.1016/j.micinf.2020.04.011.

Nikolai LA, Meyer CG, Kremsner PG, Velavan TP. Asymptomatic SARS coronavirus 2 infection: invisible yet invincible. Int J Infect Dis 2020;100:112–6, doi:http://dx.doi.org/10.1016/j.ijid.2020.08.076.

OECD. Health Risks–Overweight or Obese Population-OECD Data. The OECD; 2020. http://data.oecd.org/healthrisk/overweight-or-obese-population.htm.

Oran DP, Topol EJ. Prevalence of asymptomatic SARS-CoV-2 infection. Ann Intern Med 2020;173(5):362–7, doi:http://dx.doi.org/10.7326/M20-3012.

Peckham H, de Gruijter NM, Raine C, Radziszewska A, Ciurtin C, Wedderburn LR, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. Nat Commun 2020;11(1):1–10, doi:http://dx.doi.org/10.1038/s41467-020-19741-6.

Pourhoseingholi MA, Vahedi M, Rahimzadeh M. Sample size calculation in medical studies. Gastroenterol Hepatol Bed Bench 2013;6:14–7.

Santana-Castañeda G. Vulnerabilidad diferencial de los Estados mexicanos frente al COVID-19. Posición 2020;(3)2–24 ISSN: 2683-8915.

Savvides C, Siegel R. Asymptomatic and presymptomatic transmission of SARS-CoV-2: a systematic review. MedRxiv 2020;, doi:http://dx.doi.org/10.1101/2020.06.11.20129072 Submitted for publication.

Souch JM, Cossman JS. A commentary on rural-urban disparities in COVID-19 testing rates per 100,000 and risk factors. J Rural Health 2020;37(1):188–90, doi:http://dx.doi.org/10.1111/jrh.12450.

de Souza WM, Buss LF, Candido D da S, Carrera JP, Li S, Zarebski A, et al. Epidemiological and clinical characteristics of the early phase of the COVID-19 epidemic in Brazil. MedRxiv 2020;4:856–65, doi:http://dx.doi.org/10.1101/2020.04.25.20077396.

SSA. COVID-19–Coronavirus. 2020. https://coronavirus.gob.mx/covid-19/.

Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MM, et al. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease. Cochrane Database Systemat Rev 2020;7(7):1–96, doi:http://dx.doi.org/10.1002/14651858.CD013665.

Suárez V, Suarez Quezada M, Oros Ruiz S, Ronquillo De Jesús E. Epidemiology of COVID-19 in Mexico: from the 27th of February to the 30th of April 2020. Revista Clínica Española (English Edition) 2020;220(8):463–71, doi:http://dx.doi.org/10.1016/j.rceng.2020.05.008.

WHO. WHO Coronavirus Disease (COVID-19) Coronavirus Disease 2019 Q&Ashttps://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub[Accessed 15 November 2020]. 2020.

WHO. WHO Coronavirus Disease (COVID-19) Dashboard 2020. 2020. https://covid19.who.int.

Yu C, Zhou M, Liu Y, Guo T, Ou C, Yang L, et al. Characteristics of asymptomatic COVID-19 infection and progression: a multicenter, retrospective study. Virulence 2020;11:1006–14, doi:http://dx.doi.org/10.1080/21505594.2020.1802194.

Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, et al. Risk factors of critical & mortal COVID-19 cases: a systematic literature review and meta-analysis. J Infect 2020;81:e16–25, doi:http://dx.doi.org/10.1016/j.jinf.2020.04.021.