Socioeconomic inequalities associated with mortality for COVID-19 in Colombia: A cohort nation-wide study

Myriam Patricia Cifuentes, Laura A. Rodriguez-Villamizar, Maylen Liseth Rojas-Botero, Carlos Alvarez-Moreno, Julián A. Fernández-Niño

1 PhD, Direction of Epidemiology and Demography, Ministry of Health and Social Protection, Bogotá, Colombia
2 PhD, School of Medicine, Universidad Industrial de Santander, Bucaramanga, Colombia
3 PhD, Faculty of Medicine, Universidad Nacional de Colombia; Clínica Universitaria Colsanitas, Bogotá, Colombia

Corresponding author:
Laura A Rodriguez-Villamizar. Carrera 32 29-31 Of. 301, Bucaramanga, Colombia
Email: laurovi@uis.edu.co Telephone: +57-7-6344000 ext. 3145.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
ABSTRACT

Background: After eight months of the COVID-19 pandemic, Latin American countries have some of the highest rates in COVID-19 mortality. Despite being one of the most unequal regions of the world, there is a scarce report of the effect of socioeconomic conditions on COVID-19 mortality in their countries. We aimed to identify the effect of some socioeconomic inequality-related factors on COVID-19 mortality in Colombia.

Methods: We conducted a survival analysis in a nation-wide retrospective cohort study of confirmed cases of COVID-19 in Colombia from March 2nd to October 26th, 2020. We calculated the time to death or recovery for each confirmed case in the cohort. We used an extended multivariable time-dependent Cox regression model to estimate the hazard risk ratio (HR) by age groups, sex, ethnicity, type of health insurance, area of residence, and socioeconomic strata.

Results: There were 1,033,218 confirmed cases and 30,565 deaths for COVID-19 in Colombia between March 2nd and October 26th. The risk of dying for COVID-19 among confirmed cases was higher in males (HR=1.68 95% CI: 1.64-1.72), in people older than 60 years (HR=296.58 95% CI: 199.22-441.51), in indigenous people (HR=1.20 95% CI: 1.08-1.33), in people with subsidized health insurance regime (HR=1.89 95% CI: 1.83-1.96), and in people living in the very low socioeconomic strata (HR=1.44 95% CI: 1.24-1.68).

Conclusion: Our study provides evidence of socioeconomic inequalities in COVID-19 mortality in terms of age groups, sex, ethnicity, type of health insurance regime, and socioeconomic status.

Key words: COVID-19; Mortality; Social Inequalities; Cohort studies
INTRODUCTION

The coronavirus disease (COVID-19) is the first pandemic caused by a human coronavirus, the SARS-CoV-2. The first cluster of patients with pneumonia of unknown origin was reported in Wuhan, China in January, 2020.[1] As of October 31st, 2020, there were more than 45.5 million confirmed cases and 1.1 million deaths affecting 188 countries around the world. The region of the Americas is the most affected region accounting for more than 20.3 million confirmed cases and 636,482 deaths.[2]

Recently it has been declared that the situation due to COVID-19 corresponds to a syndemic since there is a combination between the epidemic due to the infection by SARS-CoV-2 and the epidemic due to chronic non-communicable diseases that interact in a social context of poverty and inequity.[3] There are three crisis affecting economies and societies in the region: the slow economic growth, the environmental emergency, and the growing inequality.[4] The combination of these social crises with the endemic of non-communicable diseases and the current pandemic are disproportionately affecting the region. Latin America currently holds some of the highest COVID-19 death rates in the world and is facing a humanitarian crisis powered by the longstanding inequality of its countries.[5]

COVID-19 has been recognized by some governments and media as “the great equalizer” due to its capacity to affect people of different age groups, socioeconomic conditions, or prestige.[6] While this is probably true in terms of the biological risk of infection, it is not the case for the observed risk of COVID-19 infection, severity and mortality. There is evidence of racial and socioeconomic disparities in the United States in terms of the population infected by and dying from COVID-19. [7] However, socioeconomic characteristics are not routinely collected or described in most COVID-19 analyses. [8] Therefore, there is a need for collecting and analyzing data on socioeconomic determinants of health to monitor COVID-19 inequities, identify high-risk populations, and guide the development of public health interventions within countries.

During the first wave of the pandemic by SARS-CoV-2 infection/COVID-19 in Colombia, South America, the national public health surveillance system early adapted and prepared for this new threat, being able to detect and follow-up the ongoing cases and their demographic and socioeconomic characteristics. To identify the effect of some demographic and socioeconomic inequality-related factors on COVID-19 mortality during the first eight months of the epidemic in Colombia, we conducted a survival analysis (time to death for COVID-19) using individual data from a nation-wide cohort.

METHODS

Study population
Colombia is located in the north corner of South America. According to the National Administrative Department of Statistics (DANE, for its initials in Spanish), the total population is projected by 2020 in 50,372,424 inhabitants.[9] The country is divided into 33 departments and districts which groups 1,122 municipalities. Half of the population are women (51.2%), 77.1% of people live in urban areas and 68.2% of Colombians are between 15 and 64 years old. The first case of infection for SARS-CoV-2 was confirmed on March 6th in Bogotá.

Study design and data sources
We conducted a survival analysis in a nation-wide retrospective cohort study of confirmed cases of COVID-19 in Colombia from March 2nd to October 26th, 2020. The nation-wide cohort was assembled using individual data obtained from the national public health surveillance system (SIVIGILA, for its initials in Spanish). The National Institute of Health (INS, for its initials in Spanish) compiles, verifies and adds laboratory data and other criteria for confirm or discard cases and publishes anonymized and de-identified registries as open data (www.ins.gov.co). The first day of symptoms’ onset for the first confirmed case was February 26th and there were 245 days elapsed till the end of the follow-up period.

Symptomatic and asymptomatic COVID-19 cases are confirmed in Colombia by using Reverse transcription polymerase chain reaction (RT-PCR). Starting on July 23rd, 2020, symptomatic cases can be also confirmed by using antigen-based validated tests. Deaths for COVID-19 are notified by health care services to SIVIGILA and DANE and then an individual analysis of the cases confirms, discards or keeps as suspected the reports of deaths due to COVID-19.
All procedures performed in this study followed the national and international ethical standards. Informed consent was not required due to the nature of the study and use of anonymized data from publicly available data sources.

**Outcome and predictors assessment**

The outcome event of interest for the study was COVID-19 death among confirmed cases. For deceased symptomatic cases, we computed the ‘time to event’ as the difference between the dates of symptoms onset and the date of death. For deceased asymptomatic cases, ‘time to event’ was calculated as the difference between the date of first medical appointment and the date of death. The follow-up time of asymptomatic recovered cases was the difference between the date of symptoms’ onset and date of recovering (registered as laboratory or clinical recovery). For asymptomatic recovered cases, the follow-up time was calculated as the difference between the date of the first medical appointment and the registered date of recovery. We defined censored cases as active cases for which no event (death or recovery) was confirmed by the last date of observation (October 26th). For symptomatic and asymptomatic censored cases, we used the same calculation of the follow-up time described above taking into account the date of symptoms’ onset or the date of the first medical appointment date, respectively (See Figure S1 in Supplementary material).

The exposure predictor variables of the model included the following individual demographic and socioeconomic variables that usually leads to health inequalities: age, sex, ethnicity, type of health insurance, area of residence, and socioeconomic strata. In Colombia, ethnicity minorities include “indigenous”, “African-Colombian descent”, a special group of “Raizales” which refers to descendants of the original enslaved Africans, and “Gipsy-Romany”. The type of health insurance is a proxy variable for health access. The “Contributory” type refers to job-related health insurance, the “Subsidized” type covers poor people without formal jobs which hold a subsidy paid by the government, the “Special” health regime covers few unionized workers, and the “exception” regime groups the army-related members. We included age grouped into six life-course categories: infants (0-5y), children and school-age (5-11y), adolescents (12-16y), young adults (17-44y), adults (45-69y), and seniors (60 or more years). The socioeconomic strata is a classification based on the stratification of residential properties used by DANE according to the socioeconomic resources of a census block that received public services. This classification divides houses into class levels, which range from one (very low) to six (high) being one the strata with higher socioeconomic deprivation.[10] The socioeconomic strata is used as a proxy of socioeconomic status (SES) in this study and was obtained as a self-reported variable in SIVIGILA.

**Statistical analysis**

All confirmed COVID-19 cases were included in the analysis by using the national cohort’s time to death and recovery. Exploratory customary descriptions for distributions of continuous time-to-death variable and all categorical predictors, for the different outcomes (dead, recovered and censored) included means, medians, frequencies and percentages. The modeling process included testing of proportional hazards assumption for all the predictors using hypothesis tests (p-values of terms addressing time dependent factors) and graphs (Log minus log plots and partial residuals plots from models with no interaction terms). As all predictors but sex were dependent on time, we used an extended multi-predictor time-dependent Cox regression model.[11] By using this extended model it is possible to jointly evaluate the effect of multiple time-dependent variables and their role as potential confounders or effect modifiers. We included simple product interactions between these variables and the time to event to estimate an extended Cox regression model that allows non-proportional hazards. Survival functions were calculated using the Kaplan-Meier method. As our objective was to obtain an explanatory model, we ran the multi-predictor regression models by the Enter method, therefore, the resulting equation included all variables. We assessed the coefficient signs and significance by the Wald statistic, and associations expressed as hazard ratios (HRs) with 95% confidence intervals (CIs). All tests with \( p < 0.05 \) were considered statistically significant. We performed the statistical analysis using SPSS software® version 26.

**RESULTS**

There were 1 033 218 confirmed cases and 30 565 deaths for COVID-19 in Colombia from the first day of notification, March 2nd, to October 26th. Table 1 summarizes the characteristics of the cohort of COVID-19 confirmed cases. Most confirmed COVID-19 cases were male, older than 60 years old, living in urban areas, with the contributory regime of health insurance, and living in residences that belong to the two lower levels of socioeconomic strata. Seven (0.02%) out of the 30 565 cases that end up in deaths were asymptomatic. From all 914 882 confirmed cases that end up in recovery, 11.2% were asymptomatic. The 12.1% of the 87 874 confirmed cases censored at the end of the follow-up time were asymptomatic (See Tables S1-S3 in Supplementary material).
Table 1. Sociodemographic characteristics of COVID-19 confirmed cases and deaths in Colombia up to and including October 26th, 2020

| Characteristic          | Total confirmed cases alive (N= 1 002 653) | Total confirmed deaths (N= 30 565) | Total confirmed cases (N= 1 033 218) |
|-------------------------|--------------------------------------------|-----------------------------------|--------------------------------------|
|                         | (n, % for rows)                             | (n, % for rows)                   | (n, % for columns)                    |
| Sex                     |                                            |                                   |                                      |
| Male                    | 500 820 (96.23)                            | 19 613 (3.77)                     | 520 433 (50.37)                      |
| Female                  | 501 833 (97.86)                            | 10 952 (2.14)                     | 512 785 (49.63)                      |
| Age groups              |                                            |                                   |                                      |
| Infants                 | 20 892 (99.82)                             | 38 (0.18)                         | 20 930 (2.03)                        |
| Children and school-age | 23 391 (99.95)                             | 12 (0.05)                         | 23 403 (2.27)                        |
| Adolescents             | 203 475 (99.54)                            | 231 (0.11)                        | 203 706 (19.72)                      |
| Young adults            | 430 375 (99.54)                            | 2000 (0.46)                       | 432 375 (41.85)                      |
| Adults                  | 193 142 (97.48)                            | 5000 (2.52)                       | 198 142 (19.18)                      |
| Seniors                 | 131 378 (84.95)                            | 23 284 (15.05)                    | 154 662 (14.97)                      |
| Ethnicity               |                                            |                                   |                                      |
| White, mestizo, other   | 942 712 (97.08)                            | 28 366 (2.92)                     | 971 078 (93.99)                      |
| African-Colombian       | 37 883 (96.38)                             | 1421 (3.62)                       | 39 304 (3.80)                        |
| Indigenous              | 22 011 (96.59)                             | 776 (3.41)                        | 22 787 (2.21)                        |
| Gipsy-Roman             | 34 (94.44)                                 | 2 (5.56)                          | 36 (0.00)                            |
| Raizal                  | 13 (100.00)                                | 0 (0.00)                          | 13 (0.00)                            |
| Area of residence       |                                            |                                   |                                      |
| Urban                   | 861 071 (96.87)                            | 27 844 (3.13)                     | 888 915 (86.03)                      |
| Semirural (village)     | 44 586 (97.18)                             | 1296 (2.82)                       | 45 882 (4.44)                        |
| Sparse rural            | 25 002 (96.19)                             | 991 (3.81)                        | 25 993 (2.52)                        |
| Unknown area            | 71 994 (99.40)                             | 434 (0.60)                        | 72 428 (7.01)                        |
| Type of health insurance regime |                                   |                                   |                                      |
| Contributory            | 646 670 (97.97)                            | 13 415 (2.03)                     | 660 085 (63.89)                      |
| Subsidized              | 146 290 (92.86)                            | 11 250 (7.14)                     | 157 540 (15.25)                      |
| Special                 | 16 306 (97.24)                             | 463 (2.76)                        | 16 769 (1.62)                        |
| Exception               | 47 490 (97.44)                             | 1247 (2.56)                       | 48 737 (4.72)                        |
| Uninsured               | 18 447 (97.49)                             | 474 (2.51)                        | 18 921 (1.83)                        |
| Unknown or pending insurance |                                 |                                   |                                      |
| 10 115 (97.41)          | 269 (2.59)                                  | 10 384 (1.01)                     |
| Non-registered insurance| 117 335 (91.15)                            | 3447 (2.85)                       | 120 782 (11.69)                      |
| Socioeconomic strata    |                                            |                                   |                                      |
| 1 Very low              | 183 857 (95.51)                            | 8 644 (4.49)                      | 192 501 (18.63)                      |
| 2 Low                   | 366 446 (96.80)                            | 12 113 (3.20)                     | 378 559 (36.64)                      |
| 3 Middle-low            | 200 213 (97.43)                            | 5 276 (2.57)                      | 205 489 (19.89)                      |
| 4 Middle                | 35 339 (97.47)                             | 918 (2.53)                        | 36 257 (3.51)                        |
| 5 Middle-high           | 11 857 (97.61)                             | 290 (2.39)                        | 12 147 (1.18)                        |
| 6 High                  | 6126 (96.73)                               | 207 (3.27)                        | 6333 (0.61)                          |
| Non-registered strata   | 198 815 (98.46)                            | 3117 (1.54)                       | 201 932 (19.54)                      |

Figure 1 shows the survival functions for each predictor in the model obtained from the multiple Cox Regression without time-dependent factors (See Table S4 in Supplementary material for model details). As the assumption of proportional hazards did not hold for all predictors but sex (see Figure S2-S12 in Supplementary material), we fit the multi-predictor Cox Regression for time dependent variables including the same predictors. This model was statistically significant (p < 0.001). Table 2 presents the results of our final multi-predictor time-dependent Cox regression model (See Tables S5-S6 in supplementary materials for details). The instantaneous risk of dying for COVID-19 among confirmed cases is 59% higher in males compared to females, 27% higher in indigenous people compared to whites/mestizos, and 97% higher in people with subsidized health insurance regime compared to contributory. There was evidence of a dose-response pattern by life-course age groups and SES levels. The risk of dying for COVID-19 among confirmed cases for people over 60 years is extremely higher than the risk for infants. The instantaneous risk of death for people with confirmed diagnosis of
COVID-19 living in the very low SES increases by 73% compared to the risk of people living in the high SES (HR=1.73 [95% CI: 1.48-2.04]). In contrast, living in a sparse rural area decreased the risk of mortality for COVID-19 (HR=0.83 [95% CI: 0.76-0.91]). Interactions terms between the time to the death and all the variables in the model were statistically significant (p<0.001).

Table 2. Risks of death for COVID-19 by some socioeconomic conditions in Colombia up to and including October 26th, 2020

| Socioeconomic condition               | HR (95% CI)          | P value |
|--------------------------------------|----------------------|---------|
| Sex (female as reference)            | 1.59 (1.53 – 1.65)   | <0.001  |
| Age groups (Infants as reference)    |                      |         |
| Children and school-age              | 0.42 (0.22-0.81)     | 0.009   |
| Adolescents                          | 1.16 (0.81-1.67)     | 0.414   |
| Young adults                         | 5.67 (4.00-8.04)     | <0.001  |
| Adults                               | 33.70 (23.59-48.14)  | <0.001  |
| Seniors                              | 214.31 (148.64-309.01) | <0.001 |
| Ethnicity (White/mestizo as reference) |                      |         |
| Indigenous                           | 1.27 (1.13-1.43)     | <0.001  |
| Gipsy-Roman                          | 1.56 (0.39-6.25)     | 0.530   |
| Raizal                               | 0.00 (0.00-3.44)     | 0.913   |
| African-Colombian                    | 1.01 (0.96-1.08)     | 0.613   |
| Area of residence (urban as reference) |                      |         |
| Semirural (village)                  | 0.88 (0.82-0.93)     | <0.001  |
| Sparse rural                         | 0.83 (0.76-0.91)     | <0.001  |
| Unknown area                         | 0.14 (0.12-0.16)     | <0.001  |
| Type of health insurance regime (contributory regime as reference) | | |
| Subsidized                           | 1.97 (1.89-2.04)     | <0.001  |
| Special                              | 1.29 (1.17-1.41)     | <0.001  |
| Exception                            | 1.37 (1.29-1.45)     | <0.001  |
| Uninsured                            | 1.34 (1.21-1.48)     | <0.001  |
| Unknown or pending insurance          | 1.22 (1.07-1.39)     | 0.002   |
| Non-registered insurance              | 2.57 (2.41-2.73)     | <0.001  |
| Socioeconomic strata (high SES as reference) | | |
| Very low                             | 1.73 (1.48-2.04)     | <0.001  |
| Low                                  | 1.61 (1.38-1.87)     | <0.001  |
| Middle-low                           | 1.34 (1.16-1.56)     | <0.001  |
| Middle                               | 1.16 (0.99-1.36)     | 0.059   |
| Middle-high                          | 0.94 (0.79-1.13)     | 0.531   |
| Unknown                              | 1.54 (1.30-1.83)     | <0.001  |

HR: Hazard Ratio; CI: Confidence Interval
Figure 1. Survival curves for COVID-19 by socioeconomic conditions. A) Sex; B) Age groups; C) Ethnicity; D) Area of residence; E) Health insurance regime; F) Socioeconomic status
DISCUSSION

Our results provide evidence of socioeconomic and demographic inequalities in COVID-19 mortality in Colombia. In addition to the well documented differential risk of mortality related to older age groups and male sex, this study provides evidence of socioeconomic and ethnicity inequalities in COVID-19 mortality. We identified higher mortality risks for indigenous people, people in the subsidized health regime, and those living in areas classified as very low and low SES. The risks of mortality for age groups and SES levels followed a consistent dose-response pattern.

Our findings of association between COVID-19 mortality and older age (60 year or more) and male sex are consistent with previous reports.[12] The most plausible explanation for this finding is the age-related response to sepsis in older adults with decline in the immune cell function, reduced humoral immune function, and uncontrolled production of inflammatory cytokines.[13] Our study also found an increased risk of death in men which is consistent with previous results. [14] Sex differences in COVID-19 mortality are probably explained by the increased expression in men of the angiotensin converting enzyme-2 (ACE-2), a key factor involved in the pathogenesis of COVID-19.[15]

Ethnicity disparities have been also reported in a variety of contexts. African American and Hispanic in the United States (US) are more vulnerable to COVID-19 mortality than other ethnic groups.[7] In Brazil, after age, Pardo ethnicity was the second most important risk factor for death and probable explanations are differential access to health care or susceptibility to COVID-19 infection.[16] Our study found an increased risk of COVID-19 mortality among indigenous people. Leticia, the capital of the Amazonas department with a live frontier with Brazil, holds the highest COVID-19 mortality rate across departments in Colombia. It is estimated that at least 163 indigenous communities have been infected for the SARS-CoV-2 in Latin America. Poor living and sanitary conditions combined with the burden of previous infectious diseases and malnutrition impose a higher risk to the health of individuals and entire communities.[17]

There is evidence of historical socioeconomic inequalities in previous pandemics. During the 1918 Spanish influenza pandemic there were reports that showed that mortality rates in some countries of South America was 20 times higher compared to countries in Europe.[18] The case fatality rates across countries early in the current COVID-19 pandemic showed negative correlation with countries’ Gross Domestic Product (GDP) and Human Development Index. [19,20] In the US, the COVID-19 pandemic is accelerating the health inequities by disproportionately affecting people from the most disadvantaged groups such as immigrants, people with disabilities and people in prisons and jails.[21] In Brazil, income and education inequalities were positively associated with COVID-19 incidence and mortality rates.[22,23] Our results showed that living in areas of very low or low SES is associated with higher COVID-19 mortality risk with a consistent dose-response effect pattern. These results are consistent with a previous report of inequalities in mortality by SES levels among COVID-19 confirmed cases in Bogotá.[24] These findings are also consistent with results of a nationwide ecologic study that showed increased risk of COVID-19 mortality associated with the municipalities’ multidimensional poverty index. [25] Colombia has one of the largest income gaps in Latin America and income inequalities within the country differ widely by geographic region in relation to land property, work market, and the effect of violence and armed conflict.[26] These baseline socioeconomic inequalities are translated into higher risk of exposure to and severity of COVID-19 affecting disproportionately to people in lower socioeconomic conditions in Colombia.

Social disruption stress producing pro-inflammatory gene expression has been described as a potential pathological mechanism to explain higher adverse health outcomes in populations with disadvantaged socioeconomic conditions.[27] However, the most possible explanation for the inequalities in COVID-19 mortality are the historical inequalities in terms of living and working conditions, and the unequal access to health care services. Inequalities in working conditions might explain an important part of the inequalities of COVID-19 infection and mortality. In our study, people in the subsidized health insurance regime represent people with unstable or informal work, or unemployed people who need subsidy from the government to get access to health services. Thus, the higher mortality risk observed in this group compared to the contributory health regime might be representing the social inequality related to working conditions in Colombia. People in the more disadvantaged working groups have lower-paid work and are more likely to work in key basic services (food, cleaning, delivery or public services) that require them to work in person and commute across the cities.[28] In contrast, people with higher-paid work are more likely to work from home with lower exposure to COVID-19 infection.[29]
Despite having an almost universal health insurance coverage, the Colombian health system is characterized by a strong fragmentation in the provision of health care services, an incipient primary health care, and differences in quality of health care services across regimes.[30] Therefore, differences between contributory and subsidized groups might be explained not only by underlying working conditions but also for chronic inequalities in access to high quality health care services. Limited health care services are provided in semi-rural and sparse rural areas. Our findings, however, found a potential protector effect for COVID-19 mortality for people living in those areas compared to people living in urban areas. The direction of this association might be explained by the SARS-CoV-2 transmission dynamics that favors contagion among close contacts in crowded places which is more common in urban areas.

The COVID-19 pandemic is occurring in presence of a non-communicable diseases (NCDs) epidemic and within a context of historical inequalities in the social determinants of health, which is recognized as a syndemic.[29] There are complex connections among NCDs, COVID-19 transmission dynamics and living conditions that shape disparities with higher adverse effects for disadvantaged people. People from minority ethnic groups, people living in areas with higher socioeconomic deprivation, generally have a greater number of or more severe or uncontrolled coexisting NCDs.[31] These inequalities in chronic conditions are deepened by the way people live and work which make them also more exposed to COVID-19 infection and mortality. Therefore, there is a need to measure, analyze and report demographic and socioeconomic inequities for identifying groups at higher risk for COVID-19 mortality in order to guide tailored public health interventions in countries.[8]

Our study provides strong evidence of socioeconomic inequalities in COVID-19 mortality in Colombia by using data from a nation-wide cohort of confirmed cases during the first eight months of the epidemic. However, conclusions should be carefully interpreted considering the limitations of the study. This study relies on data reported to SIVIGILA and it is possible that despite its national coverage, some degree of underreporting might be present. The probability of underreporting is higher in the sparse rural areas where most disadvantaged people live and therefore underreporting, if present, would have an attenuating effect of the effect measures. Colombia does not conduct COVID-19 mass testing and underreporting of cases and deaths might be influencing the confirmed cases captured by SIVIGILA and our results. Finally, our results are not controlled for the presence of chronic morbidities in confirmed cases so the effect of specific chronic diseases on COVID-19 mortality was not estimated and socioeconomic variables are not controlled for them.

In conclusion, our study provides evidence of demographic and socioeconomic inequalities in COVID-19 mortality in terms of age groups, sex, ethnicity, type of health insurance regime, and socioeconomic strata. Confirmed COVID-19 cases who are male, over 60 years old, indigenous, holding a government subsidized health insurance, and those living in areas classified in the lower socioeconomic strata have a higher risk of dying faster from COVID-19. Public health interventions for prevention and early detection of COVID-19 cases should be prioritized for more vulnerable groups according to the unequal mortality risks.
References

1. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med 2020;382:727–33. doi:10.1056/nejmoa2001017
2. World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard. 2020. https://covid19.who.int (accessed 17 Nov 2020).
3. Horton R. Offline: COVID-19 is not a pandemic. Lancet 2020;396:874. doi:10.1016/S0140-6736(20)32000-6
4. Economic Commission for Latin America and the Caribbean (ECLAC). Building a New Future: Transformative Recovery with Equality and Sustainability. Santiago; 2020.
5. The Lancet. COVID-19 in Latin America: a humanitarian crisis. Lancet 2020;396:1463. doi:10.1016/S0140-6736(20)32328-X
6. Mein SA. COVID-19 and Health Disparities: the Reality of “the Great Equalizer”. J Gen Intern Med 2020;35:2439–40. doi:10.1007/s11606-020-05880-5
7. Abedi V, Olulana O, Avula V, et al. Racial, Economic, and Health Inequality and COVID-19 Infection in the United States. J Racial Ethn Heal Disparities Published Online First: 2020. doi:10.1007/s40615-020-00833-4
8. Khalatbari Soltani S, Cumming RC, Delpierre C, et al. Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. J Epidemiol Community Health 2020;74:620–3. doi:10.1136/jech-2021-214297
9. Departamento Nacional de Estadística. Proyecciones de población. 2020.
10. Departamento Nacional de Estadística. Estratificación socioeconómica. 2019. https://www.dane.gov.co/index.php/sistema-estadistico-nacional-sen/69-espanol/geoestadistica/estratificacion/468-estratificacion-socioeconomica (accessed 26 Oct 2020).
11. Vittinghoff E, Gillden D V, Shiboski SC, et al. Regression Methods in Biostatistics. Linear, Logistic, Survival, and Repeated Measures Models. Second. New York: : Springer Science Business Media 2012. doi:10.1007/978-1-4614-1353-0
12. Mehrauen E, Karimi A, Barzegary A, et al. Predictors of mortality in patients with COVID-19- a systematic review. Eur J Integr Med 2020;40:101226. doi:https://doi.org/10.1016/j.eujim.2020.101226
13. Opal SM, Girard TD, Ely EW. The immunopathogenesis of sepsis in elderly patients. Clin Infect Dis 2005;41. doi:10.1086/432007
14. Nasiri MJ, Haddadi S, Tahvildari A, et al. COVID-19 Clinical Characteristics, and Sex-Specific Risk of Mortality: Systematic Review and Meta-Analysis. Front Med 2020;7:1–10. doi:10.3389/fmed.2020.00459
15. Maleki Dana P, Sadoughi F, Hallajzadeh J, et al. An Insight into the Sex Differences in COVID-19 Patients: What are the Possible Causes? Prehosp Disaster Med 2020;35:438–41. doi:10.1017/S1049023X20000837
16. Baqui P, Bica I, Marra V, et al. Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study. Lancet Glob Heal 2020;8:e1018–26. doi:10.1016/S2214-109X(20)30285-0
17. Fondo para el Desarrollo de los Pueblos Indígenas de América Latina y el Caribe. Los Pueblos Indígenas ante la pandemia del COVID-19. La Paz, Bolivia. 2020. https://indigenascontraelcov19.red/wp-content/uploads/2020/05/FILAC_PI_COVID19.pdf
18. Murray CJL, Lopez AD, Chin B, et al. Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918&amp;#x2013;20 pandemic: a quantitative analysis. Lancet 2006;368:2211–8. doi:10.1016/S0140-6736(06)69895-4
19. Shahbazi F, Khazaie S. Socio-economic inequality in global incidence and mortality rates from coronavirus disease 2019: an ecological study. New Microbes New Infect 2020;38:100762. doi:10.1016/j.mni.2020.100762
20. Asfahani S, Shahal A, Chawla G, et al. Early trends of socio-economic and health indicators influencing case fatality rate of COVID-19 pandemic. Monaldi Arch Chest Dis 2020;9:451–7. doi:10.4081/MONALDI.2020.1388
21. Okonkwo NE, Agwu UT, Jang M, et al. COVID-19 and the US response: accelerating health inequities. BMJ Evidence-Based Medicine 2020;0:bmjebm-2020-111426. doi:10.1136/bmjebm-2020-111426
22. Demenech LM, Dumith S de C, Vieira MECD, et al. Desigualdade econômica e risco de infecção e morte por COVID-19 no Brasil. Rev Bras Epidemiol 2020;23:e200095. doi:10.1590/1980-549720200095
23. Wolfenstein-Betech S, Silva AAB, Fleck JL, et al. Physiological and socioeconomic characteristics predict COVID-19 mortality and resource utilization in Brazil. PLoS One 2020;15:1–15. doi:10.1371/journal.pone.0240346
24. Grupo de Investigación en Macroeconomía de la Facultad de Economía U de los A. El patrón socioeconómico del COVID. El caso de Bogotá. 2020. https://economia.uniandes.edu.co/components/com_booklibrary/ebooks/61412.pdf (accessed 27 Oct 2020).
Rodriguez-Villamizar LA, Belalcazar-Ceron LC, Fernandez-Nino JA, et al. Air pollution, sociodemographic and health conditions effects on COVID-19 mortality in Colombia: an ecological study. Sci Total Environ Published Online First: November 2020. doi:https://doi.org/10.1016/j.scitotenv.2020.144020

Sanchez R. Desigualdad del ingreso en Colombia: un estudio por departamentos. Cuad Econ 2017;36:139–78. doi:10.15446/cuad.econ.v36n72.65880

Health I, Mattos R. Isolation, social stress, low socioeconomic status and its relationship to immune response in Covid-19 pandemic context. Brain, Behav, Immun - Heal 2020;7:100103. doi:https://doi.org/10.1016/j.bbih.2020.100103

Ribeiro F, Ribeiro F, Leist A. Who is going to pay the price of Covid-19? Reflections about an unequal Brazil. Int J Equity Health 2020;19:19–21. doi:10.1186/s12939-020-01207-2

Bambra C, Riordan R, Ford J, et al. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health 2020;74:964–8. doi:10.1136/jech-2020-214401

Giovanella L, Vega R, Tejerina-Silva H, et al. ¿Es la atención primaria de salud integral parte de la respuesta a la pandemia de Covid-19 en Latinoamérica? Trab Educ e Saúde 2021;19. doi:10.1590/1981-7746-sol00310

Di Cesare M, Khang YH, Asaria P, et al. Inequalities in non-communicable diseases and effective responses. Lancet 2013;381:585–97. doi:10.1016/S0140-6736(12)61851-0
Contributor’s statement

MPC: methodology design, verification of the underlying data, data analysis, data interpretation, writing - original draft
LAR-V: literature research, data analysis, data interpretation, writing - original draft
MLR-B: verification of the underlying data, data analysis, data interpretation, writing – review and editing
CA-M: data interpretation, writing – review and editing
JAF-N: conceptualization, methodology design, data analysis, data interpretation, writing – review and editing

Funding

This study did have a specific funding grant. Study design, data collection, data analysis, data interpretation, and writing the report were conducted as part of the work of the Direction of Epidemiology and Demography of the Ministry of Health and Social Protection of Colombia.

Competing interest

None declared.

Data Availability Statement

Data used for the current study are publicly available as open data on the government website (https://www.datos.gov.co/en/Salud-y-Protecci-n-Social/Casos-positivos-de-COVID-19-en-Colombia/gt2j-8ykr)
Supplemental Figures and Tables

Figure S1. Approaches to compute days of time to event in the cohort of COVID-19 confirmed cases, Colombia

Table S1. Descriptive statistics and distributions of the six times to death, to recovery and censored cases, each for symptomatic and asymptomatic cases.

|                  | Time to event | Time to death symptomatic | Time to death asymptomatic | Time to recovery symptomatic | Time to recovery asymptomatic | Symptomatic censored cases | Asymptomatic censored cases |
|------------------|---------------|--------------------------|---------------------------|-----------------------------|-------------------------------|---------------------------|-----------------------------|
|                  | N             | n1=30.558                | n2=7                      | n3=812.437                  | n4=102.445                    | n5=59.105                  | n6=10.666                   |
| Mean             | 25.56         | 15.97                    | 16.43                     | 25.92                       | 25.64                        | 28.00                     | 11.56                       |
| Std. Error of Mean | .01          | .07                      | 4.61                       | .01                         | .02                          | .12                       | .12                         |
| Median           | 24.00         | 14.00                    | 11.00                      | 25.00                       | 25.00                        | 14.00                     | 10.00                       |
| Mode             | 21            | 10                       | 8                          | 21                          | 25                           | 14                        | 13                          |
| Std. Deviation   | 11.81         | 11.88                    | 12.20                      | 9.37                        | 8.53                         | 29.89                     | 12.81                       |
| Variance         | 139.62        | 141.04                   | 148.95                     | 87.74                       | 72.69                        | 893.26                    | 164.19                      |
| Range            | 216           | 158                      | 34                         | 211                         | 196                          | 215                       | 213                         |
| Minimum          | 0             | 0                        | 8                          | 0                           | 2                            | 1                         | 2                           |
| Maximum          | 216           | 158                      | 42                         | 211                         | 198                          | 216                       | 215                         |
| Sum              | 26 411.767    | 488 051                  | 115                        | 21 056 624                  | 3 088 341                    | 1 655 125                 | 123 311                     |
| Percentiles      | 25            | 20                       | 8                          | 21                          | 20                           | 10                        | 7                           |
|                  | 50            | 24                       | 14                         | 11                          | 25                           | 14                        | 10                          |
|                  | 75            | 29                       | 22                         | 20                          | 29                           | 35                        | 12                          |
| IQR              | 9             | 14                       | 12                         | 8                           | 9                            | 15                        | 5                           |
Table S2. Frequencies of predictors’ categories by type of time-to-event outcome

| Predictors / categories | Time to event outcomes |
|-------------------------|------------------------|
|                         | 1   | 2     | 3   | 4   | 5   | 6   |
| **Sex**                 |     |       |     |     |     |     |
| Male                    | 19 608 | 5   | 405 391 | 59 694 | 30 547 | 5 188 |
| Female                  | 10 950 | 2   | 407 046 | 60 751 | 28 558 | 5 478 |
| **Socioeconomic status SES** |     |       |     |     |     |     |
| Unknown SES             | 3 115 | 2   | 148 867 | 33 125 | 11 229 | 5 594 |
| Very low SES 1          | 8 644 | 0   | 158 065 | 13 029 | 12 236 | 527 |
| Low SES 2               | 12 110 | 3  | 303 406 | 40 542 | 20 541 | 1 957 |
| Middle-low SES 3        | 5 275 | 1   | 160 508 | 26 072 | 11 741 | 1 892 |
| Middle SES 4            | 9 171 | 1   | 27 691  | 1 672  | 732  | 164  |
| Middle-high SES 5       | 290  | 0   | 9 289   | 1 672  | 732  | 164  |
| High SES 6              | 207  | 0   | 4 611   | 1 001  | 426  | 88   |
| **Life course age group** |     |       |     |     |     |     |
| Infants 0-5y            | 38  | 0   | 16 068  | 3 096  | 1 504 | 224  |
| Children and school-age (6-11y) | 12  | 0   | 18 147  | 3 989  | 985  | 270  |
| Adolescents (12-26y)    | 230  | 1   | 164 869 | 26 072 | 11 741 | 1 892 |
| Young adults (27-45y)   | 2 000 | 0  | 355 500 | 49 830 | 20 775 | 4 270 |
| Adults (46-59y)         | 5 000 | 0   | 156 162 | 2 2847 | 12 094 | 2 039 |
| Seniors (60y and more)  | 23 278 | 6  | 101 691 | 14 115 | 13 926 | 1 646 |
| **Ethnicity**           |     |       |     |     |     |     |
| Indigenous              | 776  | 0   | 19 502  | 1 851  | 621  | 37   |
| Romany                  | 2   | 0   | 30     | 3      | 1    | 0    |
| Raizal                  | 0   | 0   | 5      | 8      | 0    | 0    |
| Palenquero              | 0   | 0   | 0      | 0      | 0    | 0    |
| Black, African-colombian | 1 421 | 0  | 34 901  | 2 339  | 637  | 6    |
| Other: white, mestizo   | 28 359 | 7  | 757999  | 116 244 | 57 846 | 10 623 |
| **Area**                |     |       |     |     |     |     |
| City or town (Urban)    | 27 837 | 7  | 708 770 | 98 609 | 48 541 | 5 151 |
| Village (Semirural)     | 1 296 | 0   | 39 335  | 1 946  | 3226 | 79   |
| Sparse rural            | 991  | 0   | 21 322  | 673    | 2972 | 35   |
| Unknown area            | 434  | 0   | 43 010  | 19 217 | 4366 | 5 401 |
| **Health Insurance**    |     |       |     |     |     |     |
| Contributory-job related insurance | 13 411 | 4  | 527 518 | 80 960 | 33 700 | 4 492 |
| Special Insurance       | 463  | 0   | 14 274  | 1 739  | 275  | 18   |
| Exception health regime | 12 455 | 2  | 40 636  | 3 533  | 3089 | 232  |
| Gov. subsidized insurance | 11 249 | 1  | 124 099 | 8 988  | 12837 | 366  |
| Uninsured               | 474  | 0   | 13 485  | 3 671  | 1214 | 77   |
| Unknown or pending insurance | 269  | 0   | 8 821   | 778    | 505  | 11   |
| Non-registered insurance | 3 447 | 0  | 83 604  | 20 776 | 7485 | 5 470 |
| **TOTAL**               | 30 558 | 7  | 812 437 | 120 445 | 59105 | 10 666 |

¹Time to event outcome 1 refers to time to death of symptomatic cases, and outcome 2 to time to death of asymptomatic ones. Outcome 3 and 4 refer to the frequencies of time to recovery on patients with or without symptoms respectively, and outcome 5 and 6 refer to frequencies of censored cases, also with or without symptoms, respectively.
Table S3. Proportional distribution of predictors’ categories by type of time-to-event outcome

| Predictors             | Categories                     | Time to event outcomes | 1 Column % | 2 Column % | 3 Column % | 4 Column % | 5 Column % | 6 Column % |
|------------------------|--------------------------------|------------------------|------------|------------|------------|------------|------------|------------|
| Sex                    | Male                           |                        | 64.2       | 71.4       | 49.9       | 49.6       | 51.7       | 48.6       |
|                        | Female                         |                        | 35.8       | 28.6       | 50.1       | 50.4       | 48.3       | 51.4       |
| Socioeconomic status   | Unknown SES                    |                        | 10.2       | 28.6       | 18.3       | 27.5       | 19.0       | 52.4       |
|                        | Very low SES 1                 |                        | 28.3       | 0.0        | 19.5       | 10.8       | 20.7       | 4.9        |
|                        | Low SES 2                      |                        | 39.6       | 42.9       | 37.3       | 33.7       | 34.8       | 18.3       |
|                        | Middle-low SES 3               |                        | 17.3       | 14.3       | 19.8       | 21.6       | 19.9       | 17.7       |
|                        | Middle SES 4                   |                        | 3.0        | 14.3       | 3.4        | 4.2        | 3.7        | 4.2        |
|                        | Middle-high SES 5              |                        | 0.9        | 0.0        | 1.1        | 1.4        | 1.2        | 1.5        |
|                        | High SES 6                     |                        | 0.7        | 0.0        | 0.6        | 0.8        | 0.7        | 0.8        |
| Life course age group  | Infants 0-5y                    |                        | 0.1        | 0.0        | 2.0        | 2.6        | 2.5        | 2.1        |
|                        | Children and school-age (6-11y)|                        | 0.0        | 0.0        | 2.2        | 3.3        | 1.7        | 2.5        |
|                        | Adolescents (12-26y)           |                        | 0.8        | 14.3       | 20.3       | 22.1       | 16.6       | 20.8       |
|                        | Young adults (27-45y)          |                        | 6.5        | 0.0        | 43.8       | 41.4       | 35.1       | 40.0       |
|                        | Adults (46-59y)                |                        | 16.4       | 0.0        | 19.2       | 19.0       | 20.5       | 19.1       |
|                        | Seniors (60y and more)         |                        | 76.2       | 85.7       | 12.5       | 11.7       | 23.6       | 15.4       |
| Ethnicity              | Indigenous                     |                        | 2.5        | 0.0        | 2.4        | 1.5        | 1.1        | 0.3        |
|                        | Romany                         |                        | 0.0        | 0.0        | 0.0        | 0.0        | 0.0        | 0.0        |
|                        | Raizal                         |                        | 0.0        | 0.0        | 0.0        | 0.0        | 0.0        | 0.0        |
|                        | Palenquero                     |                        | 0.0        | 0.0        | 0.0        | 0.0        | 0.0        | 0.0        |
|                        | Black, African-colombian       |                        | 4.7        | 0.0        | 4.3        | 1.9        | 1.1        | 0.1        |
|                        | Other: white, mestizo          |                        |            |            |            |            |            |            |
| Area                   | City or town (Urban)           |                        |            |            |            |            |            |            |
|                        | Village (Semirural)            |                        |            |            |            |            |            |            |
|                        | Sparse rural                   |                        |            |            |            |            |            |            |
|                        | Unknown area                   |                        | 1.4        | 0.0        | 5.3        | 16.0       | 7.4        | 50.6       |
| Health Insurance       | Contributory-job related insurance |                    | 43.9       | 57.1       | 64.9       | 67.2       | 57.0       | 42.1       |
|                        | Special Insurance              |                        | 1.5        | 0.0        | 1.8        | 1.4        | 0.5        | 0.2        |
|                        | Exception health regime        |                        | 4.1        | 28.6       | 5.0        | 2.9        | 5.2        | 2.2        |
|                        | Gov. subsidized insurance      |                        | 36.8       | 14.3       | 15.3       | 7.5        | 21.7       | 3.4        |
|                        | Uninsured                      |                        | 1.6        | 0.0        | 1.7        | 3.0        | 2.1        | 0.7        |
|                        | Unknown or pending insurance   |                        | 0.9        | 0.0        | 1.1        | 0.6        | 0.9        | 0.1        |
|                        | Non-registered insurance       |                        | 11.3       | 0.0        | 10.3       | 17.2       | 12.7       | 51.3       |

1Time to event outcome 1 refers to time to death of symptomatic cases, and outcome 2 to time to death of asymptomatic ones. Outcome 3 and 4 refer to the frequencies of time to recovery on patients with or without symptoms respectively, and outcome 5 and 6 refer to frequencies of censored cases, also with or without symptoms, respectively.
Table S4. Multi-predictor Cox Regression model without terms to address time-dependent factors

| Variables in the Equation | B    | SE   | Wald  | df | Sig.  | Exp(B) | 95.0% CI for Exp(B) |
|---------------------------|------|------|-------|----|-------|--------|---------------------|
| Sex (Ref. Female)         | .53  | .01  | 1968.10 | 1  | .000  | 1.70   | 1.66-1.74          |
| Socioeconomic status SES (Ref. High SES 6) |     |      |        |    |       |        |                    |
| Unknown SES               | .01  | .07  | .02    | 1  | .899  | 1.01   | .88-1.16           |
| Very low SES 1           | .19  | .07  | 7.43   | 1  | .006  | 1.21   | 1.06-1.39          |
| Low SES 2                 | .18  | .07  | 6.92   | 1  | .009  | 1.20   | 1.05-1.38          |
| Middle-low SES 3          | .07  | .07  | 1.05   | 1  | .305  | 1.08   | .94-1.24           |
| Middle SES 4              | -.01 | .08  | .00    | 1  | .946  | .99    | .86-1.16           |
| Middle-high SES 5         | -.14 | .09  | 2.38   | 1  | .123  | .87    | .73-1.04           |
| Life course age (Infants 0-5y) |     |      |        |    |       |        |                    |
| Children and school-age (6-11y) | -1.19| .33  | 13.02  | 1  | .000  | .30    | .16-0.58           |
| Adolescents (12-26y)      | -.35 | .18  | 4.06   | 1  | .044  | .70    | .50-0.99           |
| Young adults (27-45y)     | 1.07 | .16  | 42.99  | 1  | .000  | 2.93   | 2.12-4.03          |
| Adults (46-59y)           | 2.71 | .16  | 276.42 | 1  | .000  | 14.99  | 10.89-20.63        |
| Seniors (60y and more)    | 4.40 | .16  | 735.60 | 1  | .000  | 81.74  | 59.46-112.37       |
| Ethnicity (Ref. Other, white, mestizo) |     |      |        |    |       |        |                    |
| Indigenous                | .05  | .04  | 1.51   | 1  | .219  | 1.05   | .97-1.12           |
| Romany                    | .33  | .71  | .21    | 1  | .645  | 1.39   | .35-5.54           |
| Raizal                    | -6.22| 57.1 | 1      | .01 | .913  | .00    | 8.09E-18          |
| African-Colombian         | -.03 | .03  | 1.01   | 1  | .315  | .97    | .92-1.03           |
| Area (ref. City or town (Urban)) |     |      |        |    |       |        |                    |
| Village (Semiaural)       | .07  | .09  | 7.52   | 1  | .006  | .92    | .87-0.98           |
| Sparse rural              | .05  | .42  | 4.82   | 1  | .028  | .93    | .87-0.99           |
| Unknown area              | -1.82| .05  | 1104.14| 1  | .000  | .16    | .15-0.18           |
| Health Insurance (ref. job related insurance) |     |      |        |    |       |        |                    |
| Special Insurance         | .23  | .05  | 23.06  | 1  | .000  | 1.26   | 1.14-1.38          |
| Exception health regime   | .24  | .03  | 65.57  | 1  | .000  | 1.27   | 1.20-1.35          |
| Gov. subsidized insurance | .57  | .01  | 1622.47 | 1  | .000  | 1.78   | 1.73-1.83          |
| Uninsured                 | .16  | .05  | 11.01  | 1  | .001  | 1.17   | 1.07-1.28          |
| Unknown or pending insurance | .04  | .06  | .34    | 1  | .562  | 1.04   | .92-1.17           |
| Non-registered insurance  | .74  | .02  | 1327.97| 1  | .000  | 2.10   | 2.02-2.19          |
Testing of Assumptions

Additional to overall model statistics, and to systematically explore quantitative criteria of time dependent variables in the model (see interactions between predictors and time in the final Cox Regression with time dependent variables), we used graphical criteria to assess if proportional hazards hold. The following figures present plots between log minus log transformation of survival estimates and time to event, and scatter plots between partial residuals and time to event.

**Sex**

![Log minus log function for Sex](image1)

![Simple Scatter with Fit Line of Partial residual for Sex by Time to event](image2)

Figure S2. Graphic test of proportional hazard assumption for Sex.

Left, Ln(-ln) transformation of survival estimate vs Time to event plot.

Right, Partial residual vs time to event plot.

Despite the Log minus log plot show almost parallel lines for male and female categories, the partial residual plot shows a positive trend. The interaction between Sex and time to event has a significative p-value close to zero.

**Socioeconomic status**

![LML Function for Socioeconomic Status](image3)

Figure S3. Ln(-ln) transformation of survival estimates for SES categories vs Time to event plot
Curves of Ln(-ln) transformation of survival estimates for SES categories do not overlap or cross. However, partial residual plots for all categories but Middle-high SES, have clear trends. Additionally, as interaction between SES and time to event had a p-value close to zero, we conclude SES violates the proportional hazards assumption.
Life Course age groups

Figure S5. Ln(-ln) transformation of survival estimates for Life course age groups vs Time to event plot
Similar to the former variables, the log minus log curves for categories in Life course just converge in the origin, with no further crossings. However, partial residual plots have clear trends for categories of young adults, adults and seniors, and for the interaction between Life course age variable and time to event has an overall p-value is close to zero, which confirms proportional hazards assumption do not hold.

Ethnicity

Figure S7. Ln(-ln) transformation of survival estimates for Ethnicity vs Time to event plot
Figure S8. Partial residual vs time to event plots for Ethnicity.

The shape of log minus log curves for categories in the ethnicity variable was parallel with no evident crossings. However, partial residual plots for Indigenous, Raizal and Black or African-Colombian groups, had a negative trend, and the p-value of the interaction between ethnicity and time to event close to zero, showed significant deviations from the expected assumption of proportional hazards.
Area of residence

Figure S9. Ln(-ln) transformation of survival estimates for Area vs Time to event plot

Almost parallel log minus log curves and almost flat partial residual plots contrast with the close to zero p-value of the interaction term between Area and time to event. By prioritizing the quantitative criterion, we opted to include this interaction in the final model.
Health insurance

Figure S11. Ln(-ln) transformation of survival estimates for Health insurance vs Time to event plot
Almost parallel log minus log curves for health insurance categories, which just show convergence at the origin, contrast with marked trends on partial residual plots for Subsidized insurance and the group of people with no registries of the type of health insurance. Exception regime had a slight positive trend and special regime had a slight negative trend. The overall p-value of the interaction between health insurance and time to event was also close to zero.

Figure S12. Partial residual vs time to event plots for Health insurance.
Multi-predictor Cox Regression model with terms to address time-dependent factors

Numbers at risk

To complement the numbers at risk on figures of Survival plots for each predictor in the model, we present the number of the two types of censored cases for 0, 20, 40, 60, 80, 100, 120, 140, 160, 180 and 200 days of follow-up.

Table S5. Censored symptomatic cases by follow-up time

|                  | 0 days  | 20 days | 40 days | 60 days | 80 days | 100 days | 120 days | 140 days | 160 days | 180 days | 200 days | Total    |
|------------------|---------|---------|---------|---------|---------|----------|----------|----------|----------|----------|----------|----------|
| **Sex**          |         |         |         |         |         |          |          |          |          |          |          |          |
| Male             | 20,277  | 2,355   | 1,975   | 2,522   | 2,513   | 619      | 145      | 46       | 17       | 5        | 30,547   |          |
| Female           | 20,733  | 1,845   | 1,534   | 2,026   | 1,805   | 404      | 104      | 66       | 27       | 12       | 28,558   |          |
| **Socioeconomic status** |       |         |         |         |         |          |          |          |          |          |          |          |
| SES              |         |         |         |         |         |          |          |          |          |          |          |          |
| Unknown SES      | 9,385   | 488     | 321     | 464     | 409     | 108      | 26       | 13       | 10       | 4        | 1        | 11,229   |
| Very low SES 1   | 7,174   | 1,088   | 958     | 1,362   | 1,168   | 299      | 86       | 55       | 33       | 12       | 1        | 12,236   |
| Low SES 2        | 13,344  | 1,688   | 1,458   | 1,720   | 1,761   | 390      | 94       | 53       | 18       | 10       | 5        | 20,541   |
| Middle-low SES 3 | 8,519   | 764     | 632     | 826     | 785     | 166      | 29       | 10       | 7        | 3        | 1        | 11,741   |
| Middle SES 4     | 1,715   | 108     | 95      | 106     | 124     | 36       | 8        | 5        | 3        | 0        | 0        | 2,200    |
| Middle-high SES 5| 588     | 36      | 23      | 30      | 36      | 13       | 4        | 2        | 0        | 0        | 0        | 732      |
| High SES 6       | 285     | 28      | 22      | 40      | 35      | 11       | 2        | 1        | 2        | 0        | 0        | 426      |
| **Life course age group** |       |         |         |         |         |          |          |          |          |          |          |          |
| Infants 0-5y     | 738     | 191     | 203     | 200     | 128     | 31       | 2        | 5        | 4        | 2        | 0        | 1,504    |
| Children and school-age (6-11y) | 795  | 44      | 46      | 58      | 37      | 3        | 2        | 0        | 0        | 0        | 0        | 985      |
| Adolescents (12-26y) | 8,434 | 377     | 287     | 373     | 272     | 54       | 9        | 6        | 5        | 3        | 1        | 9,821    |
| Young adults (27-45y) | 16,975| 880     | 732     | 1,067   | 910     | 160      | 19       | 17       | 7        | 6        | 2        | 20,775   |
| Adults (46-59y)  | 7,898   | 901     | 840     | 1,016   | 1,114   | 252      | 36       | 20       | 13       | 3        | 1        | 12,094   |
| Seniors (60y and more) | 6,170 | 1,807   | 1,401   | 1,834   | 1,857   | 523      | 181      | 91       | 44       | 15       | 3        | 13,926   |
| **Ethnicity**    |         |         |         |         |         |          |          |          |          |          |          |          |
| Indigenous       | 164     | 63      | 85      | 137     | 123     | 21       | 12       | 7        | 5        | 4        | 0        | 621      |
| Romany           | 0       | 0       | 0       | 1       | 0       | 0        | 0        | 0        | 0        | 0        | 0        | 1        |
| Raizal           | 0       | 0       | 0       | 0       | 0       | 0        | 0        | 0        | 0        | 0        | 0        | 0        |
| Palenquero       | 0       | 0       | 0       | 0       | 0       | 0        | 0        | 0        | 0        | 0        | 0        | 0        |
| Black, African-colombian | 216 | 48     | 34      | 45      | 170     | 72       | 22       | 21       | 6        | 3        | 0        | 637      |
| Other: white, mestizo | 40,630| 4,089  | 3,390  | 4,365  | 4,025  | 930      | 215      | 111      | 62       | 22       | 7        | 57,846   |
| **Area**         |         |         |         |         |         |          |          |          |          |          |          |          |
| City or town (Urban) | 32,904| 3,421  | 3,007  | 3,987  | 3,847  | 928      | 222      | 130      | 64       | 26       | 5        | 48,541   |
| Village (Semirural) | 2,198 | 270   | 224   | 271   | 204   | 37       | 13       | 5        | 2        | 0        | 2        | 3,226    |
| Sparse rural     | 1,903   | 339     | 231     | 234     | 201     | 45       | 8        | 3        | 6        | 2        | 0        | 2,972    |
| Unknown area     | 4,005   | 170     | 47      | 56      | 66      | 13       | 6        | 1        | 1        | 1        | 0        | 4,366    |
| **Health Insurance** |       |         |         |         |         |          |          |          |          |          |          |          |
| Contributory-job related insurance | 25,793 | 1,789  | 1,547  | 1,928  | 1,941  | 506      | 109      | 51       | 22       | 12       | 2        | 33,700   |
| Special Insurance | 132   | 13     | 10     | 28     | 59     | 26       | 4        | 2        | 0        | 1        | 0        | 275      |
| Exception health regime Gov. subsidized insurance | 2,318 | 235   | 173   | 194   | 139   | 24       | 2        | 2        | 1        | 0        | 1        | 3,089    |
| Uninsured        | 6,475   | 1,501   | 1,349   | 1,517   | 1,383   | 356      | 119      | 72       | 47       | 14       | 4        | 12,837   |
| Unknown or pending insurance | 782 | 108   | 92   | 92   | 99   | 30       | 2        | 7        | 1        | 1        | 0        | 1,214    |
| Non-registered insurance | 344 | 36     | 26     | 36     | 46     | 13       | 2        | 1        | 1        | 0        | 0        | 505      |
| **TOTAL**        | 41,010  | 4,200   | 3,509   | 4,548   | 4,318   | 4,023    | 249      | 139      | 73       | 29       | 7        | 59,105   |
Table S6. Censored asymptomatic cases by follow-up time

| Sex                                | 0 days | 20 days | 40 days | 60 days | 80 days | 100 days | 120 days | 140 days | 160 days | 180 days | 200 days | Total |
|------------------------------------|--------|---------|---------|---------|---------|----------|----------|----------|----------|----------|----------|-------|
| Male                               | 4 887  | 104     | 67      | 73      | 27      | 19       | 6        | 3        | 1        | 1        | 5 188    |       |
| Female                             | 5 279  | 73      | 47      | 46      | 11      | 9        | 8        | 3        | 1        | 1        | 5 478    |       |
| Socioeconomic status               |        |         |         |         |         |          |          |          |          |          |          |       |
| SES                                |        |         |         |         |         |          |          |          |          |          |          |       |
| Unknown SES                        | 5 531  | 26      | 13      | 14      | 4       | 1        | 3        | 1        | 0        | 0        | 1        | 5 594 |
| Very low SES 1                     | 469    | 21      | 11      | 13      | 3       | 2        | 4        | 0        | 3        | 1        | 0        | 527   |
| Low SES 2                          | 1 754  | 63      | 50      | 49      | 21      | 13       | 4        | 2        | 1        | 0        | 0        | 1 957 |
| Middle-low SES 3                   | 1 764  | 41      | 32      | 35      | 7       | 10       | 2        | 0        | 0        | 1        | 0        | 1 892 |
| Middle SES 4                       | 418    | 12      | 4       | 6       | 1       | 2        | 1        | 0        | 0        | 0        | 0        | 444   |
| Middle-high SES 5                  | 148    | 10      | 4       | 1       | 1       | 0        | 0        | 0        | 0        | 0        | 0        | 164   |
| High SES 6                         | 82     | 4       | 0       | 1       | 1       | 0        | 0        | 0        | 0        | 0        | 0        | 88    |
| Life course age group              |        |         |         |         |         |          |          |          |          |          |          |       |
| Infants 0-5y                       | 211    | 8       | 3       | 1       | 0       | 0        | 0        | 0        | 1        | 0        | 0        | 224   |
| Children and school-age (6-11y)    | 270    | 0       | 0       | 0       | 0       | 0        | 0        | 0        | 0        | 0        | 0        | 270   |
| Adolescents (12-26y)               | 2 186  | 15      | 4       | 7       | 4       | 0        | 0        | 0        | 1        | 0        | 0        | 2 217 |
| Young adults (27-45y)              | 4 180  | 31      | 23      | 26      | 4       | 1        | 1        | 0        | 2        | 2        | 0        | 4 270 |
| Adults (46-59y)                    | 1 916  | 43      | 32      | 30      | 8       | 8        | 2        | 0        | 0        | 0        | 0        | 2 039 |
| Seniors (60y and more)             | 1 403  | 80      | 52      | 55      | 22      | 19       | 11       | 3        | 0        | 0        | 1        | 1 646 |
| Ethnicity                          |        |         |         |         |         |          |          |          |          |          |          |       |
| Indigenous                         | 35     | 0       | 0       | 0       | 0       | 0        | 0        | 1        | 1        | 0        | 0        | 37    |
| Romany                             | 0      | 0       | 0       | 0       | 0       | 0        | 0        | 0        | 0        | 0        | 0        | 0     |
| Raizal                             | 0      | 0       | 0       | 0       | 0       | 0        | 0        | 0        | 0        | 0        | 0        | 0     |
| Palenquero                         | 0      | 0       | 0       | 0       | 0       | 0        | 0        | 0        | 0        | 0        | 0        | 0     |
| Black, African-columbian           | 3      | 1       | 0       | 1       | 0       | 1        | 0        | 0        | 0        | 0        | 0        | 6     |
| Other: white, mestizo              | 10 128 | 176     | 114     | 118     | 38      | 27       | 13       | 2        | 4        | 2        | 1        | 10 623|
| Area                               |        |         |         |         |         |          |          |          |          |          |          |       |
| City or town (Urban)               | 4 709  | 153     | 100     | 107     | 33      | 27       | 14       | 2        | 3        | 2        | 1        | 5 151 |
| Village (Semirural)                | 72     | 0       | 3       | 1       | 2       | 0        | 0        | 1        | 0        | 0        | 0        | 79    |
| Sparser rural                      | 32     | 0       | 1       | 0       | 0       | 1        | 0        | 0        | 1        | 0        | 0        | 35    |
| Unknown area                       | 5 353  | 24      | 10      | 11      | 3       | 0        | 0        | 0        | 0        | 0        | 0        | 5 401 |
| Health Insurance                   |        |         |         |         |         |          |          |          |          |          |          |       |
| Contributory-job related insurance | 4 141  | 121     | 87      | 94      | 18      | 23       | 6        | 1        | 0        | 0        | 1        | 4 492 |
| Special Insurance                  | 14     | 1       | 0       | 0       | 0       | 0        | 0        | 0        | 2        | 1        | 0        | 18    |
| Exception health regime            | 213    | 7       | 3       | 3       | 5       | 0        | 1        | 0        | 0        | 0        | 0        | 232   |
| Gov. subsidized insurance          | 312    | 15      | 10      | 8       | 7       | 4        | 6        | 1        | 2        | 1        | 0        | 366   |
| Uninsured                          | 71     | 1       | 2       | 0       | 1       | 0        | 1        | 0        | 0        | 0        | 0        | 77    |
| Unknown or pending insurance       | 9      | 0       | 1       | 0       | 1       | 0        | 0        | 0        | 0        | 0        | 0        | 11    |
| Non-registered insurance           | 5 406  | 32      | 11      | 14      | 6       | 1        | 0        | 0        | 0        | 0        | 0        | 5 470 |
| TOTAL                              | 10 166 | 177     | 114     | 119     | 38      | 28       | 14       | 3        | 4        | 2        | 1        | 10 666 |