

Estimating the size of high-risk populations for COVID-19 death in the Americas

Partnership between PAHO and Johns Hopkins School of Public Health

1. Introduction

A variety of **predisposing factors** have been associated with illness and death from COVID-19. Estimating the size of **high-risk populations for COVID-19 death** in the Americas allows to understand the distribution of risks associated with these factors providing opportunities for targeting interventions.

In this analysis, the distribution of risk for COVID-19 mortality is calculated by utilizing recently published estimates of risk associated with the following age groups 18-39, 40-49, 50-59, 60-69, 70-79, 80+1, sex, ethnicity², educational level, and health conditions from a UK-based study³ (Table A - Annex). These risk estimates are used to calculate the population's size exceeding defined risk-thresholds by combining the information available on the prevalence of the same factors from multiple population-based data sources and individual-level data available on nationally or locally representative surveys.

This analysis estimates the percentage of population exceeding different levels of risk for COVID-19 death and identifies population sub-groups to apply protective measures.

The results produced by this model might provide guidance to national and local policy makers to plan resources needed for shielding strategies to protect most vulnerable populations, and to better implement interventions such as social distancing, planning on vaccine procurement and allocation, maintaining the services /access to medications and treatment of the population with chronic conditions.

Furthermore, this analysis may complement other tools applied by PAHO to support the production of data related with COVID-19, such as the tool to estimate the population living with underlying health conditions that increase their risk of developing severe disease⁴.

2. Key definitions

Predisposing factors: includes social demographic characteristics such as age, sex, ethnicity, educational level, deprivation, and 11 health conditions⁵, defined as follows:

- (1) Tobacco smoking⁶: ever (current + former), never smoker.
- (2) Obesity: non-obese ($<30 \text{ kg/m}^2$), obese class I ($30-34.9 \text{ kg/m}^2$), obese class II ($35-39.9 \text{ kg/m}^2$), obese class III ($\ge 40 \text{ kg/m}^2$).
- (3) Raised blood pressure/hypertension: SBP<140 and DBP<90 and not diagnosed, SBP>=140 or DBP>=90 or diagnosed.
- (4) Raised blood glucose/diabetes: diagnosed (≤6.9 mmol/L and not on medication, ≥7.0 mmol/L or on medication), or controlled (controlled HbA1c<58 mmol/mol, uncontrolled HbA1c>=58 mmol/mol).

⁶ Tobacco smoking estimates are only reported for current and former smokers, and the results should be taken with caution due to data processing decisions reported in Williamson *et al.* 2020, page 4. The same recommendation applies to obesity (... "those with missing BMI were assumed non-obese and those with missing smoking information were assumed to be nonsmokers on the assumption that both obesity and smoking would be likely to be recorded if present").



¹ Age ranges in population-based surveys varies in the region. The age ranges presented here are from the UK study.

² Unlike the analysis of UK and US data, information on risk associated with different ethnic/racial categories and area-level measures of social deprivation may not be available and thus may not be used for the analysis of risk distribution.

³ At the time this note was developed, the following preprint version was used: Williamson E, Walker AJ, Bhaskaran KJ, et al. OpenSAFELY: factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. medRxiv 2020:2020.05.06.20092999. Available at: https://bit.ly/3fPBMn9. The final paper was published at: Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, Curtis HJ, Mehrkar A, Evans D, Inglesby P, Cockburn J. Factors associated with COVID-19-related death using OpenSAFELY. Nature. 2020 Aug;584(7821):430-6. Available at: https://www.nature.com/articles/s41586-020-2521-4.pdf

⁴ PAHO Technical Note. COVID-19 and comorbidities in the Americas. Available at: https://www.paho.org/en/documents/covid-19-and-comorbidities-americas-background-information

⁵ The health conditions listed above were selected from the original UK study based on availability of data.

- (5) Cardiovascular disease: heart attack or chest pain from heart disease (angina).
- (6) Chronic respiratory disease: asthma (diagnosed/not diagnosed), COPD (diagnosed/not diagnosed).
- (7) Cancers with direct immunosuppression (hematological malignancy).
- (8) Cancers with possible immunosuppression caused by treatment (non-hematological).
- (9) Chronic liver disease.
- (10) Chronic kidney disease.
- (11) Chronic neurological disorders (dementia, Alzheimer, stroke, etc.)

The **risk-score for an individual** is the weighted combination of various sociodemographic characteristics and predisposing health conditions, with weights defined by the relative magnitude of the contribution of these factors to the risk of death due to COVID-19.

The **risk for the underlying population** is the average of the individual risks over a representative sample from the population.

The **risk-thresholds** are defined based on risks of individuals relative to average risk for the underlying population. Specifically, size of populations exceeding 5 different risk thresholds, 1.2, 2.0, 5.0, 10.0 and 25.0, are evaluated. The population exceeding the 5.0 risk-threshold is considered to be "at high risk".

Further, an **index of excess risk** for sub-populations within a country, such as states, is defined as the average risk of individuals within the sub-population relative to that of the population of the whole country.

3. Data sources for obtaining prevalence and joint distribution of the predisposing factors

The main data sources generally available to produce these estimates are presented in the table below. Each country may assess the availability of national data prior to producing their estimates.

Table 1. Data sources for predisposing factors

Predisposing factors	Primary data sources	Complementary data sources
Age Sex Ethnicity Index of multiple deprivation (IMD) quintile		Population Census / Another national household survey
Tobacco smoking		
Obesity Reject bleed procesure/bypertension	National Health Survey / NCD survey / Local health care information system (specific	
Raised blood pressure/hypertension Raised blood glucose/diabetes Cardiovascular disease Chronic respiratory disease Chronic liver disease Chronic kidney disease	geographic area)	Global Burden of Disease
Chronic neurological disorders		2017
Cancer (non-hematological) Hematological malignancy	IARC Global Cancer Observatory (GLOBOCAN) / Local health care information system (specific geographic area)	

4. Statistical models and methods

The analysis applies recently published results from a UK study on risk of mortality associated with a variety of predisposing factors to available data on risk factor distribution in a given population (Table A - Annex). For each individual, a risk-score is defined based on weighted combination of his/her risk factor profile with weights obtained from multivariate adjusted risk factor and health condition associations reported by the UK study. The distribution of this risk score in a given population is analyzed using various health data for a specific country. Various levels of risks are defined based on relative risks of individuals compared to an "average" risk for the specific country, and the size of populations that belongs to these risk categories are estimated. When individual level data are not available, but information are available on prevalence of each predisposing factor, then a series





of approximations is used to estimate the predisposing factor distribution accounting for information on these factors from other sources. Details of these methods and their application for analysis of risk for the United States at the level of cities, counties, states and the whole country are reported elsewhere⁷.

5. Applying the methodology: data analysis exercise for Mexico

A data analysis exercise was carried out using publicly available data for Mexico, specifically the National Health and Nutrition Survey (Encuesta Nacional de Salud y Nutrición, ENSANUT) conducted in 20128. The ENSANUT 2012 was a nationally representative household survey and its main objective was to quantify the frequency, distribution and trends of health and nutrition conditions and their determinants. Although rounds of the ENSANUT were conducted in 2016 and 2018, ENSANUT 2012 was used due to its sample size and data availability at the time this exercise was conducted. Data from the International Agency for Research on Cancer (IARC) was also used.

5.1. Methods

First, the variables for analysis were identified in the dataset corresponding to adults (20 years and older). These were sex, age, body mass index (BMI) category, self-reported hypertension, smoking status, chronic heart disease, stroke, diabetes, and cancer, including hematological and non-hematological cancers. Of the selected variables, only BMI category was found to have significant missing data (26%). To account for this, values of BMI categories were imputed by using the distribution of BMI categories amongst individuals stratified by sex, age, and diabetes status. Imputation was also done for those who answered "Other" for cancer type by using IARC data for Mexico to determine whether "Other" could refer to a hematological malignancy or not. Former and current smokers were combined into the category "ever smokers". A weighted combination of the multivariable adjusted coefficients associated with current and former smokers was used (see Table A) with weights defined by relative proportion of current and former smokers in the UK study. Similarly, for the hematological and non-hematological cancer variable, a weighted combination of the coefficients associated with each of the categories defined according to the time since diagnosis was calculated: i) <1 year ii) 1-4.9 years and iii) >=5 years. In calculating risk scores of the individuals (n=46,244) and prevalences, the weights provided in the dataset were used to create weighted means of all data points.

5.2. Results

Prevalences of all the predisposing factors in the general Mexico adult population were explored (Figure 1, Table B). Around 50% of the population are 40 years or younger while around 15% of the population are older than 60 years. As expected, there are almost equal proportions of males and females. Approximately 32% of the population are obese (BMI >= 30Kg/m²) with ~22% falling in the category "Obese class I" (BMI 30-34.9Kg/m²). Hypertension and diabetes are the most prevalent conditions.

⁸ Further information on the ENSANUT 2012, including available data sets and documentation can be found at: https://ensanut.insp.mx/encuestas/ensanut2012/index.php





⁷ Jin J, Agarwala N, Kundu P, Harvey B, Zhang Y, Wallace E, Chatterjee N. Individual and community-level risk for COVID-19 mortality in the United States. Nature Medicine. 2020 Dec 11:1-6. Available at: https://www.nature.com/articles/s41591-020-01191-8#Sec25

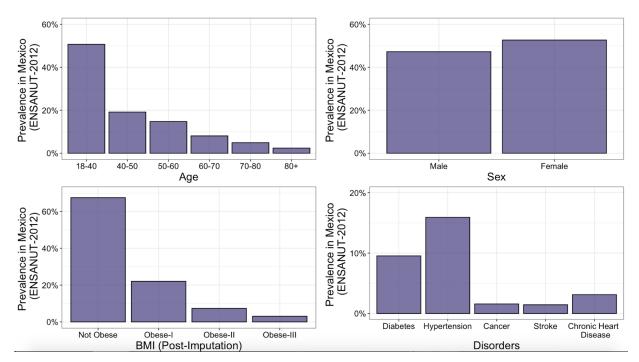


Figure 1. Prevalence of the predisposing factors in the Mexico population, ENSANUT 2012.

The risk-score for Mexico population was evaluated to explore the distribution of risk associated with the predisposing factors in this population. There is wide variation in risk across individuals in the general Mexico adult population (Figure 2). Overall, 18%, 13%, 7.3%, 5.1%, 1.8% and 0.2% of the individuals are at or above risk thresholds 1.2-fold, \geq 2-fold, \geq 3-fold, \geq 5-fold and 25-fold risk categories, respectively (Figure 2). The percentage of the population exceeding these thresholds varies strongly by age. Only a small fraction (0.1%) of the individuals who are younger than 60 years old exceed the \geq 5-fold threshold for high risk.

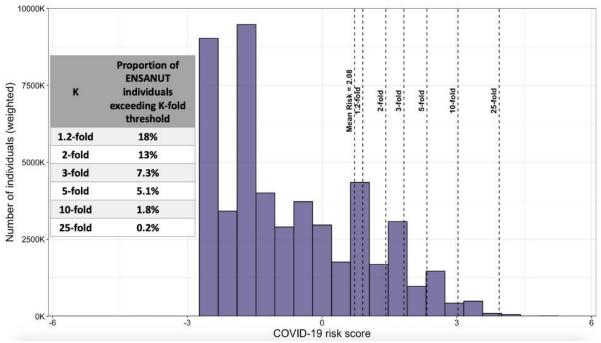


Figure 2. Distribution of risk-score across individuals who participated in the ENSANUT Mexico 2012 Survey

When exploring the distribution of various other predisposing factors among individuals in the defined high-risk groups for the general population (Figure 3), it is observed that the proportion of males, persons with hypertension, diabetes and various other health conditions are higher in higher risk groups compared to the general Mexico adult population.

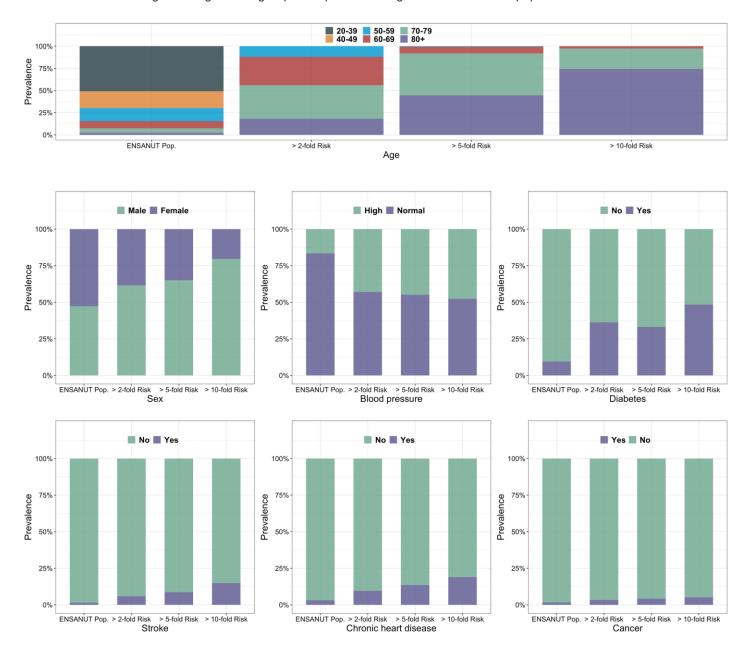


Figure 3. Distribution of predisposing factors in adult population and among individuals in different risk groups including >2-fold risk, >5-fold risk and >10-fold risk, Mexico.

There is some variation in risk due to predisposing factors across Mexico states (Figure 4). A number of major states, including Zacatecas, Ciudad de México, Veracruz, Michoacán, Durango and Morelos rank very high according to this index.

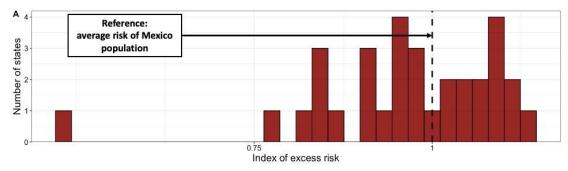


Figure 4. Distribution of the Index of Excess Risk (IER) for COVID-19 mortality across Mexico states.

The proportion of individuals crossing various risk thresholds varies even more widely across states (Table C - Annex). For example, the percentage of the adult populations in states which exceed the 5-fold risk threshold varies from 2% (Quintana Roo) to 7% (Ciudad de México).

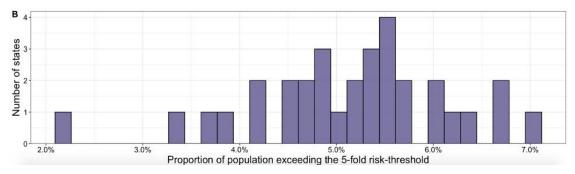


Figure 5. Projections for the proportion of high-risk population expected to occur at 5-fold risk threshold across Mexico states.

The mean risk of COVID-19 mortality at the national level and by state was also calculated and is presented in the map below (Figure 6).



Figure 6. Mean risk for COVID-19 mortality by state, Mexico.

How to do it?

- 1) Identify the sources available to produce estimates for each one of the predisposing factors that is representative of the locality and age ranges you want to produce the estimate for.
- 2) If the data is at individual-level, compute the average risk, size and proportions exceeding various risk-thresholds for the community using the R function calculateRisks.
- 3) If individual-level data is not available, produce estimates for each one of the predisposing factors for a specific population and the geographical area required (national, department, municipality, county...).
- 4) Compute the average risk, size and proportions exceeding various risk-thresholds for the geographical area required using the R function calculateRisks.

To practice this procedure using R, you can replicate the analysis presented in section 5 "Applying the methodology: data analysis exercise for Mexico", using the code available at: https://github.com/nchatterjeelab/COVID19Risk/tree/master/PAHO



ANNEX

Table A. Hazard Ratios (HRs) and 95% confidence intervals (CI) in sensitivity analyses. UK study results.9

	Fully adjusted HR and 95% CI				
	Primary analysis	Early censoring at 6/4/2020	Restricted to those with complete BMI /smoking	Adjusted for ethnicity in those where recorded	
Age					
18-<40	0.07 (0.05-0.10)	0.08 (0.05-0.13)	0.09 (0.06-0.13)	0.08 (0.06-0.11)	
40-<50	0.31 (0.25-0.39)	0.31 (0.21-0.46)	0.32 (0.25-0.41)	0.30 (0.23-0.38)	
50-<60	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	
60-<70	2.09 (1.84-2.38)	2.35 (1.90-2.91)	2.12 (1.85-2.44)	2.07 (1.79-2.39)	
70-<80	4.77 (4.23-5.38)	5.55 (4.54-6.77)	4.84 (4.25-5.50)	4.81 (4.20-5.51)	
80+	12.64 (11.19- 14.28)	13.43 (10.95- 16.45)	12.76 (11.18- 14.55)	12.04 (10.47- 13.84)	
Sex					
Female	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	
Male	1.99 (1.88-2.10)	2.18 (1.99-2.38)	2.06 (1.94-2.19)	1.93 (1.80-2.06)	
ВМІ					
Not obese	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	
Obese class I (30-34.9kg/m ²)	1.27 (1.18-1.36)	1.39 (1.25-1.54)	1.32 (1.23-1.41)	1.28 (1.18-1.38)	
Obese class II (35-39.9kg/m ²)	1.56 (1.41-1.73)	1.62 (1.39-1.90)	1.64 (1.48-1.81)	1.60 (1.43-1.80)	
Obese class III (≥40 kg/m²)	2.27 (1.99-2.58)	2.45 (2.00-3.01)	2.40 (2.10-2.74)	2.28 (1.96-2.65)	
Smoking					
Never	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	
Ex-smoker	1.25 (1.18-1.33)	1.33 (1.21-1.47)	1.24 (1.17-1.33)	1.32 (1.23-1.42)	
Current	0.88 (0.79-0.99)	0.84 (0.70-1.01)	0.92 (0.81-1.03)	0.94 (0.82-1.07)	
Ethnicity*					
White	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	
Mixed	1.64 (1.19-2.26)	1.13 (0.62-2.05)	1.58 (1.12-2.22)	1.64 (1.19-2.26)	
Asian or Asian British	1.62 (1.43-1.82)	1.77 (1.48-2.13)	1.69 (1.49-1.91)	1.62 (1.43-1.82)	
Black	1.71 (1.44-2.02)	1.90 (1.48-2.45)	1.69 (1.42-2.02)	1.71 (1.44-2.02)	
Other	1.33 (1.03-1.73)	1.81 (1.28-2.57)	1.41 (1.07-1.84)	1.33 (1.03-1.73)	
IMD quintile					
1 (least deprived)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	
2	1.13 (1.04-1.24)	1.01 (0.88-1.16)	1.12 (1.02-1.23)	1.19 (1.07-1.33)	
3	1.23 (1.13-1.35)	1.04 (0.91-1.20)	1.23 (1.12-1.35)	1.26 (1.13-1.40)	
4	1.49 (1.37-1.63)	1.27 (1.11-1.46)	1.48 (1.35-1.62)	1.53 (1.38-1.70)	
5 (most deprived)	1.75 (1.60-1.91)	1.49 (1.29-1.71)	1.72 (1.57-1.89)	1.70 (1.53-1.89)	

⁹ At the time this note was developed, the following preprint version was used: Williamson E, Walker AJ, Bhaskaran KJ, et al. OpenSAFELY: factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. medRxiv 2020:2020.05.06.20092999. Available at: https://bit.ly/3fPBMn9. The final paper was published at: Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, Curtis HJ, Mehrkar A, Evans D, Inglesby P, Cockburn J. Factors associated with COVID-19-related death using OpenSAFELY. Nature. 2020 Aug;584(7821):430-6. Available at: https://www.nature.com/articles/s41586-020-2521-4.pdf





Blood pressure				
Normal	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)
High, or diagnosed hyper- tension	0.95 (0.89-1.01)	0.94 (0.85-1.05)	0.94 (0.88-1.01)	0.97 (0.90-1.05)
Co-morbidities				
Respiratory disease ex asthma	1.78 (1.67-1.90)	1.97 (1.77-2.18)	1.74 (1.62-1.86)	1.79 (1.66-1.93)
Asthma (vs none)* ²				
With no recent OCS use	1.11 (1.02-1.20)	1.14 (1.01-1.29)	1.10 (1.02-1.20)	1.03 (0.94-1.13)
With recent OCS use	1.25 (1.08-1.44)	1.39 (1.12-1.73)	1.22 (1.05-1.42)	1.24 (1.06-1.46)
Chronic heart disease	1.27 (1.20-1.35)	1.33 (1.22-1.46)	1.27 (1.19-1.35)	1.27 (1.19-1.36)
Diabetes (vs none)* ³				
Controlled (HbA1c<58 mmol/mol)	1.50 (1.40-1.60)	1.48 (1.33-1.65)	1.47 (1.37-1.57)	1.47 (1.36-1.59)
Uncontrolled (HbA1c>=58 mmol/mol)	2.36 (2.18-2.56)	2.57 (2.27-2.91)	2.30 (2.12-2.50)	2.23 (2.03-2.45)
No recent HbA1c measure	1.87 (1.63-2.16)	1.68 (1.33-2.12)	1.85 (1.60-2.15)	1.91 (1.63-2.24)
Cancer (non- haematological, vs none)				
Diagnosed < 1 year ago	1.56 (1.29-1.89)	1.51 (1.10-2.05)	1.52 (1.24-1.86)	1.68 (1.36-2.09)
Diagnosed 1-4.9 years ago	1.19 (1.04-1.35)	1.36 (1.13-1.65)	1.20 (1.05-1.37)	1.21 (1.04-1.40)
Diagnosed ≥5 years ago	0.97 (0.88-1.06)	0.92 (0.79-1.06)	0.96 (0.87-1.05)	1.02 (0.92-1.13)
Haematological malignancy (vs none)				
Diagnosed < 1 year ago	3.52 (2.41-5.14)	2.60 (1.30-5.22)	3.77 (2.58-5.50)	3.30 (2.10-5.18)
Diagnosed 1-4.9 years ago	3.12 (2.50-3.89)	3.67 (2.66-5.06)	3.03 (2.40-3.83)	3.42 (2.67-4.38)
Diagnosed ≥5 years ago	1.88 (1.55-2.29)	1.64 (1.18-2.28)	1.90 (1.55-2.33)	1.84 (1.46-2.32)
Liver disease	1.61 (1.33-1.95)	1.86 (1.40-2.47)	1.59 (1.30-1.93)	1.61 (1.30-2.00)
Stroke/dementia	1.79 (1.67-1.93)	1.61 (1.43-1.81)	1.78 (1.65-1.92)	1.75 (1.61-1.90)
Other neurological	2.46 (2.19-2.76)	2.28 (1.88-2.76)	2.38 (2.10-2.69)	2.41 (2.11-2.76)
Kidney disease	1.72 (1.62-1.83)	1.75 (1.58-1.92)	1.71 (1.60-1.82)	1.76 (1.64-1.89)
Other immunosuppressive condition	1.69 (1.21-2.34)	2.01 (1.25-3.25)	1.52 (1.06-2.19)	1.66 (1.16-2.39)



Table B. Prevalence of predisposing factors in population 20+, Mexico.

Predisposing factors	Prevalence
Sex	
Male	47.3%
Female	52.7%
Age (in years)	
20-39	50.7%
40-49	19.2%
50-59	14.8%
60-69	8.1%
70-79	4.9%
80+	2.4%
Obesity	
Non-obese	68.1%
Obese class I (30-34.9 kg/m²)	21.8%
Obese class II (35-39.9 kg/m²)	7.1%
Obese class III (≥40 kg/m²)	3%
Diabetes	9.5%
Hypertension	16.6%
Cancer	1.6%
Stroke	1.4%
Cardiovascular disease	3.1%

Source: prepared by authors based on ENSANUT 2012 microdata.



Table C. Proportion of adults 20+ exceeding different risk-threshold by state, Mexico.

	Proportion of population exceeding K-fold risk					
State	K = 1.2	K = 2	K = 3	K = 5	K = 10	K = 25
Quintana Roo	0.104	0.063	0.030	0.022	0.008	0.000
Baja California Sur	0.141	0.096	0.054	0.033	0.011	0.003
Chiapas	0.156	0.108	0.055	0.036	0.015	0.001
Campeche	0.171	0.116	0.063	0.039	0.018	0.002
Aguascalientes	0.145	0.110	0.060	0.041	0.014	0.001
Tabasco	0.160	0.117	0.060	0.042	0.011	0.001
Baja California	0.145	0.100	0.057	0.045	0.016	0.002
Coahuila	0.163	0.122	0.065	0.045	0.011	0.001
México	0.165	0.116	0.060	0.046	0.016	0.005
Querétaro	0.134	0.099	0.059	0.046	0.021	0.001
Tlaxcala	0.169	0.124	0.060	0.048	0.020	0.003
Colima	0.177	0.129	0.070	0.049	0.021	0.003
Hidalgo	0.168	0.132	0.068	0.049	0.018	0.000
Guerrero	0.173	0.132	0.073	0.051	0.017	0.000
Chihuahua	0.158	0.119	0.066	0.052	0.018	0.001
Tamaulipas	0.174	0.124	0.077	0.052	0.019	0.003
Guanajuato	0.163	0.123	0.072	0.054	0.020	0.002
Sinaloa	0.188	0.137	0.080	0.054	0.025	0.003
Sonora	0.170	0.126	0.073	0.054	0.021	0.004
San Luis Potosí	0.185	0.139	0.078	0.055	0.022	0.000
Yucatán	0.188	0.140	0.081	0.055	0.026	0.003
Durango	0.188	0.151	0.080	0.056	0.019	0.005
Puebla	0.170	0.126	0.079	0.056	0.015	0.001
Oaxaca	0.187	0.142	0.078	0.056	0.018	0.000
Ciudad de México	0.210	0.164	0.088	0.058	0.020	0.001
Morelos	0.181	0.137	0.082	0.060	0.021	0.001
Jalisco	0.192	0.141	0.086	0.061	0.025	0.004
Nuevo León	0.169	0.128	0.084	0.062	0.025	0.004
Nayarit	0.204	0.143	0.084	0.063	0.022	0.002
Michoacán	0.185	0.146	0.091	0.067	0.024	0.003
Veracruz	0.191	0.150	0.086	0.067	0.026	0.002
Zacatecas	0.195	0.153	0.099	0.071	0.028	0.001

Source: prepared by authors based on ENSANUT 2012 microdata.





PAHO/NMH/NV/COVID-19/21-0001

© Pan American Health Organization, 2021. Some rights reserved. This work is available under license CC BY-NC-SA 3.0 IGO.

Technical assistance:

Roberta Caixeta

Advisor, NCD Surveillance, Prevention

and Control

Noncommunicable Diseases and Mental

Health Department, PAHO

Dolores Ondarsuhu

Specialist, NCD Surveillance and

Monitoring

Noncommunicable Diseases and Mental

Health Department, PAHO

Sebastián García Saiso

Director

Evidence and Intelligence for Action in

Health Department, PAHO

Adrienne Cox

Unit Chief, Health Analysis

Evidence and Intelligence for Action in

Health Department, PAHO

Prosenjit Kundu

Postdoctoral Fellow

School of Public Health, Johns Hopkins

University

Tanuj Alapati

Undergraduate Student

Johns Hopkins University

Neha Agarwala

PhD Candidate

University of Maryland, Baltimore County

Nilanjan Chatterjee

Bloomberg Distinguished Professor School of Public Health and School of Medicine, Johns Hopkins University

For further information: nmhsurveillance@paho.org



