

## 1 Estimation of COVID-19 cases in Mexico accounting for SARS-CoV-2 RT-PCR false negative results

2  
3 Isaac J. Núñez, MD  
4 Programa de especialidad en Medicina Interna, Departamento de Educación Médica, Instituto Nacional de Ciencias  
5 Médicas y Nutrición Salvador Zubirán  
6 Vasco de Quiroga # 15, Belisario Domínguez Sección XVI, Mexico City, Mexico, PC 14080

7  
8 Pablo F. Belaunzarán-Zamudio, MD  
9 National Institute of Allergy & Infectious Diseases, National Institutes of Health  
10 31 Center Dr # 7A03, Bethesda, Maryland, United States, PC 20892

11  
12 Yanink Caro-Vega, PhD  
13 Departamento de Infectología, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán  
14 Vasco de Quiroga # 15, Belisario Domínguez Sección XVI, Mexico City, Mexico, PC 14080

15  
16 Corresponding author:

17 Isaac J. Núñez  
18 isaac.nunezs@incmnsz.mx  
19 Programa de especialidad en Medicina Interna, Departamento de Educación Médica, Instituto Nacional de Ciencias  
20 Médicas y Nutrición Salvador Zubirán  
21 Vasco de Quiroga # 15, Belisario Domínguez Sección XVI, Mexico City, Mexico, PC 14080  
22 (+52) 5554870999

### 23 24 **ABSTRACT**

25 Underestimation of the number of cases during the COVID-19 pandemic has been a constant concern worldwide. Case  
26 confirmation is based on identification of SARS-CoV-2 RNA using real time polymerase chain reaction (RT-PCR) in clinical  
27 samples. However, these tests have suboptimal sensitivity, especially during the early and late course of infection. Using  
28 open data, we estimated that among 1 343 730 people tested in Mexico since February 27<sup>th</sup>, there were 838 377 (95% CL  
29 734 605 – 1 057 164) cases, compared with 604 376 considering only positive tests. ICU admissions and deaths were  
30 around 16% and 9% higher than reported. Thus, we show that accounting for the sensitivity of SARS-Cov-2 RT-PCR  
31 diagnostic tests is a simple way to improve estimations for the true number of COVID-19 cases in tested people,  
32 particularly in high-prevalence populations. This could aid to better inform public health measures and reopening policies.

### 33 34 **INTRODUCTION**

35 Around 25 485 000 confirmed cases of COVID-19 have been reported worldwide by country governments by August 31<sup>st</sup>,  
36 2020.<sup>1</sup> The Mexican government reported 599 560 confirmed cases to that date.<sup>2</sup> Case confirmation is based on  
37 identification of SARS-Cov-2 RNA using real time polymerase chain reaction (RT-PCR) in clinical samples collected through  
38 nasopharyngeal or oropharyngeal swabs, saliva, or bronchoscopy. Dealing with the disease has proven extremely  
39 challenging for governments and health systems worldwide, partially due to the difficulties in case identification. Various  
40 measures have been proposed to reduce the viruses' impact on the population, most of them rely on case identification  
41 for isolation and contact tracing.<sup>3-5</sup>

42 The World Health Organization (WHO) has highlighted the importance of generalized testing with the goal of early  
43 detection, quarantine, and contact tracing. Countries like South Korea and Iceland have been successful in implementing  
44 wide-spread testing, case-isolation and contact tracing, keeping the virus under control.<sup>6,7</sup> In countries where tests are less  
45 available, focusing this resource to high risk people was deemed reasonable as a provisional strategy, with the urge to  
46 increase testing capacity. Mexico chose a different strategy, and a decision to use testing only for surveillance purposes

47 was made early during the pandemic. Criteria for testing are applied as for the sentinel surveillance system for influenza,  
48 and the information provided allegedly used to estimate the total number of infections based on mathematical modelling.  
49 As anywhere else, underestimation of the number of cases has been a constant concern.<sup>4-7</sup>  
50 Diagnostic tests rarely, if ever, are completely reliable, and RT-PCR for SARS-CoV-2 is no exception. Specificity almost  
51 always nears 100% in this kind of tests, but poor sensitivity has been an issue. Kucirka and collaborators estimated that  
52 test sensitivity is highest at the fourth day of symptoms onset (81% [95% confidence limits; 95%CL 71-88%]).<sup>8</sup> Sensitivity is  
53 the lowest during the asymptomatic and late symptomatic periods (eg. 37%, 95%CL 26-49%) at day 21 after symptom  
54 onset. Other factors that could influence the accuracy of the test is the type of clinical specimen, severity of infection, and  
55 gene targets. A combination of these, and other factors, may account for the underestimation of the number of cases and  
56 attributable deaths worldwide.<sup>8-11</sup> Until screening and diagnostic tests performance are optimized, applying mathematical  
57 modelling strategies can aid in estimating more accurately diseases occurrence for surveillance purposes. In this study, we  
58 aimed to provide corrected estimates of the number of cases among people that were tested for SARS-Cov-2 in Mexico  
59 between February 27<sup>th</sup> and August 31<sup>th</sup>, 2020 by taking into account the probability of RT-PCR false negative tests results.

60

## 61 **METHODS**

### 62 **Study setting**

63 In Mexico, the first COVID-19 confirmed case was tested on February 27<sup>th</sup> and reported on February 28<sup>th</sup>. Community  
64 transmission was declared on March 24<sup>th</sup> and mitigation country-wide measures were announced the same day. Social  
65 distancing was urged, and non-essential businesses and activities were suspended, initially until April 14<sup>th</sup>. The testing  
66 strategy was also published that day. A case definition was developed, and testing was recommended for one in ten  
67 patients seeking care due to a mild case of an influenza like-illness in a limited number of health services previously  
68 established to monitor seasonal Influenza, and all of those requiring hospitalization. Also, heavy emphasis on voluntary  
69 quarantine if mild symptoms developed, urging people with co-morbidities and other high-risk conditions, such as older  
70 age, to search for health care. No accompanying contact tracing measure was spoken of, placing most of the responsibility  
71 at the individual level.<sup>12</sup>

72

### 73 **Data sources and selection**

74 We used the SARS-CoV-2 tests open datasets made public since April 12 by the Mexican government in their official  
75 coronavirus web page and updated daily.<sup>13</sup> The datasets include every test done at public, but not private, laboratories. It  
76 contains State and Municipality where the test was collected, sociodemographic information, dates of symptoms onset,  
77 date the patient was included in the database, ICU admission and death (if occurred) for everyone, with non-traceable,  
78 individual key identifiers. We assumed the date the patient was registered in the database was the date of testing, and we  
79 will refer to it as such from now on. We included in the analysis all tested individuals registered in the dataset between  
80 February 27<sup>th</sup> and August 31<sup>st</sup>. Patients with pending result, missing identification code, or more than 21 days with  
81 symptoms at the moment of the test were excluded.

82 We analyzed each information according to the date the tests results are reported in the database, regardless of the day it  
83 was collected to follow the format of the daily report by the Ministry of Health (See supplementary). This was not possible  
84 for people tested before April 12<sup>th</sup> and already had a result, so their result was included as the baseline count.

85

### 86 **False negative estimation model**

87 We used the method described by Kucirka and collaborators in their mathematical modelling study to calculate the false-  
88 negative rate of RT-PCR diagnostic tests. They calculated the sensitivity from day one of infection (assuming symptoms  
89 started at the 5<sup>th</sup> day of infection) until day 21 of infection. Since their estimates end at day 16 of symptoms, we replicated  
90 their analysis and estimated sensitivity up to day 21 of symptoms with 95% uncertainty bounds. Sensitivity varied  
91 depending on the day after infection, being higher during the symptomatic phase and reaching a maximum of 81% at the  
92 fourth day of symptoms (Supp Append Table S1). We used the mean estimate for the graphical representation but

93 repeated the estimation with the upper and lower confidence bounds. Specificity for every test used by the Mexican  
94 government is reportedly 100%.<sup>8, 14</sup>

95 We used the following contingency table as the basis for our analysis:

96

Test result	Person has COVID-19		Total
	Yes	No	
Positive	a	b	a + b
Negative	c	d	c + d
Total	a + c	b + d	a + b + c + d

97

98 Given a 100% of specificity, there are no false positive results then  $b=0$ , and all the positive results are true positives. From  
99 data set, each day we knew the number of true positives 'a' and the number of negative tests 'c+d'. Our interest is to  
100 estimate the daily number of false negatives 'c'. The probability of being false negative 'p' is defined in the equation (1):

$$p = \frac{c}{a + c}$$
$$c = \frac{pa}{1 - p}$$
$$c = \frac{pa}{s}$$

101

102 By reproducing Kucirca's analysis, we know 'p' and 's' (1-p) the test sensitivity for each time since symptom initiation  
103 reported. At each calendar day, we split individuals in groups, each corresponding to the number of days with symptoms  
104 when tested. For example, a day in which 100 people were tested, with 30 presenting on their 6<sup>th</sup> day and 70 in their 7<sup>th</sup>  
105 day with symptoms, two groups were created, each one with a test sensitivity (s) and the number of reported positive  
106 tests (a). Hence, we applied the equation to each group of every calendar day from February 27<sup>th</sup> to August 31<sup>st</sup> and added  
107 the false negatives calculated on every group. As the number of true COVID-19 cases is limited by the number of people  
108 tested, in case the estimation yielded a higher number of cases then the totality of people tested was used instead.

109 We also estimated the corrected number of ICU admission and the corrected number of deaths due to COVID-19 by  
110 calendar day. Assuming no difference in test precision among disease severity spectrum, we added the product of the  
111 proportion of negatives estimated to be false negatives and the number of ICU admissions or deaths among COVID  
112 negative patients. We applied the following equation to correct death and ICU admissions:

113

$$\text{Estimated true COVID19 deaths} = \text{official COVID19 positive deaths} + (\text{proportion of false negatives} * \text{official COVID19 negative deaths})$$

114

115  
116 Thus, if in a given day there were 100 deaths and 50 ICU admissions among COVID-19 negative patients, and the  
117 estimated false negative proportion using the mean estimate of the test sensitivity was of 0.40, we would add 40 deaths  
118 and 20 ICU admissions to the COVID-19 positive group for that particular date.

119

## 120 Statistical Analysis

121 Applying results derived from the false negative estimation model on the government official dataset, we estimated the  
122 daily corrected number of cases. We performed the analysis at a National level and for each one of the 31 States and the  
123 country capital (Ciudad de México, formerly known as Distrito Federal).

124 To determine if positivity rates could be due to low testing per capita, we calculated Spearman's rho of positivity rates and  
125 the number of tests done per 10 000 habitants by state. State population was obtained from the national statistics and  
126 geography institute's (INEGI) most recent published data.<sup>15</sup> We also calculated the 7-day moving average of time from

127 symptom onset to testing for the entire study period to determine if this could explain higher false negative rates during  
128 certain time periods.

129 All data analysis was done with R software version 4.0.0. The ethics committee of the Instituto Nacional de Ciencias  
130 Médicas y Nutrición Salvador Zubirán reviewed and approved the study. There was no sponsor involved in any step of the  
131 study.

132

## 133 RESULTS

134 There were 1 343 730 people tested between February 27<sup>th</sup> and August 31<sup>st</sup> according to the latest official database  
135 (August 31<sup>st</sup>). A detailed explanation of data selection is provided in supplementary materials. We included 1 280 910  
136 patients that had an available result, were tested at less than or 21 days with symptoms and had no missing ID code. Of  
137 them, 604 376 (47.2%) were SARS-CoV-2 positive and 676 534 (52.8%) negative (Table 1).

138

139

**TABLE 1. Official and estimated nation-wide COVID-19 cases by date of reporting**

Variable	Official count	Expected scenario	Best case scenario	Worst case scenario
	1 280 910	1 280 910	1 280 910	1 280 910
Positive tests	604 376 (47.2%)	838 377 (65.5%)	734 605 (57.4%)	1 057 164 (82.5%)
Negative tests	676 534 (52.8%)	442 533 (34.5%)	546 305 (42.6%)	223 746 (17.5%)
Estimated false negative tests	-	234 001 (18.3%)	130 229 (10.2%)	452 788 (35.3%)
Overall ICU count	13 038	15 085	14 167	17 008
Overall Death count	64 360	69 835	67 406	74 978

140 **Note:** Expected, best and worst-case scenario refer to estimations done with the mean test sensitivity and its upper and  
141 lower 95% confidence limits, respectively. Positive tests in the expected, best case and worst-case scenarios include  
142 estimated false negatives, while negative tests exclude them.

143

### 144 Estimated false negatives and corrected COVID-19 cases

145 We estimated a total of 838 377 (95% CL 734 605 – 1 057 164) positive cases (39% higher than official reported cases)  
146 (Figure 1). In our corrected estimates, 50 000 cases were reached by May 11<sup>th</sup> (May 7<sup>th</sup>-May 14<sup>th</sup>), while the official count  
147 reached that number at May 18<sup>th</sup>. Accumulated test positivity rate increased during the study period, being of 14.4% at  
148 March 24<sup>th</sup>, 36% at April 12<sup>th</sup>, and 50.1% by July 14<sup>th</sup>. Positivity rate also varied according to state and date (Table 2, Supp  
149 Material Figures SE). The states with the most cases were Ciudad de México, Estado de México, Guanajuato, Nuevo León,  
150 and Puebla, while those with the highest positivity rate were Veracruz, Oaxaca, Baja California, Quintana Roo and Hidalgo.  
151 Spearman's Rho for test positivity rate (taking false negatives into account) and tests performed per 10 000 people was of  
152 -0.41 (-0.40, -0.41). Time from symptom onset to testing increased over time (Supp Material Figure S1).

153

### 154 Corrected ICU hospitalization and deaths estimates

155 There were 13 038 ICU admissions and 64 360 deaths among COVID-19 positive patients during the study period (Table 1).  
156 The corrected estimate is of 15 085 ICU admissions (14 167 – 17 008) and 69 835 deaths (67 406 – 74 978). (Table 1). The  
157 magnitude of difference between official reports and corrected estimates varied between states (Table 2).

158

## 159 DISCUSSION

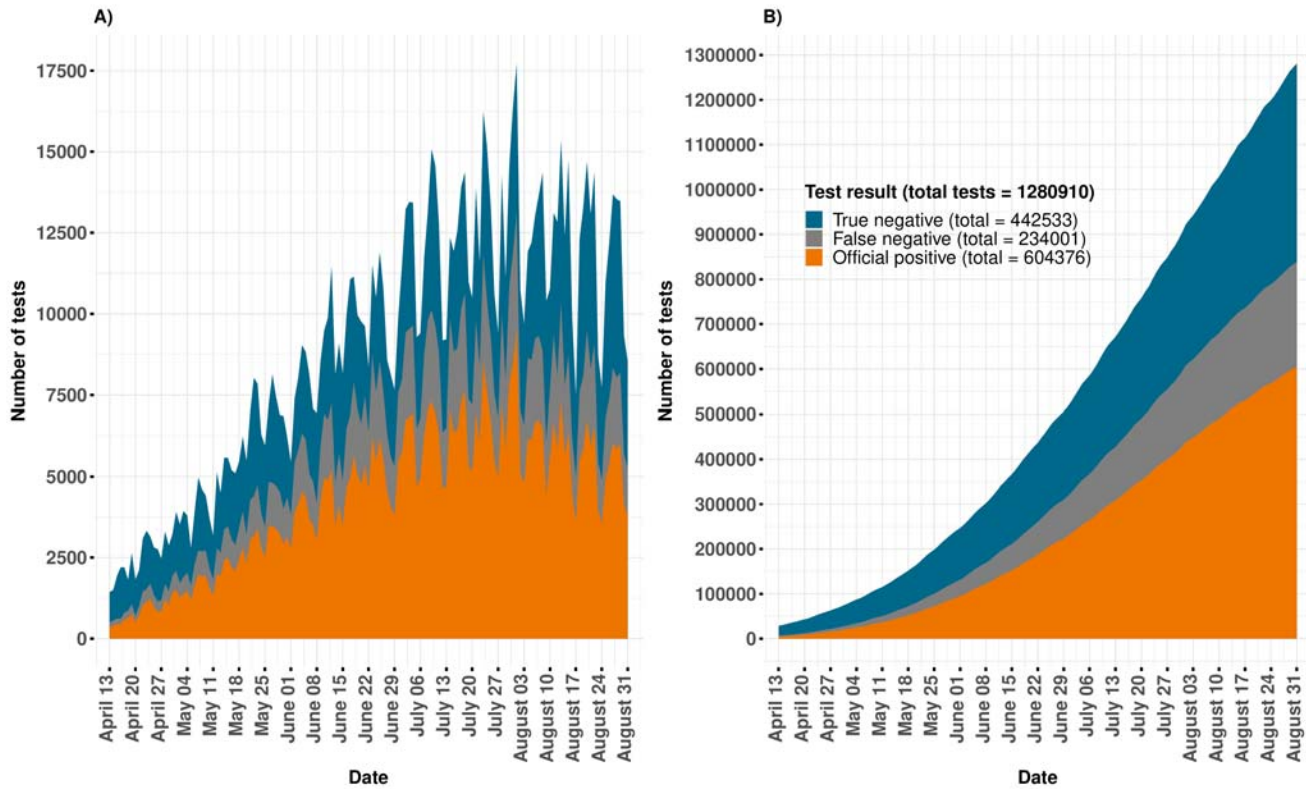
160 In this analysis, we estimated the corrected number of cases of COVID-19, ICU admissions and deaths due to COVID-19 in  
161 Mexico accounting for false negatives tests results using test sensitivities previously estimated by Kucirca and colleagues.  
162 These were estimated based on the day after the onset of symptoms when patients are tested. We identified that the  
163 number of cases of COVID-19 in Mexico based on RT-PCR testing might be almost 40% higher than currently registered,  
164 with 95% confidence limits of 21% and 75% depending on the test sensitivity. These differences vary widely by state and

165 period during the pandemic. Accordingly, the corrected number of ICU admission and deaths increased around 16% and  
166 9%, but this increase might be as high as 25% in ICU and 17% in deaths. The magnitude of these differences may require  
167 important modifications in preparedness for response, which highlights the importance of accounting for the probability  
168 of false negative tests in public health estimations.

169

170

Figure 1. Estimated proportion of tested individuals with a false negative result during the study period



171

172 A) represents new daily test results, B) represents accumulated test results

173

174 The proportion of false negatives was similar to that found in other studies.<sup>16, 17</sup> There were high heterogeneity in  
175 positivity tests proportions, corrected estimates, and confirmed and corrected estimations of ICU hospitalizations and  
176 deaths across States. A modelling study conducted in the United States showed that disease burden varied heavily among  
177 counties, both in an optimistic and in a pessimistic scenario.<sup>18</sup> A recent seroepidemiological study conducted in Spain  
178 found considerable heterogeneity in seroprevalence among provinces, with >10% in the most heavily affected ones and  
179 less than one percent in the least.<sup>19</sup> This data is consistent not only with the occurrence of “local epidemics” rather than a  
180 nation-wide epidemic; but also the fragmented response in Mexico with some regions faring better than others. On the  
181 last point, it may only reflect sound State-centered approaches, with resources being modified according to each states’  
182 needs, though.

183 We also observed a weak negative correlation between test positivity and tests-per-capita. Considering that testing per-  
184 capita is very low in the whole country, this is not surprising. States that have particularly low testing rates are Chiapas,  
185 Chihuahua, Oaxaca, Veracruz, and Queretaro. This could be responsible for the high estimated positivity rate. Sonora and  
186 Baja California have more tests done per 10 000 habitants, but their estimated positivity rate is still of ~90%. Most likely  
187 the number of cases far outnumber the number of tests, with the relatively small differences in tests-per-capita being  
188 apparently inconsequential for the positivity rate. Interestingly, the time from symptom onset to testing did not appear to

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/) .

189 change considerably over time. This suggests that increasing positivity rates most likely do not derive from changes in time  
 190 of testing after symptoms onset along time but from high disease burden and insufficient testing worsening over time.  
 191 Considering that only a small proportion of symptomatic cases who search for healthcare are tested (not even considering  
 192 asymptomatic individuals who have the virus) the true underestimation of COVID-19 cases can be huge. Given this, the  
 193 daily number of cases will most likely grow on par with the number of tests.  
 194

195 **TABLE 2 Official and estimated confirmed state-wide COVID-19 cases**

State	Reported cases (positivity rate)	Estimated cases	Estimated positivity rate (95% CL)	Reported ICU admissions	Estimated ICU admissions (95% CL)	Reported deaths	Estimated deaths (95% CL)	Total tests	Tests per 10 000 hab
VERACRUZ	28 053 (67.3%)	37 560 (32 986-41 679)	90.1% (79.1-100%)	616	745 (668-862)	915	1669 (1293-2396)	41 679	51
OAXACA	13 793 (68.9%)	17 914 (15 749-20 017)	89.5% (78.7-100%)	269	339 (297-409)	434	821 (619-1234)	20 017	50
BAJA CALIFORNIA	17 197 (63.3%)	23 009 (19 719-27 157)	84.7% (72.6-100%)	94	139 (102-182)	1036	1804 (1373-2707)	27 157	82
QUINTANA ROO	10 236 (61.5%)	13 292 (11 708-163 28)	79.9% (70.4-98.2%)	223	255 (230-305)	256	397 (315-560)	16 634	111
HIDALGO	10 095 (60.6%)	13 208 (11 517-16 365)	79.3% (69.1-98.2%)	116	135 (118-163)	348	537 (426-752)	16 666	58
SINALOA	16 191 (57.7%)	22 092 (19 179-27 795)	78.7% (68.3-99%)	597	704 (643-804)	930	1626 (1271-2293)	28 076	95
SONORA	21 210 (58.2%)	28 618 (24 925-35 917)	78.5% (68.4-98.5%)	145	180 (161-221)	884	1597 (1234-2314)	36 460	128
CHIAPAS	6148 (61%)	7819 (6803-9887)	77.6% (67.5-98.2%)	198	229 (208-281)	277	526 (361-847)	10 073	19
GUERRERO	14 659 (59.7%)	19 038 (16 860-23 298)	77.5% (68.6-94.8%)	375	436 (394-500)	604	1030 (814-1449)	24 570	70
TABASCO	28 833 (52.1%)	41 329 (35 799-52 230)	74.7% (64.7-94.4%)	238	271 (252-310)	568	878 (730-1155)	55 312	231
QUERETARO	6623 (57.2%)	8524 (7362-10 865)	73.7% (63.6-93.9%)	91	112 (94-153)	67	92 (70-133)	11 571	57
NAYARIT	5003 (57.9%)	6260 (5564-7621)	72.4% (64.4-88.2%)	163	171 (163-190)	118	163 (132-215)	8645	73
COLIMA	3798 (58.3%)	4696 (4222-5728)	72.1% (64.8-88%)	82	92 (85-112)	106	147 (122-206)	6512	92
YUCATÁN	14 973 (54.1%)	19 776 (17 272-24 996)	71.5% (62.4-90.3%)	159	180 (162-208)	206	291 (246-396)	27 668	132
CAMPECHE	5743 (51.5%)	7850 (6704-10 048)	70.4% (60.1-90.1%)	224	249 (231-280)	112	218 (160-329)	11 152	124
PUEBLA	28 183 (49.8%)	39 848 (34 192-51 630)	70.4% (60.4-91.2%)	448	506 (473-583)	840	1277 (1062-1714)	56 600	92
CHIHUAHUA	7861 (50.4%)	10 456 (8638-14 257)	67.1% (55.4-91.5%)	603	744 (638-956)	456	637 (500-909)	15 586	44
ESTADO DE MÉXICO	49 647 (47.2%)	67 627 (58 513-87 646)	64.3% (55.6-83.3%)	687	809 (734-942)	2924	4348 (3595-6065)	105 239	65
SAN LUIS POTOSÍ	18 343 (47.4%)	24 616 (21 532-30 699)	63.7% (55.7-79.4%)	151	180 (159-223)	336	472 (403-596)	38 664	142
TAMAULIPAS	24 257 (44.2%)	34 096 (28 952-45 679)	62.1% (52.7-83.2%)	129	174 (149-247)	510	806 (649-1132)	54 902	160
MORELOS	5301 (50%)	6571 (5743-8308)	61.9% (54.1-78.3%)	41	43 (41-51)	421	920 (523-1714)	10 611	56
NUEVO LEÓN	29 830 (44.6%)	40 632 (35 328-52 169)	60.8% (52.8-78%)	671	796 (731-931)	160	240 (196-329)	66 869	131
COAHUILA	22 147 (44.2%)	30 176 (25 862-39 774)	60.3% (51.6-79.4%)	58	83 (69-115)	465	684 (570-950)	50 076	169
ZACATECAS	5282 (45.3%)	6778 (5976-8289)	58.1% (51.2-71.1%)	183	194 (184-223)	75	87 (77-118)	11 665	74
BAJA CALIFORNIA SUR	7639 (45.5%)	9654 (8583-11 933)	57.4% (51.1-71%)	70	77 (70-90)	27	28 (27-41)	16 807	236
CIUDAD DE MÉXICO	118 009 (39.6%)	168 089 (144 621-220 218)	56.5% (48.6-74%)	1438	1755 (1604-2077)	2116	2829 (2490-3569)	297 668	334
GUANAJUATO	31 736 (43.1%)	41 197 (36 692-50 156)	56% (49.9-68.2%)	261	289 (267-332)	465	600 (532-739)	73 588	126
MICHOACÁN	15 346 (40.6%)	21 136 (18 357-26 676)	55.9% (48.6-70.6%)	88	108 (93-143)	305	394 (351-492)	37 810	82
JALISCO	20 525 (40.4%)	27 838 (24 089-35 906)	54.8% (47.4-70.7%)	488	567 (526-662)	523	698 (612-895)	50 792	65
TLAXCALA	5571 (34.8%)	7434 (6396-9485)	46.5% (40.5-59.3%)	441	512 (466-599)	118	140 (123-180)	15 998	126
DURANGO	6414 (34.4%)	8566 (7378-11209)	45.9% (39.5-60.1%)	80	86 (80-94)	35	38 (35-44)	18 663	106
AGUASCALIENTES	5730 (33.4%)	7699 (6516-10028)	44.8% (37.9-58.4%)	28	35 (28-56)	48	63 (51-85)	17 180	131

196  
 197 **Note: State-wise per-occurrence date analysis, study period February 27<sup>th</sup> – August 31<sup>th</sup>. Ordered by estimated**  
 198 **positivity rate (highest to lowest)**

199

200 This means that with current testing capacity it is not possible to grasp the behavior of the pandemic, as the number of  
201 tests is so small and the positivity rate so high that it would be fully dependent on them. The World Health Organization  
202 recommends a positivity rate lower than 5%, among other criteria, to commence reopening, even if a sentinel system is  
203 being used. Mexico is currently far from a safe reopening, and further still if we consider false negatives.<sup>20, 21</sup>

204 False negatives are accounted for in clinical medicine when a clinician suspects it in a patient that has a negative test but  
205 other disease indicators, such as suggestive lung images, that generate a convincing clinical scenario, and acts  
206 accordingly.<sup>22</sup> As we show here, false negatives should also be accounted for in public health estimations, and it is also  
207 possible to act accordingly. Places with low prevalence, as in the states with low positivity rate (none in our case) or with  
208 massive testing strategies, will have a small amount of cases added to their official counts. This contrasts with the picture  
209 of Mexico as a whole, where the worrisome positivity rate increases even more when false negatives are considered.<sup>23, 24</sup>

210 Currently, a reopening strategy based on a four-coloured traffic light (red, orange, yellow, and green) is being  
211 implemented, which assigns each state a colour based on several variables. Test positivity rate is among the criteria for  
212 changing the colour. Nonetheless, by not considering false negatives the positivity rate is being underestimated and it is  
213 likely this could influence the premature modification of the colour, and thus premature reopening, which could cause a  
214 new surge in cases.<sup>25-27</sup> We observed a one-week delay in reaching a similar number of cases when mean false-negative  
215 test is not considered.

216 This approach as applied for Mexico has several limitations. As only a few symptomatic people are tested we cannot  
217 estimate the corrected number of cases among the total population. This would require a vastly higher number of tests.  
218 Thus, there should be caution when interpreting our results. Our estimation presents a limited view of how much more  
219 cases there are, and thus do not represent the actual number of cases in the country; in any case our corrected estimates  
220 are in the conservative side.<sup>14, 28, 29</sup> We assumed that the date on which the information was captured in the system is the  
221 same on which the sample was procured, but there might be a small-time lag between one and another. We also  
222 identified limitations related to the diagnostic tests. For instance, there is a wide catalog of SARS-Cov-2 RT-PCR tests  
223 currently used in our country. Even if all of them have to be approved by the Institute of Diagnosis and Epidemiological  
224 Reference (INDRE) for surveillance purposes and all must comply with the Berlin protocol, we do not know how they  
225 compare to each other, and variation in diagnostic accuracy probably exists. We also do not have information on the  
226 anatomical test site, be it nasal or oropharyngeal swab, saliva or bronchoscopy sample, which could affect test sensitivity.  
227 We do not account for severity of the clinical picture on the test sensitivity. Sensitivity of the test in patients who develop  
228 pneumonia and /or critical illness could be higher than in those with less severe disease, and patients with severe disease  
229 is overrepresented in the testing strategy followed in Mexico. Even when considering all these limitations, the application  
230 of the test performance to correct the number of cases could certainly improve surveillance. The method we used can be  
231 easily adapted to other countries or areas. Our analysis can be updated if COVID-19 open data continues to be published,  
232 and thus be used to better inform decision making at the National and State level.

233

### 234 **Conclusion**

235 While it may very well be impossible to determine which patients had the false negative tests, taking the test precision  
236 into account is an effective way to improve estimations for the number of COVID-19 cases among a tested population,  
237 especially when the testing is done mainly in high-prevalence populations. We expect this could aid to better inform public  
238 health measures and reopening policies.

239

### 240 **Acknowledgements**

241 We would like to thank all the health-workers in our country for their tremendous work in caring for patients during this  
242 pandemic. Also, we thank the people behind the surveillance of the pandemic, as without them this work could not have  
243 come to be.

244

245 **Declaration of interests**

246 All authors state they have no conflict of interests.

247

248 **Author contributions**

249 Isaac J. Núñez: idea conception, literature search, study design, data curation, data analysis, data interpretation, writing

250 Pablo F. Belaunzarán-Zamudio: literature search, study design, data analysis and interpretation, writing

251 Yanink Caro-Vega: literature search, study design, data curation, data analysis, data interpretation, writing

252 **Data sharing**

253 All code utilized is available at [https://github.com/isaac-nunez/covid\\_19\\_fn\\_estimates\\_mexico](https://github.com/isaac-nunez/covid_19_fn_estimates_mexico) and all datasets are  
254 available at the official government COVID-19 webpage.<sup>13</sup>

255 **References**

- 256 1. Dong E, Du H, Gardner L. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns  
257 Hopkins University. Available from: <https://coronavirus.jhu.edu/map.html> [Accessed 4<sup>th</sup> September 2020]
- 258 2. Subsecretaría de prevención y promoción a la salud. Technical daily report on COVID-19. Mexico City, August 31<sup>th</sup>,  
259 2020. Available from: [https://www.gob.mx/salud/documentos/informacion-internacional-y-nacional-sobre-nuevo-](https://www.gob.mx/salud/documentos/informacion-internacional-y-nacional-sobre-nuevo-coronavirus-2019-ncov)  
260 [coronavirus-2019-ncov](https://www.gob.mx/salud/documentos/informacion-internacional-y-nacional-sobre-nuevo-coronavirus-2019-ncov) [Accessed 4<sup>th</sup> September 2020]
- 261 3. WHO. Coronavirus disease (COVID-19) pandemic. Available from : [https://www.who.int/emergencies/diseases/novel-](https://www.who.int/emergencies/diseases/novel-coronavirus-2019)  
262 [coronavirus-2019](https://www.who.int/emergencies/diseases/novel-coronavirus-2019) [Accessed 24<sup>th</sup> July 2020]
- 263 4. National Committee for epidemiological surveillance (CONAVE). Legal Document DGE-DG-DVEENT 02595 2020.  
264 Mexico City, March 24<sup>th</sup>, 2020.
- 265 5. COVID-19: Operational guidance for maintaining essential health services during an outbreak. World Health  
266 Organization. March 25<sup>th</sup>, 2020.
- 267 6. Her M. How Is COVID-19 Affecting South Korea? What Is Our Current Strategy? Disaster Medicine and Public Health  
268 Preparedness. Cambridge University Press; 2020;1–3.
- 269 7. Gudbjartsson DF, Helgason A, Jonsson H, et al. Spread of SARS-CoV-2 in the Icelandic Population. *N Engl J Med*  
270 2020;382:2302-15.
- 271 8. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase  
272 Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. *Ann Intern Med*, 2020;M20-1495.
- 273 9. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 result. *BMJ*, 2020;369:m1808
- 274 10. Fang FC, Naccache SN, Greninger AL. The Laboratory Diagnosis of COVID-19--Frequently-Asked Questions. *Clinical*  
275 *Infectious Diseases*, ciaa742.
- 276 11. Wikramaratna, Robert S Paton, Mahan Ghafari, Jose Lourenco Paul. Estimating false-negative detection rate of SARS-  
277 CoV-2 by RT-PCR. medRxiv 2020.04.05.20053355.
- 278 12. SSA. Covid-19 Mexico. Información General. Disponible en <https://coronavirus.gob.mx/datos/>
- 279 13. COVID-19 open data, Dirección General de Epidemiología (DGE). Available from:  
280 <https://www.gob.mx/salud/documentos/datos-abiertos-152127> [Accessed 1<sup>st</sup> September 2020]
- 281 14. Listado de pruebas moleculares útiles para el diagnóstico de SARS-Cov-2 durante la contingencia de COVID-19 en  
282 México. Subsecretaría de Prevención y Promoción de la Salud, Dirección General de Epidemiología, Gobierno de México.  
283 Available from: [https://www.gob.mx/salud/documentos/evaluaciones-provisionales-en-apoyo-a-la-emergencia-por-la-](https://www.gob.mx/salud/documentos/evaluaciones-provisionales-en-apoyo-a-la-emergencia-por-la-pandemia-de-la-enfermedad-covid-19?state=published)  
284 [pandemia-de-la-enfermedad-covid-19?state=published](https://www.gob.mx/salud/documentos/evaluaciones-provisionales-en-apoyo-a-la-emergencia-por-la-pandemia-de-la-enfermedad-covid-19?state=published) [Accessed 4<sup>th</sup> September 2020]
- 285 15. Encuesta intercensal 2015. Instituto Nacional de Estadística y Geografía (INEGI). Available from:  
286 <https://www.inegi.org.mx/programas/intercensal/2015/> [Accessed 6<sup>th</sup> September 2020]
- 287 16. Pondaven-Letourmy S, Alfin F, Bpumghit Y, Simon F. How to perform a nasopharyngeal swab in adults and children in  
288 the COVID-19 era. *Eur Ann Otorhinolaryngol Head Neck Dis*, 2020;137(4):325-327.



- 289 17. Thwe PM, Ren P. How many are we missing with ID NOW COVID-19 assay using direct nasopharyngeal swabs?  
290 Findings from a mid-sized academic hospital clinical microbiology laboratory. *Diagn Microbiol Infect Dis*,  
291 2020;98(2):115123.
- 292 18. Miller IF, Becker AD, Grenfell BT, Metcalf CJ. Disease and healthcare burden of COVID-19 in the United States. *Nat*  
293 *Med*, 2020. DOI: <https://doi.org/10.1038/s41591-020-0952-y>.
- 294 19. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, et al. Prevalence of SARS-Cov-2 in Spain (ENE-COVID): a nationwide,  
295 population-based seroepidemiological study. *Lancet*, 2020. DOI: [https://doi.org/10.1016/S0140-6736\(20\)31483-5](https://doi.org/10.1016/S0140-6736(20)31483-5)
- 296 20. Prieto-Curiel R, Robles A. ¿Qué nos podrán indicar las pocas pruebas de covid que hagamos?[Internet]. Punto decimal.  
297 2020 [cited 24 July 2020]. Available from: [https://puntodecimal.mx/politica/que-nos-podran-indicar-las-pocas-pruebas-](https://puntodecimal.mx/politica/que-nos-podran-indicar-las-pocas-pruebas-de-covid-que-hagamos)  
298 [de-covid-que-hagamos](https://puntodecimal.mx/politica/que-nos-podran-indicar-las-pocas-pruebas-de-covid-que-hagamos)
- 299 21. Public health criteria to adjust public health and social measures in the context of COVID-19. World Health  
300 Organization. May 12<sup>th</sup>, 2020.
- 301 22. Li Y, Yao L, Li J, Chen L, Song Y, Cai Z, et al. Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients  
302 clinically diagnosed with COVID-19. *J Med Virol*. 2020;92(7):903-908.
- 303 23. Versión estenográfica de la conferencia de prensa matutina, Gobierno de México. Tuesday, May 12<sup>th</sup>, 2020. Available  
304 at: [https://www.gob.mx/presidencia/es/articulos/version-estenografica-de-la-conferencia-de-prensa-matutina-martes-](https://www.gob.mx/presidencia/es/articulos/version-estenografica-de-la-conferencia-de-prensa-matutina-martes-12-de-mayo-de-2020?idiom=es)  
305 [12-de-mayo-de-2020?idiom=es](https://www.gob.mx/presidencia/es/articulos/version-estenografica-de-la-conferencia-de-prensa-matutina-martes-12-de-mayo-de-2020?idiom=es)
- 306 24. Gonzalez H. Estimando el subregistro de defunciones por COVID-19 en México [Internet]. Punto decimal. 2020 [cited  
307 24 July 2020]. Available from: [https://puntodecimal.mx/ciencia/estimando-el-subregistro-de-defunciones-por-covid-19-](https://puntodecimal.mx/ciencia/estimando-el-subregistro-de-defunciones-por-covid-19-en-mexico)  
308 [en-mexico](https://puntodecimal.mx/ciencia/estimando-el-subregistro-de-defunciones-por-covid-19-en-mexico)
- 309 25. Datos Abiertos Ciudad de México, Capacidad hospitalaria ZMVM. Gobierno de la Ciudad de México. Available at:  
310 <https://datos.cdmx.gob.mx/explore/dataset/capacidad-hospitalaria/table/> [Accessed 24<sup>th</sup> July 2020]
- 311 26. Woloshin S, Patel N, Kesselheim AS. False Negative Tests for SARS-Cov-2 Infection – Challenges and Implications. *N*  
312 *Eng J Med*, 2020. DOI: 10.1056/NEJMp2015897
- 313 27. Semáforo de riesgo epidemiológico COVID-19: indicadores y metodología. Subsecretaría de Prevención y Promoción  
314 de la Salud, Gobierno de México. June 5<sup>th</sup>, 2020. Available from: [https://coronavirus.gob.mx/wp-](https://coronavirus.gob.mx/wp-content/uploads/2020/06/Lineamiento_Semaforo_COVID_05Jun2020_1600.pdf)  
315 [content/uploads/2020/06/Lineamiento\\_Semaforo\\_COVID\\_05Jun2020\\_1600.pdf](https://coronavirus.gob.mx/wp-content/uploads/2020/06/Lineamiento_Semaforo_COVID_05Jun2020_1600.pdf)
- 316 28. Corman VM, Landt O, Kaiser M, et al. Diagnostic detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-  
317 PCR. *Euro Surveill*, 2020;25(3):pii=2000045. DOI: <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>
- 318 29. Criterios para la aceptación de pruebas moleculares útiles para el diagnóstico de SARS-Cov-2 durante la contingencia  
319 de COVID-19 en México. Subsecretaría de Prevención y Promoción de la Salud, Dirección General de Epidemiología,  
320 Gobierno de México. Mexico City, April 9th, 2020. Available at: [https://www.gob.mx/salud/documentos/evaluaciones-](https://www.gob.mx/salud/documentos/evaluaciones-provisionales-en-apoyo-a-la-emergencia-por-la-pandemia-de-la-enfermedad-covid-19?state=published)  
321 [provisionales-en-apoyo-a-la-emergencia-por-la-pandemia-de-la-enfermedad-covid-19?state=published](https://www.gob.mx/salud/documentos/evaluaciones-provisionales-en-apoyo-a-la-emergencia-por-la-pandemia-de-la-enfermedad-covid-19?state=published)
- 322