

## **The role of data collection systems on COVID-19 deaths age structure: An illustration from Latin American countries**

**Jenny García\***, **Andrés Castro\*\***, **Catalina Torres\*\*\***, **Stéphane Helleringer\*\*\*\***

\* Institute national d'études démographiques, France

\*\* Max Planck Institute for Demographic Research. Germany

\*\*\* Muséum National d'histoire naturelle, France

\*\*\*\* New York University, Abu Dhabi

### **Abstract**

The lack of timely monitoring systems of the COVID-19 lethality has been an unrelenting need in Latin America since May 2020, when the region became the pandemic's epicenter. To alleviate the pitfalls of their national vital statistics systems, particularly those related to the timely production of COVID-19 reports, Latin American countries have used different collection systems for counting COVID-19 deaths, primarily based on hospital records and surveillance systems. Little attention has been paid to the potential age bias of death counts across these reports. We examine the age distribution of COVID-19 death rates using health systems' counts and vital statistics from the year 2020 from Chile, Mexico, Ecuador, and Brazil. Next, we compare the age distribution of COVID-19 deaths to all-cause and the excess mortality in these countries. Our results indicate that incompleteness of official reports likely concentrates at older ages regardless of the data collection system. Age-reporting bias is highly related with usual mortality age patterns by place of death in each country.

## **The role of data collection systems on COVID-19 deaths age structure: An illustration from Latin American countries**

### **Introduction**

The lack of timely monitoring systems of the COVID-19 lethality has been an unrelenting need in Latin America since May 2020, when the region became the pandemic's epicenter. To alleviate the pitfalls of their national vital statistics systems, particularly those related to the timely production of COVID-19 reports, Latin American countries have used different collection systems for COVID-19 deaths, including hospital records, surveillance systems, and civil registries. Several problems regarding the completeness of COVID-19 reported death counts had been pointed out.<sup>1</sup> However, little attention has been paid to the potential age bias of death counts across these reports, which can be very consequential for understanding COVID-19 lethality among the elders.

In Mexico and Ecuador, deaths at younger ages are more likely to occur in health care facilities, and deaths at older ages are more likely to occur at home. Between 2018 and 2019, 46% of all deaths in Ecuador occurred at home; this proportion rose to 54% in the population aged 65+.<sup>2</sup> Similarly, in Mexico in 2018, 45% of all deaths occurred at home, while the corresponding proportion at ages 65+ rose to 55%.<sup>3</sup> On the contrary, in Brazil, deaths occurring outside the health system for the year 2018 represent around 33% of the total deaths, with the same proportion at age 65+. The number of hospital beds per capita, as an indicator of the capacity of national health systems, may elucidate the difference between these countries. In Brazil, the number of hospital beds per 1,000 inhabitants is around 2.1, whereas it is 1.4 in Ecuador, and 1 in Mexico.<sup>4</sup> These discrepancies are consistent with the differences mentioned above in the proportions of deaths within each country's health system.

The cumulative number of COVID-19 deaths up to the beginning of May 2021, in most Latin American countries, primarily refers to deaths occurred in hospitals, as the official COVID-19 death statistics in most countries of the region rely on institutions attached to the health system (e.g. epidemiological surveillance system and hospital records). Within those cumulative counts, the high proportion of hospital deaths is linked to the criteria implemented by the countries to attribute a death to COVID-19 in the official reports. Most deaths have been attributed to COVID-19 as the result of a positive laboratory test. Therefore, COVID-19 reported deaths are highly dependable on the testing strategies implemented in each country, as well as on the scope of national health systems. Patients who died in hospitals are usually tested, but deaths at private homes, care homes, and other institutions are being qualified in various ways and very likely to be missing from current official reports. In Mexico, for example, 91.33% of the official COVID-19 deaths (219,323 reported deaths on May 12, 2021) correspond to hospitalized cases;<sup>5</sup> whereas in Brazil, the percentage rises to 97.05% (407,815 out of the 420,180 total reported death occurred in hospitals).<sup>6</sup> Deaths occurred outside the health system will be only included in the vital registration system and with some delay. We hypothesize that, because COVID-19 mortality estimations are also based (almost exclusively) on counts reported by national health systems, these figures are biased towards younger ages, thus underestimating COVID-19 mortality among the oldest age groups of the population.

### **Data and Methods**

To illustrate this point, we examine the age distribution of COVID-19 death rates using counts from the health systems as well as civil registration and vital statistics data from Mexico,<sup>3</sup> Ecuador,<sup>2</sup> and Brazil.<sup>5</sup> In a second step, we compare the age distribution of COVID-19 deaths to all-cause and excess mortality in these countries. The

idea is to look for similarities in the mortality age patterns shown by COVID-19 and the excess mortality to the usual hospital deaths' age structure. Comparisons are made using the proportional distribution of the standardized age specific COVID-19 mortality rates proposed by Goldstein and Lee.<sup>7</sup> With this approach, we represent the relative proportional burden of COVID-19 mortality in each age group.

Comparisons between expected all-cause mortality and actual all-cause mortality are commonly used to estimate the excess mortality due to COVID-19. Expected all-cause mortality represents what would have occurred under 'normal' conditions, without COVID-19. Observed all-cause mortality can be decomposed into three additive components: the expected all-cause mortality, plus COVID-19 deaths and the COVID-19 collateral deaths. The term "collateral deaths" refers to changes in previous mortality patterns driven by COVID-19 itself, such as the effect of overwhelmed health care systems and the contingency measures taken by the countries. For example, countries with a large proportion of deaths due to external causes and experiencing strict lockdown will reduce the observed mortality compared to expected mortality. The excess mortality reflects the total burden of the pandemic that may go uncaptured due to the strict attribution to COVID-19 criteria implemented by the countries and other pandemic-related deaths.

### **Preliminary Results**

By comparing official COVID-19 deaths from the health system in Mexico and Ecuador with Brazilian civil registration, Figure 1a (in appendix) shows that the age distribution of COVID-19 deaths is substantially younger in Mexico than in Ecuador and Brazil, which may be due to undercounting deaths of old adults. Indeed, a recent Mexican report based on civil register data showed that the number of COVID-19 deaths in the country is almost 45% higher than indicated by their surveillance system.<sup>8</sup> Another element to consider is that Mexico's data includes only deaths of people with a positive laboratory test (confirmed cases/deaths herein). In contrast, Ecuador's data includes both confirmed and probable cases. Figure 1b (in appendix) shows the proportional distribution of confirmed and probable COVID-19 deaths in Ecuador. According to this figure, the confirmed cases' age structure also underestimates mortality at older ages. The evidence shown here suggests that the distribution of COVID-19 mortality is younger than expected due, in part, to age differentials in the collection system's reporting. This bias has passed unnoticed in comparative studies of COVID-19 to date.

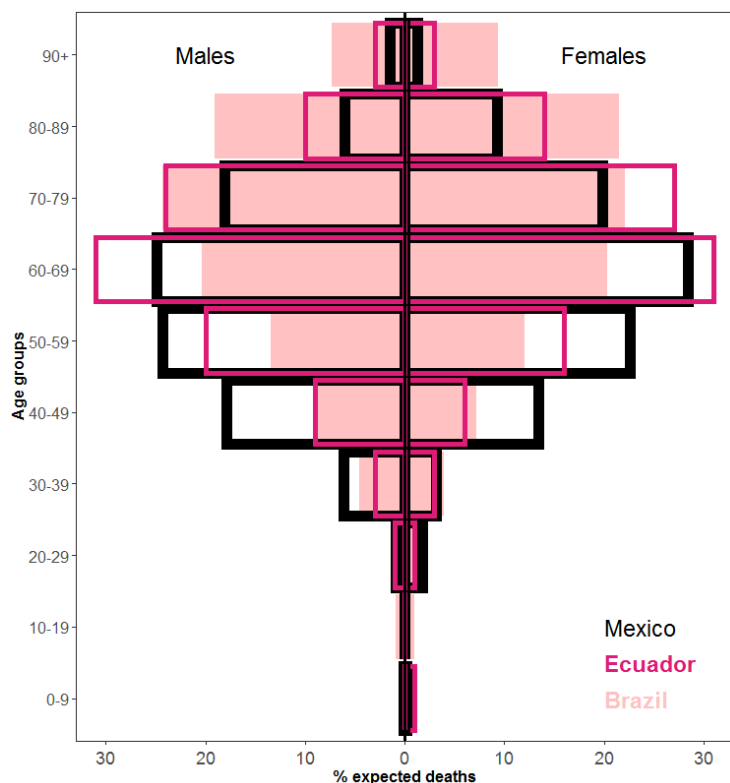
Most national reports in Latin America refer to deaths of confirmed COVID-19 cases coming solely from the health systems. Consequently, they display a younger than expected profile when contrasted to studies in other regions of the world, especially because of the high proportion of deaths at older ages occurring at home in Latin American countries, even under 'normal' circumstances.<sup>9</sup> Accurate comparisons of COVID-19 age-specific mortality across countries need to account for the potential bias across data sources. By reporting this age bias, we do not mean to diminish the severity of the COVID-19 among young adults. Our goal is to call attention to a long-standing, yet unnoticed, issue that severely limits scholars' capacity to understand the pandemic and the governments' and policy makers' ability to respond to it.

***A more comprehensive and updated analysis will be included in the final version of the paper.***

### **References**

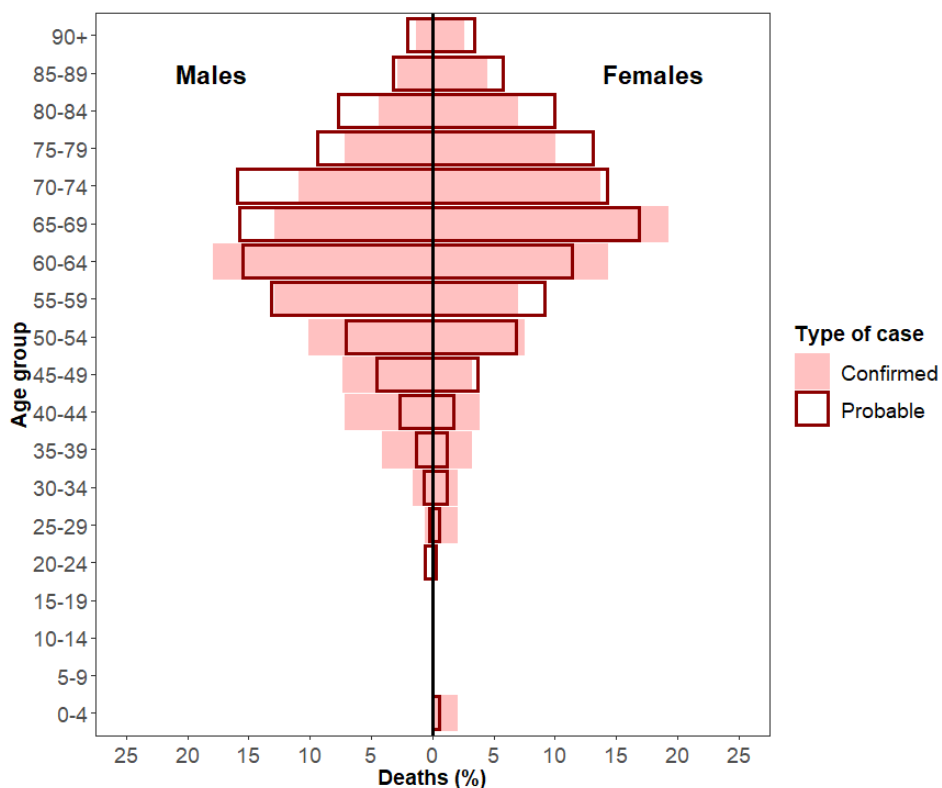
<sup>1</sup> García J., Torres C., Barbieri M., Cambois E., Camarda, C.G., Caporali A, Meslé F., Poniakina S., and Robine J.M. (2021). Differences in COVID-19 Mortality: the implications of imperfect and diverse data collection systems, *Population 2021/0* (Prepublication) , p. 35-71. <https://www.cairn-int.info/journal-population-2021-0-page-35.htm>

- <sup>2</sup> Instituto Nacional de Estadísticas y Censos (INDEC). Estadísticas de Nacimientos y defunciones Database 2018-2019. [https://www.ecuadorencifras.gob.ec/nacimientos\\_y\\_defunciones/](https://www.ecuadorencifras.gob.ec/nacimientos_y_defunciones/) [Consulted on 1 June 2020]
- <sup>3</sup> Instituto Nacional de Estadística, Geografía e Informática (INEGI), Estadísticas de defunciones. Database 2018. <https://www.inegi.org.mx/programas/mortalidad/default.html#Microdatos> [Consulted on 1 June 2020]
- <sup>4</sup> Pan American Health Organization. Health Information Platform for the Americas PLISA. Database, updated December 3, 2018. Accessed November 14, 2020: <https://www.paho.org/data/index.php/es/indicadores/visualizacion.html>
- <sup>5</sup> Dirección General de Epidemiología, Secretaría de Salud, DGE-SS (2020). *Datos Abiertos-Dirección General de Epidemiología*, accessed on 12/15/2021 <https://datos.covid-19.conacyt.mx/>
- <sup>6</sup> Cartórios de Registro Civil do Brasil. Central de Informações do Registro Civil - CRC Nacional, Database 2019. <https://transparencia.registrocivil.org.br/especial-covid> [Consulted on 29 July 2020]
- <sup>7</sup> J. R. Goldstein, R. D. Lee, Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 22035–22041 (2020).
- <sup>8</sup> Instituto nacional de estadística y geografía (INEGI). “Características de las defunciones registradas en México durante enero a agosto de 2020”, press release No. 61/21, January 27 2021. Media coverage in Aristegui noticias. “Lo que sabe hoy de las muertes por COVID-19, con datos del Inegi”, January 27 2021. <https://aristeginoticias.com/2701/mexico/lo-que-sabe-hoy-de-las-muertes-por-covid-19-con-datos-del-inegi/>
- <sup>9</sup> O’Driscoll, M. et al. Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* <https://doi.org/10.1038/s41586-020-2918-0> (2020). <https://doi.org/10.1038/s41586-020-2918-0>Nature



**Figure 1a. Proportion of expected deaths associated with COVID-19 by age group in Brazil and Mexico (accumulated counts up to April 20th, 2020) and Ecuador (accumulated counts up to April 23rd, 2020) by sex.**

*Note:* Brazil: Brazilian Civil Registry Information Center. “Deaths with suspicion or confirmation of COVID-19 by sex and age group” (deaths occurrence up to April 20th, 2020, updated on 11/15/2020); Mexico: Secretaría de Salud de México-“Open data - General directorate of epidemiology” (deaths occurrence up to April 20th, 2020, updated on 07/28/2020). Ecuador: are from the country’s Boletines epimiológicas del Ministerio de Salud Pública (deaths occurrence up to April 23rd, 2020, updated on 04/23/2020-only date by single age, sex and probable and confirmed distinction published in Ecuador up to March 28th, 2021). The expected deaths are calculated with direct standardization, using the standard population proposed by the World Health Organization ([www.who.org](http://www.who.org)).



**Figure 1b. Distribution of COVID-19 deaths by age, sex, and definition in Ecuador (deaths cumulated until April 23rd, 2020)**

*Note:* Data for Ecuador are from the country’s Boletines epimiológicas del Ministerio de Salud Pública. (deaths occurrence up to April 23rd, 2020, updated on 04/23/2020-only date by single age, sex, and probable and confirmed distinction published in Ecuador up to March 28th, 2021). “Confirmed” deaths refer to Laboratory-confirmed cases using PCR and Rapid Tests. “Probable” deaths refer to cases of people with symptoms and people who had contact with an infected person but without laboratory confirmation.