

# Estimating national excess mortality from subnational data: application to Argentina

Ariel Karlinsky<sup>1</sup>

**Suggested citation** Karlinsky A. Estimating excess mortality from subnational data: application to Argentina. Rev Panam Salud Publica. 2022;46:e19. <https://doi.org/10.26633/RPSP.2022.19>

**ABSTRACT** This paper presents a method to estimate excess mortality where national data are missing for some or all of the coronavirus disease 2019 (COVID-19) pandemic period, but subnational data exist, such as in Argentina. By making use of the stability of the regional distribution of deaths, data on deaths in Córdoba province were used to project excess deaths in Argentina from March 2020 up to the end of 2021. The number of excess deaths was estimated at 134 504, which is 14.8% higher than the reported number of COVID-19 deaths in Argentina for the same time period.

**Keywords** Excess mortality; COVID-19; Argentina.

Excess mortality, that is, the number of deaths more than expected, is widely considered the gold standard in quantifying the effects of a pandemics such as coronavirus disease 2019 (COVID-19) (1). Unfortunately, many countries have yet to report 2021 or even 2020 mortality data, thus inhibiting reliable estimates of excess mortality (2). Most of these data gaps are in Africa and Asia. In Latin America, some countries have published both 2020 and 2021 data such as Bolivia, Chile, Ecuador, Mexico, and Peru. However, some Latin American countries have not published any appropriate data such as Guyana, Honduras, and Venezuela, or have not yet published data for 2021 onwards such as Argentina, Cuba, and El Salvador. To overcome this lack of national data, a projection from subnational data to national estimates can be done. This projection makes use of the stability of the spatial distribution of deaths within countries and is appropriate where the spread and toll of COVID-19 are similar between the subnational and national data used.

## METHODS

Consider country  $c$ , which can be broken down to an  $L$  number of local units, with each unit represented by  $l$ . The number

of deaths  $D$  in the country, for every length of time, is thus a simple sum over the number of deaths in each local unit:

$$D_c = \sum_{l=1}^L D_l$$

The share of deaths in the country represented by the local-level deaths is thus  $D_l/D_c$  and it follows that for every time period:

$$D_c = D_l \times \left( \frac{D_l}{D_c} \right)^{-1}$$

In order to use  $D_l$  to project to  $D_c$  we require an estimate for the current  $D_l/D_c$ , which is unknown since we do not have  $D_c$ . Therefore,  $\hat{D}_l/D_c$  (the estimate for  $D_l/D_c$ ) is obtained by using the historical ratio. To make things concrete, we introduce another subscript for time such that our estimate for the ratio between local unit deaths and total national deaths in 2021, i.e.  $D_{l,2021}/D_{c,2021}$  might be  $D_{l,2019}/D_{c,2019}$  or  $D_{l,2020}/D_{c,2020}$ , if data exist for 2020. In simple terms, this could be from before the pandemic or from a period during the pandemic where information on both local mortality and nationwide mortality is available. We can establish bounds on this estimate by using standard

<sup>1</sup> Hebrew University, Jerusalem, Israel ✉ Ariel Karlinsky, ariel.karlinsky@mail.huji.ac.il

first-order Taylor expansion (delta method) for the variance of the ratio estimator:

$$\widehat{Var}\left(\frac{\widehat{D}_l}{\widehat{D}_c}\right) = \frac{\widehat{Var}(\widehat{D}_l) - 2\left(\frac{\widehat{D}_l}{\widehat{D}_c}\right)\widehat{Cov}(\widehat{D}_l, \widehat{D}_c) + \left(\frac{\widehat{D}_l}{\widehat{D}_c}\right)^2 \widehat{Var}(\widehat{D}_c)}{\widehat{D}_c^2}$$

Where  $\widehat{D}_l$  and  $\widehat{D}_c$  are the estimates for the total number of deaths at the local unit level and country level, respectively.

This method is appropriate given two important conditions: first, the pandemic’s spatial distribution is similar in the local unit and at the national level; and second, the share of deaths is stable, meaning movement between local units (migration) is low, so that state- or province-level local units are more suitable than village, town or city level local units, where the share of deaths is likely to be less stable.

This method was applied to Argentina and the province of Córdoba, where these conditions are satisfied. Argentina is a federation of 24 provinces, which is the local unit. The province of Córdoba is the second largest province in terms of population size. *La Voz del Interior* is a newspaper in Córdoba which published monthly number of deaths from all causes in the province from January 2019 to December 2021 (3). These data were available a long time before Argentina had released its 2020 all-cause mortality counts and are much more recent, as the latest national level data available only go up to December

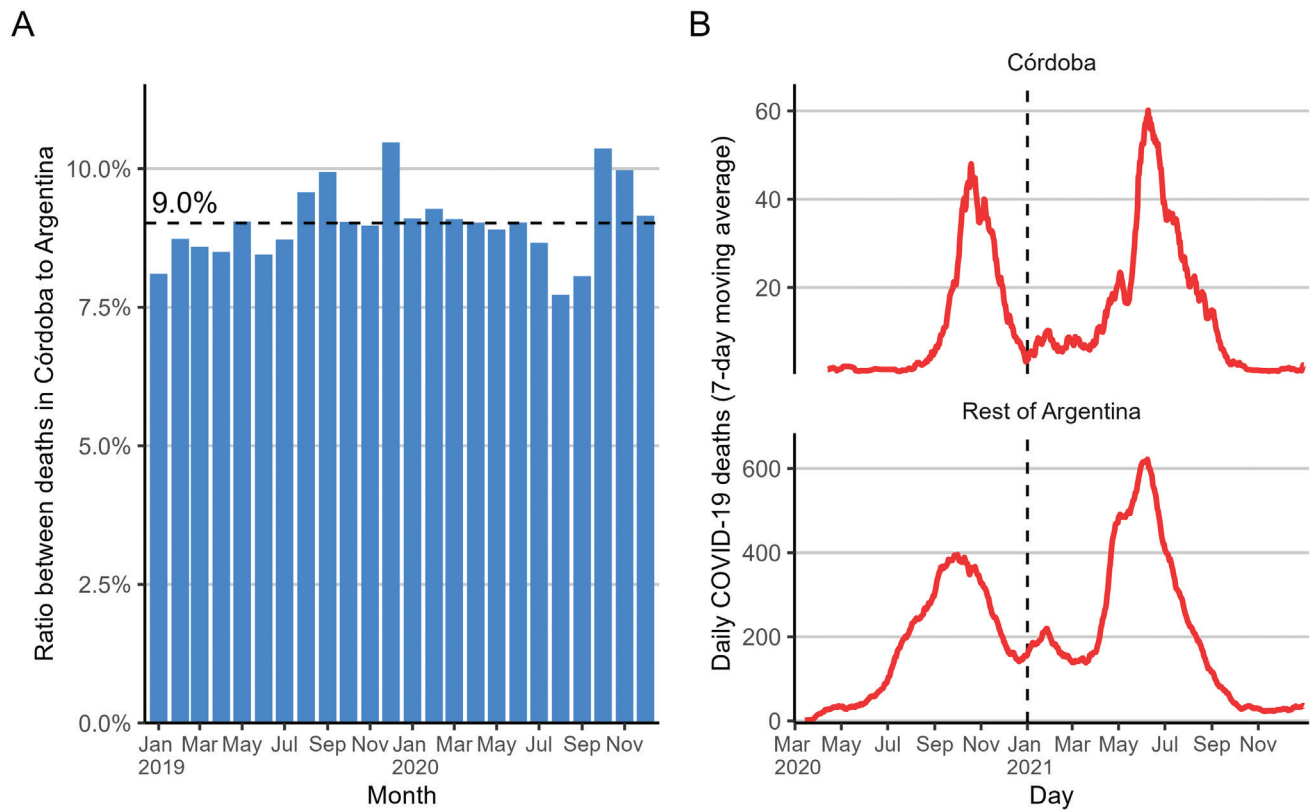
2020 (4). Additional data on the annual number of deaths by province and daily number of COVID-19 deaths by province were obtained from the open-data portal of the Argentinian Ministry of Health (5).

**RESULTS**

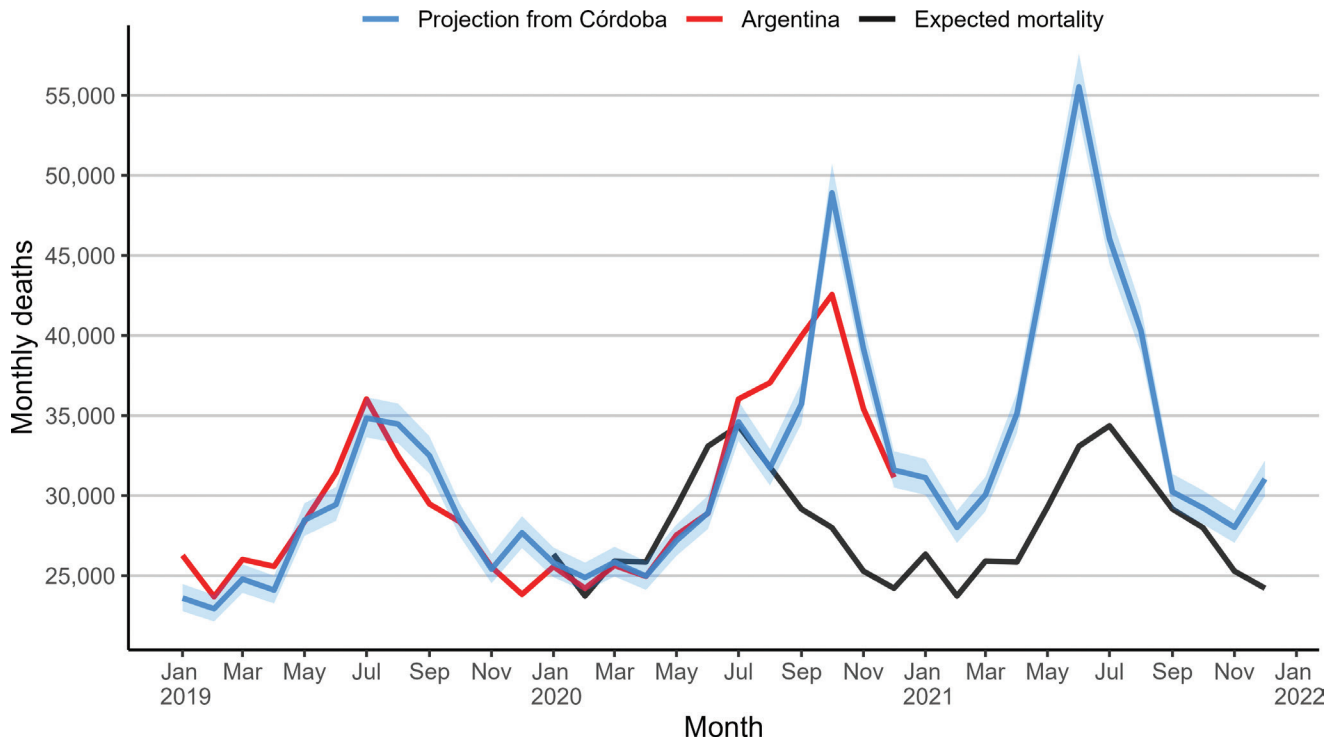
From 2005 to 2019, the share of annual all-cause deaths in Argentina accounted for by Córdoba was stable, ranging from 8.6% to 8.8%. In 2019, the share of registered monthly number of deaths in Córdoba to that of all of Argentina ranged from 8.1% (January) to 10.5% (December), with the annual share being 9.0%. In 2020, this monthly share ranged from 7.7% (August) to 10.3% (October) with the annual share also being 9.0% (Figure 1, Panel A). The spatial distribution of COVID-19 deaths is very similar between Córdoba and the rest of Argentina, with the peaks and troughs almost perfectly aligned temporally (Figure 1, panel B). Thus, the projection of national mortality from Córdoba to the rest of Argentina during the COVID-19 pandemic satisfies both conditions mentioned earlier and is likely to hold.

Estimating  $\widehat{D}_l/D_c$  using the monthly data on all-cause mortality nationally and in Córdoba (Figure 1, panel A), the share was estimated at 9.0% (95% confidence interval (CI) 8.7% to 9.3%). Multiplying the reported monthly number of deaths in Córdoba by the reciprocal of this estimated share, the projections for

**FIGURE 1. Relation between deaths in Córdoba and in all Argentina for all-cause mortality (panel A) and COVID-19 mortality (panel B)**



Note: Dashed horizontal line is the overall ratio (panel A); dashed vertical line is the start of 2021 (panel B). Source: Prepared by authors from the results.

**FIGURE 2. Monthly deaths in Argentina – observed, expected and projected from Córdoba**

**Note:** Pearson  $r = 0.91$  between projected deaths from Córdoba and observed deaths for all Argentina in 2019-2020.  
**Source:** Prepared by authors from the results.

the monthly number of deaths in all of Argentina were closely aligned with the registered national number of deaths and the projection in 2019 and most of 2020, as shown in Figure 2. There were significant differences in August and September 2020 with the total national mortality being higher, but this is then compensated for by the projected mortality being higher in October and November, such that the total annual mortality is very similar: 378 995 nationally registered deaths and 379 386 deaths as projected from Córdoba. In addition, the method used in this study estimated that the peak monthly mortality in Argentina occurred around June 2021, with about 55 500 projected deaths, which is 22 400 deaths more than expected. This corresponds with the reported COVID-19 trajectory, where deaths peaked during the same period (Figure 1, panel B).

The total number of excess deaths in Argentina from March 2020 to December 2021 (as computed by the world mortality method (2)) was 134 504 (95% CI 108 202 to 162 766), with COVID-19 deaths for the same period at 117 111, as reported to the World Health Organization (6), representing an undercount ratio of 1.15, i.e. 14.8% higher than reported.

## DISCUSSION

The main result of this study is that the number of excess deaths in Argentina, from March 2020 to the end of December 2021, as estimated by a projection using actual all-cause mortality data from the province of Córdoba, was 134 504. The number of COVID-19 deaths in Argentina reported to the World Health Organization for the same period was 117 111, representing an undercount ratio of 1.15, i.e. 14.8% higher than reported. This undercount ratio places Argentina among other Latin American

countries that have generally low undercount ratios, indicating few potentially missing COVID-19 deaths. Argentina's estimated undercount ratio is similar to Brazil (1.11), Paraguay (1.15), and Peru (1.07), and higher than Chile (0.99) and Panama (1.01). It is much lower than the undercount ratio observed in Bolivia (2.51), Ecuador (2.03), and Mexico (1.99), where proper certification of COVID-19 has been an issue due to limited testing (7).

The excess deaths estimate for Argentina in the present study is comparable to other currently available estimates from the Institute for Health Metrics and Evaluation (IHME) and The Economist (8,9), which are 125 694 and 154 403, respectively. Both the IHME and The Economist models rely only on national all-cause mortality data up to the end of 2020 in Argentina and the relation between observed all-cause mortality in other countries, and COVID-19 and socioeconomic variables, which are then used to project to countries without available all-cause mortality data.

The main limitation of the method used in this study is the requirement for a similar trajectory of COVID-19 deaths between the relevant subnational unit and the national level. Indeed, in many countries, COVID-19 has had a very different spatial spread between different regions, such that the method shown here would not be well suited to project national excess deaths. Prominent examples include the United States of America, where both reported COVID-19 and excess deaths during the first wave of the pandemic (March–April 2020) were limited to the eastern states such as New York and New Jersey. In Peru, excess deaths in the region of Lima started well before the region of Tacna (10). In Ecuador, spread began at the regional coastlines before it reached inland (7).

The method shown in this paper was used to derive an up-to-date estimate of excess mortality in Argentina, overcoming the absence of all-cause mortality data in 2021. It may be used in other settings, within Latin America and beyond, to derive similar estimates if national-level data are delayed or even do not exist for the COVID-19 period. In addition, if civil registration and vital statistics systems in some regions are disrupted or delayed, national data aggregated from regional data will be delayed as well. Using information from well functioning regional civil registration and vital statistics systems with the method shown here provides an up-to-date estimate of excess mortality. When future national-level all-cause mortality data in Argentina for 2021 are published, they may be used to validate the proposed method. Furthermore, as information on COVID-19 from Córdoba for 2022 is also expected to be available earlier than the national information for 2022, this method can continue projecting and tracking excess mortality in Argentina. In several countries, sub-national all-cause mortality data exist where national data have yet to be published or may never be published due to low coverage of national vital statistics registration. The method shown here may also be applied to countries such as India (11),

Indonesia (12), Syrian Arab Republic (13), Turkey (14), Yemen (15), and others.

**Acknowledgements.** I thank members and observers of the World Health Organization's Technical Advisory Group on COVID-19 Mortality Assessment, Dmitry Kobak, Ilya Kashnitsky and the reviewers for helpful comments and feedback, Ary Garbovetzky for his help in obtaining some of the data, and Gal Kabiri for proofreading.

**Conflict of interests.** None declared.

**Funding.** No funding was received.

**Disclaimer.** The author holds sole responsibility for the views expressed in the manuscript, which may not necessarily reflect the opinion or policy of the *Revista Panamericana de Salud Pública / Pan American Journal of Public Health* and/or those of the Pan American Health Organization.

**Availability of data.** All data and codes are available at: [https://github.com/akarlinsky/cordoba\\_argentina\\_pro](https://github.com/akarlinsky/cordoba_argentina_pro)

## REFERENCES

1. Beaney T, Clarke JM, Jain V, Golestaneh AK, Lyons G, Salman D, et al. Excess mortality: the gold standard in measuring the impact of COVID-19 worldwide? *J R Soc Med.* 2020;113(9):329–34. Available from: <https://doi.org/10.1177%2F0141076820956802>
2. Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. *Elife.* 2021;10:e69336. Available from: <https://doi.org/10.7554/eLife.69336>
3. Garbovetzky A. Impacto poblacional de la pandemia: crecieron las muertes 26% y cayeron los nacimientos 14% [Population impact of the pandemic: deaths grew 26% and births fell 14%]. *La-Voz;* 2022 [cited 2022 Jan 22]. Available from: <https://www.lavoz.com.ar/ciudadanos/impacto-poblacional-de-la-pandemia-crecieron-las-muertes-26-y-cayeron-los-nacimientos-14/>
4. Rearte A, Moisés MS, Rueda D V, Laurora MA, Flamenco MA, Pennini VA, et al. Exceso de mortalidad por todas las causas en el contexto de la pandemia de COVID-19 en Argentina, 2020 [All-cause excess mortality during the COVID-19 pandemic in Argentina, 2020]. *Rev Argent Salud Pública.* 2021; 13(Suppl):e36.
5. Datos Abiertos del Ministerio de Salud [Open data from the Ministry of Health] [internet]. Ministry of Health of Argentina; 2022 [cited 2022 Jan 22]. Available from: <http://datos.salud.gob.ar/>
6. WHO coronavirus (COVID-19) dashboard [internet]. World Health Organization; 2020 [cited 2022 Jan 22]. Available from: <https://covid19.who.int/>
7. Lima EEC, Vilela EA, Peralta A, Rocha M, Queiroz BL, Gonzaga MR, et al. Investigating regional excess mortality during 2020 COVID-19 pandemic in selected Latin American countries. *Genus.* 2021;77(1):30. Available from: <https://doi.org/10.1186/s41118-021-00139-1>
8. Estimation of excess mortality due to COVID-19. Institute for Health Metrics and Evaluation; 2021 [cited 2022 Jan 22]. Available from: <https://covid19.healthdata.org/argentina?view=cumulative-deaths&tab=trend>
9. The pandemic's true death toll. *The Economist.* 2022 [cited 2022 Jan 22]. Available from: <https://www.economist.com/graphic-detail/coronavirus-excess-deaths-estimates>
10. Sempé L, Lloyd-Sherlock P, Martínez R, Ebrahim S, McKee M, Acosta E. Estimation of all-cause excess mortality by age-specific mortality patterns for countries with incomplete vital statistics: a population-based study of the case of Peru during the first wave of the COVID-19 pandemic. *Lancet Reg Health Am.* 2021;2:100039. Available from: <https://doi.org/10.1016/j.lana.2021.100039>
11. Leffler CT, Lykins V. JD, Yang E. Preliminary analysis of excess mortality in India during the COVID-19 pandemic (update September 26, 2021). *medRxiv.* 2021. Available from: <https://doi.org/10.1101/2021.08.04.21261604>
12. Djaafara BA, Whittaker C, Watson OJ, Verity R, Brazeau NF, Widyaastuti, et al. Using syndromic measures of mortality to capture the dynamics of COVID-19 in Java, Indonesia, in the context of vaccination rollout. *BMC Med.* 2021;19(1):146. <https://doi.org/10.1186/s12916-021-02016-2>
13. Watson OJ, Alhaffar M, Mehchy Z, Whittaker C, Akil Z, Brazeau NF, et al. Leveraging community mortality indicators to infer COVID-19 mortality and transmission dynamics in Damascus, Syria. *Nat Commun.* 2021;12(1):1–10. Available from: <https://doi.org/10.1038/s41467-021-22474-9>
14. Yaman G. Excess mortality in Turkey [internet]. 2021 [cited 2022 Jan 22] Available from: <https://github.com/gucluyaman/Excess-mortality-in-Turkey>
15. Besson ESK, Norris A, Ghouth AS Bin, Freemantle T, Alhaffar M, Vazquez Y, et al. Excess mortality during the COVID-19 pandemic: a geospatial and statistical analysis in Aden governorate, Yemen. *BMJ Glob Health.* 2021;6(3):e004564. Available from: <https://gh.bmj.com/content/6/3/e004564>

---

Manuscript received on 18 October 2021. Revised version accepted for publication on 27 January 2022.

---

## Cálculo de la tasa nacional de exceso de mortalidad con datos subnacionales: el caso de Argentina

### RESUMEN

En este estudio se presenta un método para calcular el exceso de mortalidad cuando no se dispone de datos nacionales de una parte o la totalidad del período de la pandemia de la enfermedad por el coronavirus del 2019 (COVID-19), pero sí se cuenta con datos subnacionales, como es el caso en Argentina. Aprovechando la estabilidad de la distribución regional de las muertes, se emplearon los datos sobre las muertes en la provincia de Córdoba para hacer una proyección del exceso de mortalidad en Argentina desde marzo del 2020 hasta finales del 2021. Se estimó en 134 504 el número de muertes en exceso, que es 14,8% mayor que la cifra notificada de muertes por COVID-19 en Argentina en el mismo período.

**Palabras clave** Mortalidad; COVID-19; Argentina.

---

---

## Estimativa do excesso de mortalidade nacional a partir de dados subnacionais: o caso da Argentina

### RESUMO

Este artigo apresenta um método para estimar o excesso de mortalidade durante a pandemia da doença causada pelo coronavírus 2019 (COVID-19) na Argentina quando faltam dados nacionais relativos a um dado período – ou ao período inteiro –, mas existem dados subnacionais disponíveis. Diante da estabilidade da distribuição regional das mortes, foram usados os dados de mortalidade da Província de Córdoba para fazer uma projeção do excesso de mortes ocorridas no país entre março de 2020 e o final de 2021. Estimou-se um excesso de 134 504 mortes, o que corresponde a 14,8% a mais do número notificado de mortes por COVID-19 na Argentina no mesmo período.

**Palavras-chave** Mortalidade; COVID-19; Argentina.

---