**Cómo la transparencia de datos puede ayudar a combatir COVID-19**

*Para ayudar a tomar decisiones mejor informadas sobre el coronavirus, los gobiernos deberían publicar conjuntos de datos completos y legibles por máquina.*

Por [Marc Joffe](https://reason.org/author/marc-joffe/)

7 de abril de 2020

En las semanas transcurridas desde que COVID-19 se convirtió en una pandemia, [gobiernos](https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html) y entidades [no gubernamentales](http://coronavirusapi.com/) han lanzado una gran cantidad de paneles de datos. Desafortunadamente, los datos que proporcionan a menudo son incompletos e inconsistentes, lo que hace que los responsables de la formulación de políticas y el público reciban una cantidad insuficiente de información procesable. Para responder a las preguntas sobre qué medidas de salud pública son necesarias y qué medidas de protección personal son prudentes, debemos tener datos completos y estandarizados.

Las fuentes de datos más utilizadas provienen de los Centros para el Control y la Prevención de Enfermedades (CDC), junto con la visualización producida por [la Universidad Johns Hopkins](https://coronavirus.jhu.edu/map.html) , [Worldometer](https://www.worldometers.info/coronavirus/country/us/) y muchos gobiernos locales y estatales. Estas fuentes generalmente carecen de información sobre hospitalizaciones, así como datos demográficos sobre muertes y casos graves.

Una práctica común para estos sitios es publicar solo recuentos brutos de pruebas positivas y muertes atribuidas al virus. Los resultados de las pruebas no son comparables entre jurisdicciones porque los criterios para ser evaluados varían. Dado que sabemos que muchas personas que están infectadas [experimentan solo síntomas menores](https://www.sciencealert.com/here-s-what-we-know-so-far-about-those-who-can-pass-corona-without-symptoms) (si los hay), es probable que el número de casos positivos se [subestime](https://news.harvard.edu/gazette/story/2020/03/hundreds-of-u-s-coronavirus-cases-may-have-slipped-through-screenings/) en gran medida incluso en lugares que realizan pruebas de forma agresiva. Para que el número de casos positivos sea realmente significativo, la mayoría de la población en un área determinada tendría que hacerse la prueba de forma regular. Las fallas masivas en las pruebas del gobierno, desde el comienzo de esta pandemia, continúan perjudicando la capacidad de obtener una imagen completa y completa de los totales de las pruebas.

Las muertes proporcionan un indicador más confiable del impacto del coronavirus en las comunidades, pero esta estadística también tiene algunas limitaciones importantes. Primero, es un indicador rezagado: un resultado final de una enfermedad que es menos útil para predecir la trayectoria a corto plazo de la pandemia. En segundo lugar, puede haber diferencias entre jurisdicciones e incluso entre médicos forenses acerca de cómo atribuir una muerte determinada. Si bien la presencia de COVID-19 en una persona fallecida se puede determinar de manera confiable, si el virus causó la muerte es una decisión. [Se ha argumentado](https://www.reddit.com/r/Classical_Liberals/comments/fox2ll/the_threat_from_the_new_coronavirus_is_being/) , por ejemplo, que Italia tiene una tasa de mortalidad reportada más alta que los países vecinos porque no hace una distinción entre las muertes *por* COVID y las muertes *por* COVID.

Otros han sugerido que el recuento de muertes no se informa [. *El New York Times* informó que](https://www.nytimes.com/2020/04/05/us/coronavirus-deaths-undercount.html) “los funcionarios del hospital, los médicos, los expertos en salud pública y los examinadores médicos dicen que los recuentos oficiales no han logrado capturar el número real de estadounidenses que mueren en esta pandemia. El recuento insuficiente es el resultado de protocolos inconsistentes, recursos limitados y un mosaico de toma de decisiones de un estado o condado a otro”.

Una mejor medida del impacto de COVID-19 en este momento son las hospitalizaciones. Esta cifra excluye los casos asintomáticos y leves que pueden tener un impacto social menor y proporciona un indicador más en tiempo real que las muertes. Es cierto que esta medida también es vulnerable al problema de clasificación que también se aplica a las muertes.

Hablando de clasificación, solo algunas jurisdicciones desglosan los totales de COVID-19 informados por grupo de edad, sexo y presencia de comorbilidades. Se sabe que estos factores afectan la forma en que un individuo determinado experimenta el virus y, por lo tanto, estas descomposiciones son información útil tanto para el público como para los responsables de la formulación de políticas.

La ciudad de Nueva York [proporciona más datos](https://www1.nyc.gov/site/doh/covid/covid-19-data.page) que la mayoría de las jurisdicciones, lo cual es una suerte dada la gravedad de su situación. Informa sobre hospitalizaciones y también proporciona desgloses por edad y sexo. Recientemente, el tablero de la ciudad mostró una tasa de hospitalización del 0.17 por ciento para los hombres y del 0.11 por ciento para las mujeres, una diferencia significativa que se ha observado en otros lugares. Las tasas de hospitalización oscilaron entre el 0,01 por ciento para los menores de 18 años y el 0,5 por ciento para las personas de 75 años o más.

Si bien Nueva York no proporciona datos de comorbilidad para las hospitalizaciones, lo hace para las muertes. Un informe reciente [mostró](https://www1.nyc.gov/assets/doh/downloads/pdf/imm/covid-19-daily-data-summary-deaths-04042020-1.pdf) que se encontró que más del 97 por ciento de las muertes que habían sido evaluadas para la existencia de condiciones subyacentes las tenían presentes, pero el 29 por ciento de todas las muertes aún no se había evaluado. Las condiciones subyacentes incluían "diabetes, enfermedad pulmonar, cáncer, inmunodeficiencia, enfermedad cardíaca, hipertensión, asma, enfermedad renal y enfermedad gastrointestinal / hepática".

La ampliación de los informes de la ciudad de Nueva York a nivel nacional mejoraría enormemente el nivel de información procesable que tenemos. Para escalar de manera efectiva, todas las jurisdicciones deben informar sus datos en un formato estandarizado legible por máquina y enviar estos archivos de datos a un único repositorio público. Una vez más, la ciudad de Nueva York ha dado un paso importante en esta dirección al producir archivos diarios de valores separados por comas (CSV) y publicarlos en [Github](https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data) , un popular repositorio de software que también se utiliza para publicar datos públicos. En el momento de escribir este artículo, los datos de la condición subyacente solo estaban disponibles en el sitio web de la ciudad en formato PDF. Si se superaran esas limitaciones, los estados, condados y ciudades de todo el país podrían utilizar los informes de datos públicos de la ciudad de Nueva York como modelo.

COVID-19 representa una seria amenaza para la salud pública, la economía y, en última instancia, gran parte de nuestra civilización. Para atacar el problema, necesitamos tener los mejores datos posibles. Es comprensible que los hospitales, ciudades y estados abrumados no vean este como el momento ideal para cambiar la forma en que recopilan, procesan y comparten datos. Sin embargo, establecer e implementar un estándar de datos en este momento podría ayudar significativamente a los funcionarios de salud a identificar las tendencias y las mejores prácticas del coronavirus de manera más rápida y efectiva. En última instancia, el costo de un estándar de datos es mínimo en comparación con los muchos otros gastos que realizan los gobiernos en todos los niveles, pero su valor podría ser enorme.