

LOS PRONÓSTICOS MATEMÁTICOS SON HIPÓTESIS: Consideración de la incertidumbre en la presentación de datos sobre la pandemia por COVID-19

Ruben Aroca Jácome, PhD

Facultad de Jurisprudencia, Ciencias Sociales y Políticas
Carrera de Sociología - Universidad de Guayaquil

*In God we trust.
All others must bring data.
Deming, W.*

Introducción

El presente artículo es motivado por dos propósitos iniciales: a) por la necesidad de entender la naturaleza y las limitaciones de las estimaciones matemáticas y estadísticas, en especial cuando estas se vuelven un objeto noticiable (como en el caso de la información sobre el crecimiento de contagios y muertes por COVID-19); y, b) por la necesidad de diseñar ciertas indicaciones básicas para facilitar la comunicación de información sensible acerca de la situación referida a la pandemia.

Algunos antecedentes

El 16 de marzo de 2020, al inicio del Estado de Excepción para combatir la propagación del COVID-19, el presidente del Ecuador indicó que se podría llegar en el país a "800.000 contagiados" cita

que fue repetida, y no examinada, por todos los medios.

El 26 de marzo de 2020, un diario local (Diario Expreso, 2020) publicó las estimaciones de contagio y requerimiento de servicios hospitalarios realizadas por una empresa española, estudio que sostenía que en Madrid pudiera haber 1,5 millones de contagiados para ese entonces. La empresa que realizó las proyecciones no contaba con especialización en epidemiología o afines, y en España, a pesar de lo difícil de escenario actual, no se ha observado niveles de contagio de esas proporciones (Bort, 2020).

Los casos anotados no solo permiten observar que los pronósticos producidos por modelos matemáticos pueden contemplar ciertos niveles de error, sino que, en ocasiones, su uso es simplemente fallido y contraproducente.

En períodos críticos, la información debe ser exhaustivamente examinada pues con frecuencia los especialistas no son lo suficientemente claros en sus parámetros o supuestos de partida y, en lo posterior, estas omisiones adoptan otras dimensiones en los medios de comunicación, pues los periodistas suelen asumir frente a información sensible las mismas actitudes que con relación a la información producida en "periodos de normalidad". Citar declaraciones de tal o cual funcionario o científico, sin mediar un adecuado examen de datos, no constituye (al menos, no en estos casos) un papel destacable en la construcción de opinión fundamentada sino absolutamente todo lo contrario.

En este marco, se considera necesario la formación de una comprensión adecuada de la información, de sus potencialidades y limitaciones y, claro está, de su relación con la realidad. Para ello, resulta útil formular algunas recomendaciones que tanto especialistas como periodistas puedan tener en cuenta a la hora de comunicar resultados o publicar noticias en un tiempo tan crítico como el de la pandemia por COVID-19.

Estas recomendaciones se basan especialmente en: Goodson (2020), Ordway (2020), Red Etica (2020) y Paullus (2000).

Los componentes de la información sobre COVID-19

A mediados de febrero de 2020, en medio de purgas internas en la cúpula provincial del partido por sospechas de un manejo inadecuado de la crisis, en la provincia Hubei se modificó el criterio de detección de nuevos casos de infección por COVID-19 mediante la inserción de una nueva categoría denominada "casos clínicos". En resumidos términos, esto implicó contabilizar como casos no sólo aquellos en los que practicaba el test para la detección del virus, sino también a todos aquellos que presentaban los síntomas característicos (Feng). El resultado de dicha decisión hizo que Hubei pase de un reporte de 1638 casos positivos el 11 de febrero a otro con 14840 casos positivos al siguiente día (véase: gráfico 1).

El efecto en la representación estadística (gráfico 1) puede describirse como un "disparo" del tamaño de la población contagiada que al publicarse tal cual, por los servicios de noticias, implicó la producción de una mayor inquietud en los públicos interesados. No obstante, lo que había variado era solamente una definición conceptual respecto a qué considerar un "caso real" de contagio, y *no la realidad del contagio en sí*.

En una buena parte de los países de occidente, se considera como caso diagnosticado como positivo a todo caso confirmado por laboratorio, independiente de los signos clínicos o síntomas (Ministerio de Sanidad - Gobierno de España).

Esta es la situación de países como España, Colombia o Ecuador. Sin embargo, los datos podrían modificarse también si consideramos, tal como lo hizo la provincia de Hubei en su momento, las siguientes definiciones (WHO):

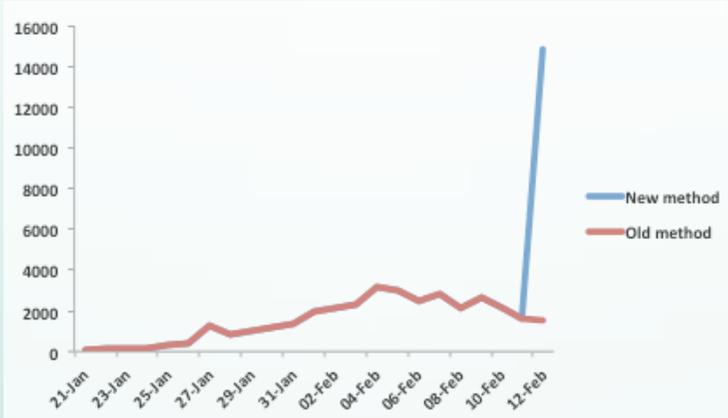


Gráfico N° 1: Hubei: Reporte diario de casos confirmados positivos por COVID 19. Fuente: Feng, 2020.

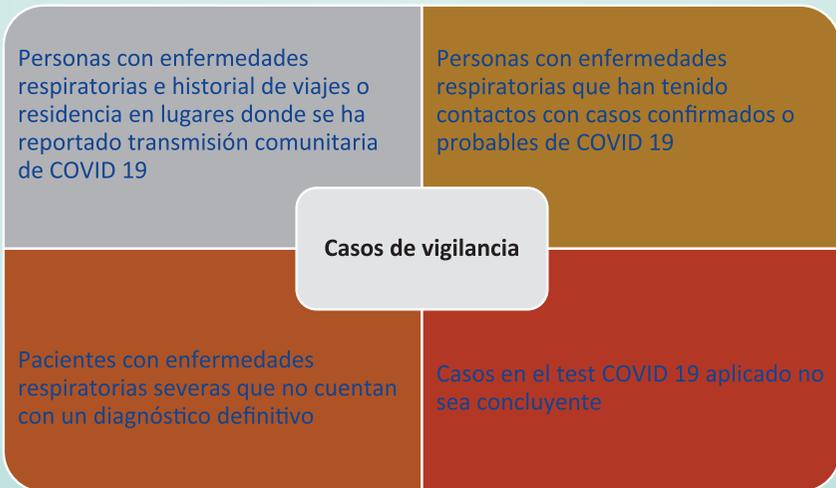


Ilustración N° 1: Definiciones de casos para vigilancia. Guidance on global surveillance for COVID-19, WHO, 27 February 2020.

Cualquiera de las cuatro tipologías descritas (véase: ilustración 1) implican, por un lado, el desarrollo de procesos de observación y registro que probablemente no siempre se ajustan a las recomendaciones de la OMS; y, en segundo lugar, la transformación de la información que representa el crecimiento diario o periódico de los casos de COVID-19. Realizar estas estimaciones toma tiempo e integran, claro está, un cierto nivel de incertidumbre.

La Organización Mundial de la Salud ha solicitado a los organismos competentes del área de salud en cada país, la generación de un reporte confirmatorio de casos de COVID-19 dentro de las 48 horas siguientes a su detección. El reporte requerido se compone de un total de 82 campos o variables que requieren de un seguimiento detallado de cada paciente y de personal adiestrado para el registro de la información.

La producción de datos sobre el crecimiento del contagio y su letalidad resulta, por tanto, un problema aparte. Es comprensible que, entre los tiempos de aplicación de test, la obtención de los resultados, su validación y registro de información sumen varios días o semanas hasta su publicación. En el caso de Gran Bretaña, se estima que este proceso puede tomar

hasta 4 (Doyle) días mientras que en Colombia puede tomar entre dos y tres semanas (Lopez).

En conclusión, en primer lugar, los reportes diarios que emiten los países no representan el crecimiento que el contagio adquiere de un día al siguiente, ni la forma "natural" en que la epidemia se comporta. Representa, eso sí, el esfuerzo de cada país tanto por otorgar respuestas ante el fenómeno como por registrar dicho proceso. Y es justo aquí donde incertidumbre y estadísticas cobran una mayor importancia.

En segundo lugar, toda proyección estadística del número de casos infectados por COVID-19 tiene sesgos que dependen de:

- a) Número de test aplicados en cada país;
- b) Tiempo transcurrido entre la detección de un caso y su ingreso al sistema de información nacional correspondiente;
- c) Tiempo transcurrido entre la situación descrita en *b* y su comunicación a sus distintos usuarios.

Presentando información crítica: algunas recomendaciones

Las recomendaciones que se incluyen a continuación se formulan en consideración de la necesidad de

entender la naturaleza y las limitaciones de las estimaciones matemáticas y estadísticas, en la perspectiva de facilitar la comunicación de información sensible acerca de la pandemia por COVID-19. Estas reglas, que pretenden encontrar utilidad cuando se trate de comunicar datos o publicar noticias en un tiempo tan crítico como el presente, consideran que sus usuarios potenciales son tanto los investigadores especialistas como los informadores de medios periodísticos.

Primera recomendación: sin indicaciones sobre su procedencia, es imposible conocer la validez estadística. Corrija, por tanto, la abstracción desmesurada de las representaciones estadísticas.

Desconocemos el contexto de casi todas las estadísticas que leemos u oímos comentar, y esto implica no solo a quienes se supone que las estadísticas representan sino también a quienes las construyen, a sus antecedentes y procedimientos.

Sin una ambientación, sin un trasfondo y sin indicaciones sobre su procedencia es imposible conocer la validez estadística. Por otro

lado, en literatura o en narrativa, el foco de atención se concentra casi siempre en los individuos y no en el análisis, las argumentaciones y las medidas: es decir, casi siempre es concreto y rara vez abstracto. Un enfoque así es necesario para corregir la abstracción desmesurada y mantener la estadística dentro de la perspectiva humana.

Segunda recomendación: los modelos funcionan, en cierto modo, a la manera de hipótesis. Por tanto, científicos y periodistas deben expresar el nivel de incertidumbre asociado al pronóstico.

Todo pronóstico matemático se formula en calidad de hipótesis, la misma que posee un componente de incertidumbre asociado a la proyección o pronóstico. Una hipótesis es una afirmación con cierta probabilidad de ocurrir, pero lo que hace que una hipótesis sea considerada como tal es que su probabilidad de ocurrencia es indeterminada. La indeterminación o incertidumbre puede formularse de varios modos, ya sea como rango de error o como la forma en que deben ser consideradas las afirmaciones del pronóstico. Tomemos en consideración el siguiente ejemplo:

Si consideramos que la tasa de mortalidad por COVID 19 detectada por la OMS oscila entre el 2% y 3%, de este modo, teniendo en cuenta el número de muertes por COVID 19 reportados por España y Ecuador al 05/04/2020, entonces:

Assumptions	España		Assumptions	Ecuador	
	2% death rate	3% death rate		2% death rate	3% death rate
Total Cases	632,050	421,367	Total Cases	9,000	6,000

Ilustración N° 2: Consideración de la tasa de mortalidad como supuesto para estimar nivel de contagio de la población

La estimación anterior (véase: ilustración 2) calcula el nivel de contagio por COVID-19 no a través de la progresión de casos diarios, sino a través de la tasa de defunción. Esto parte de *considerar como supuesto* hallazgos de ciertos estudios (Zunyou, 2020) que sitúan la mortalidad por COVID-19 ente el 2% y 3% de los casos de contagio. Este sería la base del pronóstico y, como tal, puede resultar muy cuestionable.

Tercera recomendación: escoja cuidadosamente los expertos a los que consulta.

Todos tenemos o experimentamos la ansiedad por acceder a información adecuada. En las redacciones, esta necesidad puede sentirse como apremio, presión por la competencia, etc. Esto puede producir que periodistas urgidos por el cierre de la edición pregunten a profesionales no competentes dejándose guiar por criterios poco pertinentes y obtengan, por este vía, opiniones inexactas o inválidas. Esto es particularmente crítico cuando se trata de temas de salud pública.

Antes de seleccionar a un experto, es necesario revisar su experiencia o lo que ha escrito respecto al tema objeto de análisis o su vinculación a un área muy relacionada a dicho objeto. Una vez que haya lo

grado dar este paso, recuerde: un experto expone sus pronósticos o los resultados de su trabajo, pero los mismos deben ser sometidos a observación de otros.

Cuarta recomendación: los periodistas deben consultar y citar otros expertos antes de publicar un pronóstico.

Con referencia a lo anterior, los resultados producidos por la aplicación de un modelo deben ser sometidos a evaluación de especialistas, pues es de este modo en que se puede dejar en claro la actitud que se debe asumir frente a dichos resultados. Debe evitarse relacionar estos productos con explicaciones basadas en opiniones o con nociones del sentido común.

Quinta recomendación: los científicos e investigadores deben describir con claridad los parámetros críticos y supuestos que utilizan en sus modelos o pronósticos. De lo contrario, usted tendrá que hacer que los pongan en claro.

¿Cuáles son las condiciones que deben cumplirse para que los pronósticos obtenidos por la aplicación de un modelo puedan considerarse como probables? Si la tasa de contagio por COVID-19 está en función tanto de la capacidad de

transmisión del virus en sí, como de la efectividad de las políticas de aislamiento y capacidad hospitalaria, es necesario comprender el rango en que “juegan” estos supuestos.

Es muy comentado, por estos tiempos, el caso del tratamiento que el diario Financial Times hizo de la presentación del modelo producido a partir del paper conocido como “Estudio Oxford”. Este modelo, desarrollado por el Imperial College de Londres, describe una forma de actuación implementada en la mayor parte de las naciones de occidente para “aplanar la curva” (Ferguson, N. et al., 2020), que consiste en imponer medidas de alejamiento social más extremas cada vez que los ingresos en las unidades de cuidados intensivos (UCI) empiezan a aumentar, y suavizarlas al reducirse la cantidad de las personas ingresadas. Parte del funcionamiento

del modelo matemático implica considerar escenarios de alto contagio y bajo contagio de la población. El diario Financial Times (véase: ilustración 3) tomó dichos parámetros como realidades potenciales e informó que la mitad de la población del Reino Unido se encontraba “ya contagiada” (Cookson, 2020).

En el caso anterior lo que correspondía, en primer lugar, era una revisión exhaustiva de los supuestos y componentes del modelo e informar en qué consistían y estimar qué tan lejos o distante se encontraba la población de dicho escenario. En segundo lugar, habría sido de gran utilidad relacionar estos resultados con la aplicación de otros modelos matemáticos sobre el problema. En este caso, las analogías con simulaciones realizadas en otros países, como China, podrían haber resultado útiles (Trilla, 2020).

The screenshot shows the Financial Times website interface. At the top, there is a navigation bar with the FT logo and a search icon. Below the navigation bar, there is a main headline: "Coronavirus may have infected half of UK population – Oxford study". To the right of the headline, there is a graphic showing the FT website on multiple devices (laptop, tablet, smartphone). Below the headline, there is a sub-headline: "Make informed decisions with the FT". Below the sub-headline, there is a brief description of the FT's global reporting: "Keep abreast of significant corporate, financial and political developments around the world. Stay informed and spot emerging risks and opportunities with independent global reporting, expert commentary and analysis you can trust."

Ilustración N° 3: Sección de la nota del diario FT.

Véase: <https://www.ft.com/content/5ff6469a-6dd8-11ea-89df-41bea055720b>

Sexta recomendación: ser lo más transparente posible.

Si se va a hablar de los resultados de la aplicación de un modelo o una proyección estadística, todo el modelo o proyección, así como sus hojas de datos, estructura de cálculo o códigos, deben ser liberadas para observación (véase:

ilustración 4). Esto implica una exigencia doble: por un lado, que el especialista esté dispuesto a poner a disposición sus datos para que otros puedan examinarlos y, por otro, que los periodistas tengan la oportunidad de también recoger los criterios derivados de evaluaciones de pares.

Master CoVidActNow CoVid-19 Model													
Model Inputs													
1	Model Inputs												
2	Last Updated Date	3/27/20											
3	Estimated Initial R0	2.4											
4	Hospitalization Rate	7.3%											
5	Case Fatality Rate	1.1%											
6	Fatality Rate Increase If Hospitals Overloaded		1.0%										
7	Population	731,545											
8	Hospital Beds	1,669	0.22%										
9	Hospital Bed Utilization	60%											
10	Emergency Capacity Build	207.9%											
11													
12													
13	Notes	Date	Effective R0	Beginning Susceptible	New Infected	Current Infected	Recovered or Died	Ending Susceptible (Right Axis)	Actual Reported Hospitalized? (Current Trends)	Predicted Hospitalized (Current Trends)	Cumulative Infected (Right Axis)	Cumulative Deaths	Available Hospital Beds
14													
15													
16													
17													
18		3/3/20		731,545	1	0	0	731,544	0.1				
19		3/7/20	2.4	731,544	2	1	0	731,541	0.2	0	2	0	547
20		3/15/20	2.4	731,541	8	4	0	731,538	0.3	0	8	0	547
21		3/15/20	2.4	731,538	14	9	0	731,521	0.8	1	23	0	547
22		3/19/20	3.0	731,521	43	23	1	731,478	2.3	3	65	1	676
23		3/23/20	0.4	731,478	19	63	4	731,459	1.0	1	85	1	869

Ilustración N° 4: Modelo Master CoVidActNow CoVid-19 Model, totalmente liberado. Véase: <https://docs.google.com/spreadsheets/d/1YEj4Vr6lG1jQ1R3LG6fri-jJYNynKcgTjzo2n0FsBwZA/edit#gid=1579455912>

Séptima recomendación: los tomadores de decisión deben utilizar múltiples modelos para fundamentar el diseño de una política.

Como se indicó en el apartado 5, al hablar de las medidas que toman para enfrentar la pandemia por COVID-19, varios gobiernos argumentan desde el modelo desarrollado por el Imperial Colle-

ge (Ferguson, N. et al., 2020). Sin embargo, se conoce bien de la existencia de otros modelos (CoVidActNow, 2020). Es pertinente, por tanto, preguntar en qué medida se ha procedido a comparar los modelos existentes, sus resultados y su congruencia.

Tomar decisiones sobre la consideración de un solo modelo de pronóstico no sólo que no resulta

científica y técnicamente razonable, sino que, en este caso, puede hacer la diferencia en la tasa de contagio y defunciones (véase: Ilustración 3).

Octava recomendación: indicar cuando el modelo o pronóstico es producido por una fuente sin formación en enfermedades infecciosas.

Un centro de investigaciones especializado en modelos generados por inteligencia artificial o redes neuronales de análisis puede propiciar nuevos recursos para procesar la información, incluso si los resultados de sus estudios no resultan plausibles. Por ello, si se está frente a un modelo diseñado por un organismo o universidad sin experiencia en epidemiología o enfermedades infecciosas, es necesario tener una doble actitud. Por un lado, poner los resultados de ese ejercicio bajo escrutinio de los epidemiólogos de modo que sea adecuadamente valorado, pero asegurando, por otro lado, brindar oportunidad suficiente a otros ángulos de análisis científico que podrían resultar de importancia.

Conclusiones

El presente artículo partió de haber identificado la necesidad de: 1) entender la naturaleza y las limitaciones de las estimaciones

matemáticas y estadísticas; y, 2) diseñar, en consecuencia, ciertas reglas básicas para facilitar la comunicación de información sensible acerca de la situación referida a la pandemia por COVID-19.

Se considera que sin indicaciones sobre la forma de producción de datos y procedencia, los usuarios de la información (gestores de salud, periodistas, políticos y públicos) no podrán determinar: a) el nivel de validez asociado a la estimación, cuestión que se debe tener a cuenta en la perspectiva de corregir la abstracción desmesurada y mantener la estadística dentro de la perspectiva humana; b) el componente de incertidumbre metodológicamente asociado al pronóstico matemático, el mismo que generalmente se formula como rango de error o como en relación a determinadas condicionalidades.

Esto suele no considerarse adecuadamente, generalmente por omisiones expositivas de tipo metodológico o por el apremio por el cierre de la edición. Esto es particularmente crítico cuando se trata de temas de salud pública. Por ello, debe comprenderse que, una vez que los expertos han expuesto sus pronósticos o los resultados de su trabajo, estos deben ser sometidos a observación de otros, en la perspectiva de dejar en claro la

actitud que se debe asumir frente a dichos resultados. Como condición adicional, se recomienda que todo el modelo o proyección, así como sus hojas de datos, estructura de cálculo o códigos, deben ser liberadas para observación.

La aplicación de las medidas enunciadas produce la posibilidad de comparar varios modelos existentes, sus resultados y congruencia, en la perspectiva de evitar la toma de decisiones sesgadas por los *outputs* de un solo modelo de pronóstico. El impacto de esta actitud

puede hacer la diferencia en los ritmos de la tasa de contagio y defunciones.

Finalmente, se sostiene que las reglas enunciadas no deben inhibir al investigador o informador en preservar su apertura frente a fuentes emergentes de producción de datos, suponiendo, por un lado, que los resultados de ese ejercicio deben situarse bajo escrutinio especializado, pero asegurando, por otro lado, oportunidad suficiente a una pluralidad de ángulos de análisis.

Referencias

- Bort, C. (22 de Marzo de 2020). *Coronavirus: algoritmos para dimensionar los recursos sanitarios. Estamos a tiempo*. Obtenido de Medium: <https://medium.com/@carlosbort/coronavirus-algoritmos-para-dimensionar-los-recursos-sanitarios-estamos-a-tiempo-98e4df7f59b8>
- Cookson, C. (24 de Marh de 2020). *Coronavirus may have infected half of UK population — Oxford study*. Obtenido de Financial Times: <https://www.ft.com/content/5-ff6469a-6dd8-11ea-89df-41bea055720b>
- CoVidActNow. (27 de March de 2020). *COVID-19 Data Model*. Obtenido de Master CoVidActNow CoVid-19 Mode: <https://docs.google.com/spreadsheets/d/1YEj4-Vr6lG1jQ1R3LG6frijJYNynKcgTjzo2n0FsBwZA/edit#gid=1579455912>
- Diario Expreso. (26 de Marzo de 2020). *Actualidad*. Obtenido de Expreso.ec: <https://www.expreso.ec/actualidad/coronavirus-ecuador-carlos-bort-estadisticas-contagios-7860.html>
- Doyle, L. (Apr 7, 2020). *Coronavirus test: How long does it take to get your test back?* *Express*, <https://www.express.co.uk/life-style/health/1253887/coronavirus-test-how-long-to-get-coronavirus-test-back>.
- Feng, E. (February 13, 2020). *A Change In How 1 Chinese Province Reports Coronavirus Adds Thousands Of Cases*. *NPR*, <https://www.npr.org/sections/goatsandsocka/2020/02/13/805519117/a-change-in-how-one-chinese-province-reports-coronavirus-adds-thousands-of-cases>.
- Ferguson, N. et al. (2020). *Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*. *Imperial College COVID-19 Response Team*.

- Goodson, M. (2020). All models are wrong. *Royal Statistical Society*, RSS Data Science Section.
- Lopez, C. (30 de Marzo de 2020). Entrevista a Claudia López, Alcaldesa Mayor de Bogotá. (N. Yamid, Entrevistador)
- Ministerio de Sanidad - Gobierno de España. (11 de abril de 2020). *Procedimiento de actuación frente a casos de Infección por el nuevo Coronavirus (SARS - CoV-2)*. Madrid: DIRECCIÓN GENERAL DE SALUD PÚBLICA, CALIDAD E INNOVACIÓN.
- Ordway, D. (2020). Covering COVID-19 and the coronavirus: 5 tips from a Harvard epidemiology professor. *Journalist's Resource*.
- Paullus, J. (2000). *Erase una vez un número*. Madrid: Tusquets Editores S.A.
- Red Etica. (2020). Coronavirus: 10 virtudes del periodismo responsable ante la pandemia. *Fundación Red Etica*.
- Trilla, A. (01 de Febrero de 2020). *Un mundo, una salud: la epidemia por el nuevo coronavirus COVID-19*. Obtenido de Elsevier.es: <https://www.elsevier.es/es-revista-medicina-clinica-2-avance-resumen-un-mundo-una-salud-epidemia-S002577532030141X>
- WHO. (27 February 2020). Global surveillance for COVID-19 caused by human infection with COVID-19 virus . *World Health Organization*, [https://www.who.int/publications-detail/global-surveillance-for-human-infection-with-novel-coronavirus-\(2019-ncov\)](https://www.who.int/publications-detail/global-surveillance-for-human-infection-with-novel-coronavirus-(2019-ncov)).
- Zunyou, W. e. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China. *JAMA*.