



Ejes

# Estrategias de análisis de información ante la pandemia de 2020

HÉCTOR BENÍTEZ-PÉREZ\*,  
FABIÁN GARCÍA-NOCKETTI\*,  
ELENA LARRAGA\*,  
WILLIAM LEE\*,  
MANUEL SUÁREZ-LASTRA\*,  
JORGE VELASCO\*,  
JESÚS M. SIQUEIROS-GARCÍA\*

## EL CONTEXTO

Queremos empezar este texto agradeciendo a la revista *Ciencia UANL* el espacio y la invitación para presentar esta reseña de varias acciones en torno a la pandemia, bajo la perspectiva de la información sobre COVID-19 en un amplio espectro. En abril del presente año, al interior del Subsistema de la Investigación Científica de la UNAM, en la Coordinación de la Investigación Científica, se definieron diversas acciones en apoyo a la sociedad en el contexto de la contingencia sanitaria por la que estamos pasando; particularmente en lo relativo a estrategias de insumos, equipamientos o suministros, en el sector biológico y de la salud, en el terreno del estudio ambiental y por supuesto en lo referente al análisis de la información sobre el comportamiento del virus, contagios y enfermedad. Es en este último donde se concentrará este documento.

\* Universidad Nacional Autónoma de México.  
Contacto: hector.benitez@iimas.unam.mx



Como bien es sabido, a partir de la detección del primer caso de COVID-19 se han generado datos tanto de los infectados, de los susceptibles, de los internados, de los recuperados y de los fallecidos. Con dicha información diversos sectores de la sociedad han construido representaciones que les permiten, como organizaciones o individuos, dar un sentido de certeza a la situación que se está viviendo desde varios niveles: desde el nacional hasta el nivel local. El objetivo ha sido buscar respuestas acerca del impacto de la pandemia derivada de la enfermedad por coronavirus iniciada en 2019 (COVID-19) en distintos ámbitos: personal, familiar, de recreación, escolar, laboral, comunitario, municipal, estatal y por supuesto entender el efecto que tendrá en el orden mundial.

Diversos grupos académicos, con un profundo sentido civil, y con distintas formaciones académicas, han construido pequeñas redes, en las que se han articulado para generar diversas formas de representación que permitan entender de manera inicial esta realidad tan compleja que nos ha tocado vivir. Hemos decidido dividir el presente trabajo en cuatro grandes partes: Acopio de datos, Generación de modelos, Integración de la información, Perspectiva a largo plazo y algunas Conclusiones generales.

## ACOPIO DE DATOS

El primer punto fundamental es encontrar y acceder a la información. Existen diversas fuentes de datos públicas proporcionadas por el Gobierno de la Ciudad de México, algu-

nos gobiernos locales y la Secretaría de Salud Federal. Dicha información nos permite ubicar a pacientes desde el punto nominal, su estado, su evolución, y en ciertos casos sus probables diagnósticos y tratamientos en el mismo sentido. En dichas fuentes de datos, siempre existe un aspecto de incertidumbre debido a errores locales en la captura de la información, propia de la dinámica de cómo fluye la información en distintos puntos, como el Triage, Urgencias, Hospitalización, UCI, entre tantos otros.

Los datos obtenidos son multifactoriales, lo que nos arroja una alta dimensionalidad que va desde el tipo de paciente, su ubicación en el hospital, hasta la disponibilidad de servicios en diversas regiones en el país. El interés se centra en entender la realidad de los datos a los que se tiene acceso, con la finalidad de contar con información limpia y completa que permita un mejor uso de la misma. Para este propósito se requiere comprender las diversas fuentes de información y la realidad de los datos, sus incertidumbres y las posibles carencias que existen en las distintas regiones del país para la toma de información. A partir de este contexto, es posible entonces iniciar un preprocesamiento de la información, que implica en primer lugar entender en qué áreas puede ser de interés o utilidad y cómo los diversos grupos de trabajo pueden acceder a la misma.

En esta dirección, lo importante es establecer grupos de trabajo transdisciplinarios, cuyos integrantes sean expertos en las áreas de estudio para entender los datos desde diversas perspectivas: la epidemiológica, la



geográfica, la matemática, la social, la analítica, hasta la limpieza de los propios datos. De esa manera se pueden tomar en cuenta diferentes formas de visualización, de representación, de pronósticos y de ponderación, e incluso hasta de movilidad, lo que resulta fundamental para la toma de decisiones y el entorno donde se vive esta pandemia.

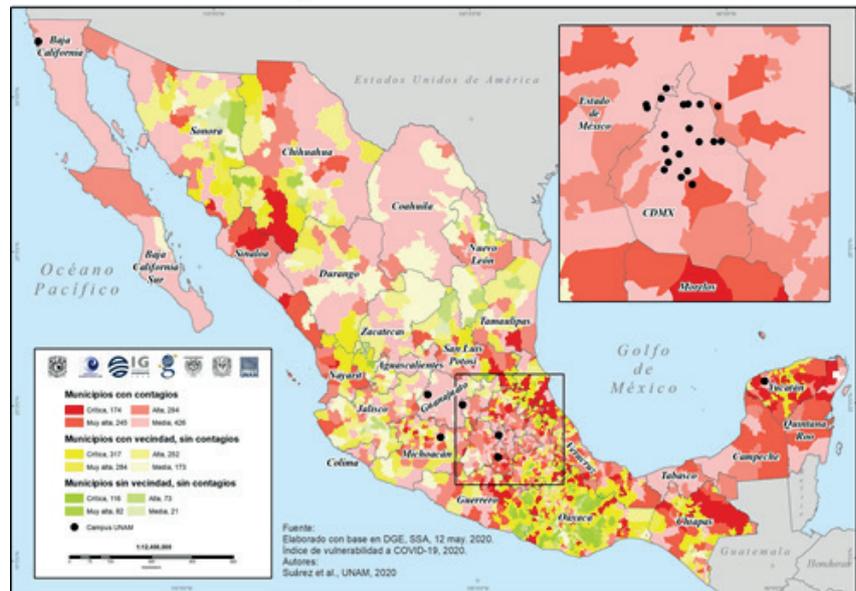
Cabe señalar que en esta búsqueda de datos, no sólo es de interés el dato generado por la fuente de salud, también es de vital importancia conocer el contexto socioeconómico, el demográfico y el contexto del sistema de salud como infraestructura asequible. Otra fuente de información tan o más valiosa en este propósito es la generación de encuestas a nivel nacional, con el objeto de poder medir diversas variables, que en un principio no se pueden contemplar, como las percepciones ante una situación nunca antes experimentada; las sensaciones de cercanía emocional, la construcción de una realidad distinta, tanto individual como colectiva, y el hecho de las comorbilidades, desde el punto de vista del propio individuo que se declara con alguna condición crónico-degenerativa. En este contexto tan complejo e interrelacionado, se constituyen grupos de académicos bajo la perspectiva ambiental, sociodemográfica, biológica, biomatemática, de movilidad, de manejo de riesgos, entre algunas otras. El resultado es un principio de representatividad y construcción de modelos fundamentales.

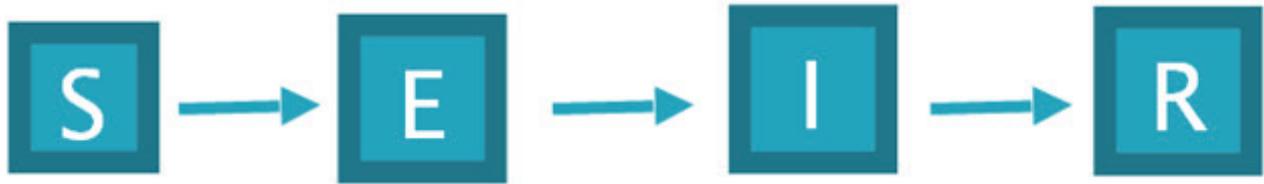
## GENERACIÓN DE MODELOS

Entendiendo los primeros análisis de datos como preprocesamiento de la información, es importante entender las diversas dimensiones y escalas de la pandemia, la representación geográfica, la generación de indicadores con diversas fuentes de información de forma georreferenciada, la generación de encuestas, la obtención de datos clínicos, estructurales, sobre movilidad, sobre condiciones de comorbilidad, sobre vulnerabilidad, sobre condiciones de vida local, regional y nacional. Bajo estas diversas fuentes de información, distintos modelos se pueden generar en la inter-

pretación, caracterización o pronóstico de nuestra realidad. Así, el gran reto de la multidisciplinaria en diversos frentes está en el desarrollo de modelos que ayuden a interpretar, referenciar, escalar y pronosticar las conductas de diversos grupos humanos, ambientales, entre otros. La diversidad de modelos es necesaria para entender la pandemia, evaluar las estrategias que se establezcan, los efectos socioeconómicos, ambientales, emocionales y poblacionales: modelos matemáticos epidemiológicos, de movilidad, de representación a través de patrones de riesgo, de indexación, de generación de métricas, entre tantos otros.

**México: municipios con contagios, vecinos y sin vecindad en relación a su grado de vulnerabilidad a COVID-19.**





Particularmente, los modelos matemáticos son un medio para analizar, evaluar y predecir la evolución de la epidemia y el impacto de las medidas para su control. Los modelos clásicos para el análisis epidemiológico de la dinámica de una enfermedad son los llamados compartamentales. En éstos, la población en riesgo se subdivide o compartamentaliza en grupos, dependiendo de su estatus respecto de la enfermedad: susceptibles, expuestos, infecciosos, recuperados. A estos compartimientos básicos se les pueden añadir aquéllos que sean necesarios para enfocar aspectos específicos de la enfermedad. En general, dados los casos confirmados o estimaciones de los casos probables, las tasas de intercambio entre los compartimientos son estimadas (Angulo *et al.*, 2020; Acuña-Zegarra, Santana-Cibrian y Velasco-Hernández, 2020; Mena *et al.*, 2020). Dichas tasas juegan un papel fundamental para el entendimiento de cambio en el número de infectados, suscepti-

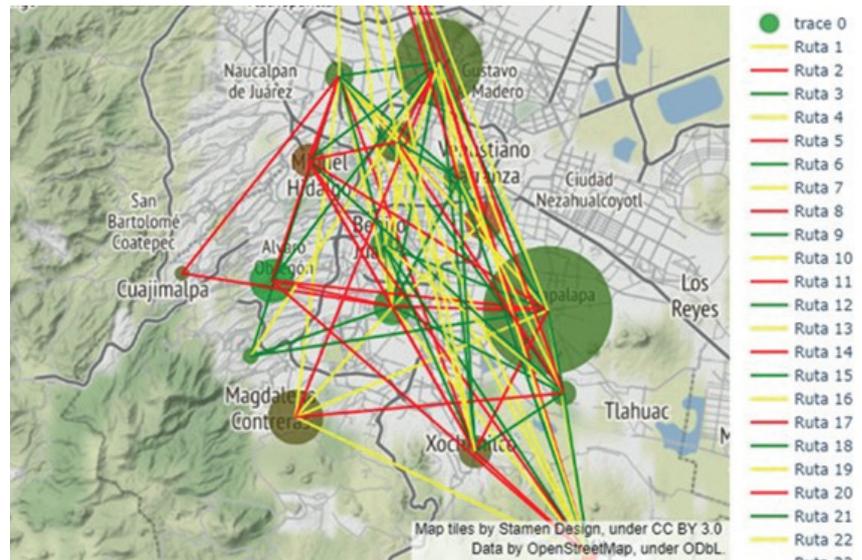
bles, confinados y recuperados. De hecho, un factor muy importante es el número reproductivo básico que se define como el número de infecciones secundarias que un individuo infeccioso típico generará en una población completamente susceptible. Después, conforme la epidemia se desarrolla, se puede calcular el R, número reproductivo instantáneo que identifica sobre la población diaria internada, así como la infectada, el grado de contagio que se tiene en la epidemia; si este valor es mayor a uno, la población volverá a estar en potencial riesgo, por esto es importante tener dicho valor menor a uno puesto que la población infectada tendrá un potencial hacia el decrecimiento, de otra forma tenderá a mantener un porcentaje activo de infectados. Una vez hecho esto, y después de que cada factor ha sido ponderado y analizado desde la perspectiva epidemiológica (no estadística) se pueden hacer proyecciones de escenarios.

Por otro lado, se determinaron los índices de vulnerabilidad para los diferentes municipios del país, donde se puede observar el efecto de datos socioeconómicos y demográficos en la infraestructura de salud (Mena *et al.*, 2020). De hecho, con base en este índice de vulnerabilidad, se ha podido analizar el impacto combinado entre cada una de las figuras de mérito antes señaladas a distintas escalas. Actualmente se trabaja en verles como series de tiempo con base en la incertidumbre generada por los diversos estadios que estamos enfrentando durante esta pandemia. Con base en dicha información, se han señalado varias formas de visualización que van desde la construcción de mapas interactivos, diversas formas de datos, hasta la de la interacción de variables en principio no interconectadas como las ambientales, a través de mecanismos de predicción muy potente como EPI-SPECIES (UNAM, 2020). La siguiente figura muestra un extracto del manejo gráfico sobre el índice de vulnerabilidad.



Por otro lado, la aglomeración de las personas en el transporte público fomenta la transmisión de la enfermedad y a medida que viajan más personas y los viajes son más largos, esta probabilidad de transmisión se incrementa. En esta dirección, se ha desarrollado un modelo en el ámbito de los sistemas dinámicos discretos metapoblacionales, que toma en cuenta patrones de movilidad a través del transporte público para entender el papel de estos patrones en la propagación de esta enfermedad. Así, el modelo permite evaluar la efectividad de diferentes estrategias para alentar la dispersión de COVID-19 y las estrategias de retorno a las actividades normales considerando los efectos del uso masivo del transporte público. Dicho modelo permite construir patrones de movilidad, de riesgo por contacto y de movilidad con la búsqueda en la disminución de dichos riesgos. La siguiente figura muestra un extracto del índice de movilidad (actualmente en construcción).

Ahora bien, como se ha visto en varias fuentes de análisis, el COVID-19 deriva en una situación sistémica donde varios órganos se ven afectados en distinta escala, lo cual hace muy complicado poder tratar en ciertos estadios del desarrollo de la misma. Entonces, para entender el riesgo que conllevan diversas variables clínicas, tanto sistémicas como de análisis bioquímico, se está construyendo un sistema de análisis mediante un generador de patrones de riesgo, que permite determinar cuáles son los factores más importantes para que una persona en particular, según su condición y estado de salud, cuando se le toma una muestra de información relevante. Dicho análisis multivaria-



Conexiones entre las diferentes delegaciones de la Ciudad de México, los colores de las líneas indican el riesgo de contagio y los círculos la incidencia al interior de una delegación.



do requiere de una constante supervisión en la lectura de la información generada, puesto que pueden llegar a malinterpretarse dichos patrones de riesgo.

No menos importante es el análisis de la expansión de la información y por lo tanto de su limpieza, de su estructura, de su integración, de la generación de un directorio de datos; en síntesis, de la generación de una base

de datos consistente. Dicho esfuerzo, aunque pareciera una tarea poco atractiva, implica un proceso de desarrollo de modelos de suma importancia y diametral interés, puesto que la sola visualización de la información, la compartición simple de la misma y el contraste a partir de diversos accesos, presenta una contribución para el entendimiento de las diversas variables que entran en juego en cualquiera de los modelos antes mencionados.

# INTEGRACIÓN DE LA INFORMACIÓN

En principio, se han diseñado varias estrategias con base en el modelo SEIR; de éstas, diversos resultados son de interés, por ejemplo, establecer todos los posibles escenarios para que a partir de datos reales se puedan ir discriminando posibles tendencias o modelos que permitan establecer estrategias de movilidad a partir de análisis previos por zonas metropolitanas, con el fin de entender el transcurso de la propia infección.

En este aspecto, es importante definir estrategias de visualización para entender las diversas dimensiones de los datos, asimismo, entender el impacto entre la observación y la relación que pueden guardar a diversas escalas.

## PERSPECTIVA DE LARGO PLAZO

El trabajo realizado hasta ahora y el aprendizaje que ello conlleva constituyen la base y el fundamento para consolidar una forma nueva de investigación que atienda de manera permanente los riesgos y retos que presenta el COVID-19, así como nuevas posibles situaciones similares en el futuro. Si algo ha puesto de manifiesto esta situación es la importancia del análisis de datos y de contar con los mecanismos actualizados para su recolección de la manera más eficiente y limpia posible. Asimismo, la

emergencia de grupos transdisciplinarios de investigación que buscan entender la complejidad de la crisis de salud en la que nos encontramos, experimentan un reto no trivial que se manifiesta en la interacción entre distintas tradiciones de trabajo y de investigación. Estas experiencias de colaboración son un conocimiento muy valioso para la conformación de nuevos grupos de investigación que puedan enfrentar problemas de complejidad similar a los actuales relativos al COVID-19, como, en general, los problemas que típicamente incluyen la interacción intensa entre distintas dimensiones, entre ellas la epidemiológica, la ambiental y la social. Estos esfuerzos deben dar fruto a un proyecto permanente que trascienda las coyunturas del momento.

La pandemia representa el gran reto, nos ha abierto la posibilidad de interactuar entre áreas que de manera natural sólo observamos, ya sea como un caso de estudio o como un área muy abstracta. En este momento es posible establecer estrategias de trabajo en la medida que los objetivos sean concretos, aunque retadores, para los diversos colegas que trabajan en alguna iniciativa.

## CONCLUSIONES

Es de fundamental interés el poder mostrar la convergencia de diversos grupos de trabajo que dan como resultado distintos puntos de vista de un fenómeno complejo. Aun dicho experimento social no está exento de grandes retos que van desde la comunicación y la transferencia de conocimientos, hasta la probable interpre-

tación de resultados tan importantes para la sociedad en su conjunto. La actual pandemia será estudiada durante muchos años desde diversas perspectivas que, interactuando en ámbitos interdisciplinarios, seguramente nos proveerán de un conocimiento más profundo que nos permita aminorar los efectos de futuros fenómenos similares.

## REFERENCIAS

- Acuña-Zegarra, M.A., Santana-Cibrian, M., y Velasco-Hernández, J.X. (2020). Modeling behavioral change and COVID-19 containment in México: A trade-off between lockdown and compliance. *Mathematical Biosciences*.325.
- Angulo, M.T., Castaños, F., Velasco, J.X., et al. (2020). A simple criterion to design optimal nonpharmaceutical interventions for epidemic outbreaks. *medRxiv preprint*. Doi: <https://doi.org/10.1101/2020.05.19.20107268>
- Mena, R.H., Velasco-Hernández, J.X., Mantilla-Beniers, N.B., et al. (2020). Using posterior predictive distributions to analyse epidemic models: COVID-19 in México City. *Research Gate*. Disponible en: <https://www.researchgate.net/publication/341162160>
- Suárez, M., Valdez, C., Galindo, C., et al. (2020). Índice de vulnerabilidad ante el COVID-19 en México. *Investigaciones Geográficas*. [S.I.] Disponible en: <http://www.investigaciones-geograficas.unam.mx/index.php/rig/article/view/60140>
- UNAM. (2020). Nicho epidemiológico. Plataforma de exploración de datos epidemiológicos. Disponible en: [http://species.c3.unam.mx/covid19/geoportal\\_v0.1.html](http://species.c3.unam.mx/covid19/geoportal_v0.1.html)