



A través de datos colectados en todo el país, por todas las unidades hospitalarias del sector salud y privados, se concentra la información de los pacientes que han solicitado consulta con sospecha de estar infectados

Sistema de información nacional depurado sobre la evolución de la pandemia COVID-19

Antonio Loza, Luis Pineda, Dwight Daniel Dyer Leal, Héctor Benítez, Ricardo Ciria, Ricardo Cruz, Diego Palomares, Zian Fanti y Rosa María Gutiérrez-Ríos

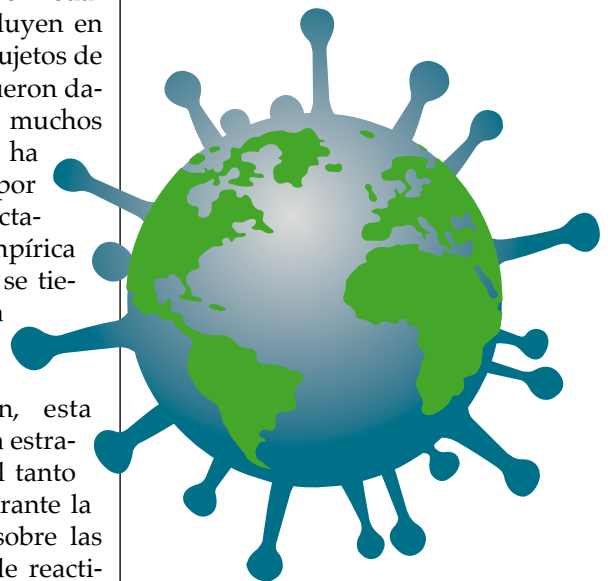
A finales del mes de diciembre del 2019 en Wuhan, China surgieron casos de una neumonía de causas desconocidas y para la cual los tratamientos usuales no ejercían efecto [1]. Su origen se vinculó con alimentos de origen animal no procesados, provenientes de un mercado local en la provincia de Hubei, la cual se volvió el epicentro del brote [2,3]. El padecimiento cobró entonces gran importancia internacional y no solo en China: las autoridades de salud de ese país, como una primera estrategia, identificaron a los pacientes infectados, localizando después a sus contactos, lo que permitió iniciar un compendio organizado de datos clínicos y epidemiológicos. Al mismo tiempo, a través de técnicas basadas en biología molecular se identificó al agente infeccioso, que resultó ser un nuevo coronavirus —denominado entonces 2019-nCoV— causante del Síndrome Respiratorio Agudo (SARS-CoV) [1,4]. Para enero del presente año, a la enfermedad se le dio el nombre de COVID-19, pero aún se ignoraba qué tan contagiosa era, ya que no se había detectado contagio entre individuos. A partir

de enero, el número de enfermos se incrementó exponencialmente en China, cobrándose además cientos de vidas a su paso. Se describieron además una serie de síntomas que incluían fiebre, tos seca, diarrea, dolor de garganta y dependiendo de la región, una mortalidad que oscilaba entre el 2 al 37%, teniendo mayor impacto entre los individuos de la tercera edad. Como China ya había estado expuesta a otros coronavirus en el 2013, que infectaron a más de 8 mil personas y causaron la muerte a 774, el país estaba preparado con un sistema de monitoreo capaz de responder a las pandemias [1]. El SARS-CoV-2, como ahora conocemos al virus causante de la enfermedad COVID-19, se propagó a Japón y Tailandia, por lo que desde el 30 de enero del 2020, la Organización Mundial de la Salud, declaró la situación como de emergencia epidemiológica [5]. Desde esa fecha la enfermedad se propagó fuera de China y en febrero del 2020, llegó a nuestro país [ver *Biotecnología en Movimiento* No. 21]. Desde entonces a nivel mundial se comenzaron a crear bases de datos, que son sistemas que almacenan de forma organizada la información,



que pudieran dar cuenta de la evolución de la COVID-19 en los distintos países del mundo. México no ha sido la excepción y a raíz de la epidemia causada por el virus de la influenza A(H1N1), se crea el antecedente de lo que es el sistema de vigilancia epidemiológica (SNAVE) [6], que ha tenido como misión, entre otras, en la actual pandemia, llevar un registro exhaustivo de los casos de la COVID-19 a nivel nacional. A través de datos colectados en todo el país, por todas las unidades hospitalarias del sector salud y privados, se concentra la información de los pacientes que han solicitado consulta con sospecha de estar infectados. La Dirección General de Epidemiología (DGE) de la Secretaría de Salud (SALUD) cuenta con la base de datos del Sistema de Vigilancia Epidemiológica de Enfermedades Respiratorias Virales (SISVER) que incluye los registros individuales de quienes se han presentado en las unidades de salud en México, por presentar sintomatología asociada a la enfermedad. Estas unidades incluyen, unidades de salud monitoras de enfermedad respiratoria (parte del modelo 'Centinela' de vigilancia epidemiológi-

ca) y unidades no monitoras que complementan el sistema de atención médica para COVID-19. El recurso incluye las referencias generales de los pacientes, sus síntomas y comorbilidades —es decir, la presencia de otras enfermedades que se ha demostrado influyen en la vulnerabilidad—; si fueron sujetos de muestreo y su resultado, y si fueron dados de alta o fallecieron, entre muchos otros datos. La información ha sido y sigue siendo recabada por profesionales de la salud directamente y constituye la fuente empírica original del conocimiento que se tiene de quienes han padecido la enfermedad en nuestro país, además de ser la única de alcance nacional. Por esta razón, esta base de datos tiene importancia estratégica y de seguridad nacional tanto para la toma de decisiones durante la fase aguda de la pandemia, sobre las medidas para los protocolos de reactivación económica por entidad federativa, así como para los análisis y los estudios epidemiológicos y farmacológicos que se llevarán a cabo muy posiblemente durante los próximos años.





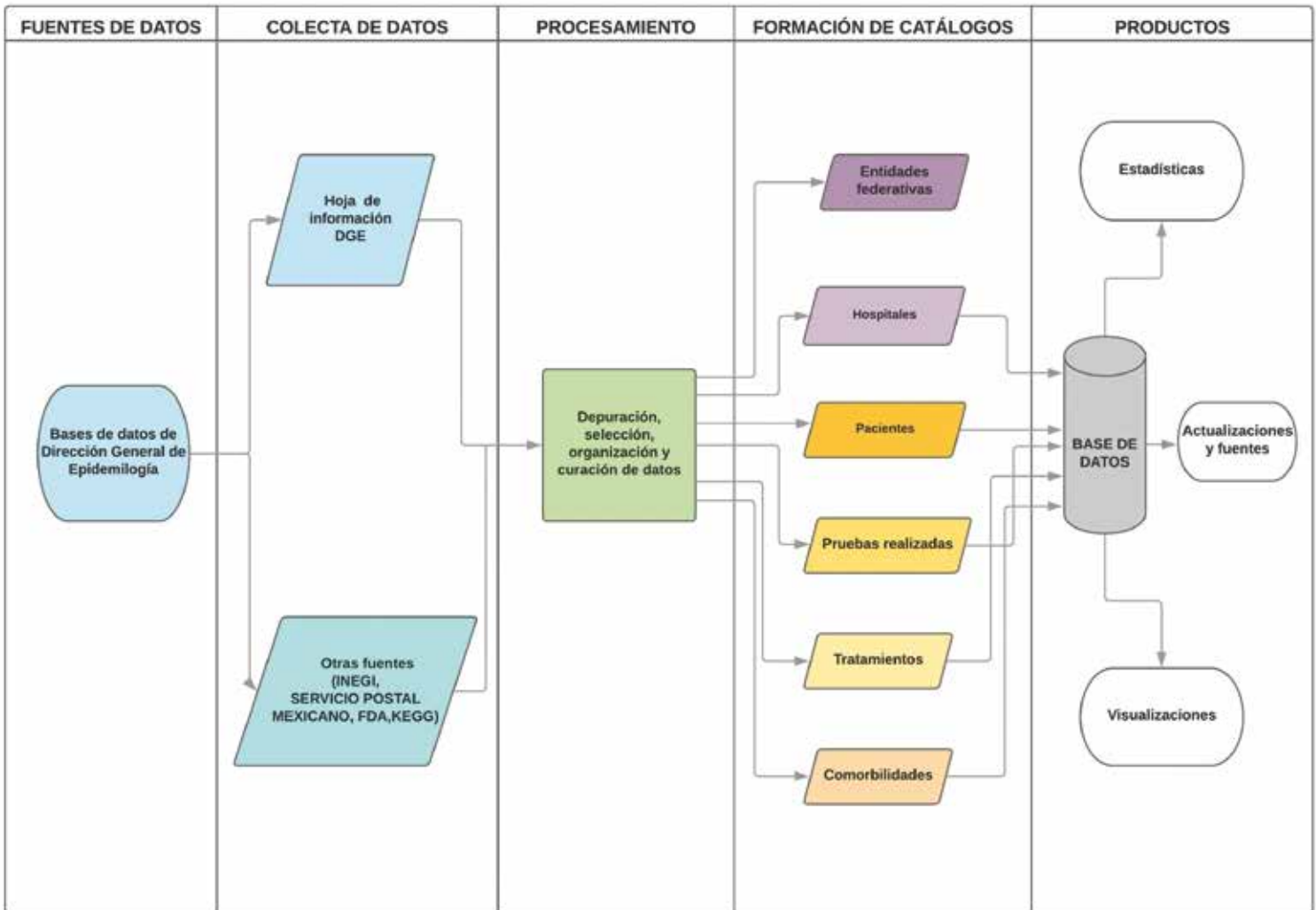
La base de datos tiene registrados aproximadamente dos millones de casos analizados y se ha concluido una primera versión del sistema de gestión de la información que se extenderá a futuro para desplegar información georreferenciada

Figura 1. Modelo de la base de datos COVID-19 IBT-IIMAS

Hacia un sistema de inteligencia epidemiológica

Sin embargo, dada la infraestructura informática con la que cuenta el sector de la salud actualmente, así como las muy diversas dificultades que se han presentado con la pandemia, la información está sujeta a errores e inconsistencias. Por esta razón, se requiere depurar la información para contar con un recurso útil y confiable. La base de datos está disponible para la UNAM mediante un convenio de colaboración y confidencialidad firmado con la Secretaría de Salud y se ha utilizado para realizar diversos análisis y evaluaciones de manera muy intensa durante la crisis provocada por la pandemia. Es importante mencionar que este recurso contiene información muy sensible que se ha mantenido reservada y cuyo acceso sólo será posible acatando los compromisos de confidencialidad que se definan para tal efecto. En este contexto la UNAM, a través de la Coordinación de la Investigación Cientí-

fica (CIC), acordó apoyar a la Dirección General de Epidemiología (DGE) de la Secretaría de Salud para depurar la información y hacerla accesible a usuarios registrados, conforme a un compromiso de confidencialidad y a los lineamientos de uso que actualmente se definen entre SALUD y la UNAM. Para este efecto, académicos e investigadores del Instituto de Biotecnología (IBt) y del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS) de la UNAM desarrollan una base de datos depurada y un sistema de visualización para el seguimiento y análisis de la evolución de la epidemia a nivel nacional, con las características mencionadas. Actualmente el sistema consiste en la base de datos alojada en el IBt y con la interfaz habilitada en el IIMAS. La interfaz se accede en la dirección de Internet <http://covid-19.iimas.unam.mx/>, tanto desde equipos de cómputo como teléfonos celulares, aunque es de acceso controlado y se requiere un identificador de usuario y contraseña. Dentro de la interfaz se



han habilitado módulos para consultas de interés con diversas opciones; las respuestas se presentan de manera gráfica así como a través de archivos de datos que se pueden utilizar para otros análisis de interés.

Desarrollo del sistema de información COVID-19 IBT-IIMAS

El sistema de información lo comprende un sistema de gestión de bases de datos relacionales de código abierto conocido, que organiza la información en una o más tablas de datos. Los tipos de datos pueden estar relacionados entre sí; estas relaciones ayudan a estructurar los datos, para su posterior consulta. Los métodos desarrollados hasta el momento incluyen un proceso de curación de los contenidos proporcionados por la DGE, con lo cual se verifica su consistencia; la clasificación de datos conforme atributos que permiten agruparlos en tablas; la división y organización de la información en tablas usando 'scripts' desarrollados por el grupo de trabajo como se muestra en la Figura 1.

La construcción y mantenimiento del sistema, —cuya página de acceso y un ejemplo de consulta se muestra en la Figura 2— forma parte de la línea de desarrollo académico “Modelos y Datos” que pertenece al grupo de cuatro grandes líneas que propuso la UNAM a través de la Coordinación de la Investigación Científica para el combate a la COVID-19.

El sistema desarrollado por la UNAM permite obtener información por entidad federativa, edad y sexo, por lo cual, el nivel de detalle de la información que puede ser consultada supera ampliamente al que se presenta en otros portales dedicados a COVID-19 (veáse por ejemplo: *WorldoMeter* consultable en <https://www.worldometers.info/coronavirus/>; la de la Universidad Johns Hopkins, consultable en <https://coronavirus.jhu.edu/map.html>). Más aún, la información que se presenta en el sistema de consulta es la punta de un iceberg, en el sentido de que constituye solo una pequeña fracción de las posibilidades de análisis que el sistema permite. Como por ejemplo, consultas complejas hechas a la medida, con cruces de variables específicas o incluso con resolución



espacial a nivel de colonia, lo cual puede solicitarse por grupos de investigación, bajo el convenio de confidencialidad firmado entre la UNAM y SALUD.

Al momento de la redacción de este artículo, la base de datos tiene registrados aproximadamente dos millones de casos analizados y se ha concluido una primera versión del sistema de gestión de la información que se extenderá a futuro para desplegar información georreferenciada. A pesar de los problemas en la cobertura, oportunidad y calidad en el registro derivados de las diversas fuentes de origen, el sistema ha logrado reflejar de manera muy clara tendencias ya descritas como la vulnerabilidad ante la COVID-19 de ciertos grupos de pacientes con comorbilidades, o grupos de edad, que se asemejan a las observadas en otras regiones en el mundo. Adicionalmente, el sistema tiene el potencial de brindar alguna luz sobre fenómenos como la dinámica derivada de coexistencia de la epidemia de COVID-19 con otras enfermedades virales como por ejemplo influenza, y, eventualmente, podría incluso ayudar a resolver la incógnita sobre la duración de la inmunidad y efectividad de las distintas vacunas probadas. Es por ello, que contar con bases de datos nacional confiables, derivada de fuentes oficiales es un recurso fundamental para la toma de decisiones y la generación de estrategias que permitan a las autoridades contender contra la actual epidemia y problemas de salud pública presentes y futuros.

Contacto: rosa.gutierrez@mail.ibt.unam.mx

Figura 2. Página de inicio y ejemplo de gráfica de distribución de pacientes por fecha de ingreso a la base de datos DGE

Referencias

1. Wang C, Horby PW, Hayden FG, Gao GF (2020), A novel coronavirus outbreak of global health concern. *Lancet* **395**:470–3.
2. Lu H, Stratton CW, Tang YW (2020), Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle, *J Med Virol* **92**:401–2.
3. Bogoch II, Watts A, Thomas-Bachli A, et al. (2020), Pneumonia of unknown aetiology in Wuhan, China: Potential for international spread via commercial air travel, *J Travel Med* **27**:1–3.
4. Kour JM, Hirschhaut M (2020), Reseña histórica del COVID-19 ¿Cómo y por qué llegamos a esta pandemia? *Acta Odontol. Venez* **58**:1–9.
5. Singhal T (2020), Review on COVID19 disease (COVID-19), *Indian J Pediatr* **87**:281–6.
6. OPS, OMS. (2014) Guía operativa para la vigilancia nacional de la Infección Respiratoria Aguda Grave (IRAG), pp. 1–67. A, Disponible en: http://www.who.int/influenza/surveillance_monitoring/fluid/en/

El Dr. Antonio Loza, el MC. Ricardo Ciria, la Dra. Rosa María Gutiérrez-Ríos y el Lic. Diego Palomares pertenecen al Departamento de Microbiología Molecular del IBT-UNAM. Los Dres. Luis Pineda, Héctor Benítez, Ricardo Cruz y Zian Fanti forman parte del personal académico del IIMAS-UNAM, y el Dr. Dwight Daniel Dyer Leal a la Dirección General de Promoción de la Salud, de la Secretaría de Salud.