



EPIDEMIOLOGÍA DIGITAL

DIGITAL EPIDEMIOLOGY

David Pastor Escuredo

Centro de Innovación en Tecnología para el Desarrollo

Humano-itdUPM

david.pastor@upm.es

Fecha recepción artículo: 01/06/2020 • Fecha aprobación del artículo: 17/08/2020

RESUMEN

La epidemiología es una disciplina que ha experimentado grandes avances basados en modelos computacionales y que cada vez incrementa su alcance gracias a nuevas fuentes de datos. Además de las variables tradicionales consideradas en epidemiología, ahora se integra, en tiempo real, el análisis de patrones sociales a gran escala con datos locales, permitiendo estudios multi-escala. En un mundo cada vez más conectado, los modelos y análisis de las interacciones y comportamientos sociales son clave para entender y frenar epidemias. El Big Data está permitiendo confrontar y refinar modelos con datos reales, además de permitir nuevas aplicaciones como el mapeo y seguimiento de enfermedades en tiempo real u optimizar la distribución de recursos como las vacunas o la realización de tests.

La epidemiología digital es, por tanto, clave para el control de enfermedades y diseño de protocolos y políticas de acción. En esta revisión abordamos las áreas de estudio que están configurando esta disciplina: el análisis de propagación basado en modelos que integran redes de personas y rastreo de contactos, el análisis de movilidad y la propagación espacio-temporal de enfermedades infecciosas y el emergente análisis de la *infodemia* entendida como el estudio de los problemas relacionados con la información y conocimiento relacionado con las epidemias. La epidemiología digital permite generar y refinar nuevos mecanismos operacionales como son la prevención y mitigación, el monitoreo de la evolución de la epidemia y su impacto y la evaluación de medidas sanitarias y políticas (farmacéuticas y no-farmacéuticas) para parar epidemias. Las epidemias no pueden ser tratadas como fenómenos únicamente sanitarios, están intrínsecamente relacionadas con fenómenos socio-económicos y las dinámicas de otros sectores como el turismo o la agricultura, por tanto, debemos aproximar la epidemiología desde los sistemas complejos y con una visión sistémica. Además de los avances que se han acelerado por la pandemia COVID-19, discutimos sobre las oportunidades y retos para luchar epidemias de manera más eficaz y humana.

Palabras clave: Epidemiología, Redes, Propagación, Impacto, Infodemia, Movilidad, Big data, Seguimiento de contactos, Confinamiento, privacidad, Ética, Vigilancia.



ABSTRACT

Epidemiology has recently witnessed great advances based on computational models. Its scope and impact are getting wider thanks to the new data sources feeding analytical frameworks and models. Besides traditional variables considered in epidemiology, large-scale social patterns can be now integrated in real time with multi-source data bridging the gap between different scales. In a hyper-connected world, models and analysis of interactions and social behaviors are key to understand and stop outbreaks. Big Data along with apps are enabling for validating and refining models with real world data at scale, as well as new applications and frameworks to map and track diseases in real time or optimize the necessary resources and interventions such as testing and vaccination strategies.

Digital epidemiology is positioned as a discipline necessary to control epidemics and implement actionable protocols and policies. In this review we address the research areas configuring current digital epidemiology: transmission and propagation models and descriptions based on human networks and contact tracing, mobility analysis and spatio-temporal propagation of infectious diseases and the emerging field of *infodemics* that comprises the study of information and knowledge propagation related to epidemics. Digital epidemiology has the potential to create new operational mechanisms for prevention and mitigation, monitoring of the evolution of epidemics, assessing their impact and evaluating the pharmaceutical and non-pharmaceutical measures to fight the outbreaks. Epidemics should not be treated as solely a health crisis, epidemics are interconnected with socio-economics and the dynamics of other sectors such as tourism, commerce or agriculture. Thus, epidemics have to be approached from the lens of complexity and require systemic solutions. The field has experienced unprecedented advances because of the COVID-19 pandemic. Opportunities and challenges to tackle epidemics more effectively and with a human-centered vision are here discussed.

Keywords: Epidemiology, Contact tracing, Networks, Propagation, Impact, Infodemics, Mobility, Big Data, Lockdown, Privacy, Ethics, Surveillance.

David Pastor Escuredo es PhD, Master e Ingeniero de Telecomunicación por la Universidad Politécnica de Madrid (UPM). Ha sido investigador asociado de la UC San Diego, el CNRS y la University of Washington. Ha colaborado y trabajado como consultor para varias agencias de Naciones Unidas: UN Global Pulse, UNHCR, UNICEF y WFP. Actualmente es fellow de Catalyst Europe, programa del MIT y el EIT Health para digitalización de la salud, y lidera proyectos de innovación digital para el desarrollo en el itdUPM.

Correo electrónico: david.pastor@upm.es



INTRODUCCIÓN

La epidemiología es la disciplina encargada de identificar la distribución, incidencia y etiología de enfermedades humanas (Green, Freedman et al. 2000; Salathe, Bengtsson et al. 2012). Aunque los datos y los modelos (Anderson, Anderson et al., 1992) han sido siempre parte de la epidemiología, la aparición de nuevas fuentes de Big Data y tecnologías (Bell, Hey et al. 2009) han posibilitado que la computación abra nuevas oportunidades para incrementar el impacto y el conocimiento (Salathe, Bengtsson et al. 2012). En paralelo a la aparición y uso de nuevas fuentes de datos, el crecimiento de la Inteligencia Artificial, y especialmente las técnicas de Machine Learning (LeCun, Bengio et al., 2015; Bullock, Pham et al., 2020), están dando lugar a muchas metodologías y aplicaciones que pueden clasificarse como la emergente epidemiología digital.

Los estudios de epidemiología se han basado fundamentalmente en datos recogidos en la práctica clínica o en trabajos de campo (Hernán and Robins, 2006). Los primeros pasos de la epidemiología digital se pudieron dar gracias a la información en Internet, especialmente datos de búsquedas (e.g. Google flue trends) por parte de la población (Brownstein, Freifeld et al., 2009; Eysenbach, 2009; Ginsberg, Mohebbi et al., 2009, Cervellin, Comelli et al., 2017), lo cual también generó las primeras dificultades metodológicas y epistemológicas de esta nueva disciplina (Cook, Conrad et al., 2011; Olson, Konty et al., 2013; Lazer, Kennedy et al. 2014). Actualmente contamos con muy diversas fuentes como las redes sociales, los móviles, las apps u otros servicios tecnológicos que generan datos (Salathé, 2018). Esta tendencia en la digitalización se está agudizando precisamente por la convivencia con una pandemia, COVID-19 causada por el virus SARS-CoV-2, que está acelerando la adopción digital de todos los sectores de la sociedad que hasta ahora habían llevado una transición tecnológica más lenta.

Por tanto, la epidemiología digital es la encargada de entender la dinámica de los patrones, tanto sociales como clínicos y de salud, de las poblaciones afectadas por una enfermedad y las causas de estos patrones (Salathé, 2018). Acorde con la definición de la Organización Mundial de la Salud (OMS), la epidemiología es el estudio de la distribución y los determinantes de estados o eventos (en particular de enfermedades) relacionados con la salud y la aplicación de esos estudios al control de enfermedades y otros problemas de salud. Por tanto, la epidemiología tiene una dimensión realmente práctica cuyos resultados están destinados a mejorar los sistemas de respuesta frente a epidemias, tanto a prevención, gestión de la epidemia, mitigación y preparación para futuras epidemias u oleadas. Además, la epidemiología, por su importancia y complejidad, está haciendo avanzar a las disciplinas teóricas y técnicas. El hecho de que la epidemiología sea cada vez más digital no tiene implicaciones meramente tecnológicas, sino que es lo que permite aumentar el propio alcance de la epidemiología para gestionar la complejidad de las enfermedades y sus factores biológicos, ambientales y sociales. Este alcance es mayor porque se usan datos de diversas fuentes, incluso datos que no estaban recolectados o diseñados para aplicaciones de la salud (Salathé, 2018). En este documento vamos a revisar las áreas de trabajo científico y técnico y las contribuciones más importantes que van a dar forma a la epidemiología del presente y del futuro, en la que la actual pandemia del COVID-19 ha supuesto un antes y un después (Vespignani, Tian et al. 2020).



MODELOS Y REDES

Existen diversos modelos epidemiológicos que se clasifican fundamentalmente en dos tipos: modelos basados en ecuaciones y modelos basados en agentes. Los modelos basados en ecuaciones asumen homogeneidad y un comportamiento en masa (Anderson, Anderson et al., 1992). El progreso en sistemas de computación ha permitido generar modelos basados en agentes que pueden modelar la heterogeneidad de comportamientos presentes en una epidemia (Aleta, Martín-Corral et al., 2020; Espana, Cavany et al., 2020). Ambos tipos de modelos se basan en una conceptualización de la enfermedad mediante diversos estados de las personas, siendo el modelo más sencillo el marco SIR (Susceptible, Infectado, Recuperado). Existen diversas adiciones a este marco para introducir complejidades de ciertas epidemias, notablemente, el modelo SEIR introduce el estado Expuesto que engloba las personas contagiadas en proceso de incubación. Cada estado se parametriza con el fin de cuantificar la transición entre estados teniendo en cuenta criterios biológicos y sociales que suelen ser derivados de estudios clínicos y de encuestas. Los modelos típicamente dan como resultado una estimación del ritmo reproductivo R que acopla las características de la enfermedad con los mecanismos de transmisión, así como la temporalidad y volumen de la evolución.

Un elemento clave de los modelos epidemiológicos más sofisticados es introducir la complejidad de los contactos entre personas en forma de redes. El comportamiento de algunas enfermedades es sistémico y por tanto su propagación depende de la estructura del sustrato en el que se propagan. Aunque todas las escalas, desde lo bioquímico a lo social, están profundamente interrelacionadas, el estudio a escala de las redes sociales es útil para predecir la propagación de enfermedades y actuar en consecuencia. Las redes permiten modelar la componente de comportamiento de la transmisibilidad de las enfermedades mediante la caracterización de la topología de la red y también su dinámica.

Dependiendo de la topología de la red, una enfermedad puede propagarse a diferentes velocidades y alcances (Pastor-Satorras and Vespignani, 2001; Moreno, Pastor-Satorras et al. 2002; de Arruda, Petri et al., 2020). Varios estudios han partido de los modelos de red “scale-free” donde la distribución del número de conexiones de cada nodo sigue una ley de potencia, que representa adecuadamente redes de personas reales y de ahí su utilidad. El estudio epidemiológico con este modelo demuestra la ausencia de un umbral epidemiológico y también una gran heterogeneidad de comportamiento (Pastor-Satorras and Vespignani, 2001). Modelos que limitan el tamaño de las redes “scale-free” muestran la aparición de un umbral epidemiológico muy bajo y un comportamiento heterogéneo incluso en redes de pequeño tamaño (Pastor-Satorras and Vespignani, 2002). Ello implica que las redes “scale-free” son propensas a la propagación independientemente de la tasa de propagación de la enfermedad (Pastor-Satorras and Vespignani, 2001). Además, la velocidad de propagación de una epidemia es inversamente proporcional a las fluctuaciones del grado, ya que es fácil llegar a nodos muy conectados que propagan la enfermedad en un efecto cascada (Barthélemy, Barrat et al., 2005).

Además de la estructura, cuantificar la dinámica de las redes ha sido útil para entender epidemias, fundamentalmente a través de los procesos de percolación y difusión. En el caso de enfermedades, la percolación se entiende por el proceso por el cual los nodos de la red pasan de estado sano a estado contagio, lo cual depende de la compartimentación de la red (Newman, 2002; Meyers, Newman et al., 2006). La difusión permite cuantificar la potencial propagación entendida como múltiples caminos que se establecen dentro de la red (Buono, Alvarez-Zuzek et al., 2014). Ambos procesos determinan el umbral epidemiológico.



Las redes forman parte de modelos epidemiológicos mediante matrices de contactos. Estas matrices pueden ser estratificadas y multi-capa (Aleta, Martín-Corral et al., 2020), es decir que los contactos dependen de los grupos de edad pero también de diversas capas de interacción social (e.g. trabajo, hogar escuelas, etc). Recientes avances metodológicos en redes comprenden el uso de hipergrafos mediante enlaces que conectan varios nodos en lugar de relaciones entre pares (de Arruda, Petri et al., 2020).

Las redes son además útiles para sistemas de predicción de riesgo y propagación, por ejemplo, mediante el análisis de la red de *hubs* de transporte (Colizza, Barrat et al., 2007), confirmando que la heterogeneidad favorece la propagación. Por tanto, es lógico que las acciones para prevenir, contener y detener epidemias deban buscar reducir y homogeneizar el grado de los nodos en todas las escalas para que la epidemia sea más fácilmente predecible y controlable (Chinazzi, Davis et al., 2020; Martín-Calvo, Aleta et al., 2020; Zhang, Litvinova et al., 2020).

Además de una larga tradición, el estudio de redes es de total actualidad gracias a las nuevas fuentes de datos, adquiridos vía diversas tecnologías como el Bluetooth o la geolocalización de alta resolución de los dispositivos móviles, que permiten realizar “contact tracing” con detalle y obtener matrices de contacto dinámicas y de alta resolución (Ferretti, Wymant et al., 2020). Metodológicamente, las redes de diversas fuentes sirven para crear redes multipartitas que representan las interacciones entre personas y lugares.

Las redes también son útiles para entender los procesos de recuperación y la resiliencia, la cual es favorecida por procesos heterogéneos de recuperación dentro de la red (Boy, Pastor-Escuredo et al., 2019; de Arruda, Petri et al., 2020). Otra aplicación es entender la interacción entre enfermedades concurrentes (Sanz, Xia et al., 2014).

MOVILIDAD Y PROPAGACIÓN

La movilidad tiene un impacto directo en la propagación de enfermedades de transmisión aérea o a través de vectores y desde hace varias décadas se ha convertido en un aspecto central de la epidemiología, aunque su análisis fue limitado originalmente por la falta de datos dinámicos para diferenciar entre diversos tipos de movilidad incluyendo migraciones y circulación (Prothero, 1977). En los últimos años hemos podido presenciar nuevas formas de movilidad: turismo, eventos, negocios, laboral a largo plazo o movilidad de estudiantes. Ya no es posible caracterizar la movilidad mediante encuestas y datos estáticos, el análisis de la movilidad y la distribución espacial de factores de propagación depende de la disponibilidad y resolución de datos longitudinales (Eubank, Guclu et al., 2004; Riley, 2007; Gonzalez, Hidalgo et al., 2008).

La movilidad humana presenta diversas escalas de organización, tanto temporales como espaciales (Candia, González et al., 2008; Gonzalez, Hidalgo et al., 2008; Zufria, Pastor-Escuredo et al., 2018; Pastor-Escuredo and Frias-Martinez, 2020). La movilidad se segmenta en diversas capas interconectadas y cada una de estas capas es un medio de propagación con su propia dinámica, pero que a la vez están interconectadas (Balcan, Colizza et al., 2009). Esta estructura de la movilidad tiene un efecto amplificador de la dispersión de enfermedades más allá del entorno inicial de la misma. Los primeros estudios en modelar la movilidad como factor epidemiológico se centraron en la escala global basándose en datos demográficos y de estadísticas de movilidad internacional (Balcan, Gonçalves et al., 2010). La resolución temporal, sin embargo, era muy estática, permitiendo sólo introducir variabilidad estacional (Balcan, Hu et al., 2009). Los modelos de



movilidad más usados son el modelo de gravitación y el modelo de radiación (Simini, González et al., 2012). Sin embargo, estos modelos sólo funcionan en zonas concretas con fuertes asunciones por lo que es difícil generalizarlos en la práctica (Perrota, 2018).

En enfermedades transmitidas por vectores, como por ejemplo la malaria o el dengue, o por medios físicos como el aire o el agua, la movilidad a pequeña escala afecta a la exposición de las personas a ser infectados y la movilidad a mayor escala puede producir la introducción, reinserción y circulación de infectados e incluso la propagación global (Buscarino, Fortuna et al., 2008; Stoddard, Morrison et al., 2009; Lynch and Roper, 2011). En muchos casos, las enfermedades generan cambios sistémicos en la movilidad de la población con un impacto incontrolable (Meloni, Perra et al., 2011). Por ello, se ha identificado la necesidad de crear mecanismos de monitorización de movilidad de alta resolución posibles por datos de móviles.

Los datos de móviles son datos generados por las operadoras de telecomunicaciones y contienen geolocalización de llamadas. También los servicios de geolocalización de “smartphones” permiten recoger trazas de movilidad. La resolución temporal de estos datos es muy alta (Pulse, 2012; Blondel, Decuyper et al., 2015) y requieren pasar por un proceso de anonimización y agregación para respetar la privacidad (De Montjoye, Hidalgo et al., 2013; Pulse, 2015).

Los datos longitudinales de alta resolución permitieron la caracterización de “hotspots” y optimizar la localización de acciones para prevenir y parar las enfermedades (Bejon, Williams et al., 2010; Dolgin, 2010). Sin embargo, ha sido gracias a los datos de móviles cuando se ha producido una auténtica revolución de la movilidad para epidemias (Bengtsson, Lu et al., 2011; Wesolowski, Eagle et al., 2012; Tatem, Huang et al., 2014; Tizzoni, Bajardi et al., 2014; Bengtsson, Gaudart et al., 2015; Wesolowski, Qureshi et al., 2015). Un aporte fundamental de los datos de móviles ha sido poder caracterizar los flujos de personas entre zonas con diferentes riesgos de contagio y prevalencia de la enfermedad para caracterizar la dinámica del riesgo. Por ejemplo, la caracterización de “sources” y “sinks” ayuda a entender qué zonas pueden producir contagios aunque no haya una densidad alta del vector (Wesolowski, Eagle et al., 2012). Sin embargo, un análisis más detallado más allá de regiones geográficas basada en los flujos de persona a través de las redes (Pastor-Escuredo and Frias-Martinez, 2020), puede ayudar a entender mejor la dinámica del riesgo y los fenómenos de super-propagación. Estas técnicas, sin embargo, requieren de un manejo muy cuidadoso del análisis y los resultados para evitar mecanismos invasivos de la privacidad y de las libertades de los individuos.

Los eventos culturales en muchas regiones también han sido analizados con estos datos mostrando su impacto en las epidemias (Finger, Genolet et al., 2016). El análisis a largo plazo de la movilidad y los diferentes perfiles de movilidad de la población es una herramienta útil para entender la dinámica de las epidemias de forma desagregada (United Nations International Children’s Emergency Fund [UNICEF], 2020). La desagregación del análisis por grupos poblacionales diversos es clave ya que las epidemias están intrínsecamente relacionadas con la diversidad socio-económica (Martín-Calvo, Aleta et al., 2020) y las dinámicas de otros sectores como son el trabajo, el turismo (Belderok, Rimmelzwaan et al., 2013; Chinazzi, Davis et al., 2020) o la agricultura y las migraciones rural-urbanas (Zufiria, Pastor-Escuredo et al., 2018). Los datos de móviles pueden ser combinados con datos de encuestas para ofrecer una descripción de alta resolución espacio-temporal pero también demográfica (Wesolowski, Stresman et al., 2014).

En la actual pandemia del COVID-19, el análisis de movilidad ha ido un paso más allá en cuanto a casos de uso y nivel de agregación utilizado (Oliver, Lepri et al., 2020). Gran parte de las aplicaciones y estudios se han



enfocado en las medidas de confinamiento: controlar a la población y sus movimientos para asegurar que se cumpliera con el confinamiento, medir el impacto del confinamiento en el avance de la pandemia y evaluar las medidas de desescalamiento (Martín-Calvo, Aleta et al., 2020).

INFODEMIA vs INFORMACIÓN

Un aspecto clave durante una pandemia es que toda la población tome las mejores decisiones posibles para evitar el escalado de efectos negativos. La asimetría del impacto negativo de una epidemia requiere de medidas adicionales en la gestión del riesgo y del incremento de precaución (Taleb, Bar-Yam et al., 2014; Taleb, 2019; Norman, Bar-Yam et al., 2020). Para poder tomar buenas decisiones es necesario tener la información adecuada en el momento adecuado (Greenwood, Howarth et al., 2017), por ello, la propagación de la información durante la pandemia se ha convertido en un foco de total relevancia como ha declarado la propia Secretaría General de las Naciones Unidas. La actualidad viene marcada por la propagación de muchas noticias falsas, contaminantes y generadoras de división (Shu, Sliva et al., 2017; Bakir and McStay, 2018; Lazer, Baum et al., 2018). Esta situación se agrava exponencialmente en momentos de crisis, en paralelo a una pandemia se puede vivir una infodemia (Hua and Shaw, 2020; Vaezi and Javanmard, 2020; Zarocostas, 2020).

Multitud de trabajos han estudiado la propagación de la información, y en concreto los rumores y las noticias falsas, a través de redes con un paralelismo a cómo se propaga una epidemia en un red compleja (Nekovee, Moreno et al., 2007; Miritello, Moro et al., 2011; Morales, Borondo et al., 2014). A través del análisis de datos de redes sociales se pueden identificar quiénes son los líderes de la red y su influencia en la propagación de información (Newman, Barabási et al., 2006; Bodendorf and Kaiser, 2009; Pastor-Escuredo and Tarazona, 2020) y cuantificar la emergencia de comportamiento viral y de “spreaders”. Una aproximación viral del comportamiento en la red puede servir para identificar “spreaders” (Borge-Holthoefer, Meloni et al., 2013). Además del análisis de la propagación, el análisis de contenido semántico es una herramienta útil para clasificar texto. Las nuevas herramientas de Deep Learning (LeCun, Bengio et al., 2015) están haciendo esta tarea escalable y precisa (Ruchansky, Seo et al., 2017; Singhanian, Fernandez et al., 2017; Popat, Mukherjee et al., 2018).

Ante el riesgo de mensajes de odio y ruptura, se destaca la necesidad de usar la información de forma positiva, ayudando a controlar el riesgo, generando información constructiva a través de las redes. La información puede empoderar a la población para tomar mejores decisiones individuales y colectivas, y ayudar a mantener a las personas y su entorno seguros frente a epidemias, además del potencial para construir resiliencia e inyectar un refuerzo socio-económico basado en la inteligencia colectiva (Luengo-Oroz, Pham et al., 2020).

La información recogida por parte de los ciudadanos también sirve para gestionar el riesgo y entender mejor la epidemia a falta de confirmación de casos en información clínica (Leung and Leung, 2020; Sun, Chen et al., 2020), alimentando a nuevos modelos que gestionen probabilidades en función de datos más allá de los demográficos y los clínicos. En ese sentido, nuevos sensores para poder monitorear variables y cambios de estado en la población serán necesarios y ya hay algunas herramientas para, por ejemplo, clasificar enfermedades analizando la tos (Imran, Posokhova et al., 2020) o la pérdida de olfato (Menni, Valdes et al., 2020). Finalmente, nuevos canales de información entre autoridades y población son necesarios para generar confianza y mejorar la respuesta.



SISTEMAS OPERACIONALES POTENCIADOS CON INTELIGENCIA ARTIFICIAL

PREDICCIÓN Y PREVENCIÓN

Una respuesta temprana y rápida a una epidemia posibilita controlarla y mitigar su impacto. Por ello, los métodos presentados tienen el objetivo principal de predecir la evolución de las epidemias. La predicción de estos modelos se basa fundamentalmente en las dimensiones geográfica (área de expansión), temporal (curva de la epidemia) y de transmisión (R y pico de la curva). Las variables utilizadas para caracterizar los modelos son fundamentalmente demográficas (densidad de población, edad, sexo). Esto conlleva serias limitaciones en cuanto al entendimiento de las epidemias y la habilidad de los modelos para predecir la misma propagación y su impacto. Variables como la vulnerabilidad, la desigualdad socioeconómica o las infraestructuras y saneamiento son clave para tener modelos más eficaces. Las epidemias son procesos complejos que se dan en sistemas sociales que son a su vez complejos. Además, la especificidad y complejidad de estos sistemas tienen una importante diversidad en función de las diferentes zonas geográficas donde ocurren las epidemias lo cual dificulta la aplicación de modelos. Este hecho es incluso más relevante en pandemias con alcance global donde una misma enfermedad tiene dinámicas distintas a lo largo del espectro de regiones. Algunos esfuerzos se han realizado en esta dirección, desarrollando estrategias de “clustering” geográfico para tener perfiles epidemiológicos (Carrillo-Larco and Castillo-Cara, 2020; Hartono, 2020; Hu, Ge et al., 2020). Además, es necesario introducir componentes dinámicas en tiempo real que permitan reajustar los modelos en función de las medidas que se toman en ámbito local e internacional. Esto, sin duda, conlleva mejorar infraestructuras de datos y también el propio diseño de los modelos, que deben pasar de fases de investigación y testeo a fases de implantación y producción para ser realmente operativos. Además, es necesaria la creación de repositorios estructurados que recojan las diversas parametrizaciones de los modelos y las evidencias clínicas que soportan tales configuraciones de parámetros.

Además de los avances en modelado de propagación de las epidemias basados en movilidad y redes de contactos, la crisis del COVID-19 ha propiciado la aplicación de las herramientas contemporáneas de “Machine Learning” a series temporales para predecir el comportamiento de las curvas (Zou and Hastie, 2005; Liu, Clemente et al., 2020). El uso de diversas arquitecturas de Deep Learning permite entender los patrones temporales en los datos (Bandyopadhyay and Dutta, 2020; Huang, Chen et al., 2020). Estos modelos tienen potenciales implicaciones en el diseño de políticas y estructuración de las fases de confinamiento y desescalada. Para poder entrenar estos modelos de forma adecuada es necesario disponer de datos de entrenamiento que no siempre están disponibles (Fong, Li et al., 2020). Por ejemplo, para la crisis del COVID-19 se han usado datos de gripe de varios años siendo una enfermedad con un comportamiento muy distinto, datos de búsquedas en internet y redes sociales (Yang, Santillana et al., 2015; Lu, Hattab et al., 2019) o “small data” epidemiológico (Bullock, Pham et al., 2020).

Algunas enfermedades tienen un alto componente ambiental en su propagación, ya sea por condiciones para que los virus habiten o por favorecer una alta densidad de vectores (Wesolowski, Stresman et al., 2014; Rajarethinam, Ong et al., 2019). Por ello, es importante integrar datos ambientales y sociales en los modelos con alta resolución y actualizados en tiempo real, comparando con series temporales de base. Para enfermedades en las que la vía de propagación principal es de persona a persona, se ha demostrado que



es necesario modelar los casos asintomáticos y su contribución a la propagación (Mizumoto, Kagaya et al., 2020). Cada vez hay más investigación orientada a identificar biomarcadores y tener estudios clínicos para poder controlar casos asintomáticos o entender diferentes respuestas inmunes, paso necesario para configurar parámetros clave en los modelos (Nicholas, Petra et al., Kermali, Khalsa et al., 2020; Shi, Wang et al., 2020). El fenómeno de super-propagación ha cobrado vital importancia en la pandemia COVID-19 ya que la alta capacidad de infección y la variabilidad de la distribución de R (dispersión k) en la población da lugar a eventos de super-propagación y también a un rol clave de super-propagadores individuales.

La sociedad necesita nuevas herramientas para gestionar el riesgo sistémico que suponen las epidemias. Esto implica usar la información de forma más inteligente y aprovechando la complejidad de los sistemas sociales a favor del control de la propagación, en un proceso de acción inverso al de la propia propagación. El riesgo es multidimensional y aunque la dimensión sanitaria es la más prioritaria durante muchas fases de la pandemia, es necesario considerar el riesgo en lo económico, en la desigualdad social, en las libertades y derechos o en la educación y desarrollo cognitivo y psicológico de gran parte de la población.

SEGUIMIENTO Y MEDICIÓN DE IMPACTO

La pandemia de COVID-19 ha globalizado las estrategias no farmacéuticas utilizadas para frenar epidemias como son el confinamiento y el seguimiento de contactos y movilidad (“contact tracing”) (Eames and Keeling, 2003). El objetivo de estas estrategias es reducir el ritmo reproductivo básico (R_0 media de R) por lo cual es el resultado básico de los modelos predictivos (Dietz, 1993; Liu, Ajelli et al., 2018), aunque los fenómenos de super-propagación suponen un reto adicional crítico para estas estrategias. Estas estrategias son más o menos restrictivas socialmente dependiendo de la morbilidad, medios de transmisión y mortalidad de la enfermedad (Glass, Glass et al., 2006). Para hacer el seguimiento, además de fuentes epidemiológicas tradicionales y datos clínicos, se pueden utilizar muy diversas fuentes de datos entre las que destacan los datos de móviles (Oliver, Lepri et al., 2020), datos de búsquedas y redes sociales (Lamos, Moura et al., 2020) u otros datos generados mediante dispositivos móviles. En general, estas iniciativas siguen en un estado embrionario de investigación y no han madurado para tener impacto claro en la toma de decisiones, aunque se están adoptando con rapidez en la lucha contra la COVID-19. Algunas organizaciones internacionales están planteando una visión global de análisis para entender el impacto en los más vulnerables a través de diversos sistemas sociales para entender mejor cómo una pandemia impacto en la población mundial (UNICEF, 2020).

Como se ha mencionado, uno de los sistemas que se han desarrollado durante la pandemia del COVID-19 es el “contact tracing” que puede inferirse a través de Bluetooth o de proximidad física basada en coordenadas GPS. Existen diversas arquitecturas, especialmente la centralizada (como PEPP-PT) y la descentralizada (DP-3T) que gestionan de forma diferente el cálculo de los contactos creados de personas que han sido infectadas y que por tanto crean alta probabilidad de contagio. A pesar del gran potencial para control epidemiológico (por ejemplo creando matrices de contactos desagregadas), el uso de la tecnología tiene que venir acoplado con un sistema de gobernanza y seguridad de los datos y los resultados que sea apropiado para respetar los derechos de las personas a nivel individual y colectivo.

Gran parte del seguimiento de una epidemia se basa en la información de diversas fuentes. Como se detalla anteriormente, el filtrado de la información es clave para un seguimiento adecuado y evitar efectos nocivos



(World Health Organization [WHO], 2020). Este efecto se amplifica por las nuevas plataformas digitales que permiten no sólo propagar la información, sino también usarla para generar opinión o con fines comerciales (Mejova and Kalimeri, 2020; Singh, Bansal et al., 2020). Al respecto, se están desarrollando acciones reparadoras concretas como la curación de contenidos o chatbots de ayuda (Bullock, Pham et al., 2020).

EVALUACIÓN DE MEDIDAS

Poder evaluar las medidas frente a epidemias es clave para una mejora continua de los sistemas de respuesta a corto, medio y largo plazo. También es clave para fines políticos y por tanto es importante que haya mecanismos transparentes y fiables para medir el impacto de tales medidas. La pandemia del COVID-19 ha propiciado mucho avance e innovación en esta área debido a la severidad de las medidas y el consecuente impacto social y económico. Algoritmos de Deep Learning se han usado con datos epidemiológicos para modelar diversos escenarios de la pandemia a nivel global (Hu, Ge et al., 2020) y medir el efecto de la cuarentena (Dandekar and Barbastathis, 2020). Los datos de móviles también se han usado con éxito para medir los efectos de las medidas de distanciamiento social en diferentes áreas y grupos poblacionales (Martín-Calvo, Aleta et al., 2020; UNICEF, 2020). Un aspecto clave en la pandemia de la COVID-19, que podrá ser aplicado a otras pandemias, es entender la contribución de medidas no-farmacéuticas en cada región e intentar aislar el impacto de cada medida de mitigación (Cowling, Ali et al., 2020; Davies, Kucharski et al., 2020; Flaxman, Mishra et al., 2020; Lai, Ruktanonchai et al., 2020). En la actual pandemia, es de vital relevancia como las medidas iniciales tendrán que ser adaptadas o complementadas para controlar potenciales futuros brotes y olas (Aleta, Martín-Corral et al., 2020).

PERSPECTIVAS

Las enfermedades tienen varias dimensiones: molecular, genética, clínica, social y política. Debido a su complejidad, la dimensión social modula la propagación y sus consecuencias. Un mejor entendimiento de los sistemas humanos es clave para generar mecanismos sociales, políticos y sanitarios para luchar contra enfermedades de forma colectiva. Los factores que determinan la propagación e impacto de las epidemias son muy diversos y variables en cada sociedad, pero un desarrollo más sostenible puede hacernos más fuertes, robustos, resilientes y anti-frágiles a epidemias. Por ejemplo, la desigualdad puede resultar en una menor eficacia de las medidas de mitigación. Este hecho no tiene sólo una fundamentación social y humana, sino también matemática y científica: el riesgo de epidemias es sistémico y la vulnerabilidad de una pequeña parte del sistema genera un riesgo global difícil de controlar y que afecta al sistema en su conjunto (Taleb, Read et al., 2014, Norman, Bar-Yam et al., 2020). El riesgo también viene dado por la potencial variabilidad inmune de las personas: individuos y eventos super-propagadores depende de cada enfermedad y supone un reto adicional en sistemas de prevención (Lloyd-Smith, Schreiber et al., 2005; Kupferschmidt 2020).

Debido a la variabilidad de factores y dinámicas de las epidemias, es necesario tener mecanismos de actuación a diversos niveles, desde la ayuda asistencial local a personas vulnerables a mecanismos de gobernanza globales (Mello and Wang, 2020). Por ejemplo, se ha puesto de manifiesto la necesidad de crear “Puntos de Atención” locales que faciliten tests rápidos, especialmente en regiones con sistemas sanitarios débiles o muy deslocalizados o en casos de mucha incertidumbre sobre el estado epidemiológico de la



población. También las técnicas de diagnóstico remoto pueden suponer una ventaja para la comprensión de la epidemia en tiempo real. Es necesario plantear el debate sobre la coordinación internacional para tomar ciertas medidas sincronizadamente y con ello aumentar los impactos positivos de medidas severas como el confinamiento o el cierre de fronteras.

Existe una gran problemática epistemológica, operativa e incluso política para entender cómo las diferentes medidas tomadas para frenar a una epidemia realmente contribuyen a esta misión. Es considerablemente difícil aislar los diversos factores que influyen en la modificación de la curva epidemiológica que son biológicos, farmacológicos, sociales y políticos (Flaxman, Mishra et al., 2020; Lai, Ruktanonchai et al., 2020). Teniendo en cuenta la diversidad de políticas y estrategias sanitarias, cada vez más, parece necesario un sistema de medición de estos factores que pueda ser implementado gracias a esta propia diversidad de estrategias recopilando datos detallados tanto clínicos, de movilidad, socio-económicos y moleculares. Estos sistemas de análisis son clave para ofrecer las evidencias que permitan avanzar hacia mecanismos de respuesta más efectivos y ágiles. El reto científico es importante, poder inferir relaciones entre diversos factores y los hechos es altamente complejo y es necesario desarrollar protocolos de experimentación y medición.

Es necesario evaluar y diseñar un ecosistema de datos que nos permita con rigor dar respuesta a las cuestiones de la epidemiología digital, lo que implica diseñar mejores estrategias de adquisición de datos. Además, es necesario alinear datos de diversa naturaleza para poder tener mecanismos más descriptivos y que permitan afinar modelos (Wesolowski, Buckee et al., 2016). La pandemia del COVID-19 ha supuesto una revolución en cuanto a los métodos de recolección de datos (Ferretti, Wymant et al., 2020). El siguiente reto internacional es el control de la epidemia a través de las fronteras, lo cual supone retos tecnológicos y regulatorios (Kuhn, Beck et al., 2020; Mello and Wang, 2020; Vinuesa, Theodorou et al., 2020).

Potenciar las interacciones entre la investigación molecular, la práctica y los estudios clínicos y lo social es clave para una gobernanza efectiva (Luengo-Oroz, Pham et al., 2020). Detalles sobre avances recientes en investigación molecular y clínica también son de gran importancia para hacer un acoplamiento entre las medidas epidemiológicas, las capacidades del sistema sanitario, la investigación fundamental y las consecuencias clínicas, epidemiológicas y sociales (Bullock, Pham et al., 2020). Los mecanismos de comunicación entre organizaciones y expertos deben ser ágiles y soportados por evidencias en tiempo real. A su vez, es deseable que haya nuevos canales de confianza entre autoridades y la población. La Inteligencia Artificial también puede ser clave en estructurar toda la información e investigación que se genera. Por ejemplo, la pandemia del COVID-19 ha motivado la aparición de “dashboards” de datos de diversa naturaleza, sin embargo, salvo excepciones (Dong, Du et al., 2020), son iniciativas fragmentadas y que no tienen un modelo de uso claro. La Inteligencia Artificial debe ir un paso más allá, y realmente analizar diversas fuentes de datos, modelar el conocimiento, estructurar mensajes y facilitar la toma de decisiones tanto de gestores, políticos, empresarios, trabajadores y sociedad civil en general. La respuesta de la sociedad a esta clase de enfermedades tiene que aprovechar los procesos y las tecnologías exponenciales (Luengo-Oroz, Pham et al., 2020) y la inteligencia colectiva (Aleks Berditchevskaia, 2020). Es clave madurar soluciones para que puedan convertirse en verdaderos sistemas operativos.

El uso radical e invasivo de la tecnología durante el COVID-19 ha puesto de manifiesto varias cuestiones éticas de vital importancia que ya resonaban en algunos ámbitos (Berman, Carter et al., 2020; Pastor-Escuredo, 2020; Vinuesa, Theodorou et al., 2020). Un aspecto primordial es la privacidad y derechos fundamentales de



las personas y cómo puede verse afectada por el uso sin límite de sus datos justificado por las acciones para frenar la pandemia (Mello and Wang, 2020). Las buenas prácticas internacionales promueven principios como la minimización de datos, la descentralización y la concreción (en uso y alcance temporal), además de que dicho uso sea trazable y consentido por las personas. Es importante también reflexionar sobre la colaboración público-privada en torno a los datos y cómo problemas como la vigilancia masiva y agresiva pueden mermar la libertad de las personas o incrementar la brecha digital si no se definen los marcos adecuados (Mello and Wang, 2020). La urgencia y la necesidad de usar datos para controlar epidemias no debe desembocar en un escenario digital nocivo para el desarrollo social y tecnológico de la sociedad.

Todavía es pronto para predecir cómo será el futuro tras la COVID-19 (Sohrabi, Alsafi et al., 2020), es muy probable que, además del impacto económico, tenga efectos a largo plazo en usos y hábitos sociales: las formas de movilidad, medidas de higiene e incluso en relaciones sociales. Estos cambios pueden ir incluso en contra de tendencias que han ido emergiendo en los últimos tiempos. La digitalización sin duda se acelerará y ello tendrá consecuencias en todos los niveles, entre las cuales destaca el futuro del trabajo (Malone, 2004).

Las tecnologías y métodos propuestos deben ser parte de estrategias centradas en las personas, que permitan controlar las epidemias y que también ayuden a generar protección colectiva digital como complemento a otras medidas y a tratamientos farmacéuticos. Es momento de actuar y mejorar nuestros mecanismos de respuesta frente a pandemias futuras y avanzar hacia una sociedad más preparada y sostenible (General, 2019).



REFERENCIAS BIBLIOGRÁFICAS

- Aleks Berditchevskaia, K. P. (2020). "Coronavirus: seven ways collective intelligence is tackling the pandemic." Recuperado de <https://theconversation.com/coronavirus-seven-ways-collective-intelligence-is-tackling-the-pandemic-133553>.
- Aleta, A., et al. (2020). "Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19." *Nature Human Behaviour*: 1-8.
- Anderson, R. M., et al. (1992). *Infectious diseases of humans: dynamics and control*, Oxford university press.
- Bakir, V. and A. McStay (2018). "Fake news and the economy of emotions: Problems, causes, solutions." *Digital journalism* 6(2): 154-175.
- Balcan, D., et al. (2009). "Multiscale mobility networks and the spatial spreading of infectious diseases." *Proceedings of the National Academy of Sciences* 106(51): 21484-21489.
- Balcan, D., et al. (2010). "Modeling the spatial spread of infectious diseases: The GLObal Epidemic and Mobility computational model." *Journal of computational science* 1(3): 132-145.
- Balcan, D., et al. (2009). "Seasonal transmission potential and activity peaks of the new influenza A (H1N1): a Monte Carlo likelihood analysis based on human mobility." *BMC medicine* 7(1): 45.
- Bandyopadhyay, S. K. and S. Dutta (2020). "Machine learning approach for confirmation of covid-19 cases: Positive, negative, death and release." medRxiv.
- Barthélemy, M., et al. (2005). "Dynamical patterns of epidemic outbreaks in complex heterogeneous networks." *Journal of theoretical biology* 235(2): 275-288.
- Bejon, P., et al. (2010). "Stable and unstable malaria hotspots in longitudinal cohort studies in Kenya." *PLoS medicine* 7(7).
- Belderok, S.-M., et al. (2013). "Effect of travel on influenza epidemiology." *Emerging infectious diseases* 19(6): 925.
- Bell, G., et al. (2009). "Beyond the data deluge." *Science* 323(5919): 1297-1298.
- Bengtsson, L., et al. (2015). "Using mobile phone data to predict the spatial spread of cholera." *Scientific reports* 5: 8923.
- Bengtsson, L., et al. (2011). "Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti." *PLoS medicine* 8(8).
- Berman, G., et al. (2020). *Digital contact tracing and surveillance during COVID-19. General and child-specific ethical issues*.
- Blondel, V. D., et al. (2015). "A survey of results on mobile phone datasets analysis." *EPJ data science* 4(1): 10.
- Bodendorf, F. and C. Kaiser (2009). *Detecting opinion leaders and trends in online social networks. Proceedings of the 2nd ACM workshop on Social web search and mining*.
- Borge-Holthoefer, J., et al. (2013). "Emergence of influential spreaders in modified rumor models." *Journal of Statistical Physics* 151(1-2): 383-393.
- Boy, J., et al. (2019). *Towards an understanding of refugee segregation, isolation, homophily and ultimately integration in Turkey using call detail records. Guide to Mobile Data Analytics in Refugee Scenarios*, Springer: 141-164.
- Brownstein, J. S., et al. (2009). "Digital disease detection—harnessing the Web for public health surveillance." *The New England journal of medicine* 360(21): 2153.



- Bullock, J., et al. (2020). "Mapping the landscape of artificial intelligence applications against COVID-19." arXiv preprint arXiv:2003.11336.
- Bueno, C., et al. (2014). "Epidemics in partially overlapped multiplex networks." PloS one 9(3): e92200.
- Buscarino, A., et al. (2008). "Disease spreading in populations of moving agents." EPL (Europhysics Letters) 82(3): 38002.
- Candia, J., et al. (2008). "Uncovering individual and collective human dynamics from mobile phone records." Journal of physics A: mathematical and theoretical 41(22): 224015.
- Carrillo-Larco, R. M. and M. Castillo-Cara (2020). "Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: An unsupervised machine learning approach." Wellcome Open Research 5(56): 56.
- Cervellin, G., et al. (2017). "Is Google Trends a reliable tool for digital epidemiology? Insights from different clinical settings." Journal of epidemiology and global health 7(3): 185-189.
- Chinazzi, M., et al. (2020). "The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak." Science 368(6489): 395-400.
- Colizza, V., et al. (2007). "Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions." PLoS medicine 4(1).
- Cook, S., et al. (2011). "Assessing Google flu trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic." PloS one 6(8).
- Cowling, B. J., et al. (2020). "Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study." The Lancet Public Health.
- Dandekar, R. and G. Barbastathis (2020). "Neural Network aided quarantine control model estimation of global Covid-19 spread." arXiv preprint arXiv:2004.02752.
- Davies, N. G., et al. (2020). "Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study." The Lancet Public Health.
- de Arruda, G. F., et al. (2020). "Social contagion models on hypergraphs." Physical Review Research 2(2): 023032.
- de Arruda, G. F., et al. (2020). "Impact of the distribution of recovery rates on disease spreading in complex networks." Physical Review Research 2(1): 013046.
- De Montjoye, Y.-A., et al. (2013). "Unique in the crowd: The privacy bounds of human mobility." Scientific reports 3: 1376.
- Dietz, K. (1993). "The estimation of the basic reproduction number for infectious diseases." Statistical methods in medical research 2(1): 23-41.
- Dolgin, E. (2010). Targeting hotspots of transmission promises to reduce malaria, Nature Publishing Group.
- Dong, E., et al. (2020). "An interactive web-based dashboard to track COVID-19 in real time." The Lancet Infectious Diseases 20(5): 533-534.
- Eames, K. T. and M. J. Keeling (2003). "Contact tracing and disease control." Proceedings of the Royal Society of London. Series B: Biological Sciences 270(1533): 2565-2571.
- Espana, G., et al. (2020). "Impacts of K-12 school reopening on the COVID-19 epidemic in Indiana, USA." medRxiv.
- Eubank, S., et al. (2004). "Modelling disease outbreaks in realistic urban social networks." nature 429(6988): 180-184.



- Eysenbach, G. (2009). "Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet." *Journal of medical Internet research* 11(1): e11.
- Ferretti, L., et al. (2020). "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing." *Science* 368(6491).
- Finger, F., et al. (2016). "Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks." *Proceedings of the National Academy of Sciences* 113(23): 6421-6426.
- Flaxman, S., et al. (2020). "Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe." *nature*: 1-5.
- Fong, S. J., et al. (2020). "Finding an accurate early forecasting model from small dataset: A case of 2019-nCoV novel coronavirus outbreak." *arXiv preprint arXiv:2003.10776*.
- General, U. S. (2019). *The age of digital interdependence, Report of the UN Secretary-General's High-Level Panel on Digital Cooperation*.
- Ginsberg, J., et al. (2009). "Detecting influenza epidemics using search engine query data." *nature* 457(7232): 1012-1014.
- Glass, R. J., et al. (2006). "Targeted social distancing designs for pandemic influenza." *Emerging infectious diseases* 12(11): 1671.
- Gonzalez, M. C., et al. (2008). "Understanding individual human mobility patterns." *nature* 453(7196): 779-782.
- Green, M. D., et al. (2000). "Reference guide on epidemiology." *Reference Manual on Scientific Evidence* 2: 638.
- Greenwood, F., et al. (2017). "The signal code: A human rights approach to information during crisis." Harvard, MA.
- Hartono, P. (2020). "Generating Similarity Map for COVID-19 Transmission Dynamics with Topological Autoencoder." *arXiv preprint arXiv:2004.01481*.
- Hernán, M. A. and J. M. Robins (2006). "Estimating causal effects from epidemiological data." *Journal of Epidemiology & Community Health* 60(7): 578-586.
- Hu, Z., et al. (2020). "Artificial intelligence forecasting of COVID-19 in China." *arXiv preprint arXiv:2002.07112*.
- Hu, Z., et al. (2020). "Forecasting and evaluating intervention of Covid-19 in the World." *arXiv preprint arXiv:2003.09800*.
- Hua, J. and R. Shaw (2020). "Coronavirus (Covid-19) 'infodemic' and emerging issues through a data lens: The case of China." *International journal of environmental research and public health* 17(7): 2309.
- Huang, C.-J., et al. (2020). "Multiple-Input Deep Convolutional Neural Network Model for COVID-19 Forecasting in China." *medRxiv*.
- Imran, A., et al. (2020). "AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app." *Informatics in Medicine Unlocked*: 100378.
- Kermali, M., et al. (2020). "The role of biomarkers in diagnosis of COVID-19—A systematic review." *Life Sciences*: 117788.
- Kuhn, C., et al. (2020). "Covid Notions: Towards Formal Definitions--and Documented Understanding--of Privacy Goals and Claimed Protection in Proximity-Tracing Services." *arXiv preprint arXiv:2004.07723*.
- Kupferschmidt, K. (2020). "Why do some COVID-19 patients infect many others, whereas most don't spread the virus at all?" *Science*.



- Lai, S., et al. (2020). "Effect of non-pharmaceutical interventions to contain COVID-19 in China."
- Lampos, V., et al. (2020). "Tracking COVID-19 using online search." arXiv preprint arXiv:2003.08086.
- Lazer, D., et al. (2014). "The parable of Google Flu: traps in big data analysis." *Science* 343(6176): 1203-1205.
- Lazer, D. M., et al. (2018). "The science of fake news." *Science* 359(6380): 1094-1096.
- LeCun, Y., et al. (2015). "Deep learning." *nature* 521(7553): 436-444.
- Leung, G. M. and K. Leung (2020). "Crowdsourcing data to mitigate epidemics." *The Lancet Digital Health* 2(4): e156-e157.
- Liu, D., et al. (2020). "A machine learning methodology for real-time forecasting of the 2019-2020 COVID-19 outbreak using Internet searches, news alerts, and estimates from mechanistic models." arXiv preprint arXiv:2004.04019.
- Liu, Q.-H., et al. (2018). "Measurability of the epidemic reproduction number in data-driven contact networks." *Proceedings of the National Academy of Sciences* 115(50): 12680-12685.
- Lloyd-Smith, J. O., et al. (2005). "Superspreading and the effect of individual variation on disease emergence." *nature* 438(7066): 355-359.
- Lu, F. S., et al. (2019). "Improved state-level influenza nowcasting in the United States leveraging Internet-based data and network approaches." *Nature communications* 10(1): 1-10.
- Luengo-Oroz, M., et al. (2020). "Artificial intelligence cooperation to support the global response to COVID-19." *Nature Machine Intelligence*: 1-3.
- Lynch, C. and C. Roper (2011). "The transit phase of migration: circulation of malaria and its multidrug-resistant forms in Africa." *PLoS medicine* 8(5).
- Malone, T. W. (2004). *The future of work*, Audio-Tech Business Book Summaries, Incorporated.
- Martín-Calvo, D., et al. (2020). Effectiveness of social distancing strategies for protecting a community from a pandemic with a data driven contact network based on census and real-world mobility data, Working paper, <https://covid-19-sds.github.io>. Consultado el 18 de abril del 2020.
- Mejova, Y. and K. Kalimeri (2020). "Advertisers jump on coronavirus bandwagon: Politics, news, and business." arXiv preprint arXiv:2003.00923.
- Mello, M. M. and C. J. Wang (2020). "Ethics and governance for digital disease surveillance." *Science*.
- Meloni, S., et al. (2011). "Modeling human mobility responses to the large-scale spreading of infectious diseases." *Scientific reports* 1: 62.
- Menni, C., et al. (2020). "Loss of smell and taste in combination with other symptoms is a strong predictor of COVID-19 infection." medRxiv.
- Meyers, L. A., et al. (2006). "Predicting epidemics on directed contact networks." *Journal of theoretical biology* 240(3): 400-418.
- Miritello, G., et al. (2011). "Dynamical strength of social ties in information spreading." *Physical Review E* 83(4): 045102.
- Mizumoto, K., et al. (2020). "Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020." *Eurosurveillance* 25(10): 2000180.
- Morales, A. J., et al. (2014). "Efficiency of human activity on information spreading on Twitter." *Social Networks* 39: 1-11.



- Moreno, Y., et al. (2002). "Epidemic outbreaks in complex heterogeneous networks." *The European Physical Journal B-Condensed Matter and Complex Systems* 26(4): 521-529.
- Nekovee, M., et al. (2007). "Theory of rumour spreading in complex social networks." *Physica A: Statistical Mechanics and its Applications* 374(1): 457-470.
- Newman, M. E. (2002). "Spread of epidemic disease on networks." *Physical Review E* 66(1): 016128.
- Newman, M. E., et al. (2006). *The structure and dynamics of networks*, Princeton university press.
- Nicholas, G. D., et al. (2020) "Age-dependent Effects in the Transmission and Control of COVID-19 Epidemics." *Nature medicine*.
- Norman, J., et al. (26 de enero del 2020). "Systemic Risk of Pandemic via Novel Pathogens-Coronavirus: A Note." *New England Complex Systems Institute*.
- Oliver, N., et al. (2020). *Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle*, American Association for the Advancement of Science.
- Olson, D. R., et al. (2013). "Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales." *PLoS computational biology* 9(10).
- Pastor-Escuredo, D. (2020). "Ethics in the digital era." *arXiv preprint arXiv:2003.06530*.
- Pastor-Escuredo, D. and E. Frias-Martinez (2020). "Flow descriptors of human mobility networks." *arXiv preprint arXiv:2003.07279*.
- Pastor-Escuredo, D. and C. Tarazona (2020). "Characterizing information leaders in Twitter during COVID-19 crisis." *arXiv preprint arXiv:2005.07266*.
- Pastor-Satorras, R. and A. Vespignani (2001). "Epidemic dynamics and endemic states in complex networks." *Physical Review E* 63(6): 066117.
- Pastor-Satorras, R. and A. Vespignani (2001). "Epidemic spreading in scale-free networks." *Physical review letters* 86(14): 3200.
- Pastor-Satorras, R. and A. Vespignani (2002). "Epidemic dynamics in finite size scale-free networks." *Physical Review E* 65(3): 035108.
- Perrota, D. (2018). "Can Mobile Phone Traces Help Shed Light on the Spread of Zika in Colombia?". Recuperado de <https://www.unglobalpulse.org/2018/04/can-mobile-phone-traces-help-shed-light-on-the-spread-of-zika-in-colombia/>.
- Popat, K., et al. (2018). "DeClarE: Debunking fake news and false claims using evidence-aware deep learning." *arXiv preprint arXiv:1809.06416*.
- Prothero, R. M. (1977). "Disease and mobility: a neglected factor in epidemiology." *International journal of epidemiology* 6(3): 259-267.
- Pulse, U. G. (2012). "Big data for development: Challenges & opportunities." Naciones Unidas, Nueva York.
- Pulse, U. G. (2015). "Mapping the risk-utility landscape: mobile data for sustainable development and humanitarian action." *Global Pulse Project Series* no18.
- Rajarethinam, J., et al. (2019). "Using human movement data to identify potential areas of Zika transmission: case study of the largest Zika cluster in Singapore." *International journal of environmental research and public health* 16(5): 808.
- Riley, S. (2007). "Large-scale spatial-transmission models of infectious disease." *Science* 316(5829): 1298-1301.



- Ruchansky, N., et al. (2017). Csi: A hybrid deep model for fake news detection. Proceedings of the 2017 ACM on Conference on Information and Knowledge Management.
- Salathé, M. (2018). "Digital epidemiology: what is it, and where is it going?" Life sciences, society and policy 14(1): 1.
- Salathe, M., et al. (2012). "Digital epidemiology." PLoS computational biology 8(7).
- Sanz, J., et al. (2014). "Dynamics of interacting diseases." Physical Review X 4(4): 041005.
- Shi, Y., et al. (2020). COVID-19 infection: the perspectives on immune responses, Nature Publishing Group.
- Shu, K., et al. (2017). "Fake news detection on social media: A data mining perspective." ACM SIGKDD Explorations Newsletter 19(1): 22-36.
- Simini, F., et al. (2012). "A universal model for mobility and migration patterns." nature 484(7392): 96.
- Singh, L., et al. (2020). "A first look at COVID-19 information and misinformation sharing on Twitter." arXiv preprint arXiv:2003.13907.
- Singhanian, S., et al. (2017). 3han: A deep neural network for fake news detection. International Conference on Neural Information Processing, Springer.
- Sohrabi, C., et al. (2020). "World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)." International Journal of Surgery.
- Stoddard, S. T., et al. (2009). "The role of human movement in the transmission of vector-borne pathogens." PLoS neglected tropical diseases 3(7).
- Sun, K., et al. (2020). "Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study." The Lancet Digital Health.
- Taleb, N. N. (2019). The Statistical Consequences of Fat Tails, STEM Publishing.
- Taleb, N. N., et al. (2014). "The precautionary principle: fragility and black swans from policy actions." NYU Extreme Risk Initiative Working Paper: 1-24.
- Taleb, N. N., et al. (2014). "The precautionary principle (with application to the genetic modification of organisms)." arXiv preprint arXiv:1410.5787.
- Tatem, A. J., et al. (2014). "Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning." Malaria journal 13(1): 52.
- Tizzoni, M., et al. (2014). "On the use of human mobility proxies for modeling epidemics." PLoS computational biology 10(7).
- United Nations International Children's Emergency Fund (2020). "Magic Box COVID-19 report." Recuperado de <https://www.unicef.org/innovation/magicbox/covid>.
- Vaezi, A. and S. H. Javanmard (2020). "Infodemic and risk communication in the era of CoV-19." Advanced Biomedical Research 9.
- Vespignani, A., et al. (2020). "Modelling COVID-19." Nature Reviews Physics: 1-3.
- Vinuesa, R., et al. (2020). "A socio-technical framework for digital contact tracing." arXiv preprint arXiv:2005.08370.
- Wesolowski, A., et al. (2016). "Connecting mobility to infectious diseases: the promise and limits of mobile phone data." The Journal of infectious diseases 214(suppl_4): S414-S420.
- Wesolowski, A., et al. (2012). "Quantifying the impact of human mobility on malaria." Science 338(6104): 267-270.



- Wesolowski, A., et al. (2015). "Impact of human mobility on the emergence of dengue epidemics in Pakistan." *Proceedings of the National Academy of Sciences* 112(38): 11887-11892.
- Wesolowski, A., et al. (2014). "Quantifying travel behavior for infectious disease research: a comparison of data from surveys and mobile phones." *Scientific reports* 4: 5678.
- World Health Organization (2020). "Infodemic management - Infodemiology;" Recuperado de <https://www.who.int/teams/risk-communication/infodemic-management>.
- Yang, S., et al. (2015). "Accurate estimation of influenza epidemics using Google search data via ARGO." *Proceedings of the National Academy of Sciences* 112(47): 14473-14478.
- Zarocostas, J. (2020). "How to fight an infodemic." *The Lancet* 395(10225): 676.
- Zhang, J., et al. (2020). "Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China." *Science*.
- Zou, H. and T. Hastie (2005). "Regularization and variable selection via the elastic net." *Journal of the royal statistical society: series B (statistical methodology)* 67(2): 301-320.
- Zufria, P. J., et al. (2018). "Identifying seasonal mobility profiles from anonymized and aggregated mobile phone data. Application in food security." *PloS one* 13(4): e0195714.