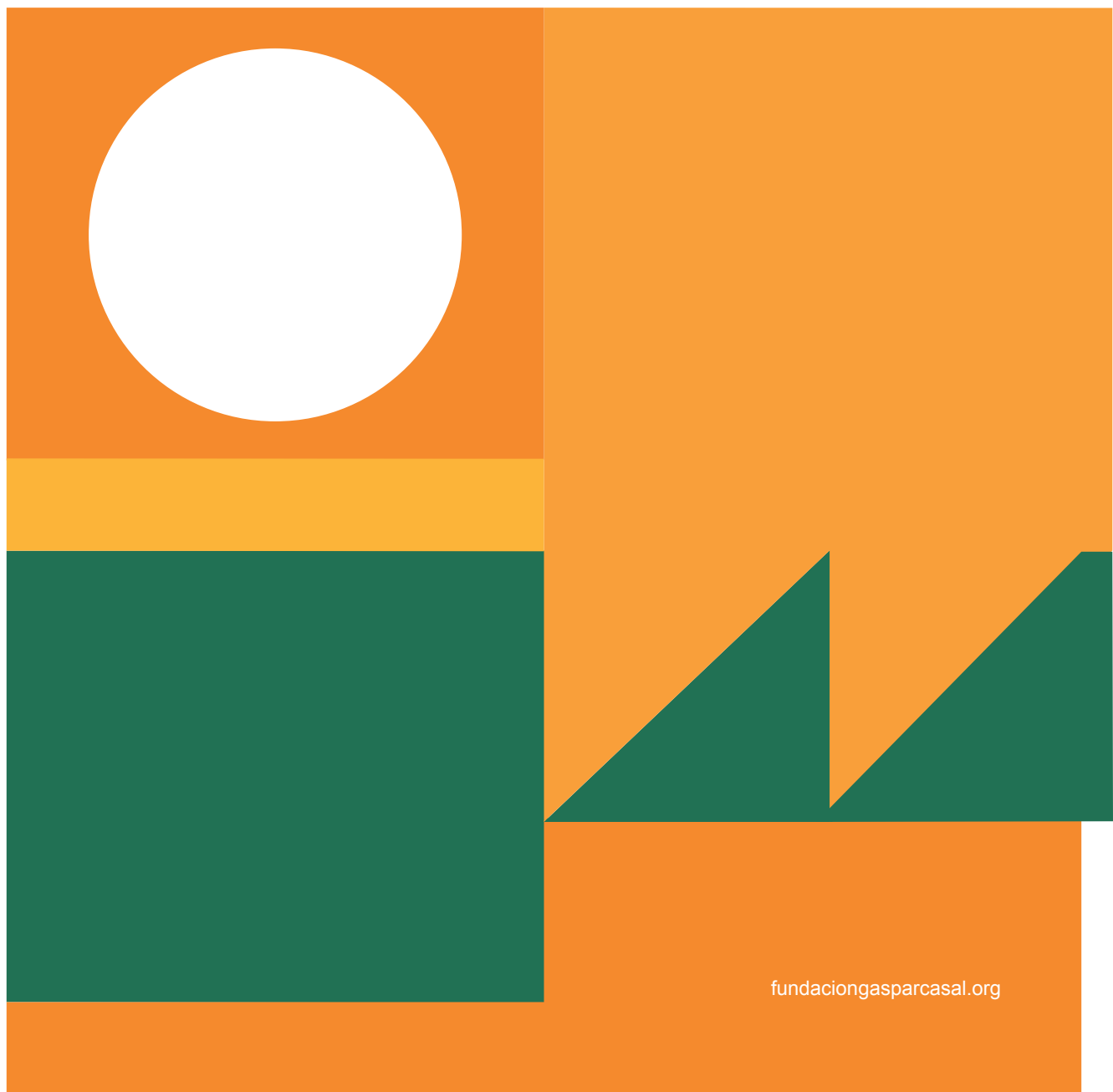


Salud Pública de Precisión



fundaciongasparcasal.org

SALUD PÚBLICA, DE PRECISIÓN

SALUD PÚBLICA, DE PRECISIÓN



Salud Pública de Precisión
GSK
Fundación Gaspar Casal

NP-ES-VX-IMU-210001

ISBN: 978-84-0000-00-0

Depósito legal: M-00000-2021

Materias: Medicina, Matemáticas y Ciencia, Salud pública, Tecnología, Ingeniería, Computación e Informática.

Formato: 170 × 240 mm

Páginas: 174

Coordinadores: Alicia del Llano Núñez-Cortés, Roberto Nuño Solinís
y Juan E. del Llano Señarís

© 2021 Fundación Gaspar Casal
www.fundaciongasparcasal.org
C/ General Díaz Porlier 78, 8A
28006 Madrid

Edita: Libroacadémico, S.L.
Tel.: 91 550 02 60
info@libroacademico.com
www.libroacademico.com

A efectos de transparencia, le informamos que GSK ha colaborado en la financiación de la presente publicación. Su contenido refleja las opiniones, criterios, conclusiones y/o hallazgos propios de los autores, los cuales pueden no coincidir necesariamente con los de GSK. GSK recomienda siempre la utilización de sus productos de acuerdo con la ficha técnica aprobada por las autoridades.

Los panelistas del grupo de discusión no recibieron honorarios por parte de GSK por su participación en el trabajo.

ÍNDICE

Autores	9
Introducción	11
1. Conceptos fundamentales	11
2. Un nuevo paradigma en Salud Pública.....	13
3. Hacia el camino de la Salud Pública de Precisión.....	17
4. De la teoría a la práctica: áreas de mayor desarrollo de la Salud Pública de Precisión	19
5. Conclusión	22
Revisión Sistemática de la Literatura	29
1. Introducción	29
2. Métodos	30
3. Búsqueda bibliográfica	30
4. Resultados	32
5. Discusión	35
6. Conclusión	37

Informe de Experto: el papel de la epidemiología genética	59
1. Diferencias entre la tradicional y la nueva salud pública	59
2. Determinantes de la salud pública	61
3. Nuevas tecnologías en la Salud Pública de Precisión	63
4. Uso de la epidemiología genética en la Salud Pública de Precisión	68
5. Conclusión y recomendaciones futuras.....	79
Análisis del grupo de discusión	89
1. Justificación y objetivos	89
2. Métodos y participantes	90
3. Preguntas.....	91
4. Transcripciones	92
5. Resultados del análisis focal	154
Hoja de ruta.....	163

AUTORES

Diego Benavent

Reumatólogo, Hospital Universitario La Paz de Madrid

Juan Ramón González

Associate Research Profesor, ISGlobal

Juan del Llano

Director de la Fundación Gaspar Casal

INTRODUCCIÓN

1. CONCEPTOS FUNDAMENTALES

Salud Pública: disciplina que tiene como objeto de estudio la salud de la población. Se considera la ciencia encargada de prevenir la enfermedad y la discapacidad, prolongar la vida y fomentar la salud física y mental mediante esfuerzos llevados a cabo en la comunidad. De este modo, trata de mantener un nivel de vida adecuado a través de actividades organizadas por las administraciones públicas y con la participación de la sociedad.

Medicina de Precisión: enfoque emergente para el tratamiento y prevención de enfermedades que tiene en cuenta las variaciones en los genes, el medio ambiente y el estilo de vida para llevar a cabo intervenciones adecuadas a cada paciente concreto en cada momento. No sólo engloba diagnóstico y terapéutica, sino también puede ayudar a la prevención de las enfermedades.

Macrodatos o Big data: conjunto de nuevas tecnologías y arquitecturas diseñadas para la obtención de valor mediante el análisis y procesamiento de grandes volúmenes de datos de muy diversas características y de una forma rápida (1). Tradicionalmente, los datos se presentan con cuatro dimensiones (*volumen, variedad, velocidad y veracidad*) y pueden ser tanto estructurados como no estructurados.

Inteligencia artificial o Artificial Intelligence (AI): rama de la informática cuyo objetivo es la simulación de tareas tradicionalmente relacionadas con la inteligencia humana, que son llevadas a cabo por parte de ordenadores. Incluye procesos como aprendizaje (integración de nueva información y reglas para su manejo), razonamiento o mejora mediante retroalimentación.

Genómica: campo de la biología molecular encargado del estudio de los genes, sus funciones y su influencia en el crecimiento, desarrollo y funcionamiento del organismo. Para ello usa una amplia variedad de técnicas de estudio del ADN pero, a diferencia de la genética, que estudia genes de forma individual, la genómica estudia el conjunto de todos los genes que se encuentran en un organismo.

Prevención individualizada: forma de prevención en salud pública en la que los profesionales de la salud tienen en cuenta las características, estilo de vida, antecedentes familiares, anamnesis, factores de riesgo y medicación del individuo para poder llevar a cabo propuestas que tengan como objetivo mantener o mejorar la calidad de vida del individuo (2).

Prevención personalizada: forma de prevención en salud pública que incluye las actividades de prevención individualizada y en la que también se consideran aspectos biológicos como biomarcadores, genética, transcriptómica, proteómica y metabolómica del individuo para llevar a cabo propuestas que tengan como objetivo mantener o mejorar su calidad de vida.

Prevención de precisión: forma de prevención en salud pública que incluye las actividades de prevención personalizada y en la que también se consideran la situación socioeconómica y los datos de comportamiento del individuo para llevar a cabo propuestas que tengan como objetivo mantener o mejorar su calidad de vida.

Epidemiología molecular: rama de la epidemiología que incluye medidas biológicas a nivel molecular para valorar los factores de riesgo genéticos y ambientales, y su contribución a la etiología, distribución y prevención de enfermedades en la población. Sus técnicas se pueden emplear para medir la exposición a factores de riesgo, la respuesta biológica o las características del huésped que influyen en su susceptibilidad (3).

Sistema de información geográfica (GIS, por sus siglas en inglés): marco de trabajo para reunir, gestionar y analizar datos mediante la integración de diversas fuentes. Gracias al análisis de la ubicación espacial y la organización de capas de información para su visualización, se pueden crear mapas y escenas tridimensionales para revelar el conocimiento de patrones y relaciones en los datos (4).

2. UN NUEVO PARADIGMA EN SALUD PÚBLICA

Medicina y salud pública son conceptos que presentan una estrecha relación y que mantienen un objetivo en común, que es el abordaje de los problemas de salud de la población. Mientras la medicina se centra en la atención sanitaria de los individuos, la salud pública se centra en las poblaciones como sujeto de acción para promover la salud y la prevención de las enfermedades. Para ello, aborda causas sociales y ambientales de falta de salud, así como la ausencia de equidad sanitaria en la población.

En esta búsqueda de la mejora de la salud de la población por parte de medicina y salud pública, se pueden llevar a cabo diferentes acciones sobre los individuos y la población. Las acciones de menor impacto en la población son las acciones clínicas y de atención de la salud; por el contrario, ciertas acciones de salud pública como el abordaje de los factores socioeconómicos (vivienda, educación, etc) son las que tienen un mayor impacto en la salud poblacional (5). En ese sentido, parece razonable centrar los esfuerzos en la implementación de estrategias que aborden dichos factores poblacionales como medio de prevenir enfermedades, y así mejorar la salud global. Las nuevas tecnologías de la información y la comunicación, incluyendo la capacidad de análisis de datos y gestión de la información, abren una importante ventana de oportunidad para poder enfocarse en esta línea.

A lo largo de este libro analizaremos cómo el uso de ingentes cantidades de datos, incluyendo análisis genómicos, clínicos sociales y ambientales, puede llevarnos a la mejora de la salud de la población.

¿QUÉ ES LA SALUD PÚBLICA DE PRECISIÓN?

El concepto de la Salud Pública de Precisión (SPP) fue acuñado por Dr. Tarun Weeramanthri en 2013 y se introdujo en la literatura académica en los años

siguientes en una búsqueda de establecer una modernización de la vigilancia, la epidemiología y los sistemas de información, con el objetivo de implementar intervenciones dirigidas que se llevaran a cabo con una perspectiva de salud de la población (6). Posteriormente, los doctores Khoury y Galea describieron la SPP como *“la mejora de la capacidad para prevenir enfermedades, promover la salud y reducir las diferencias de salud en la población mediante la aplicación de métodos y tecnologías emergentes que midan la enfermedad, los patógenos, las exposiciones, los comportamientos y la susceptibilidad en la población; y el desarrollo de políticas y programas de aplicación específicos con el objetivo de mejorar la salud”* (7).

La salud pública tradicional se basa principalmente en datos epidemiológicos con el objetivo de establecer prioridades en política sanitaria, impulsar la mejora de la salud global en la población y centrarse en los grupos desfavorecidos. Para ello, se apoya en la vigilancia de factores de riesgo y enfermedades, cribado, desarrollo de intervenciones y evaluación de la implementación de estas.

La SPP va un paso más allá, mediante la aplicación de tecnologías que describan y analicen con mayor exactitud a los individuos y su entorno a lo largo de la vida, con el mismo objetivo de adaptar las intervenciones preventivas a ciertos grupos de riesgo y mejorar la salud general poblacional (8). Para ello, el uso de los macrodatos y las nuevas tecnología informáticas son fundamentales, por lo que algunos autores sugieren que estos avances son la característica definitoria del enfoque de precisión de la salud pública (9). Los macrodatos, que pueden provenir de diferentes fuentes de información, permiten desarrollar tareas de computación masiva mediante complejos sistemas de gestión de información y análisis. Si bien no aportaría demasiado valor el disponer únicamente de estos datos brutos, la información que se extrae tras su procesamiento supone un cambio de paradigma en la investigación gracias a la mejora de las capacidades de análisis descriptivo, predictivo y prescriptivo. En los últimos años, la convivencia de diferentes tecnologías está siendo fundamental para maximizar el valor de los macrodatos, en especial las que llegan de la mano de la inteligencia artificial (AI, del inglés *Artificial Intelligence*). Los sistemas de AI se entrenan mediante programación escrita por humanos que genera algoritmos capaces de actualizar sus parámetros de partida y, con ello, modificar su arquitectura original en función de los datos a los que sean expuestos (10). Con este cambio

de paradigma se abre la puerta a una nueva dimensión cognitiva de los profesionales de la salud, en que las tecnologías utilizadas buscan proporcionar “la intervención adecuada, en el momento adecuado, a la población adecuada” (11).

Muchas estrategias de prevención y salud pública antes de la llegada del concepto de SPP ya se basaban en principios de selección y estratificación, en la posibilidad de analizar perfiles genéticos personalizados, o en la búsqueda de datos precisos de cada individuo (12). A pesar de que la atención a los determinantes relacionados con los datos y la precisión no sean nuevos, son evidentes los beneficios que pueden ser obtenidos tanto para el individuo como para la población con el aumento de información en la última década (13). Entre otros, el uso de los macrodatos pretende mejorar la detección de pacientes con factores de riesgo sociales o conductuales cuyo abordaje podría reducir la carga de la enfermedad y mejorar la gestión sanitaria; aumentar la precisión diagnóstica y terapéutica, mediante una mejor estratificación del riesgo y una mayor atención al contexto de cada paciente en las decisiones clínicas; o mejorar los sistemas de salud para adaptar los servicios a las necesidades de la población a la que sirven.

Por concretar estos objetivos con un ejemplo, entre otros muchos que comentaremos a lo largo del libro, el impacto de la SPP se puede apreciar en una investigación que evidencia que hasta el 90% de las enfermedades infecciosas relacionadas con los mosquitos en todo el mundo (como el dengue o el chikungunya) podrían abordarse centrándose en sólo el 14% de las zonas geográficas de alto riesgo (14). En una época profundamente marcada por la pandemia producida por la COVID-19, el uso de los datos para optimizar intervenciones que benefician a la población permite vislumbrar el profundo impacto que la SPP puede tener a nivel global.

MEDICINA DE PRECISIÓN VS SALUD PÚBLICA DE PRECISIÓN

Según entendemos actualmente el concepto de SPP se asemejaría al de medicina de precisión, pero a nivel poblacional. Como medicina de precisión entendemos el uso de herramientas diagnósticas y de tratamientos que se dirigen de forma específica a las necesidades individuales de los pacientes, basándose en características genéticas, psicosociales y biomarcadores. Aunque existen muchas descripciones de medicina de precisión, la que probablemente tiene

mayor alcance es la desarrollada por el National Institutes of Health (NIH), que la describe como *“abordaje emergente para la prevención y tratamiento que tiene en cuenta la variabilidad individual, ambiental y el estilo de vida de cada persona”*. A pesar de que en dicha definición se mencione la prevención de enfermedades de cada persona, debemos tener en cuenta que la evidencia emergente apoya la idea de que los factores de riesgo individuales solo explican de forma limitada la variación del riesgo de enfermedad a nivel de la población (15) (16). Las diferencias en la salud de los individuos parecen ser atribuibles en gran medida a sus condiciones de vida de forma diaria que se acumulan a lo largo de la vida (17), existiendo un limitado poder explicativo de los factores de riesgo a nivel individual. Así, independientemente de que la medicina de precisión logre su objetivo a nivel del paciente individual, es probable que no se produjeran los cambios esperados en la salud poblacional solo con ello.

Para ilustrar este paradigma, imaginemos que se hubieran aplicado las herramientas de la medicina de precisión hace más de dos siglos, en el contexto del descubrimiento del agente causante del escorbuto. Puede que el médico naval que describió la enfermedad, el Dr. James Lind, hubiera buscado tratamiento específico para cada marino según su susceptibilidad genética a la enfermedad, implicando complejos protocolos de diagnóstico y tratamiento que hubiera que adaptar a cada marino para proporcionar con la dosis exacta adecuada de ácido ascórbico para no desarrollar los síntomas del escorbuto. Este complejo proceso no solo hubiera sido difícil de implementar, sino que probablemente hubiera dejado de lado a los estratos sociales más desfavorecidos, sin poder recomendarles la solución que requerían para prevenir su enfermedad. Sin embargo, mediante una intervención más sencilla consistente en el cambio de la alimentación de la tripulación, añadiendo en el cargamento del barco algunas frutas y alimentos con alto contenido en vitamina C, se consiguieron cambiar las circunstancias para que no se desarrollase el escorbuto de manera general. Por tanto, parece razonable buscar el equilibrio entre la atención individual y los factores de riesgo poblacionales evitables que aparecen en las condiciones de la vida diaria.

Mientras que la medicina de precisión requiere datos genéticos, de estilo de vida y ambientales para cumplir los objetivos de tratamiento potencialmente individualizados, la SPP se refiere a una mayor granularidad en la definición de las cohortes públicas para la implementación de intervenciones específicas en

ellas (18). En uno de los primeros artículos en que se describe la SPP, Khoury y cols (6) proporcionan algunos ejemplos del uso de nuevas tecnologías para mejorar la precisión en salud pública, como la genómica de patógenos, la mejora de la vigilancia epidemiológica o las intervenciones selectivas a miembros de la población en riesgo. Sin embargo, la SPP implica algo más que el uso de tecnologías. Se trata de utilizar como unidad de intervención a las “poblaciones” que utilizan esas tecnologías que alimentan la medicina de precisión. Por ello, es fundamental mantener el enfoque de que estamos tratando de mejorar la salud a nivel de la población, utilizando datos más precisos para orientar la prevención y el control de las enfermedades y así mejorar la salud global.

3. HACIA EL CAMINO DE LA SPP

IMPORTANCIA DE LOS MACRODATOS Y LA INTELIGENCIA ARTIFICIAL EN LA SPP

La digitalización de los sistemas de salud está generando inmensas cantidades de datos de carácter clínico y demográfico, para cuyo análisis se requiere de sistemas de computación cada vez más complejos y software especializado. Con ello se está propiciando un escenario en que el análisis de datos biomédicos está permitiendo la implementación de metodologías de AI, con la perspectiva de optimizar procesos, diagnósticos y tratamientos.

Recapitulando las características de los macrodatos, es importante remarcar sus características de *volumen*, *variedad* y *velocidad* para entender su impacto en la SPP. Así, el *volumen* ha propiciado avances médicos en áreas que requieren mucho input de datos para poder extraer información relevante, como son la genómica (19), la explotación de imágenes médicas (20) o las enfermedades con mayor impacto en la salud de la población (21). El hecho de permitir que se integren una gran *variedad* de datos para conseguir una visión más completa del paciente o de la población, ha permitido investigaciones en diversas áreas, como el análisis de la calidad del aire (22) o los *wearables* (23). La mayor *velocidad* de recogida de datos ha facilitado el desarrollo en áreas que se pueden beneficiar de la disponibilidad continua de estos, como el diagnóstico o la predicción clínica (24).

Son muchas las disciplinas dentro del espectro de la salud pública que están aprovechando la potencial vigilancia de la salud de la población a gran escala

y en tiempo real. Algunos casos de uso de macrodatos en salud pública incluyen las enfermedades infecciosas (25), la ciencia de la salud ambiental (26), la salud comunitaria (27), la epidemiología (28), la salud materno-infantil (29), o la seguridad laboral (30).

FUENTES DE (MACRO)DATOS EN SPP

Una de las áreas donde mejor se puede apreciar la importancia de los macrodatos en el desarrollo de la SPP es la genómica. De hecho, la salud pública y la genómica han presentado una relación de sinergia en las últimas décadas; así, la salud pública actuó de catalizador en la aplicación de la medicina genómica en el examen de los recién nacidos en muchos países del mundo, incluso antes de la aparición del concepto de medicina de precisión. El cribado de recién nacidos es probablemente el programa de SPP más grande que existe en el mundo, en el que se determinan las condiciones genéticas o metabólicas de los recién nacidos de forma masiva. Con ello se consigue diagnosticar a los bebés en las primeras etapas de su vida de diversas condiciones que presentan alto riesgo de morbilidad y discapacidad prevenibles, como fibrosis quística o anemia drepanocítica, mediante intervenciones que se inicien en una etapa temprana de la vida. Antes de que existiera este cribado de recién nacidos muchos de estos bebés se perdían sin poder llegar a tener acceso a una intervención precoz para su enfermedad.

Debido al importantísimo desarrollo de las técnicas de secuenciación en las últimas dos décadas y su consecuente disminución del coste, se han desarrollado varias iniciativas de generación de datos genómicos a gran escala, como el Exome Aggregation Consortium (ExAC) (31), que incluye la secuenciación de más de 60,000 genomas, o el atlas del genoma del cáncer (TCGA), donde se han secuenciado miles de muestras de tumores (32). Se ha estimado que hasta una cuarta parte de la población de los países desarrollados probablemente tendrá secuenciado del genoma en el año 2025 (33). Este volumen de información abre la puerta a la comprensión de nuevos mecanismos de enfermedad, así como a la predicción del efecto de intervenciones preventivas, lo que mantiene una firme promesa de revolucionar la SPP.

Una aplicación de la genómica específica en SPP es la genómica de patógenos, que se ha convertido en una poderosa herramienta con numerosas

aplicaciones en el seguimiento y el control de los brotes de enfermedades infecciosas (34). Un ejemplo de la utilización de esta genómica de los patógenos durante la pandemia por COVID-19, es la llevada a cabo en los Países Bajos, con la combinación de análisis de secuencias del genoma completo (whole-genome sequence analyses) y datos epidemiológicos para proporcionar evaluaciones fiables de las pautas de transmisión del SARS-CoV-2 en la comunidad. Con estos datos sobre la transmisión se pudieron fundamentar las decisiones relativas a la cancelación de las reuniones masivas y las recomendaciones relativas a los aislamientos (35).

Más allá de la genómica, los datos de vigilancia epidemiológica detallados son esenciales para orientar las intervenciones de salud pública. De este modo, la información geográfica puede hacerse más precisa mediante el uso de nuevas tecnologías para identificar los puntos conflictivos y dirigir las intervenciones. Como ejemplo, durante la evaluación de la mortalidad infantil en África, se identificaron importantes disparidades por regiones gracias a mapas que incluyeron una resolución espacial elevada; estas diferencias se habrían pasado por alto con un análisis a nivel nacional con métodos tradicionales (11). Del mismo modo, la iniciativa norteamericana “End the HIV Epidemic” se centra en 50 regiones de Estados Unidos donde se producen más del 50% de los nuevos diagnósticos de VIH (36).

4. DE LA TEORÍA A LA PRÁCTICA: ÁREAS DE MAYOR DESARROLLO DE LA SPP

Para el desarrollo de la SPP se requieren datos de múltiples disciplinas –ómicas, y nuevos métodos estadísticos para su procesamiento (37). Algunos de estos métodos computacionales como el modelado geoespacial, la epidemiología molecular o el análisis de clases latentes, se benefician de los macrodatos para identificar subpoblaciones concretas. A continuación, se exponen estudios donde se han utilizado estas metodologías en diversos ámbitos; el alcance limitado de esta síntesis se verá complementado con la revisión sistemática de la literatura que se realizará de forma posterior.

VIGILANCIA DE ENFERMEDADES

El uso de grandes datos en la SPP ha demostrado su utilidad en la vigilancia de enfermedades. Algunos ejemplos son el control de resistencia antimicrobiana y

la vigilancia de brotes de cólera o dengue (38) (39) (40). Para ello, se realiza seguimiento de las personas afectadas por dicha enfermedad (portadores o pacientes), cuyo movimiento se considera un componente crítico en la dinámica de transmisión de muchos patógenos transmitidos por vectores (41). Los datos a gran escala facilitan este seguimiento y las nuevas tecnologías apoyan métodos alternativos para rastrear en tiempo real a los individuos afectados (25). Mecanismos como la utilización de datos de teléfonos móviles, las búsquedas en la web, el rastreo mediante geoposicionamiento, Twitter o el seguimiento de datos de pasajeros en vuelos han demostrado ser prometedores para delinear ese movimiento en diversas enfermedades (42) (43) (44) (45).

ESTRATIFICACIÓN Y PREDICCIÓN DEL RIESGO

La predicción de riesgos con el objetivo de establecer intervenciones preventivas es uno de los campos más prometedores en SPP. Tras prometedores proyectos que no se convirtieron en éxitos prácticos, donde destaca la sonada decepción inicial con Google Flu Trends, se ha ganado conciencia de las barreras y los riesgos de los macrodatos. Así, se tratan de evitar modelos que solo utilicen datos de Internet o de redes sociales, o que presenten sobreajuste por disponer de poca heterogeneidad de casos para el aprendizaje (46). Se han presentado algunos ejemplos de estudios que buscan la predicción del riesgo en enfermedades como el ébola, la diabetes o el tabaquismo (47). Por ejemplo, en este último caso se investigó las relaciones condicionales de ciertas variaciones genéticas de las cuales se habían descrito que presentaban efectos en la continuación del tabaquismo en los jóvenes. Así, se observó que la relación de los *clusters* que habían sido descritos como protectores o predictores de uso de tabaco, estaban condicionados a diferentes niveles de consumo de cigarrillos de base (48). Con ello se entiende que juntar varios estratos de datos de diferentes niveles puede proporcionar información más precisa para la predicción de riesgos de enfermedades.

INTERVENCIONES TERAPÉUTICAS DIRIGIDAS

La aplicación de intervenciones terapéuticas a grupos con características comunes dentro de una población heterogénea se ha realizado en salud pública desde hace muchos años. Un ejemplo clásico se publicó ya en los años 90, con la descripción de un modelo que identificaba a personas con alto riesgo de contraer una infección de transmisión sexual, que eran por tanto susceptibles de seguir una intervención intensiva para reducir ese riesgo (49). Otros ejemplos

más recientes son la identificación de subgrupos de mujeres con mayor riesgo de no seguir las directrices de realización de screening de cáncer de mama, o la identificación de “puntos calientes” de transmisión de paludismo para intervenir en ellos y reducir la transmisión de la enfermedad (50) (51).

Mediante la utilización de los macrodatos han surgido distintas iniciativas de SPP que han permitido una mayor precisión de estas intervenciones terapéuticas a grupos concretos. Por ejemplo, en población infantil cada vez es más aceptado que en lugar de la existencia de una única curva de desarrollo en la población, los niños pertenecen a subpoblaciones distintas con diferentes trayectorias de desarrollo. Por ello, se realizó un estudio en que utilizaron nuevas técnicas estadísticas para identificar las trayectorias del Índice de Masa Corporal (IMC) en niños y analizar las asociaciones entre estas trayectorias distintas y las actitudes alimentarias, sedentarias y de actividad física. En él, se concluyó que se podrían optimizar las intervenciones centrándose en los comportamientos de alto riesgo en la infancia precoz, especialmente en los niños de baja estatura, los niños de etnias no occidentales y los niños cuya madre presentaba sobrepeso (52).

EPIDEMIOLOGÍA DE LAS ENFERMEDADES

La disponibilidad de grandes cantidades de datos de gran variedad está posibilitando la descripción precisa de ciertas enfermedades y facilitando así su comprensión. Esta caracterización precisa se ha utilizado en áreas diversas dentro de la SPP como la diabetes o la adicción a opioides (53) (54). En estos casos se utilizó información genómica para describir variantes genéticas con consecuencias en los distintos fenotipos de cada enfermedad. Otro ejemplo es el estudio del golpe de calor, para lo cual se estudiaron las olas de calor mediante el análisis de visitas a urgencias, ingresos hospitalarios y datos de mortalidad. En este estudio, se concluyó que las hospitalizaciones y las visitas a urgencias relacionadas con el calor son las dos medidas más sensibles para evaluar el impacto de una ola de calor en la salud poblacional (55).

5. CONCLUSIÓN

La implementación de las estrategias de la SPP supone la aparición de nuevos retos que implican a la sociedad en su conjunto. Así, según se puede interpretar

de algunos de los ejemplos ya analizados, la SPP se muestra como una prometedora disciplina para los ciudadanos y los sistemas sanitarios. No obstante, se deberá encontrar el equilibrio entre riesgo y beneficio en su aplicación, ya que se presentan importantes implicaciones éticas, sociales y jurídicas. A lo largo de este libro analizaremos la más reciente evidencia en relación con la SPP para poder contextualizar su uso en la sociedad actual.

REFERENCIAS

1. Cuadernos de Gestión Sanitaria Número 4: Datos de Vida Real y Macro-Datos: Su Incorporación a la Evaluación - Fundación Gaspar Casal. Available at: <https://fundaciongasparcasal.org/publicaciones/cuadernos-de-gestion-sanitaria-numero-4-datos-de-vida-real-y-macro-datos-su-incorporacion-a-la-evaluacion/>. Accessed February 12, 2021.
2. Bíró K, Dombrádi V, Jani A, Boruzs K, Gray M. Creating a common language: defining individualized, personalized and precision prevention in public health. *J Public Health (Oxf)* 2018.
3. La “epidemiología molecular”: nueva ruta de investigación o compañero de viaje? Available at: <https://iris.paho.org/handle/10665.2/15552>. Accessed February 12, 2021.
4. Jia P, Cheng X, Xue H, Wang Y. Applications of geographic information systems (GIS) data and methods in obesity-related research. *Obes Rev* 2017;18(4):400–11.
5. Frieden TR. The Future of Public Health. *N Engl J Med* 2015;373:1748–1754. Available at: <http://www.nejm.org/doi/10.1056/NEJMsa1511248>. Accessed February 12, 2021.
6. Khoury MJ, Iademarco MF, Riley WT. Precision Public Health for the Era of Precision Medicine. *Am J Prev Med* 2016.
7. Khoury MJ, Galea S. Will precision medicine improve population health? *JAMA - J Am Med Assoc* 2016;316(13):1357–8.

8. Weeramanthri TS, Dawkins HJS, Baynam G, Bellgard M, Gudes O, Semmens JB. Editorial: Precision Public Health. *Front Public Heal* 2018.
9. Dowell SF, Blazes D, Desmond-Hellmann S. Four steps to precision public health. *Nature* 2016.
10. Mesko B. The role of artificial intelligence in precision medicine. *Expert Rev Precis Med Drug Dev* 2017;2:239–241. Available at: <https://www.tandfonline.com/doi/full/10.1080/23808993.2017.1380516>.
11. Khoury MJ, Engelgau M, Chambers DA, Mensah GA. Beyond Public Health Genomics: Can Big Data and Predictive Analytics Deliver Precision Public Health? *Public Health Genomics* 2019.
12. Taylor-Robinson D, Kee F. Precision public health - The Emperor's new clothes. *Int J Epidemiol* 2019.
13. Diez Roux A V., Katz M, Crews DC, Ross D, Adler N. Social and Behavioral Information in Electronic Health Records: New Opportunities for Medicine and Public Health. *Am J Prev Med* 2015.
14. Dowell SF, Blazes D, Desmond-Hellmann S. Four steps to precision public health. *Nature* 2016;540:189–191. Available at: <http://www.nature.com/news/four-steps-to-precision-public-health-1.21089>. Accessed February 12, 2021.
15. Marmot MG, Rose G, Shipley M, Hamilton PJS. Employment grade and coronary heart disease in British civil servants. *J Epidemiol Community Health* 1978.
16. Marmot MG, Stansfeld S, Patel C, North F, Head J, White I, et. al. Health inequalities among British civil servants: the Whitehall II study. *Lancet* 1991.
17. Marmot M, Friel S, Bell R, Houweling TA, Taylor S. Closing the gap in a generation: health equity through action on the social determinants of health. *Lancet* 2008.
18. Dolley S. Big data's role in precision public health. *Front Public Heal* 2018.

19. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, *et. al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015.
20. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology* 2016.
21. Hripcsak G, Ryan PB, Duke JD, Shah NH, Park RW, Huser V, *et. al.* Characterizing treatment pathways at scale using the OHDSI network. *Proc Natl Acad Sci U S A* 2016.
22. Zheng Y, Liu F, Hsieh HP. U-Air: When urban air quality inference meets big data. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*; 2013.
23. Chen M, Zhang Y, Li Y, Hassan MM, Alamri A. AIWAC: Affective interaction through wearable computing and cloud technology. *IEEE Wirel Commun* 2015.
24. Ahmad S, Ramsay T, Huebsch L, Flanagan S, McDiarmid S, Batkin I, *et. al.* Continuous multi-parameter heart rate variability analysis heralds onset of sepsis in adults. *PLoS One* 2009.
25. Hay SI, George DB, Moyes CL, Brownstein JS. Big Data Opportunities for Global Infectious Disease Surveillance. *PLoS Med* 2013.
26. Braem B, Latré S, Leroux P, Demeester P, Coenen T, Ballon P. Designing a smart city playground: Real-time air quality measurements and visualization in the City of Things testbed. In: *IEEE 2nd International Smart Cities Conference: Improving the Citizens Quality of Life, ISC2 2016 - Proceedings.*; 2016.
27. Consolvo S, McDonald DW, Toscos T, Chen MY, Froehlich J, Harrison B, *et. al.* Activity sensing in the wild: A field trial of UbiFit Garden. In: *Conference on Human Factors in Computing Systems - Proceedings.*; 2008.
28. Lardon J, Abdellaoui R, Bellet F, Asfari H, Souvignat J, Texier N, *et. al.* Adverse drug reaction identification and extraction in social media: A scoping review. *J Med Internet Res* 2015.

-
29. Daley D, Bachmann M, Bachmann BA, Pedigo C, Bui MT, Coffman J. Risk terrain modeling predicts child maltreatment. *Child Abus Negl* 2016.
 30. Bragazzi NL, Dini G, Toletone A, Brigo F, Durando P. Leveraging big data for exploring occupational diseases-related interest at the level of scientific community, media coverage and novel data streams: The example of silicosis as a pilot study. *PLoS One* 2016.
 31. Anon. ExAC Browser | re3data.org. Available at: <https://www.re3data.org/repository/r3d100012122>. Accessed February 12, 2021.
 32. Anon. The Cancer Genome Atlas Program - National Cancer Institute. Available at: <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>. Accessed February 12, 2021.
 33. Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, *et. al.* Big Data: Astronomical or Genomical? *PLoS Biol* 2015;13:e1002195.
 34. Rasmussen SA, Khoury MJ, Rio C Del. Precision Public Health as a Key Tool in the COVID-19 Response. *JAMA - J Am Med Assoc* 2020.
 35. Oude Munnink BB, Nieuwenhuijse DF, Stein M, O'Toole Á, Haverkate M, Mollers M, *et. al.* Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. *Nat Med* 2020.
 36. Katz I, Jha AK. HIV in the United States: Getting to Zero Transmissions by 2030. *JAMA - J Am Med Assoc* 2019.
 37. Dolley S. Big data's role in precision public health. *Front Public Heal* 2018;6:1. Available at: www.frontiersin.org. Accessed February 12, 2021.
 38. MacFadden DR, Fisman D, Andre J, Ara Y, Majumder MS, Bogoch II, *et. al.* A platform for monitoring regional antimicrobial resistance, using online data sources: Resistanceopen. *J Infect Dis* 2016.
 39. Chunara R, Andrews JR, Brownstein JS. Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. *Am J Trop Med Hyg* 2012.

40. Gomide J, Veloso A, Meira W, Almeida V, Benevenuto F, Ferraz F, *et. al.* Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. In: *Proceedings of the 3rd International Web Science Conference, WebSci 2011.*; 2011.
41. Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Soldan VP, Kochel TJ, Kitron U, *et. al.* The role of human movement in the transmission of vector-borne pathogens. *PLoS Negl Trop Dis* 2009.
42. Finger F, Genolet T, Mari L, Magny GC De, Manga NM, Rinaldo A, *et. al.* Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. *Proc Natl Acad Sci U S A* 2016.
43. Wesolowski A, Qureshi T, Boni MF, Sundsøy PR, Johansson MA, Rasheed SB, *et. al.* Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proc Natl Acad Sci U S A* 2015.
44. Huff A, Allen T, Whiting K, Breit N, Arnold B. FLIRT-ing with Zika: A Web Application to Predict the Movement of Infected Travelers Validated Against the Current Zika Virus Epidemic. *PLoS Curr* 2016.
45. Vazquez-Prokopec GM, Bisanzio D, Stoddard ST, Paz-Soldan V, Morrison AC, Elder JP, *et. al.* Using GPS Technology to Quantify Human Mobility, Dynamic Contacts and Infectious Disease Dynamics in a Resource-Poor Urban Environment. *PLoS One* 2013.
46. Gandon S, Day T, Metcalf CJE, Grenfell BT. Forecasting Epidemiological and Evolutionary Dynamics of Infectious Diseases. *Trends Ecol Evol* 2016.
47. Phan TP, Alkema L, Tai ES, Tan KHX, Yang Q, Lim W-Y, *et. al.* Forecasting the burden of type 2 diabetes in Singapore using a demographic epidemiological model of Singapore. *BMJ Open Diabetes Res Care* 2014.
48. Pugach O, Cannon DS, Weiss RB, Hedeker D, Mermelstein RJ. Classification Tree Analysis as a Method for Uncovering Relations Between CHRNA5A3B4 and CHRNA3A6 in Predicting Smoking Progression in Adolescent Smokers. *Nicotine Tob Res* 2017.

-
49. Richert CA, Peterman TA, Zaidi AA, Ransom RL, Wroten JE, Witte JJ. A method for identifying persons at high risk for sexually transmitted infections: Opportunity for targeting intervention. *Am J Public Health* 1993.
 50. Gomez S, Tan S, Keegan THM, Clarke CA. Disparities in mammographic screening for Asian women in California: A cross-sectional analysis to identify meaningful groups for targeted intervention. *BMC Cancer* 2007.
 51. Bousema T, Griffin JT, Sauerwein RW, Smith DL, Churcher TS, Takken W, *et. al.* Hitting hotspots: Spatial targeting of malaria for control and elimination. *PLoS Med* 2012.
 52. Koning M, Hoekstra T, Jong E De, Visscher TLS, Seidell JC, Renders CM. Identifying developmental trajectories of body mass index in childhood using latent class growth (mixture) modelling: associations with dietary, sedentary and physical activity behaviors: a longitudinal study. *BMC Public Health* 2016.
 53. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, *et. al.* An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* 2017.
 54. Smith AH, Jensen KP, Li J, Nunez Y, Farrer LA, Hakonarson H, *et. al.* Genome-wide association study of therapeutic opioid dosing identifies a novel locus upstream of OPRM1. *Mol Psychiatry* 2017.
 55. Xiao J, Spicer T, Jian L, Yun GY, Shao C, Nairn J, *et. al.* Variation in population vulnerability to heat wave in Western Australia. *Front Public Heal* 2017.

REVISIÓN SISTEMÁTICA DE LA LITERATURA

1. INTRODUCCIÓN

La salud pública de precisión (SPP) es un nuevo campo dentro del área de la salud pública que aprovecha los avances tecnológicos para realizar descripciones y análisis precisos de individuos en grupos de población, con el objetivo de mejorar su salud global. La expansión de los datos clínicos a gran escala junto con las nuevas técnicas de procesamiento de estos, están siendo determinantes en el desarrollo de herramientas que permitan obtener conclusiones que conduzcan a prácticas médicas y de salud pública más precisas, a través de la mejora de la prevención, la detección, el diagnóstico o el tratamiento de enfermedades (1) (2).

Recientemente, el concepto de SPP ha sido objeto de un gran debate, habiendo cobrado protagonismo incluso en las revistas médicas de mayor impacto (3). Si bien hay defensores y detractores del término, lo cierto es que las nuevas metodologías de las que hace uso la SPP están trayendo un cambio de paradigma con un impacto indiscutible.

Junto con las grandes expectativas generadas por la SPP, llegan muchos interrogantes sobre la factibilidad de las estrategias de análisis y su implementación. En esta revisión sistemática de la literatura enmarcamos los artículos que hacen uso del término de “salud pública de precisión” para realizar

investigaciones clínicas o epidemiológicas, centrándonos en las diversas aplicaciones de la SPP y la metodología usada para gestión y el análisis de los grandes datos de diferentes fuentes.

2. MÉTODOS

Se ha llevado a cabo una revisión sistemática de la literatura (RSL). El proceso se realizó en acuerdo a los estándares de la Colaboración Cochrane (Cochrane Handbook for Systematic Reviews of Interventions) y en acuerdo con la declaración PRISMA (PRoposal to Improve the publication of Systematic reviews and Meta-Analyses).

La localización de la evidencia científica de la revisión sistemática se ha desarrollado en 4 pasos:

- Desarrollo de la estrategia y búsqueda bibliográfica.
- Selección de los artículos relevantes.
- Lectura crítica de los artículos finalmente seleccionados.
- Tabla de resultados resumen de los estudios analizados

3. BÚSQUEDA BIBLIOGRÁFICA

Se utilizan para la localización de los artículos relevantes las bases de datos de la Web Of Science (WOS), Google y Medline mediante el motor de búsqueda OVID SP.

Criterios para considerar los estudios en la presente revisión:

- Criterios de inclusión:

Población del estudio: se seleccionaron aquellos estudios originales cuya población comprende participantes humanos de cualquier edad y sexo.

Exposición: se localizaron aquellos estudios que emplearan Big Data en sector sanitario y cuyos autores incluyeran el término “salud pública de

precisión”, indistintamente de que abordasen aspectos clínicos o financieros-administrativos, investigadores o éticos-sociales.

Resultados: se incluyen aquellos estudios que incluyan resultados con impacto en diferentes áreas de la salud de la población general o de la comunidad.

Se seleccionaron ensayos clínicos, estudios observacionales, análisis de evaluación económica y series de casos. Se analizan los resultados obtenidos dentro del periodo comprendido entre 2013 y 2020.

- **Criterios de exclusión:**

Se excluyeron revisiones narrativas, editoriales y cualquier tipo de artículo en que no se explicitara una metodología de investigación.

Se excluyen también aquellos estudios no directamente relacionados con el sector de servicios sanitarios, investigación sanitaria o la industria farmacéutica o de dispositivos sanitarios.

Se excluyen trabajos no publicados en lengua inglesa o en español.

- **Estrategia de búsqueda**

Se utilizó como descriptor o palabras claves, “precision public health” en el título de los trabajos publicados. Los filtros metodológicos disponibles en cada una de las bases de datos consultados cribaron los resultados obtenidos limitando el periodo temporal propuesto en los últimos 7 años.

- **Selección de los estudios**

Tras el proceso de búsqueda de los estudios dos autores revisaron de forma independiente los resultados (Benavent D, Quecedo L).

En una primera fase, se eliminaron los duplicados y se revisaron los estudios en formato de título y *abstract* en función de los criterios de elegibilidad preestablecidos. En una segunda fase, los estudios fueron revisados nuevamente por los dos autores a texto completo. Los desacuerdos se resolvieron mediante discusión o con la participación de un tercer autor

(Gol-Monserrat J). El resultado de esta fase determina el número total de estudios incluidos.

El resumen del proceso de búsqueda se publica mediante el esquema PRISMA. Tras la lectura detallada de los artículos seleccionados se evalúa la calidad de las evidencias, utilizamos las recomendaciones Prisma para las revisiones sistemáticas de la literatura y SRQR para los estudios cualitativos.

4. RESULTADOS

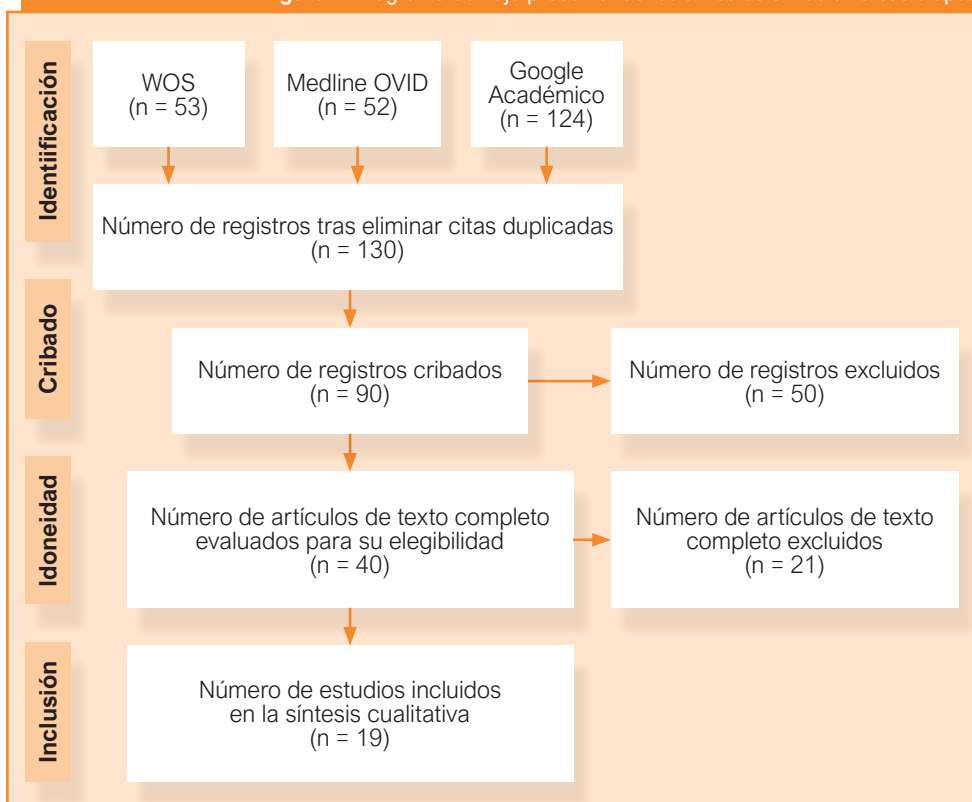
La estrategia inicial arrojó un total de 229 resultados tras la búsqueda en las bases de datos electrónicas: Medline 52, Google 124 y WOS 53 artículos. Tras la exclusión de duplicados se mantuvieron un total de 130 trabajos. Tras el análisis de los títulos y abstracts se excluyeron 90 trabajos y tras la lectura en texto completo otros 21 fueron descartados, al detectarse incumplimientos en al menos una de las condiciones anteriormente citadas, en los criterios de inclusión o exclusión, quedando finalmente seleccionados 19 estudios. El resultado de la búsqueda está recogido en el diagrama PRISMA adjunto (Figura 1).

Debido a la heterogeneidad en las publicaciones, no fue posible realizar un metaanálisis de los resultados, realizando una revisión cualitativa de los artículos. Debido a que los estudios han sido realizados con una extensa variedad metodológica y empleando distintos tipos de herramientas de análisis de datos e inteligencia artificial, se presentan los resultados según la disciplina predominante a la que se circunscribe el trabajo.

Estudios incluidos:

En la Tabla 1 pueden encontrarse de forma detallada los 19 estudios incluidos. Estos estudios fueron principalmente de tipo observacional. Algunos de los trabajos se plantearon en un marco teórico, sin llegar a implementarse en una cohorte de pacientes, si bien la mayoría son estudios retrospectivos sobre grandes áreas poblacionales.

Tras la revisión en texto completo de los artículos, se realizó una estratificación de los estudios según la disciplina predominante sobre la que se enfoca el

Figura 1: Diagrama de flujo presentando los artículos en las diversas etapas

estudio de investigación. Así, se estratificaron los trabajos según las siguientes categorías: patología infecciosa, endocrinología-nutrición, obstetricia, cardiología, oncología y farmacología. En la [Figura 2](#) se pueden ver los estudios incluidos estratificados por área médica, siendo medicinas infecciosas la disciplina que agrupó el mayor número de trabajos.

Según el año de publicación, se incluye 1 estudio publicado en el año 2016, 2 en el 2017, 3 en el 2018, 4 en el 2019 y 9 en el 2020.

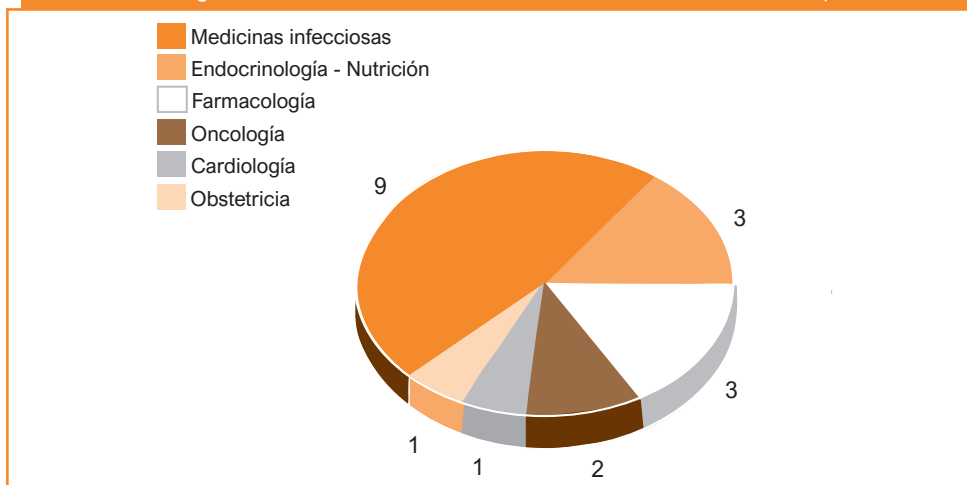
La población incluida en los estudios varía entre 2901 casos de infecciones respiratorias agudas graves de *Cummings et al*, para su vigilancia y caracterización, a los 15,072,746 niños cuyos datos se analizan en el estudio de *Reiner et. al* con el objetivo de estimar las distribuciones de prevalencia, incidencia y mortalidad de diarrea entre 2000 y 2017 en diversos niveles subnacionales.

Las fuentes de obtención de los datos son variadas. Así, se obtiene información de bases nacionales o registros de salud en 7 de los artículos, de visita médica o programas de salud en 6, y de herramientas de medida tipo cuestionario en 6 artículos más; no obstante, en la mayoría de los artículos se obtienen datos de múltiples fuentes. En varios artículos se utiliza información de bases de datos de aspectos como variables climáticas (*Ferrao et al*, 2018), genéticas (*Nagar et al*) o de actividad de búsqueda en Internet (*Liu et al*). Además, en 11 de los artículos se hace referencia a parámetros geográficos. Si bien la metodología más utilizada son los modelos de geoestadística Bayesiana, la metodología es muy heterogénea, desde la estadística descriptiva en diferentes regiones, estadística espacial (análisis de puntos calientes) o algoritmos de *clustering*.

En 5 de los estudios se utilizan modelos predictivos. *Franchini et. al.* desarrollan un algoritmo para la predicción del diagnóstico de insuficiencia cardiaca en la población; *Helfinstein et. al.* identifican predictores del retraso en la búsqueda de atención sanitaria en pacientes con tuberculosis; *Huang et. al.* analizan los factores asociados con el parto en un centro sanitario; *Zhou et. al.* utilizan algoritmos predictivos para analizar la variabilidad genética de la dihidropirimidina deshidrogenasa y la tiopurina metiltransferasa; *Liu et. al.* desarrollan un algoritmo para la predicción en tiempo real de la evolución de la actividad de COVID-19.

Dada la gran variedad en la metodología y en las áreas de enfoque de los estudios que se incluyen en la revisión, no se realiza una comparativa entre los mismos.

Figura 2: Estudios incluidos en la revisión sistemática, estratificados por área médica



5. DISCUSIÓN

En esta revisión sistemática de la literatura se repasan los trabajos publicados en relación con el uso de macrodatos y tecnologías basadas en inteligencia artificial para analizar cualquier tipo de intervención que afecte la salud de la población general o comunidad y cuyos autores utilizaran el término SPP.

Como se puede ver en los resultados, se han desarrollado trabajos en diversas disciplinas con una importante variedad en las fuentes de obtención de los datos, en los modelos de Inteligencia Artificial y en la metodología empleada. Llama la atención que, a pesar de incluirse la revisión de la literatura de los últimos 7 años, la mayoría de los estudios se agrupa en los tres últimos años del estudio (2018-2020). El concepto de SPP se propuso inicialmente por el Dr. Tarun Weeramanthri en 2013 basado en su experiencia en genómica, tecnología espacial en la salud y vinculación de datos (24). No obstante, su introducción en la literatura académica fue posterior, según se puede comprobar en los datos obtenidos de esta revisión. Así, fue en 2016 cuando *Khoury et. al.* plantearon el término con el objetivo de implementar intervenciones dirigidas que se llevaran a cabo con una perspectiva de salud de la población (25).

En los estudios incluidos, la población y el territorio analizado son muy variables, acorde a los objetivos de cada estudio concreto. Así, el rango de poblaciones en los estudios analizados va desde los 2901 casos de infecciones respiratorias agudas graves de origen incierto (5) en Uganda al análisis de 15,072,746 niños menores de 5 años en 94 países (21). Además, en uno de los artículos se realiza una simulación con la población total de India, para estimar la propagación del virus del Ébola en caso de un hipotético brote (22).

Destaca la diversidad de fuentes de obtención de variables que se emplean en los estudios. Las fuentes más empleadas son bases de datos nacionales o registros de salud, datos de visita médica o cuestionarios, junto con datos geográficos obtenidos por diversas metodologías. Si bien no todos los estudios incluyen variables geográficas, en más de la mitad de los artículos (11) sí se hace algún tipo de análisis que relacione la situación geográfica con los resultados obtenidos.

Una de las metodologías usada en varios de los trabajos emplea modelos geoestadísticos que ofrecen análisis de datos detallados sobre poblaciones a escalas geográficas muy precisas con el fin de orientar las intervenciones de salud pública en las comunidades.

Los sistemas de información geográfica (GIS, del inglés Geographic Information Systems) emplean sofisticadas plataformas de software y hardware para analizar y cotejar datos sobre la distribución geoespacial de la incidencia de las enfermedades, los riesgos y los resultados de la salud (26). Mientras que la tecnología y sus aplicaciones están dando sus primeros pasos, los estudios preliminares han demostrado que el GIS puede ser una modalidad útil para guiar la política de salud pública y obtener datos de alta precisión sobre los riesgos de la población, las comorbilidades, los cambios en las tasas de incidencia, la eficacia de los programas de tratamiento y los factores socioeconómicos asociados al riesgo (27,28). Hay varios ejemplos ilustrativos de estudios incluidos que utilizan técnicas de geolocalización. Así, en el estudio de *Ferrao et. al.* de 2018 (6), emplean modelización espacial basadas en GIS junto con variables climáticas, sociodemográficas y clínicas, para crear un mapa con áreas de riesgo de malaria. *Lynch et. al.* utilizan un geolocalizador basado en GIS para determinar la incidencia ajustada de cáncer de hígado según datos censales de Pensilvania, Estados Unidos (15). Mediante el desarrollo de un marco geoestadístico basado en modelos bayesianos, *Reiner et. al.* relacionan variables geoespaciales con probable implicación en la diarrea, para estimar las distribuciones de prevalencia, incidencia y mortalidad de diarrea entre 2000 y 2017 en diversos niveles subnacionales (21).

Si bien los métodos estadísticos para el análisis de datos en los distintos estudios presentan un amplio rango, desde la estadística descriptiva a complejos algoritmos de predicción basados en machine learning, la mayoría de los trabajos utilizan análisis de agrupaciones de casos para establecer asociaciones. Estos análisis pueden realizarse por diferentes métodos, pero lo más habitual en los trabajos encontrados es la utilización de *clustering*, un tipo de inteligencia artificial no supervisada que trata de agrupar los casos por similitud en relación con determinadas variables. Gracias al poder de estas agrupaciones, se obtienen conclusiones como que los clústeres de SARI (del inglés Severe acute respiratory infection) tienen más probabilidades de producirse en zonas

urbanas con mayor densidad de población y durante las dos temporadas de lluvia anuales (5); la estimación del riesgo de paludismo según diferentes áreas en Chimoio (7); agrupaciones de la población que difieren en la búsqueda de atención por tuberculosis (TBC) (9); desigualdades en las tasas de notificación de enfermedades de transmisión sexual (ETS) en distintas regiones de Portugal (12); disparidades regionales en el retraso del crecimiento en Argentina (18); regiones con distinta mortalidad por diarrea en distintas partes del mundo (21).

Dado el impacto que dichas metodologías de extracción y análisis de datos pueden tener en la patología infecciosa, una de las potenciales aplicaciones es en el estudio y seguimiento de epidemias. Son varios los artículos incluidos que analizan la propagación de virus; concretamente, *Liu et. al.* utilizan el término SPP en un artículo que tiene como objetivo la predicción en tiempo real de la evolución de la actividad del COVID-19 (14). En dicho estudio se utilizan datos de diversas fuentes, como informes médicos, datos del Centers for Disease Control and Prevention (CDC) de infecciones de COVID-19, actividad de búsqueda en Internet relacionada con COVID-19 (Baidu), actividad en los medios de comunicación, previsiones diarias de la actividad de COVID-19, y previsiones del modelo de simulación epidemiológica GLEAM (Global Epidemic And Mobility). Mediante técnicas de agrupamiento, se obtiene una visualización gráfica de resultados, que permite la previsión de la evolución del COVID-19.

6. CONCLUSIÓN

Más allá de la nomenclatura precisa y del uso del término SPP, parece claro el reciente aumento de estudios que utilizan avances tecnológicos y metodologías novedosas para realizar descripciones y análisis precisos de individuos en grupos de población con el objetivo de mejorar la salud global. Así, está abriendo un nuevo campo de investigación en salud, que permitirá realizar la intervención adecuada, en la población adecuada y en el momento adecuado.

Tabla 1: Tabla de evidencia. Principales características de los estudios incluidos.

#	Estudio	Población	Objetivo de la investigación
1	Anderson LN <i>et. al.</i> (4)	5185 niños, 3-11 años	Identificar clases de niños en base a variables relacionadas con la obesidad y evaluar las asociaciones entre las clases identificadas y sobrepeso u obesidad
2	Cummings MJ <i>et. al.</i> (5)	2901 casos de SARI. Análisis de secuenciación en 176 pacientes	Vigilancia sindrómica de SARI en Uganda para la detección y caracterización de patógenos respiratorios virales con potencial epidémico
3	Ferrao JL <i>et. al.</i> (2018) (6)	Población de Chimoio, Mozambique. 174 km ² 324,816 habitantes, en 2017	Creación de un mapa para identificar y modelizar áreas de riesgo de malaria
4	Ferrao JL <i>et. al.</i> (2016) (7)	237,497 habitantes de Chimoio, Mozambique, en 2007. Registro de 9 años (2006-2014).	Establecer la espacialidad y temporalidad del paludismo y describir las características sociodemográficas de los pacientes con dicha enfermedad en Chimoio

	Variables	Metodología de análisis de datos	Medidas de resultado
	15 variables relacionadas con familia, metabolismo, comportamientos de salud y actividad escolar	Análisis de clases latentes. Regresión logística multinomial	Clases de riesgo según los factores para obesidad. Magnitud del riesgo (OR) en función del número de características de riesgo de cada clase
	Grupos de SARI de origen incierto. Muestras nasofaríngeas para la captura de secuenciación de oligonucleótidos (VirCapSeq-VERT) y vinculación con estrategias epidemiológicas convencionales	Análisis espacio-temporal y secuenciación genómica	Detección de clústeres y su distribución geográfica y temporal. Relación de clústeres con genotipo viral.
	Variables climáticas, sociodemográficas y clínicas: temperatura, precipitaciones, altitud, distancia al agua y carreteras, índice de vegetación, utilización de la tierra, prevalencia del paludismo y densidad de población. Estos recursos se obtuvieron de: Bioclimatic (1950 hasta 2000) (WorldClim), ArcGIS 10.2, imágenes Landsat, modelo de elevación digital	Proceso analítico jerárquico. Modelización espacial	Peso relativo de cada factor individual para el mapa. Mapa con áreas de riesgo de malaria
	Datos semanales sobre casos de paludismo (incluye datos demográficos básicos). Datos del área de residencia de los pacientes	Análisis de tendencias lineales, y ANOVA multifactorial. Análisis de clústeres temporal y espacial	Tasa de incidencia. Fracción Atribuible de malaria (número de casos de malaria/ total de visitas por centro)

(continúa en la pág. siguiente)

5	Franchini M <i>et. al.</i> (8)	153,393 personas mayores de 14 años que acudieron al médico de atención primaria	Establecer y validar un algoritmo para la identificación de pacientes que presentan o están en riesgo de padecer insuficiencia cardíaca
6	Helfinstein S <i>et. al.</i> (9)	Encuesta de 84,625 hogares en Chennai, India. 1667 personas con síntomas de TBC entre 2018-2019.	Identificar perfiles de personas en riesgo de retrasar la búsqueda de atención sanitaria.
7	Huang VS <i>et. al.</i> (11)	15,120 miembros de familias (5968 mujeres, 4199 hombres y 4953 matriarcas)	Identificar los factores que impulsan la prestación de partos en centros sanitarios para el establecimiento de prioridades y objetivos de intervención.
8	Jardim C <i>et. al.</i> (12)	Datos de 308 municipios en Portugal con una población media de 33,575 habitantes (rango 459-504471)	Trazar un mapa de los patrones espaciales de clamidia, gonorrea y sífilis; identificar las zonas de alta tasa de ETS; determinar su asociación con la situación socioeconómica, el nivel de urbanización y la densidad de población.

Archivo de todos los residentes en la Toscana que reciben asistencia por parte del sistema nacional de salud italiano. Informes de alta del hospital. Recetas de medicamentos Certificaciones de enfermedades crónicas. Información de consultas externas de visitas, pruebas de laboratorio, imagen y prescripción	Algoritmo basado en códigos de fármacos y códigos CIE-9 como marcadores de IC. Validación doble	Validación de algoritmo CARPEDIEM, con el reporte de métricas de predicción del diagnóstico del algoritmo tanto con el de los médicos de cabecera como con el estándar clínico.
El cuestionario incluyó ítems que evaluaban la búsqueda de atención sanitaria, las percepciones generales de servicios de salud; el apoyo social; el estado general de salud, las creencias y los prejuicios; los canales de información; el conocimiento de las enfermedades, las creencias y el estigma; y el perfil sociodemográfico.	Análisis multivariado y <i>clustering</i> no supervisado para identificar predictores de retrasar búsqueda de atención sanitaria.	Características asociadas con una mayor probabilidad de haber buscado atención de un médico (OR=desplazamiento de 1 DE de la media). División en clústeres, porcentaje de personas que no buscan atención médica diagnosticadas con TBC en cada segmento.
Encuesta con: – Variables demográficas – Comportamiento durante el embarazo (controles pre-parto, toma de suplementos de hierro y ácido fólico, etc). – Planificación de nacimientos, barreras de identificación, opinión sobre servicios e infraestructuras, percepción del riesgo, planificación financiera para el parto.	Estadística descriptiva. Modelo predictivo (regresión logística). Redes Bayesianas causales (ML). Árboles de decisión	Factores asociados con tener parto en un centro (OR= probabilidad de parto en centro sanitario [público o privado] vs en casa) Árbol de decisión que dividen a las mujeres en grupos según sus características
Notificaciones de ETS a nivel municipal para el período de 2015 a 2017 del Sistema Nacional de Vigilancia Epidemiológica de Portugal	Modelos de geoestadística Bayesiana.	Tasas de notificación normalizadas. Zonas de tasas bajas y altas de riesgo. Riesgo relativo (RR) con Intervalos de Credibilidad (CrI)

(continúa en la pág. siguiente)

9	Joost S <i>et. al.</i> (13)	15,423 adultos de una población transversal en Ginebra	Investigar la existencia de agrupaciones espaciales de frecuencia de ingesta de bebidas azucaradas y del IMC, y su posible superposición espacial.
10	Liu <i>et. al.</i> (14)	32 provincias en China con casos de COVID-19 registrados entre el 10 de enero y el 21 de febrero de 2020.	Desarrollar una metodología que utilice unas cantidades de datos muy limitadas y procedentes de múltiples fuentes para llevar a cabo predicción en tiempo real de la evolución de la actividad de COVID-19.
11	Lynch SM <i>et. al.</i> (15)	9,460 casos de cáncer hepático del registro de cáncer de Pensilvania	Identificar áreas que necesiten prevención de cáncer hepático mediante la combinación de datos de vigilancia existentes, comparando las recomendaciones para las áreas geográficas objetivo usando sólo datos censales con un enfoque que une otros datos geoespaciales
12	Makurumidze R <i>et. al.</i> (16)	60 distritos de Zimbawe	Evaluar el desempeño a nivel de distrito en la atención del VIH (logro de los objetivos de pruebas de VIH, vinculación con el tratamiento y cobertura) con el fin de formular recomendaciones específicas a los distritos

	Cuestionario auto-reportado FFQ, evalúa la ingesta y frecuencia de 97 alimentos, incluyendo bebidas azucaradas, en las cuatro semanas anteriores Datos demográficos, de salud (IMC) y geográficos del estudio transversal Bus-Santé	Índices espaciales Getis-Ord Gi	Identificación de clústeres espaciales de IMC y frecuencia de ingesta de bebidas azucaradas
	Informes médicos, datos del CDC de infecciones de COVID-19. Actividad de búsqueda en Internet relacionada con COVID-19 (Baidu). Actividad en los medios de comunicación. Previsiones diarias de la actividad de COVID-19, del modelo de simulación epidemiológica GLEAM (Global Epidemic And Mobility).	<i>Clustering</i> para la información geoespacial de COVID-19. Técnica de aumento de datos para aumentar observaciones. Regresión LASSO	Previsión de COVID-19 con 2 días de antelación. Visualización gráfica de resultados del modelo
	Censo de Estados Unidos de 2010. Tramos censales en el percentil 80 de 3 variables demográficas: % hispanos, % negros, % nacidos 1950-1959. Geolocalizador ArcGIS Estadísticas por escaneo espacial (SatScan) con tasas elevadas de cáncer hepático	Análisis espaciales para incidencia ajustada de cáncer de hígado. Regresión logística	SatScan identifica tramos censales con tasas de incidencia de cáncer hepático. Se calcula S, E y VPP de tramos censales que se ubican en un grupo de riesgo elevado.
	Número de personas sometidas a pruebas de VIH, número de personas que dieron positivo e iniciaron el tratamiento antirretroviral, número total de pacientes que recibieron tratamiento antirretroviral a finales de 2018. Información obtenida del sistema de información de salud de distritos del Ministerio de Salud.	Estadística descriptiva	Porcentajes de los distintos objetivos en cada distrito. Mapas que muestran prevalencia del VIH, el logro de objetivo de tests, vinculación con terapia antirretroviral y cobertura del tratamiento

(continúa en la pág. siguiente)

13	Nagar SD <i>et. al.</i> (17)	2 poblaciones vecinas con ascendencia distinta: Antioquia y Chocó. 28,137,656 SNPs, 1995 pharmaSNPs	Estudiar a nivel mundial la distribución de las variantes farmacogenómicas y las diferencias de frecuencia de los alelos farmacogenómicos entre las dos poblaciones colombianas.
14	Nuñez PA <i>et. al.</i> (18)	110,633 niños menores de 5 años con 724 872 medidas antropométricas en Misiones, Argentina	Estimar las tendencias de la prevalencia de retraso de crecimiento infantil y contrastar la hipótesis de su reducción diferencial según ubicación geográfica y nivel socioeconómico
15	Persmark A <i>et. al.</i> (19)	6,846,106 adultos en Suecia	Obtener un mapeado de la heterogeneidad de los riesgos y las desigualdades socioeconómicas en la recepción de recetas de opiáceos
16	Preston MA <i>et. al.</i> 2018 (20)	32,569 respondedores del sistema de vigilancia comportamental de Estados Unidos de 1997 a 2014	Examinar las variaciones de los mandatos estatales y de la utilización de los médicos en la moderación de las tasas de screening de cáncer colorrectal.
17	Reiner RC <i>et. al.</i> (21)	15,072,746 niños menores de 5 años	Estimar las distribuciones de prevalencia, incidencia y mortalidad de diarrea entre 2000 y 2017 en diversos niveles subnacionales.

Polimorfismos farmacogenómicos de un solo nucleótido (pharmaSNPs), de Pharmacogenomic Knowledgebase. Datos de la variación de la secuencia del genoma humano tomados, del proyecto 1000 Genomes	Análisis de la secuencia completa del exoma (WES)	Patrones de pharmaSNPs en el mundo y en Colombia. Frecuencia media del alelo menor. Índice de fijación. Magnitud del efecto de pharmaSNPs en distintos fármacos
Medidas antropométricas de dos programas de salud. Geolocalización de centros de salud. Condiciones socioeconómicas y ambientales según la influencia de cada centro de salud	Modelos de geoestadística Bayesiana. Regresión lineal	Prevalencias de retraso del crecimiento por clústeres. Riesgo relativo espaciotemporal. Disminución anualizada de prevalencia.
Estratos a partir de combinaciones de género, edad, ingresos, estado de convivencia y presencia o ausencia de malestar psicológico.	Regresión logística multinivel. Análisis multinivel de heterogeneidad individual y exactitud discriminatoria	Coefficiente de Partición de Varianza (VPC) y área bajo la curva ROC (AUC), para calcular la exactitud discriminatoria (DA) de los estratos para discernir las prescripciones de opiáceos.
Estado de screening de cáncer colorrectal. Características demográficas, de visita médica, estado de salud autorreportado, estado de salud real, seguro médico, la fecha de prueba de detección de cáncer colorrectal más reciente y la póliza de seguro.	Modelos logísticos multivariados con variables de tendencia temporal. Modelo Diferencias en diferencias en diferencias (DDD)	La estimación del DDD es descrita por la variable de interés denominada Compliance.
Datos de 466 encuestas en 94 países de ingresos bajos y medios, incluyendo preguntas sobre diarrea, típicamente en las dos semanas previas. 15 covariables geoespaciales a nivel subnacional potencialmente relacionadas con la prevalencia de diarrea	Modelos de geoestadística Bayesiana	Estimaciones de prevalencia, incidencia y mortalidad de la diarrea local, mostrando los datos espaciales y de variación temporal mediante mapas.

(continúa en la pág. siguiente)

18	Sau A <i>et. al.</i> (22)	Simulación con la población total de India en 01.01.2017	Estimar la propagación de la enfermedad del virus del Ébola en caso de un hipotético brote en la India
19	Zhou Y <i>et. al.</i> (23)	Datos de 138,842 personas en 8 poblaciones.	Analizar la variabilidad genética de la DPD y TPMT

OR: odds ratio; SARI: infecciones respiratorias agudas graves, del inglés severe acute respiratory infection; S: sensibilidad; VPP: valor predictivo positivo; VPN: valor predictivo negativo; TBC: tuberculosis; ML: machine learning; IMC: índice de masa corporal; IU: intervalo de incertidumbre; FFQ: Food Frequency Questionnaire; LASSO: least absolute shrinkage and selection operator; SNP: single nucleotide polymorphism; pharmaSNPs: pharmacogenomic SNPs; WES: whole exome sequence análisis; STEM: Spatiotemporal Epidemiological Modeler; ETS: enfermedades de transmisión sexual; DDD: diferencias en diferencias en diferencias; BRFSS: Behavioral Risk Factor Surveillance System; GIS: geographical information system; SEIR (Susceptible-Expuesto-Infeccioso-Recuperado). DPD: dihidropirimidina deshidrogenasa; TPMT: tiopurina metiltransferasa.

Tabla 2: Principales resultados de los estudios incluidos.

#	Estudio	Resultados
1	Anderson LN <i>et. al.</i>	Se identificaron seis clases basadas en características relacionadas con obesidad (número de clase: definición de esta; “nombre”, porcentaje del total de la muestra): Clase 1: más probabilidades de tener padres con sobrepeso u obesidad, bajos ingresos familiares, y no cumplir con las pautas de bebidas azucaradas o de tiempo de sedentarismo en pantallas; “Familia y conductas de riesgo para la salud” (20%) Clase 2: alta probabilidad de LDL y no HDL; “Riesgo metabólico” (7%) Clase 3: alta probabilidad de LDL, no HDL, triglicéridos elevados, tiempo de sedentarismo en pantallas, requerimiento de recursos escolares complementarios y preocupación escolar; “Alto riesgo” (6%) Clase 4: triglicéridos elevados; “Triglicéridos elevados (21%) Clase 5: alto tiempo de sedentarismo con pantallas, recursos escolares complementarios y preocupación escolar; “Comportamientos de riesgo para la salud y el desarrollo” (22%) Clase 6: menor probabilidad de casi todos los indicadores de riesgo en comparación con las otras clases; “Sano” (24%).

(continúa en la pág. siguiente)

Perfil epidemiológico del virus del Ébola, que se transmitió por Sierra Leona entre 2014-2016.	Simulación basada en el modelo compartimental SEIR.	Distribución espaciotemporal de la población expuesta, infectada y recuperada de Ébola en un intervalo de 4 meses,
Datos de las variantes de un solo nucleótido en DPD y TPMT	Algoritmos de predicción basados en ML	Cálculo de frecuencias de fenotipo del metabolizador DPD y TPMT

		Los niños de las clases 1-5 tenían mayores probabilidades de tener sobrepeso y obesidad en comparación con la clase 6 ("sana"). La clase 3 "Alto riesgo" fue la más fuertemente asociada con sobrepeso infantil (OR 1.9, 95% CI 1.2, 3.2) y obesidad (OR 3.3, IC 95% 1.7, 6.7).
2	Cummings MJ <i>et. al.</i>	Se detectaron 9 clusters, que representan 301 casos, suponiendo 10,4% del total de 2901 casos de SARI de origen desconocido. Los clusters tenían más probabilidades de producirse: en zonas urbanas con mayor densidad de población durante las dos temporadas de lluvia anuales. Dentro de los clusters detectados, se identificó un brote de SARI asociado al sarampión (n=18); el análisis de la secuencia implicó la cocirculación del genotipo B3 (13 casos), endémico en África del este, y el genotipo D4 (5 casos) probablemente importado de Inglaterra. Se detectó un probable grupo de SARI nosocomial asociado a un nuevo picobirnavirus muy relacionado con los virus porcinos y de dromedarios.
3	Ferrao JL <i>et. al.</i> (2018)	Previo a la elaboración del mapa de riesgo de paludismo, se calculan los pesos relativos de cada factor utilizado en el modelo espacial. La temperatura media (22,4%) y la precipitación (20,8%) presentaron los pesos más altos. La fórmula del mapa de riesgo de la malaria que usan: [(Temperatura 0,224) + (precipitación 0,208) + (altitud 0,104) + (pendiente 0,073) + uso de tierra 0.082) + (distancia a agua 0,123) + (distancia a carretera 0,038) + (densidad de población 0,051) + (prevalencia de malaria 0.051) + (índice de vegetación 0.047)].

(continúa en la pág. siguiente)

		En el artículo se presenta un mapa que muestra que el 96% de áreas en Chimoio presentan riesgo moderado, el 4% alto riesgo y el 0 % bajo riesgo. El mapa muestra que las “zonas residenciales” del centro y el suroeste, presentan un alto riesgo de paludismo, mientras que el resto de las zonas residenciales presentan un riesgo moderado.
4	Ferrão JL <i>et. al.</i> (2016)	Entre 2006 y 2014 se identificaron 490,555 casos de malaria en Chimoio. Media de incidencia de malaria anual: 20,1%. Fracción atribuible de malaria: 16%. Hubo diferencias en la incidencia semanal y anual de la malaria a lo largo del período; no entre hombres y mujeres. Los niños menores de 5 años de edad fueron tres veces más propensos a contraer malaria que los adultos ($p < 0,05$). Se identificaron tres grupos de malaria según su relación temporal (nombre del grupo: semanas de aparición, promedio de casos semanales ($m \pm SD$): Grupo 1: semanas 25-47, 618 $\pm$ 251,9 Grupo 2, semanas 18-24 y 48-51, 1066 $\pm$ 317,4 Grupo 3, semanas 1-17 y 52, 1587 $\pm$ 722,4 Se identificaron tres grupos según las zonas residenciales (nombre del grupo: distribución geográfica, porcentaje del total): Grupo 1: principalmente urbano, 10% Grupo 2, principalmente suburbano, 22% Grupo 3: principalmente rural, 28%
5	Franchini M <i>et. al.</i>	Se realiza una validación del algoritmo con subpoblación de 389 pacientes muestreados al azar, comparando los resultados con: 1) diagnóstico de IC por médicos de cabecera, 2) diagnóstico de IC por cardiólogos. Medidas psicométricas del algoritmo CARPEDIEM respecto al estándar clínico: S: oscila entre el 72 y el 92% según el incremento de la edad, con un valor medio del 88%. VPP: 65,1% VPN: 65 % Medidas psicométricas del algoritmo CARPEDIEM comparado con el estándar de los médicos generalistas: S: oscila entre el 34,1 y el 81, con un valor medio del 76.1%. VPP: 13.6% VPN: 99%

(continúa en la pág. siguiente)

		Se muestra evidencia de la relación entre la severidad del estado clínico y la validez del algoritmo CARPEDIEM en las métricas, que también se puede inferir del grado de concordancia con las puntuaciones clínicas evaluadas dentro de la cohorte de validación, que aumenta de las condiciones menos severas a las peores (Kappa de 0.44 a 0.70).
6	Helfinstein S <i>et. al.</i>	El 57% de los inscritos (n=942) habían buscado atención médica para sus síntomas en el momento de la entrevista inicial. Los encuestados que informaron de fiebre (OR: 2.23, 95% CI 1.17 - 4.48) o pérdida de peso no intencionada (OR: 1.69, 95% CI 1.18 - 2.42) tenían más probabilidades de buscar atención sanitaria. Características asociadas al retraso de búsqueda de atención sanitaria: fumar (OR: 0.60, 95% CI 0.41 - 0.87), beber (OR: 0.53, 95% CI 0.35 - 0.80), preferir instalaciones diferentes a las de la comunidad (OR: 0.66, 95% CI 0.48- 0.90) o creer que se tiene un riesgo bajo de TBC (OR: 0.84, 95% CI 0.75 -0.94). El análisis de clústeres reveló siete segmentos de la población que difieren en la búsqueda de atención, desde una agrupación de jubilados, desempleados y discapacitados, en la que el 70% buscó atención sanitaria de forma precoz, hasta una agrupación de hombres empleados bebedores y fumadores, en la que sólo el 42% lo hizo. Los modelos mostraron que más de la mitad de las personas diagnosticadas de TBC que retrasan la búsqueda de atención médica pertenecerían a este último clúster.
7	Huang VS <i>et. al.</i>	El modelo causal mostró que los factores más asociados a un parto en centro sanitario fueron tener un plan para el parto (OR=6.1, 95% CI 6.0 a 6.3), creer que el hospital es más seguro que el hogar (OR=5.4, 95% CI 5.1 a 5.6) y tener conocimiento de incentivos financieros (OR=3.4, 95% CI 3.3 a 3.5). Los factores individuales contribuyeron al 69% de la variación en el lugar del parto.

(continúa en la pág. siguiente)

		El análisis de segmentación mostró cuatro subgrupos distintos diferenciados por la percepción del riesgo del parto en centros sanitarios, la paridad y la planificación. El modelo clasificó correctamente el 81.9% de los casos; la sensibilidad (es decir, la correcta clasificación del parto en centros) fue del 93.3% y la especificidad (es decir, la correcta clasificación del parto domiciliario) fue del 35.4%.
8	Jardim C <i>et. al.</i>	En el período de estudio se notificaron 4819 casos de clamidia, gonorrea y sífilis, que representaron el 15.3%, el 33.2% y el 51.5% de las notificaciones, respectivamente. Se observaron amplias desigualdades espaciales en las tasas de notificación de ETS, que se concentraron predominantemente en las 2 áreas metropolitanas de Oporto y Lisboa y se dispersaron concéntricamente alrededor de ellas. Las tasas de notificación de las 3 ETS presentaron fuerte correlación ($r > 0.8$). Las tasas de gonorrea (Q1-densidad más baja vs. Q5-recuperación más alta, 2.10; CrI del 95%, 1.08-4.25) y sífilis (RR, 3.16; CrI del 95%. 2.00-5.13) se asociaron con la densidad de población. Las notificaciones de clamidia (Q1-menos urbana vs. Q5-más RR, 9.64; 95% CrI, 1.73-93.59) y sífilis (RR, 1.92; 95% CrI, 1.30-2.88) aumentaron con el nivel de urbanización. Las tasas de notificación de gonorrea se asociaron con la privación socioeconómica (Q1-menos vs. Q5-más privado RR, 1.75; 95% CrI, 1.07-2.88).
9	Joost S <i>et. al.</i>	El 13.2% ($n = 2034$) de los participantes se clasificaron en los grupos de mayor frecuencia de ingesta de bebidas azucaradas y el 10.7% ($n = 1651$) en los grupos de menor frecuencia de ingesta de bebidas azucaradas. En el análisis se evidencian varios clusters significativos de IMC y frecuencia de ingesta de bebidas azucaradas. Se identifican clusters superpuestos de frecuencia de ingesta de bebidas azucaradas e IMC en áreas específicas donde residía el 11.1% ($n = 1719$) de los participantes. Después del ajuste para covariables individuales (nivel de educación, sexo, edad, nacionalidad e ingresos medios a nivel de vecindario), los clusters identificados persisten y sólo se atenuaron ligeramente, lo que indica que los determinantes adicionales a nivel de vecindario influyen en la distribución espacial de la frecuencia de ingesta de bebidas azucaradas e IMC.

(continúa en la pág. siguiente)

10	Liu C <i>et. al.</i>	El modelo entrenado para la predicción de COVID-19 es capaz de producir pronósticos estables y precisos con 2 días de antelación y en tiempo real; así supera una colección de modelos tomados como referencia en 27 de las 32 provincias chinas. En el artículo se incluye una visualización gráfica de las estimaciones obtenidas por ARGONet + GLEAM, comparando el número de casos confirmados de COVID-19, según lo informado por el CDC, con las estimaciones anticipadas de ARGONet + GLEAM.
11	Lynch SM <i>et. al.</i>	Entre 3217 tramos censales, 1596 fueron positivos para al menos una de las 3 variables estándar (hispanos, negros y nacidos entre 1950-1959) Entre los 9460 casos de cáncer hepático se identificaron 5 grupos de alto riesgo significativo en 402 tramos censales. 324 tramos censales fueron positivos para un cluster de alto riesgo y una variable estándar (S = 92%; E = 37%; VPP = 17.4%). Cuando se incorporan 3 variables socioeconómicas a una variable estándar (%negros), mejoran las propiedades del algoritmo: (S = 93%; E = 63%; VPP = 26.3%).
12	Makurumidze R <i>et. al.</i>	De los 60 distritos analizados, cuarenta y ocho (80.0%) superaron el 90% de sus objetivos de pruebas de VIH en 2018. La vinculación con el tratamiento antirretroviral fue menor del 90% en 40 distritos (83.3%). Treinta distritos (50.0%) tuvieron una cobertura de tratamiento antirretroviral superior al 90%. De los otros 30 distritos con una cobertura de tratamiento antirretroviral subóptima (< 90%), 18 habían alcanzado el objetivo de tests de VIH pero con una vinculación subóptima con el tratamiento antirretroviral, 6 habían alcanzado objetivos de tests y una alta vinculación con el tratamiento, 4 tenían resultados subóptimos tanto en los objetivos de tests y en la vinculación con el tratamiento antirretroviral y 2 tenían resultados de objetivos de tests subóptimos y una alta vinculación con el tratamiento antirretroviral.

(continúa en la pág. siguiente)

13	Nagar SD <i>et. al.</i>	A nivel global, aparecieron variantes farmacogenómicas con frecuencias de alelos menores inusualmente elevadas con importantes diferencias de distribución en la población. Algunas variantes farmacogenómicas también muestran frecuencias alélicas anómalas dentro de las dos poblaciones colombianas, con diferencias asociadas a sus distintos perfiles de ascendencia genética. Por ejemplo, el alelo C del SNP rs4149056 [Miembro de la familia del transportador de aniones orgánicos portadores de solutos 1B1 (SLCO1B1)5], que está asociado con un mayor riesgo de toxicidad para una estatina de uso habitual, se encuentra con una frecuencia relativamente alta en Antioquia y está asociado con la ascendencia europea. Además de los alelos farmacogenómicos relacionados con el aumento del riesgo de toxicidad, se muestra evidencia de que ciertos alelos relacionados con la dosis y el metabolismo tienen grandes diferencias de frecuencia entre las dos poblaciones, al estar asociadas con sus ancestros específicos.
14	Nuñez PA <i>et. al.</i>	Encuentran una disminución de la prevalencia del retraso en el crecimiento en general. Aparecen disparidades regionales con valores de prevalencia más altos en las regiones septentrional y nororiental. En esas zonas, el retraso del crecimiento disminuyó en mayor medida durante el período estudiado, aunque el patrón espacial permaneció. El retraso en el crecimiento aumentó en las zonas urbanas periféricas y en las zonas rurales dispersas que son vulnerables desde el punto de vista socioeconómico.
15	Persmark A <i>et. al.</i>	El estudio de intersecciones de MAIHDA detectó disparidades socioeconómicas en el riesgo absoluto de recibir opioides, con un riesgo de prescripción de opioides que osciló entre el 2.77% (IC del 95%: 2.69-2.86) entre los hombres de bajos ingresos de 18 a 34 años, que vivían solos, sin malestar psicológico; y el 28.25% (IC del 95%: 27.95-28.56) entre las mujeres de ingresos medios de 65 años o más, que vivían solas, con malestar psicológico. Sin embargo, los hallazgos revelan que estas disparidades no son unidimensionales, sino más bien interseccionales y complejas.

(continúa en la pág. siguiente)

		En un modelo que combinaba los efectos aditivos y de interacción, los estratos interseccionales tenían un DA justo para distinguir a los usuarios de opiáceos de los no usuarios (VPC = 13.2%, AUC = 0.68). Sin embargo, en el modelo que descomponía los efectos totales en efectos aditivos y de interacción, el VPC era muy bajo (0.42%), lo que indicaba la existencia de pequeños efectos de interacción para varios de los estratos interseccionales.
16	Preston MA <i>et. al.</i> 2018	La utilización de atención médica aumentó la probabilidad de realizarse screening frente al incumplimiento en 9.9% puntos ($p=0.007$), según indicaron los efectos marginales medios ajustados del modelo DDD. Esto sugiere que si se realizaran visitas médicas de rutina y mandatos en las pólizas, unos 8.85 millones de personas adicionales elegibles por edad recibirían un screening de cáncer colorrectal. Este análisis apoya la aplicación de pólizas de seguros que consideren las visitas médicas y políticas más firmes que aumenten los exámenes médicos en general.
17	Reiner RC <i>et. al.</i>	Se evidencian mayores disminuciones de la mortalidad por diarrea entre 2000 y 2017 en Asia meridional y sudoriental y en América del Sur, donde el 54.0% (IU del 95% 38.1-65.8), 17.4% (7.7-28.4) y 59.5% (34.2-86.9) de las subunidades nacionales, respectivamente, registraron disminuciones de las muertes por diarrea superiores al 10%. Las regiones con mayor número de muertes se encuentran fuera de África, y las subunidades nacionales de mayor mortalidad están situadas en Pakistán. No obstante, los niños de gran parte de África siguen teniendo un alto riesgo de morir por causa de diarrea. Indonesia mostró la mayor desigualdad geográfica dentro de un mismo país; algunas regiones tenían tasas de mortalidad casi cuatro veces superiores a la tasa media del país. La mayoría de las zonas de alto riesgo tenían una cobertura deficiente en materia de agua, saneamiento e higiene, un alto nivel de retraso del crecimiento infantil o una baja cobertura de terapia de rehidratación oral.

(continúa en la pág. siguiente)

18	Sau A <i>et. al.</i>	Se expone en varios mapas la dinámica de la hipotética transmisión de Ébola en la India del 01.01.2017 al 01.01.2019. El 01.01.2017 se notificó un solo caso de Ébola en Nueva Delhi y desde entonces, se propagó hipotéticamente según el modelo determinístico de SEIR. La distribución espacio-temporal de la población expuesta, infectada y recuperada de Ébola en un intervalo de 4 meses está representada en varias figuras. Se estima que si no se interviniera para detener la propagación, en un plazo de dos años, casi la mitad del país se verá afectada por la enfermedad del Ébola y el número acumulado de personas expuestas, personas infectadas y muertes será de 106,947,760, 30,651,674 y 18,391,005, respectivamente.
19	Zhou Y <i>et. al.</i>	Se muestra como algoritmos no específicos de farmacogenética comúnmente utilizados que funcionan bien con datos de enfermedades tienen un poder predictivo discreto para las variantes DPYD y TPMT (62.1-78.6%). Por el contrario, un clasificador basado en machine learning, entrenado exclusivamente en datos de variantes DPYD, muestra mayor precisión. El algoritmo fue capaz de predecir las consecuencias de las variantes DPYD y TPMT con una precisión in vivo del 91.4%, en comparación con el 95.3% in vitro. La deficiencia de DPD mostró una alta complejidad genética, lo que abogaba por un perfil DPYD basado en la secuenciación, mientras que el genotipado de cuatro variantes en el TPMT era suficiente para explicar >95% de su variabilidad fenotípica. Se encontraron importantes diferencias interétnicas en la frecuencia y la constitución genética de la deficiencia de DPD y TPMT.

REFERENCIAS

1. Baynam G, Bauskis A, Pachter N, Schofield L, Verhoef H, Palmer RL, *et. al.* 3-Dimensional facial analysis-facing precision public health. *Front Public Heal* 2017;10:31.
2. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng* 2018;719–31.
3. Chowkwanyun M, Bayer R, Galea S. “Precision” Public Health — Between Novelty and Hype. *N Engl J Med* 2018;379:1398–1400.
4. Anderson LN, Sandhu R, Keown-Stoneman CDG, Rubeis V De, Borkhoff CM, Carsley S, *et. al.* Latent class analysis of obesity-related characteristics and associations with body mass index among young children. *Obes Sci Pract* 2020;6:390–400.
5. Cummings MJ, Tokarz R, Bakamutumaho B, Kayiwa J, Byaruhanga T, Owor N, *et. al.* Precision Surveillance for Viral Respiratory Pathogens: Virome Capture Sequencing for the Detection and Genomic Characterization of Severe Acute Respiratory Infection in Uganda. *Clin Infect Dis* 2019;68:1118–1125.
6. Ferrao JL, Niquisse S, Mendes JM, Painho M. Mapping and modelling malaria risk areas using climate, socio-demographic and clinical variables in Chimoio, Mozambique. *Int J Environ Res Public Health* 2018;15:1–15.
7. Ferrão JL, Mendes JM, Painho M, João SZ. Spatio-temporal variation and socio-demographic characters of malaria in Chimoio municipality, Mozambique. *Malar J* 2016;15:1–11.
8. Franchini M, Pieroni S, Passino C, Emdin M, Molinaro S. The CARPEDIEM Algorithm: A Rule-Based System for Identifying Heart Failure Phenotype with a Precision Public Health Approach. *Front Public Heal* 2018;6:1–10.
9. Helfinstein S, Engl E, Thomas BE, Natarajan G, Prakash P, Jain M, *et. al.* Understanding why at-risk population segments do not seek care for tuberculosis: A precision public health approach in South India. *BMJ Glob Heal* 2020;5:1–11.

10. Kavanaugh A, Lee SJ, Weng HH, Chon Y, Huang XY, Lin SL. Patient-derived joint counts are a potential alternative for determining Disease Activity Score. *J Rheumatol* 37:1035–1041. Available at: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=med8&AN=20156946> http://rd8hp6du2b.search.serialssolutions.com//?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&rft_id=info:sid/Ovid:med8&rft.genre=article&rft_id=info:
11. Huang VS, Morris K, Jain M, Ramesh BM, Kemp H, Blanchard J, *et. al.* Closing the gap on institutional delivery in northern India: A case study of how integrated machine learning approaches can enable precision public health. *BMJ Glob Heal* 2020;5:1–14.
12. Jardim Santos C, Gomes B, Ribeiro AI. *Mapping Geographical Patterns and High Rate Areas for Sexually Transmitted Infections in Portugal: A Retrospective Study Based on the National Epidemiological Surveillance System.*; 2020.
13. Joost S, Ridder D De, Marques-Vidal P, Bacchilega B, Theler JM, Gaspoz JM, *et. al.* Overlapping spatial clusters of sugar-sweetened beverage intake and body mass index in Geneva state, Switzerland. *Nutr Diabetes* 2019;9. Available at: <http://dx.doi.org/10.1038/s41387-019-0102-0>.
14. Liu D, Clemente L, Poirier C, Ding X, Chinazzi M, Davis J, *et. al.* Real-time forecasting of the COVID-19 outbreak in Chinese provinces: Machine learning approach using novel digital data and estimates from mechanistic models. *J Med Internet Res* 2020;22:e20285.
15. Lynch SM, Wiese D, Ortiz A, Sorice KA, Nguyen M, González ET, *et. al.* Towards precision public health: Geospatial analytics and sensitivity/specificity assessments to inform liver cancer prevention. *SSM - Popul Heal* 2020;12.
16. Makurumidze R, Decroo T, Lynen L, Chinwadzimba ZK, Damme W Van, Hakim J, *et. al.* District-level strategies to control the HIV epidemic in Zimbabwe: A practical example of precision public health. *BMC Res Notes* 2020;13:1–6. Available at: <https://doi.org/10.1186/s13104-020-05234-8>.

17. Nagar SD, Moreno AM, Norris ET, Rishishwar L, Conley AB, O'Neal KL, *et. al.* Population pharmacogenomics for precision public health in Colombia. *Front Genet* 2019;10:1–14.
18. Nuñez PA, Fernández MS, Turjanski P, Pérez A, Rivero MR, Angelo C De, *et. al.* Substantial reduction in child stunting is differentially associated to geographical and socioeconomic disparities in Misiones Province, Argentina. *Trop Med Int Heal* 2020;25:874–885.
19. Persmark A, Wemrell M, Zettermark S, Leckie G, Subramanian S V., Merlo J. Precision public health: Mapping socioeconomic disparities in opioid dispensations at Swedish pharmacies by Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy (MAIHDA). *PLoS One* 2019;14:1–21. Available at: <http://dx.doi.org/10.1371/journal.pone.0220322>.
20. Preston MA, Ross L, Chukmaitov A, Smith SA, Odum ML, Dahman B, *et. al.* Health Insurance Coverage Mandates: Colorectal Cancer Screening in the Post-ACA Era. *Cancer Prev Res* 2020:canprevres.0028.2020.
21. Reiner RC, Wiens KE, Deshpande A, Baumann MM, Lindstedt PA, Blacker BF, *et. al.* Mapping geographical inequalities in childhood diarrhoeal morbidity and mortality in low-income and middle-income countries, 2000-17: Analysis for the Global Burden of Disease Study 2017. *Lancet* 2020;395:1779–1801.
22. Sau A. A Simulation Study on Hypothetical Ebola Virus Transmission in India Using Spatiotemporal Epidemiological Modeler (STEM): A Way towards Precision Public Health. *J Environ Public Health* 2017;2017.
23. Zhou Y, Dagli Hernandez C, Lauschke VM. Population-scale predictions of DPD and TPMT phenotypes using a quantitative pharmacogene-specific ensemble classifier. *Br J Cancer* 2020:1–8. Available at: <http://dx.doi.org/10.1038/s41416-020-01084-0>.
24. Collins FS, Varmus H. A New Initiative on Precision Medicine. *N Engl J Med* 2015;372:793–5.

25. Khoury MJ, Galea S. Will precision medicine improve population health? *JAMA - J Am Med Assoc* 2016;316(13):1357–8.
26. Jia P, Cheng X, Xue H, Wang Y. Applications of geographic information systems (GIS) data and methods in obesity-related research. *Obes Rev* 2017;18(4):400–11.
27. Fleisch AF, Rifas-Shiman SL, Koutrakis P, Schwartz JD, Kloog I, Melly S, et. al. Prenatal exposure to traffic pollution: Associations with reduced fetal growth and rapid infant weight gain. *Epidemiology* 2015;26:43–50.
28. MacQuillan EL, Curtis AB, Baker KM, Paul R, Back YO. Using GIS Mapping to Target Public Health Interventions: Examining Birth Outcomes Across GIS Techniques. *J Community Health* 2017.

EL PAPEL DE LA EPIDEMIOLOGÍA GENÉTICA

1. DIFERENCIAS ENTRE LA TRADICIONAL Y LA NUEVA SALUD PÚBLICA

La salud está influenciada por muchos factores que pueden ser organizados, de forma genérica, en cinco categorías que se encuentran interconectadas: la genética, el comportamiento, el ambiente, el sistema sanitario y los factores sociales.¹ Estas condiciones se encuentran altamente relacionadas por la posición socioeconómica que viene determinada por varios factores, como su educación, ocupación, ingresos o incluso por cuestiones políticas.² and that there has been a rapid emergence of theoretical models and frameworks and rise in the volume of literature in a relatively small period, substantial ambiguities exist around this concept (2 Todos estos determinantes sociales impactan la salud y el bienestar de las personas y las comunidades con las que interactúan. De esta forma, una intervención de salud pública (SP) es un acto realizado por, o en favor de una persona o población con el propósito de mejorar, mantener, promover o modificar la salud o el funcionamiento de las condiciones de salud.³

La práctica tradicional de la SP se ha centrado en tratar aquellos determinantes de la salud modificables, es decir, todos menos los genéticos, llevando a cabo distintas iniciativas entre las que se incluyen: permitir el acceso universal a la atención sanitaria, realizar campañas de concienciación de estilos de vida

saludables tanto en escuelas como dirigidas a la población general, o restringir la exposición laboral a sustancias tóxicas.⁴ La base de estas acciones se sustenta básicamente en el conocimiento de factores de riesgo para enfermedades que se obtienen de estudios epidemiológicos. Este conocimiento justifica la implementación de intervenciones que lleven a una reducción de la prevalencia de estos factores a nivel poblacional como podría ser el caso del hábito tabáquico o la obesidad.

Recientemente, se ha propuesto el uso de tecnologías modernas, que permiten analizar de forma más precisas a los individuos durante lo largo de su vida, permitiendo el uso de factores genéticos que podría hacer más eficientes las iniciativas que se han llevado a cabo en la SP tradicional. Esta hipótesis se basa en el hecho de que la población podría estratificarse según sus características individuales, de forma que las políticas de salud podrían hacerse de forma dirigida a grupos de individuos más vulnerables o que tuvieran una mejor predisposición a responder a dichas intervenciones.⁵ Este nuevo concepto de SP se conoce como Salud Pública de Precisión (SPP).

La SPP podría considerarse como una extensión de la medicina de precisión. Mientras que la medicina de precisión intenta de dar cuenta de la variabilidad de la enfermedad a nivel individual, la salud pública de precisión lo hace a nivel de poblaciones.⁶ El concepto de precisión se basa en una caracterización a profundidad de los individuos, en un caso, y las poblaciones, en el otro, por medio de la obtención masiva de datos y su procesamiento. De tal forma que la SPP pretende encontrar las características de una población sobre la cual una determinada intervención sería óptima, y así aplicar la intervención más adecuada en el momento más adecuado a la población más adecuada. La colección de datos masivos, sea a nivel individual o de población, pretende capturar algunas de las condiciones que a la postre guíen la práctica clínica y las intervenciones de salud pública. Sin embargo, la obtención de esos datos se hace usualmente de forma agnóstica, insesgada, o sin hipótesis previas, por lo cual el reto es descubrir las condiciones específicas de riesgo o tratamiento de los individuos y las poblaciones en los datos masivos.

Las principales aplicaciones de la SPP se basan en el poder del análisis masivo de datos. De acuerdo con lo que se ha expuesto en la revisión sistemática de la literatura acerca de la SPP, el análisis de datos genómicos es una de las

principales aportaciones de la SPP. Sin embargo, la explotación de grandes bases de datos biomédicos, sociodemográficos, geográficos o ambientales mediante el uso de métodos computacionales avanzados (aprendizaje profundo) también permiten mejorar la salud pública mediante intervenciones más precisas.⁷ Por ejemplo, el mapeo de medidas de salud a nivel geográfico y su análisis temporal permitiría la identificación de regiones vulnerables que podrían requerir intervenciones dirigidas.⁸

Cabe destacar que, mucho antes de la explosión del análisis masivo de datos, y en particular de los datos genómicos, ya se propuso la elaboración de perfiles genéticos y epigenéticos a distintas escalas en estudios de cribado y prevención en salud pública. De esta forma, aunque las ideas subyacentes en la SPP no sean del todo novedosas, la promesa de la SPP es que su implementación a escala global permitirá una traslación efectiva y responsable de la ciencia de datos y en particular de la epidemiología genética en enormes beneficios de la salud poblacional. El hecho de disponer de información más precisa permitirá a los consumidores tomar decisiones informadas sobre su salud de forma individualizada. Un ejemplo de la ventaja que pretende la SPP frente a la SP tradicional es que las intervenciones tengan una mayor adherencia en la población. Existen evidencias de que aquellos individuos que son conscientes de tener un mayor riesgo de sufrir un evento adverso por su riesgo genético, como podría ser en el caso en la enfermedad coronaria, tienen una mayor adherencia a estilos de vida saludable que se traduce en un menor riesgo de sufrir eventos coronarios.⁹ Finalmente, dado que la salud pública se preocupa por los segmentos más vulnerables de la población, la salud pública de precisión también permitirá el tratamiento de las disparidades de salud en la implementación de la genómica en la práctica.

2. DETERMINANTES DE LA SALUD PÚBLICA

Un gran esfuerzo de los estudios en SP se basa en la identificación de los factores de enfermedad modificables, que sean susceptibles de intervención, de tal forma que estas puedan ser de carácter preventivo o terapéutico, beneficiando a la población en general. En primera línea están los factores medio ambientales, comportamentales y laborales. Es así como el proyecto de Carga Global de las Enfermedades (GBD en inglés: Global Burden of Disease)

considera en sus estimaciones 86 exposiciones de este tipo como factores de riesgo.¹⁰ Los factores de riesgo son diferentes en cada país y dependen de su ingreso económico per-cápita. Por ejemplo, el principal factor de riesgo en países con altos ingresos económicos es el tabaquismo, en Latinoamérica y el Caribe es el alto índice de masa corporal y en el África sub-Sahariana es la práctica del sexo inseguro, mientras que la exposición al alcohol en los países del Oriente Próximo y el norte de África es muy baja. Como consecuencia, las políticas de salud pública deben diferir de acuerdo a la población a la que se dirigen.

Dentro de estos determinantes de enfermedades detectados en el proyecto GDB, 26 son medio ambientales. Sin embargo, un reciente estudio que compila numerosos meta-análisis identificó 197 asociaciones entre 69 exposiciones ambientales en 83 enfermedades y distintas causas de muerte, reportadas en 103 publicaciones.¹¹ Mas específicamente, en esta revisión se encontraron 91 asociaciones con contaminación del aire, seguidas de 24 asociaciones relacionadas con tabaco. Los compuestos químicos resultaron ser el tercer grupo con 19 asociaciones reportadas. De esta forma, acciones dirigidas a reducir la exposición a estos factores o compuestos químicos ayudarían a reducir de forma drástica la incidencia y mortalidad de muchas enfermedades.

Sin embargo, existen factores no modificables que deben ser tenidos en cuenta en el diseño e implementación de las intervenciones a nivel poblacional. En esta categoría encontramos los factores genéticos cuya variabilidad sustenta una heterogeneidad intrínseca en la población. El sustrato genético de los individuos no solo los hace diferentes entre ellos, sino que mantiene diferencias en el riesgo a enfermedades. En las últimas décadas, la epidemiología genética ha impulsado el estudio detallado de la influencia de las variantes genómicas, como los cambios a nivel nucleotídico, genómico o cromosómico. Todos estos cambios han resultado fundamentales para incrementar el conocimiento preciso del desarrollo y riesgo a múltiples enfermedades comunes pero complejas, y han permitido identificar grupos de riesgo sobre los cuales una intervención o tratamiento pueden ser más efectivos. A continuación, se describen las nuevas tecnologías relacionadas con el desarrollo de la epidemiología genética y su influencia en la salud pública de precisión.

3. NUEVAS TECNOLOGÍAS EN LA SALUD PÚBLICA DE PRECISIÓN

La SPP se beneficia de la aparición de nuevas tecnologías que incorporan información de biomarcadores que permiten estratificar a la población de forma más precisa que lo hacía la SP tradicional. El uso de información biológica o genética ayuda a considerar grupos de individuos más homogéneos a los cuales se pueden dirigir las intervenciones que han mostrado una mayor eficacia para ese subgrupo de individuos. A diferencia de la medicina personalizada, donde los biomarcadores se usan para mejorar la asignación de terapias que aumenten la supervivencia de los individuos, el uso de la epidemiología genética en SP se hace para maximizar la probabilidad que las intervenciones tengan el mayor impacto en salud.

Los biomarcadores tienen el potencial de identificar estadios tempranos de la enfermedad, por ejemplo, enfermedad preneoplásica o etapas muy tempranas del cáncer, son muy prometedores para mejorar la supervivencia y la calidad de vida de nuestra población. Existen tecnologías como la biopsia líquida que podrían usarse para la detección precoz de enfermedades como el cáncer. Esta metodología se basa en la recolección y el análisis mínimamente invasivo de moléculas que pueden aislarse de los fluidos corporales, principalmente sangre completa, suero, plasma, orina y saliva entre otros. Una gran cantidad de moléculas circulantes, entre ellas las células tumorales en ADN (ctDNA por sus siglas en inglés), células libres en ADN (cfDNA) o vesículas extracelulares, más específicamente exosomas, pueden ser identificadas con este método y determinar de forma precoz la existencia de un cáncer.¹² Si bien el uso de esta tecnología está más extendido en la práctica clínica, recientes estudios han mostrado que la detección de mutaciones en cfDNA puede ser usado en personas sanas como un cribado poblacional.¹³ cell-free DNA (cfDNA) También se conoce que las alteraciones genéticas y epigenéticas, incluida la metilación del ADN y la expresión alterada de microARN, podrían estar contribuyendo a varias enfermedades autoinmunes, cáncer, y enfermedades infecciosas, por lo que se podrían usar para la evaluación de riesgo y la protección de salud pública.¹⁴

Estos biomarcadores han demostrado una gran eficacia a nivel clínico^{15,16} ya que pueden mejorar la eficiencia de ciertos fármacos cuando los pacientes se seleccionan según este tipo de marcadores.¹⁷ Sin embargo, el uso de los mismos aún está lejos de poder ser usados de forma masiva a nivel poblacional.

Más prometedor es el uso de *score* genéticos (puntuaciones genéticas). Durante los últimos años los resultados de los estudios de asociación de todo el genoma (GWAS, del inglés genome-wide association study) se han utilizado para generar modelos de riesgo poligénico para diversas enfermedades complejas basados en estos *scores*. Estos modelos combinan variantes de muchos genes para predecir el riesgo relativo heredado, en contraste con los métodos más tradicionales que solo consideran el papel de uno o unos pocos genes en el riesgo de enfermedad. Si bien el desarrollo y la validación de *scores* de riesgo poligénicos pueden ser muy prometedores para impactar la prevención de enfermedades y la promoción de la salud en el futuro, el valor actual de aplicar tales modelos en la medicina clínica y la SP aún sigue siendo discutible, aunque cada vez menos. Por ejemplo, ya existen evidencias muy claras de su utilidad en algunas patologías como las enfermedades cardiovasculares o el cáncer de mama, donde la estratificación de la población con riesgos más altos tiene una utilidad demostrada.^{18,19} El acceso a grandes bases de datos poblacionales como el caso del biobanco de Reino Unido (UKB) donde se dispone de información genética y fenotípica para más de medio millón de personas está permitiendo proponer nuevos y más precisos *scores* genéticos. Usando este enorme recurso, recientemente se ha creado un biomarcador para mortalidad general que permite dar consejos en individuos jóvenes con un mayor riesgo para que modifiquen su estilo de vida.²⁰ Más adelante describiremos con más detalle el uso de estos biomarcadores en SP que auguran un futuro muy prometedor para la SPP.

Aunque no forme parte primordial de este informe, cabe destacar que los biomarcadores digitales también han entrado a formar parte de la nueva era de datos masivos y pueden ser utilizados en SPP. Un biomarcador digital se obtiene a partir de datos fisiológicos y de comportamiento que son cuantificados y recopilados mediante dispositivos digitales. Los datos recopilados se utilizan normalmente para explicar, influir y/o predecir resultados relacionados con la salud.²¹ Cada vez más, muchas aplicaciones de teléfonos inteligentes también están disponibles para la gestión de la salud.²² Hay aprox. 300.000 aplicaciones de salud y más de 340 dispositivos disponibles en la actualidad y la cantidad de aplicaciones se duplica cada 2 años.²³ and it is imperative to leverage new technologies to generate new data and support the advent of precision medicine (PM Actualmente, también disponemos de una nueva clase de dispositivos

portátiles acoplados a teléfonos inteligentes (relojes inteligentes) que pueden monitorizar el estado de salud de la población. La principal ventaja de estas nuevas tecnologías es que permiten llevar a cabo intervenciones para promover la salud de forma muy precisa ya que la recomendación puede llevarse a cabo no sólo en función del perfil genético del individuo, sino también utilizando información actual de su estado de salud medido mediante estos dispositivos.²⁴ Un excelente ejemplo del aprovechamiento de las tecnologías digitales está siendo el apoyo a la respuesta del problema de salud pública propiciado por el COVID-19 a nivel mundial, incluyendo la vigilancia de la población, la identificación de casos, el rastreo de contactos y la evaluación de intervenciones sobre la base de datos de movilidad y comunicación con la población.²⁵

El acceso a estas nuevas tecnologías, junto con el acceso a medicamentos y vacunas, son cruciales para la SP. Nuestra principal preocupación durante los últimos años ha sido garantizar el acceso a los medicamentos esenciales y llevar a cabo investigaciones para abordar las enfermedades desatendidas. Con la SPP nos surge una nueva necesidad y es asegurar que el acceso a las nuevas tecnologías sea equitativo. Para ello, los responsables de la creación de políticas de salud deben comprender claramente los procesos de innovación que conducen a nuevas tecnologías y las formas en que estas tecnologías se difunden en los sistemas de salud.²⁶

Los determinantes para garantizar un acceso equitativo a estas nuevas tecnologías deberían equipararse a los determinantes que, según la Organización Mundial de la Salud (OMS), se establecen para los medicamentos esenciales: la selección y uso racional, precios asequibles, financiación sostenible y disponer de sistemas sanitarios y de suministro fiables.²⁶ Quizás el punto más crítico para las nuevas tecnologías radique en el coste. El precio para la obtención de biomarcadores genéticos ha descendido enormemente durante los últimos años. Esto ha sido gracias a que el coste de secuenciación se ha reducido drásticamente y a la posibilidad de secuenciar un panel de biomarcadores que reduce tanto el tiempo de análisis como el precio. Sin embargo, llevar a cabo estas iniciativas en las poblaciones más desfavorecidas podría suponer una enorme tensión financiera continua en los presupuestos familiares, a menudo necesitando gastos muy elevados en salud y por lo tanto pudiendo llevar a las familias por debajo del umbral de pobreza.²⁷

En los siguientes años se nos plantea otro reto clave para la SPP que consiste en saber qué tecnologías debemos priorizar. Según el libro verde presentado al Parlamento de los Estados Unidos por el Secretario de estado de Salud Pública y Atención Primaria, las medidas basadas en las nuevas tecnologías deberían centrarse en abordar la susceptibilidad individual, especialmente en brindar a las personas información e intervención personalizada.²⁸ En este sentido, el uso de *scores* genéticos parece ser la mejor solución para ser utilizados a nivel poblacional. Existen estudios empíricos donde se ha genotipado de forma masiva a la población obteniendo tanto resultados beneficiosos para la salud como mejoras en el crecimiento económico del país.²⁹ No obstante, llevar a cabo programas de cribado preventivo en población adulta joven, que es el grupo al que se le supone un mayor beneficio para este tipo de intervenciones, resulta altamente costoso para un sistema único de pago o para los sistemas públicos de salud.³⁰ Dado los numerosos beneficios de este tipo de intervenciones, debemos plantearnos cómo lograr la prevención mediante estas nuevas tecnologías que tengan una escalabilidad adecuada para los servicios de salud.

Si se pretende tener un impacto real en la salud de la población y evitar desigualdades, será necesario aplicar estas nuevas tecnologías a los determinantes más amplios de la salud que no sean únicamente los factores genéticos. Esto podría incluir, por ejemplo, invertir en la capacidad de los equipos de salud pública locales para comprender y seguir los cambios en los determinantes más amplios de su población. El acceso a datos masivos sobre temas como la contaminación del aire, la accesibilidad de los espacios verdes, los datos de todos los servicios gubernamentales y problemas agudos como el consumo de drogas o los cambios en el mercado laboral son cada vez más fáciles de recolectar a escala y todos podrían ser útiles.

De manera similar, los nuevos datos podrían ayudar a que los responsables de la formulación de políticas locales y nacionales comprendan mejor los determinantes comerciales, como la disponibilidad y el precio de los alimentos y las bebidas. Mientras tanto, hay una serie de preocupaciones emergentes de salud pública que son fundamentalmente problemas digitales, como los juegos de azar en línea, el impacto de las redes sociales en la salud mental y las dudas sobre las vacunas. También es probable que estos necesiten soluciones digitales a nivel de población.

Centrarse en la susceptibilidad individual puede parecer una solución más rápida y sencilla que abordar los problemas estructurales subyacentes. Los formuladores de políticas a menudo se sienten atraídos, comprensiblemente, por intervenciones que son relativamente fáciles de implementar y que probablemente tengan un impacto a corto plazo. Sin embargo, para lograr mejoras significativas a largo plazo en la salud, el gobierno necesita una comprensión sofisticada de los determinantes más amplios. Y aprovechar al máximo estas oportunidades requiere invertir en herramientas técnicas que permitan la recopilación, difusión y análisis de datos, así como en formar a nuestros futuros investigadores en las habilidades necesarias para el análisis de datos.

El sector privado también tiene un papel importante ya que suele ser el primero en acceder a estas tecnologías tan avanzadas, permitiéndole ofrecer un producto novedoso y con gran impacto en salud. Este sector suele hacer uso de la información genómica para sus propios procesos internos. Por ejemplo, las aseguradoras privadas podrían utilizar el cribado genómico para segmentar sus clientes en función del riesgo de padecer según qué enfermedades. Esta práctica es conocida en economía de la salud como selección de riesgos y se utiliza para recalcular el importe de las primas en función del nivel de riesgo; de esta manera, la empresa trata de controlar el efecto de transferencia de renta entre los asegurados de menos riesgo y los asegurados de más riesgo y así incorporar a su cartera de clientes, principalmente los de bajo riesgo. Esta manera de actuar tiene potenciales ramificaciones perversas, como la imposición de precios abusivos a personas que presentan factores de riesgo no modificables.

Por otro lado, otras entidades sanitarias privadas como clínicas y hospitales están posicionados para beneficiarse del uso de diagnóstico precoz en base a biomarcadores y datos genómicos y epigenómicos porque las poblaciones a tratar son pequeñas y pueden asumir el riesgo de incurrir en ineficiencias económicas si sirve para mejorar su calidad asistencial.

La SPP también tiene algunas desventajas y muchos retos que debemos solventar. Estos están relacionados principalmente con el consentimiento informado y la confidencialidad de los datos, entre otros.³¹ Otras desventajas y desafíos ya son conocidos por la experiencia adquirida durante los últimos años en la medicina personalizada y las soluciones adoptadas en esta área de la medicina

podrían ser utilizados en la SPP. Estos aspectos incluyen temas relacionados con la incertidumbre en torno a la validez de las pruebas genómicas dada la complejidad, por ejemplo, de la expresión génica. También se deben tener en cuenta aspectos como los incentivos financieros para las empresas privadas por la comercialización excesiva de sus servicios el posible mal uso de la información genómica por parte de empresas privadas y discriminación basada en los datos genómicos (por parte de las compañías de seguros, el sistema de salud y los empleadores). Las preocupaciones clave destacadas fueron el acceso de los médicos a las pautas clínicas y las oportunidades de capacitación para el uso de pruebas genéticas y la interpretación de datos. A pesar estos desafíos, el principal consenso entre la comunidad médica es que hay un gran interés en utilizar la información genética en la práctica si tienen acceso a los conocimientos y herramientas necesarias.³²

4. USO DE LA EPIDEMIOLOGÍA GENÉTICA EN LA SALUD PÚBLICA DE PRECISIÓN

A lo largo de este documento ya ha quedado claro que la SPP basada en la colección y análisis de datos genómicos está sin duda a la vanguardia de la búsqueda de condiciones individuales, en este caso de tipo hereditario, que permiten definir más precisamente una enfermedad y su evolución a nivel individual. Debido a que uno de los factores que permiten diferenciar las poblaciones es su sustrato genético, o su historia hereditaria, las características genéticas comunes de un grupo de pacientes con un riesgo dado a una enfermedad, conduce naturalmente a considerar esta población para una intervención adecuada. Así pues, la epidemiología genética ejerce una influencia de tipo reduccionista o “*bottom-up*” al desarrollo de la salud pública de precisión.

Algunas características poblacionales (historia familiar, estilo de vida, ...) son más influyentes en la incidencia de enfermedades comunes como por ejemplo la diabetes, la obesidad, el cáncer o la enfermedad cardiovascular. La genética sólo es un factor y no puede ser utilizada en aislamiento a otros determinantes de salud como son los socioeconómicos, incluyendo vivienda, educación y acceso a salud. Por lo tanto, la salud pública de precisión también requiere otros datos de seguimiento y aplicación de análisis sofisticados para seguir la distribución temporal y geográfica de las enfermedades, y así poder actuar de forma acorde. Si bien, debemos tener esto en cuenta, a continuación, describiremos

la translación de la epidemiología genética, incluyendo la genómica, transcrip-tómica y epigenómica, a la salud pública de precisión.

La aplicación de la epidemiología genética, por ejemplo, en cáncer, conlleva frecuentemente a la identificación de grupos muy reducidos de individuos, con mutaciones raras, que tendrán poco impacto en el desarrollo de políticas de prevención y tratamiento de la enfermedad.⁷ Por lo tanto, el primer aspecto que debemos tener en cuenta para la SPP es que las condiciones genéticas de los individuos deben tener una frecuencia alta en la población para poder ser consideradas dentro de una estrategia de SP. Asimismo, el sustrato genético de los individuos es un factor no modificable y por lo tanto no prevenible. Sin embargo, la epidemiología genética permite identificar grupos en riesgo o identificar factores modificables por medio de mecanismos etiológicos que involucran el sustrato genético de los individuos.

Así es como al investigar las diferencias en implementación de tratamientos y en sus resultados entre los pacientes, los proveedores de salud y políticos pueden refinar sus decisiones más allá de los determinantes sociales clásicos como el género, edad y etnicidad. En el contexto de salud pública, los biomarcadores moleculares tienen pues la función de determinar la heterogeneidad de la enfermedad y predecir la respuesta a las intervenciones, a nivel poblacional, de las políticas de salud pública.³³

EPIDEMIOLOGÍA GENÉTICA

La epidemiología genética pretende contribuir a entender las causas y el efecto de intervenciones sobre las enfermedades dentro de una población determinada. El conocimiento de factores más precisos mediante el uso de nuevas tecnologías puede mejorar de manera sustancial el diagnóstico de las enfermedades en estado preclínico abriendo la puerta a la anticipación de intervenciones y tratamientos. De forma indirecta, la epidemiología genética también puede ayudar a establecer determinantes ambientales de enfermedades por medio de técnicas como la aleatorización mendeliana.³⁴

La epidemiología genética usa, como biomarcadores, variantes estructurales comunes en la secuencia del DNA de los individuos. El genoma de una persona se entiende como la secuencia total de nucleótidos del DNA que reside en los cromosomas. Los datos genómicos son pues, datos de alta dimensionalidad

que sobre-determinan al individuo, por lo que la dimensión máxima posible de los datos es el número total de nucleótidos. Sin embargo, es la comparación entre genomas de los individuos lo que da lugar a variables genéticas de interés. Las variantes genómicas más comunes son los cambios en un nucleótido (SNP, del inglés single-nucleotide polymorphisms), que pueden ser detectados con gran eficacia en un barrido sobre el genoma. Los SNPs son pues cambios en una posición del genoma de un nucleótido por otro, de 4 posibles (adenina, guanina, citosina, uracilo), para un conjunto de individuos en la población. En un estudio epidemiológico típico se puede disponer de la detección de hasta 10 millones de SNPs por individuo usando métodos de imputación.³⁵

Los estudios de asociación genética (GWAS) tienen como objetivo la identificación de variantes genéticas que se asocien con la aparición de una enfermedad. El procedimiento más común es el de demostrar la asociación de la enfermedad con cada una de las variables genómicas y recoger aquellas con una asociación estadísticamente significativa, teniendo en cuenta el número total de variables genómicas estudiadas. Este proceso produce una estratificación en la población por cada variable (SNP) significativa. Inicialmente, se esperaba que las variantes identificadas lograran explicar un porcentaje elevado del riesgo genético, pero con el tiempo resultó evidente que el sustrato genético de la mayoría de las enfermedades complejas es explicado por el efecto acumulativo de muchas variantes de bajo riesgo. Sin embargo, numerosas variantes genómicas han sido detectadas con riesgos moderados, indicando tanto su rol en el desarrollo de la enfermedad como su importancia para estratificar la población y ayudar a cribar los individuos de acuerdo con su riesgo. Estudios de secuenciación completa del ADN o de variantes previamente hipotetizados han logrado identificar variantes raras con un alto efecto en el desarrollo de enfermedades complejas, por lo que también estas variantes pueden ser utilizadas como predictores de riesgo y ser utilizadas para estratificar a la población.

Los estudios de asociación genética han revelado variantes genéticas asociadas funcionalmente a enfermedades concretas.³⁶ Ejemplos paradigmáticos son el efecto de variantes en el gen *FTO* y su asociación al índice de masa corporal, así como variaciones en el gen *C4* y su rol en esquizofrenia, cuyos efectos han sido validados experimentalmente. Sin embargo, estos estudios también

han revelado, inclusive en estos casos, que el sustrato genético de las enfermedades con mayor prevalencia en la población viene explicado por la suma independiente de muchas variantes genéticas comunes con un efecto muy pequeño. Como consecuencia, no existen variantes nucleotídicas que por sí solas tengan efectos funcionales determinantes en el desarrollo de las enfermedades complejas, y por lo tanto puedan ser utilizadas de forma independiente para el diagnóstico precoz.³⁶ Sin embargo, como sus efectos son aditivos, es posible utilizar la suma de los efectos de las variantes de riesgo como una nueva cantidad genética con mayor poder predictivo que cada variante por separado (*score* genético). Desde esta perspectiva el resultado de los estudios de asociación es usado como un compendio de variantes de riesgo por su capacidad predictiva, dejando de lado la interpretación causal de las asociaciones, ya que es probable que cada una de las variantes tenga un rol en diferentes procesos biológicos que conduzcan al desarrollo de la enfermedad. Los *scores* genéticos son poderosos predictores, con poca capacidad de explicación etiológica. En el contexto de la salud pública de precisión, el objetivo general para la aplicación de los *scores* genéticos es en situaciones donde individuos con alto *score* se puedan beneficiar de intervenciones específicas como cambios en conducta, por ejemplo, con respecto al ejercicio, o tratamientos dirigidos como la administración de fármacos más precisos.

Si bien los *scores* genéticos están focalizados a la predicción, dejando de lado la complejidad de las explicaciones causales de las variantes involucradas, el objetivo por determinar los mecanismos que dan lugar a las enfermedades complejas por medio de los estudios de asociación es defendido por el modelo “omnigénico” de Boyle *et. al.*³⁷. Este modelo defiende la existencia de genes centrales en las funciones biológicas afectadas por las enfermedades. Sin embargo, debido a la interconexión de las redes de regulación genética, genes fuera de estas funciones todavía mantienen una capacidad de influencia, aunque sea mínima. De tal forma, que la mayoría de la contribución genética se encontrará fuera de los genes en funciones centrales, por la adición de numerosos efectos pequeños e independientes. Como consecuencia, la capacidad predictiva mejorará en la medida en la que se tengan en cuenta más efectos pequeños, en contraposición con la explicación de los efectos causales, que se mejorará al demarcar los genes centrales para la enfermedad. Esto marca dos estrategias diferentes para la salud pública de precisión: una, que se focaliza en

el diagnóstico temprano para mejorar la evolución y tratamiento de la enfermedad, y otra, que intenta descubrir las causas para identificar factores modificables de las funciones biológicas afectadas por la enfermedad.

En general, la capacidad de predicción de los *scores* genéticos es bastante buena. Por ejemplo, individuos con un *score* cardiovascular por encima del 8% de la población (es decir, tener un valor del *score* superior al percentil 92 observado en la población general), tienen un riesgo relativo de 3 de desarrollar la enfermedad, con un riesgo ligeramente menor para el riesgo de fibrilación auricular, cáncer de mama y diabetes tipo 2.³⁸ La magnitud de este efecto podría ser suficiente para su uso en un cribado preclínico de pacientes, a los cuales se les podría administrar tratamientos preventivos y controles médicos frecuentes. Sin embargo, los tratamientos preventivos generalmente consisten en recomendaciones como comer sano y dejar de fumar. Debido a que toda la población se vería beneficiada de estas recomendaciones, estas dejarían de ser recomendaciones precisas a un grupo de riesgo. Si la recomendación es sobre el tipo de medicación, queda por saber si esta consiste en cambiar la dosis o en dejar de tomarla. Sin embargo, individuos con un alto *score* podrían beneficiarse de un incremento a la adherencia al tratamiento derivada al conocimiento de sus condiciones genéticas (mayor riesgo).

Es importante considerar que el riesgo genético de los individuos forma parte del conjunto de factores para una enfermedad dada. En varios casos, el sustrato genético puede ser independiente de otros determinantes. Por ejemplo, *Kera et. al.* mostraron la independencia entre el *score* genético para la enfermedad de las arterias coronarias y factores de riesgo relacionados al estilo de vida.⁹ El *score* genético mostró una alta asociación con eventos coronarios, validando estudios anteriores, e indicando la importancia de la susceptibilidad genética para identificar grupos en riesgo. Así mismo, la asociación del *score* genético fue de la misma magnitud que la asociación con los factores de estilo de vida: riesgos relativos entre 1.70 y 2.2 para ambos factores en tres estudios diferentes. El estudio muestra que, si bien toda la población puede beneficiarse de un cambio en el estilo de vida, los individuos con mayor riesgo genético tienen una mayor adherencia a estos cambios. De tal forma, que la precisión en la intervención consiste en la administración del diagnóstico genético a una población específica. Claramente, un diagnóstico generalizado podría tener

efectos contrarios en los individuos que no presenten la susceptibilidad genética, pero esto ya formaría parte de otro gran tema relacionado con el consejo genético.

No todas las enfermedades complejas, que se benefician de un diagnóstico precoz, poseen *scores* genéticos con capacidad predictiva robusta. En el caso de la diabetes de tipo 2 los estudios de asociación han detectado por lo menos 40 variantes genéticas.³⁹ Sin embargo, *score* genéticos basados estas variantes no superan la capacidad predictiva de mediciones clínicas comunes, como historia familiar, índice de masa corporal o niveles de azúcar en sangre. La capacidad predictora de los *scores*, en términos de razón de probabilidades (odds ratio en inglés), no supera el 15%, por lo que su aplicación en salud pública de precisión es limitada.

Además, un elemento clave en el desarrollo de los *scores* genéticos es la precedencia ancestral de la población. En resumen, *scores* genéticos desarrollados para individuos europeos, tienden a perder predictibilidad en otras poblaciones. Por lo tanto, es importante tener en cuenta que los *scores* genéticos pueden terminar siendo herramientas de diagnóstico aptas únicamente para poblaciones específicas y no generalizables.

EL TRANSCRIPTOMA

En las últimas décadas los estudios transcriptómicos han sido ampliamente usados para investigar las causas de enfermedad a nivel molecular.⁴⁰ Los datos transcriptómicos miden el nivel de expresión de los genes en el genoma. De forma muy resumida, estos datos miden el nivel de actividad de los genes de un individuo, en un tejido concreto. Estos datos también son de alta dimensionalidad y al ser funcionales, en contraposición a los datos genómicos que son estructurales, dependen de múltiples factores como el tejido y las condiciones del individuo en las que fueron obtenidos (hora, temperatura, estrés, etc).

Reacciones complejas a nivel bioquímico influyen la decodificación, o transcripción, de las secuencias de DNA. El producto de estas reacciones son moléculas de RNA, algunas de las cuales son usadas en la producción de proteínas, los elementos básicos de la fisiología celular. Los datos transcriptómicos son obtenidos de un barrido insesgado del repertorio de las moléculas de RNA presentes en una muestra biológica en un momento dado. Estas moléculas

proviene de aproximadamente unos 20,000 genes. Si bien la dimensionalidad de los datos transcriptómicos es inferior a la de los datos genómicos, como estos datos son dinámicos y cambian a través del organismo, se requieren de estudios más sofisticados para poder encontrar biomarcadores que puedan usarse a nivel poblacional.

Los datos transcriptómicos son complejos y permiten observar más que el nivel de transcripción de los genes. Por ejemplo, pueden ser usados para determinar isoformas, fusiones y estudiar el rol de moléculas de RNA que no se transcriben en proteínas. Los datos transcriptómicos han sido utilizados en proyectos importantes como ENCODE (Encyclopedia of the regulatory elements) y TCGA (The Cancer Genome Atlas), para ofrecer una fotografía general de la transcripción en las líneas celulares de miles de tumores primarios.

Los estudios de asociación a nivel transcriptómico han sido usados para estudiar las diferencias transcripcionales entre individuos con diferentes estados de enfermedad. Por ejemplo, se han detectado grupos de pacientes de cáncer de mama con diferencias en pronóstico por medio de la transcripción de genes involucrados en la señalización del estrógeno.⁴¹ En un estudio reciente usando una detección insesgada de RNA en 2,400 células, se logró detectar repertorios de nuevos tipos de células dendríticas y monocitos que podrán ser usados para monitorizar el progreso inmunológico de pacientes.⁴² phagocytosis, and antigen presentation and consist of multiple specialized subtypes. However, their identities and interrelationships are not fully understood. Using unbiased single-cell RNA sequencing (RNA-seq La capacidad de monitorizar el transcriptoma es una herramienta poderosa para obtener información sobre los mecanismos de un gran número de enfermedades. De hecho, varios ensayos clínicos usan estos datos para estratificar mejor los individuos que responden a tratamientos (<https://clinicaltrials.gov>). También se usa para comparar el efecto de diferentes intervenciones farmacológicas como por ejemplo en psoriasis.⁴³ El objetivo fundamental es, pues, usar estos datos para encontrar las funciones biológicas centrales que son afectadas en enfermedades complejas, y así determinar los mecanismos etiológicos para identificar nuevas herramientas de diagnóstico y tratamiento.

Los datos transcriptómicos son sensibles a muchos factores, por lo que es necesario estandarizar su precisión y reproducibilidad, así como su especificidad

y sensibilidad, de tal forma que su utilidad clínica aumente.⁴⁴ Así mismo es necesario el desarrollo de herramientas de diagnóstico que utilicen un número reducido de genes para reducir su coste y facilitar su aplicación diagnóstica. Su utilidad en salud pública de precisión, aunque podría ser más precisa que los datos genómicos por su dinamismo (permitiría detectar estados más cercanos a la enfermedad), está aún por demostrar.

EL EPIGENOMA

Si bien la salud pública de precisión se beneficia de la identificación de grupos de riesgo, para saber a quién intervenir, y del conocimiento más preciso de las causas de las enfermedades, para precisar el objetivo de la intervención, es claro que su poder reside en saber qué factores intervenir. Algunos factores de riesgo ambientales dejan su huella en el DNA de los individuos, alterando la forma en que los genes modulan su actividad. La accesibilidad del ADN es esencial para la expresión de los genes. Uno de los mecanismos que la célula usa para regular esta accesibilidad es por medio de la adhesión de un grupo metilo a la cadena de ADN involucrada en la expresión de un gen en particular. La metilación de la cadena de ADN promueve el silenciamiento de los genes a los que afecta, afectando a su vez las funciones de la célula en las que las proteínas codificadas por esos genes participan. La metilación del ADN es un fenómeno epigenético, es decir que es una característica adquirida y a su vez heredable. Es así como se explica la herencia en el tipo celular de las células madre a las células hijas. La importancia de la metilación para la salud pública de precisión es que esta es altamente influenciada por factores ambientales. De tal forma, los efectos de un factor de riesgo ambiental pueden ser detectados a nivel genético, y, en consecuencia, se podrá determinar su rol en las funciones centrales afectadas por la enfermedad.

Los datos epigenómicos se refieren a la colección insesgada de estados de metilación en el genoma. Estos estados dependen, entre otras cosas, del tejido, edad y sexo del individuo. La dimensionalidad de estos datos depende del número de sitios en donde se puede metilar el DNA, que es aproximadamente un 1% del genoma. De la misma forma que los datos transcriptómicos son menos grandes que los genómicos y dinámicos, pero en escalas de tiempo mayores. A diferencia de los datos genómicos que son estáticos en tiempo, los cambios epigenómicos son plásticos, lo que los convierte no sólo en biomarcadores de

exposiciones, sino en dianas para modificación mediante intervenciones preventivas o tratamientos terapéuticos.¹⁴

Los estudios de asociación epigenómica (EWAS, del inglés epigenome-wide association studies) son más comunes en estudios que comparan tejido canceroso a tejido sano. Sin embargo, estudios pioneros epigenéticos, como los de *Perera et. al.*⁴⁵, encontraron que los hijos de mujeres embarazadas, expuestas a hidrocarburos aromáticos policíclicos (PHA en inglés), tenían más frecuencia de aductos de ADN en sangre de cordón umbilical, bajo peso al nacer y reducida circunferencia craneal. Mas recientemente se ha visto que el bajo peso al nacer, asociado a numerosas enfermedades a lo largo de la vida, está fuertemente correlacionado con la metilación del ADN. En un meta-análisis de 24 cohortes y 8825 neonatos, 914 sitios de metilación se correlacionaron con diferencias en el peso al nacer.⁴⁶

Exposiciones intrauterinas, como el tabaquismo, el índice de masa corporal antes del embarazo, hiperglicemia, hipertensión, niveles de folatos o desnutrición de las madres están asociadas con el peso al nacer de sus hijos. Estudios epidemiológicos han mostrado que el bajo peso al nacer es un determinante de enfermedades posteriores, como enfermedades cardio-metabólicas, mentales, algunos cánceres y mortalidad. En estas asociaciones de largo alcance temporal, el peso al nacer actúa como un proxy de las exposiciones intrauterinas. Si bien los niveles de metilación en los pacientes se refieren a eventos ya pasados, su conocimiento es esencial para determinar los factores modificables que pueden ser objeto de las políticas de salud pública. Sin embargo, apenas se empiezan a considerar los factores éticos relacionados con la responsabilidad sobre las decisiones de las mujeres durante el embarazo en la salud de sus hijos mucho después de nacer.

Varios estudios en animales y humanos ofrecen evidencias sobre la influencia de las exposiciones prenatales, incluyendo productos químicos y condiciones nutricionales, en el desarrollo de enfermedades durante la adultez. Uno de los compuestos químicos presentes en la dieta, por medio del uso de ciertos biberones, y recipientes de alimentos es el bisfenol A. Estudios recientes sugieren que las vías de acción de este compuesto químico son por medio de la metilación del ADN.⁴⁷ De esta forma, los datos epigenéticos pueden definir con mayor precisión las exposiciones y las funciones biológicas afectadas por una

determinada exposición ambiental, que puede ser modificada por políticas de salud pública de prevención y ser un arma clave en la SPP.⁴⁸

APLICACIONES DE BIOMARCADORES MOLECULARES EN SALUD PÚBLICA

La utilización de variantes genómicas para cribar pacientes o individuos en riesgo es ilustrada por el estudio genético de la hemocromatosis hereditaria. Esta enfermedad que afecta el metabolismo del hierro y su pronóstico aumenta sustancialmente con un diagnóstico precoz. Si bien la enfermedad es causada en un 80% de los casos por una mutación nucleotídica C282Y en el gen *HFE* en ambos cromosomas, no se puede descartar su presencia en individuos con una sola mutación C282Y y una variante en otra posición del gen. De tal forma que el cribado de pacientes basado en C282Y, a pesar de ser viable económicamente y preciso, no ha sido adoptado como herramienta de prevención, y el diagnóstico de la enfermedad todavía se realiza por medio de cuadros clínicos. Cribados de pacientes con respecto a mutaciones de alto riesgo que ocurran en tan sólo uno de los cromosomas (dominantes) incluyen mutaciones nucleotídicas en los genes *BRCA1* y *BRCA2* con un incremento importante en los casos de cáncer de mama. El riesgo asociado puede ser del 5% para una enfermedad con una incidencia muy alta y grandes cargas al sistema de salud. Por lo tanto, la detección de esta variante antes de la manifestación clínica de la enfermedad puede ser usada para intervenciones de prevención como controles frecuentes o incluso mastectomías profilácticas. A nivel poblacional el cribado puede ser difícil de implementar debido a su coste³⁴.

Los atributos para un cribado genético de pacientes deben incluir pocas variantes genéticas que expliquen una proporción alta de casos. Mas de un millón de personas tendrían que ser cribadas por el factor V de Leiden, cuya mutación incrementa la coagulación de la sangre, para evitar una sola muerte. El coste del cribado debe ser, así mismo, bajo y su administración simple. A nivel de salud pública es importante que los cribados estén dirigidos a mitigar y prevenir enfermedades con coste social alto, cuyas medidas de prevención primarias estén ya implementadas. De esta forma, es posible ofrecer tratamientos e intervenciones preventivas adecuadas a los individuos detectados por la mutación. Wald ha identificado tres elementos para la eficacia de un cribado genético de la población: 1) La identificación de personas con alto riesgo de

una enfermedad concreta de tal forma que pruebas adicionales y medidas de prevención le puedan ser administradas; 2) llegar a poblaciones en riesgo que no han buscado atención médica y; 3) disponer de medidas de seguimiento de las que se puedan beneficiar los individuos cribados.⁴⁹

Los estudios de asociación genética han identificado variantes genómicas comunes en la población asociadas con enfermedades importantes. Sin embargo, el riesgo asociado a estas variantes es como mucho moderado, y en muchos casos no es clara la aportación que obtenga una persona sana del conocimiento de estas variantes. Este es el caso de la variante común e4 en el gen *APOE* que se asocia con un incremento en los casos de la enfermedad de Alzheimer. La variante e4 tienen una influencia mucho más baja en el desarrollo de la enfermedad que variantes más raras en genes *APP* y *PSEN* asociadas con una transmisión hereditaria alta. Los tests genéticos tienen como finalidad saber si un pariente de Alzheimer, al heredar las variantes genéticas, desarrollará la enfermedad. De tal forma que la detección de la variante e4, a pesar de ser común, no presenta el valor predictivo de las variantes raras. Sin embargo, si se ofrece la prueba para parientes que decidan saber su riesgo genético se debe también ofrecer el asesoramiento adecuado para los casos positivos. Así mismo se debe evitar que esta información pueda ser usada por aseguradoras para discriminar individuos con predisposición genética a padecer Alzheimer (<https://www.alzheimers.org.uk/>).

La correcta comunicación de riesgos genéticos y su gestión es particularmente importante en el ámbito de la salud pública. Los riesgos genéticos no son modificables por lo que es importante establecer las ventajas específicas en el cribado genético de individuos sanos³⁴. Aún si la asociación entre las variantes genéticas y el riesgo a una enfermedad se confirma, este conocimiento puede promover otras vías prevención diferentes al cribado. Por ejemplo, la mutación 677TT en el gen *MTHFR* está asociado a un incremento en el 20% de casos de la enfermedad de las arterias coronarias, mediante incrementos de homocisteína, cuyos niveles pueden ser disminuidos con la suplementación de folato. De hecho, hay otras variantes genéticas que pueden afectar los niveles de homocisteína, por lo que es más conveniente medir sus niveles en sangre que hacer test genéticos. De tal forma que las asociaciones genéticas no solo son pueden ser usadas para determinar poblaciones en riesgo sino también para

detectar factores modificables, que puedan ser incorporados en un tratamiento preventivo o primario de la enfermedad.

5 CONCLUSIÓN Y RECOMENDACIONES FUTURAS

Millones de biomarcadores moleculares pueden ser obtenidos para un individuo. Hemos visto como estos biomarcadores, dependiendo de su naturaleza biológica pueden ser utilizados en el contexto de la salud pública de precisión; es decir en la aplicación de una intervención adecuada a una población adecuada en el instante adecuado. Los biomarcadores genéticos, que consisten en sustituciones puntuales de nucleótidos en la cadena de ADN, tienen como función principal identificar grupos con diferente riesgo genético a enfermedades complejas. La forma más efectiva de estratificar la población es por medio de los *scores* genéticos, que son contruidos a partir de la suma de numerosas variantes genéticas. Estos *scores* tienen una capacidad predictiva, no modificable, de indicar qué individuos están en mayor riesgo. De tal forma que los biomarcadores genéticos son utilizados en mayor medida para el cribado de pacientes en mayor riesgo o que respondan mejor a un tratamiento. Sin embargo, la utilidad del cribado depende de la calidad de la predicción y del tipo de atención médica o de prevención que se le pueda ofrecer a individuos de alto riesgo.

Es de suma importancia destacar que los aspectos éticos son cruciales para llevar a cabo algunas de las intervenciones de salud pública de precisión anteriormente descritas, por ejemplo, cuando estén relacionados con el cribaje asociado a coberturas médicas por aseguradoras o las condiciones que el cribaje pueda suponer en el entorno laboral. A modo de ilustración, se sabe que la enzima N-acetiltransferasa está involucrada en el metabolismo de las arilaminas, que está presente en industrias como el secado en seco. Como las arilaminas están correlacionadas con mayor incidencia en cáncer de vejiga, que es más pronunciado en individuos con ciertas mutaciones genéticas, es pues posible que los empleadores de estas industrias estén interesados en crear programas de cribado genético para discriminar empleados, creando una disparidad en oportunidades de empleo.⁵⁰

Por otro lado, al identificarse una población potencialmente vulnerable en términos genéticos, si se trata de una patología de etiología compleja y con

estrategias preventivas conocidas (dieta y ejercicio, por ejemplo) se puede valorar, si se tiene la información, el ponerla en conocimiento del paciente, para que actúe como un potenciador de la adherencia a comportamientos que mejoren su pronóstico. En otras palabras, aún no existiendo tratamientos específicos, la información genómica podría actuar como un factor protector en sí misma, si se utiliza para maximizar la prevención no farmacológica donde sea posible.

El objetivo de la salud pública de precisión es la aplicación de medidas que mitiguen o prevengan la exposición a factores modificables. Si bien los marcadores genéticos no son modificables, estos pueden ser utilizados para identificar factores modificables asociados las enfermedades.

La noción de qué variantes genéticas pueden ser indicadores de factores de exposición ambiental ha sido estudiada en muchos contextos. Por ejemplo, es sabido que los individuos intolerantes a la lactosa, debido a una mutación dominante, beben menos leche. Por otro lado, no es claro si las asociaciones entre consumo de leche y reducción al riesgo de osteoporosis se deben a otros factores sociales relacionados con la ingesta de leche, como el acceso a este alimento. Sin embargo, la asociación entre la mutación causante de la intolerancia a la lactosa y la osteoporosis, sí indica que la leche es un factor de protección, pues los otros factores relacionados con el consumo de leche no comparten esta vía de exposición.³⁴ Este tipo de estudios se denominan de aleatorización mendeliana y se diferencian de los estudios convencionales de identificación de riesgos y caracterización de poblaciones vulnerables.

En conclusión, la utilización de datos genómicos y posteriormente incluso epigenómicos o transcriptómicos, presenta un gran potencial siempre que se den las condiciones adecuadas en términos de coste-efectividad, coste-utilidad y capacidad para detectar correlaciones fuertes y causales. Idealmente, la salud pública de precisión permitirá revitalizar una disciplina como es la epidemiología genética que nunca había sido tan importante para promover la salud de nuestras comunidades más vulnerables y excluidas. El desarrollo de estas tecnologías permitirá la detección de poblaciones de riesgo para patologías con etiologías complejas y diversas y la implementación de estrategias preventivas sólo en las personas que realmente lo necesitan, aligerando la carga asistencial futura para el sistema sanitario y ahorrándole al potencial paciente los males derivados de padecer una enfermedad.

BIBLIOGRAFÍA

1. Closing the gap in a generation: health equity through action on the social determinants of health - Final report of the commission on social determinants of health. Available at: <https://www.who.int/publications/i/item/WHO-IER-CSDH-08.1>. (Accessed: 7th February 2021)
2. Islam, M. M. Social determinants of health and related inequalities: Confusion and implications. *Frontiers in Public Health* **7**, 11 (2019).
3. International Classification of Health Interventions (ICHI). Available at: <https://www.who.int/standards/classifications/international-classification-of-health-interventions>. (Accessed: 12th February 2021)
4. Satcher, D. & Higginbotham, E. J. The public health approach to eliminating disparities in health. *American Journal of Public Health* **98**, 400–403 (2008).
5. Weeramanthri, T. S. *et. al.* Editorial: Precision Public Health. *Front. Public Heal.* **6**, (2018).
6. Chowkwanyun, M., Bayer, R. & Galea, S. “Precision” Public Health — Between Novelty and Hype. *N. Engl. J. Med.* **379**, 1398–1400 (2018).
7. Khoury, M. J., Engelgau, M., Chambers, D. A. & Mensah, G. A. Beyond Public Health Genomics: Can Big Data and Predictive Analytics Deliver Precision Public Health? *Public Health Genomics* **21**, 244–249 (2019).
8. Seeking precision in public health. *Nature Medicine* **25**, 1177 (2019).
9. Khera, A. V. *et. al.* Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary Disease. *N. Engl. J. Med.* **375**, 2349–2358 (2016).
10. Stanaway, J. D. *et. al.* Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Stu. *Lancet* **392**, 1923–1994 (2018).
11. Rojas-Rueda, D. *et. al.* Environmental Risk Factors and Health: An Umbrella Review of Meta-Analyses. *Int. J. Environ. Res. Public Health* **18**, 704 (2021).

12. Ignatiadis, M., Sledge, G. W. & Jeffrey, S. S. Liquid biopsy enters the clinic — implementation issues and future challenges. *Nat. Rev. Clin. Oncol.* 1-16 (2021). doi:10.1038/s41571-020-00457-x
13. Alborelli, I. *et. al.* Cell-free DNA analysis in healthy individuals by next-generation sequencing: a proof of concept and technical validation study. *Cell Death Dis.* **10**, 1–11 (2019).
14. Rozek, L. S., Dolinoy, D. C., Sartor, M. A. & Omenn, G. S. Epigenetics: Relevance and implications for public health. in *Annual Review of Public Health* **35**, 105–122 (Annual Reviews Inc., 2014).
15. Quezada, H., Guzmán-Ortiz, A. L., Díaz-Sánchez, H., Valle-Rios, R. & Aguirre-Hernández, J. Omics-based biomarkers: current status and potential use in the clinic. *Boletín Medico del Hospital Infantil de Mexico* **74**, 219–226 (2017).
16. Lambert, S. A., Abraham, G. & Inouye, M. Towards clinical utility of polygenic risk scores. *Human Molecular Genetics* **28**, R133–R142 (2019).
17. Boessen, R. *et. al.* Improving clinical trial efficiency by biomarker-guided patient selection. *Trials* **15**, 103 (2014).
18. Mavaddat, N. *et. al.* Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am. J. Hum. Genet.* **104**, 21–34 (2019).
19. Khera, A. V. *et. al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics* **50**, 1219–1224 (2018).
20. Meisner, A. *et. al.* Combined Utility of 25 Disease and Risk Factor Polygenic Risk Scores for Stratifying Risk of All-Cause Mortality. *Am. J. Hum. Genet.* **107**, 418-431 (2020).
21. Coravos, A., Khozin, S. & Mandl, K. D. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *npj Digital Medicine* **2**, 1-5 (2019).

22. Kamel Boulos, M. N., Brewer, A. C., Karimkhani, C., Buller, D. B. & Dellavalle, R. P. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online J. Public Health Inform.* **5**, (2014).
23. Seyhan, A. A. & Carini, C. Are innovation and new technologies in precision medicine paving a new era in patients centric care? *Journal of Translational Medicine* **17**, 114 (2019).
24. Bert, F., Giacometti, M., Gualano, M. R. & Siliquini, R. Smartphones and health promotion: A review of the evidence. *J. Med. Syst.* **38**, (2014).
25. Budd, J. *et. al.* Digital technologies in the public-health response to COVID-19. *Nature Medicine* **26**, 1183–1192 (2020).
26. Promoting access to medical technologies and innovation: Intersections between public health, intellectual property and trade. Available at: <https://www.who.int/publications/i/item/9789241504874>. (Accessed: 9th February 2021)
27. Niëns, L. M. *et. al.* Quantifying the impoverishing effects of purchasing medicines: A cross-country comparison of the affordability of medicines in the developing world. *PLoS Med.* **7**, (2010).
28. Advancing our health: prevention in the 2020s – consultation document - GOV.UK. Available at: <https://www.gov.uk/government/consultations/advancing-our-health-prevention-in-the-2020s/advancing-our-health-prevention-in-the-2020s-consultation-document>. (Accessed: 10th February 2021)
29. Séguin, B., Hardy, B. J., Singer, P. A. & Daar, A. S. Genomic medicine and developing countries: Creating a room of their own. *Nature Reviews Genetics* **9**, 487–493 (2008).
30. Zhang, L. *et. al.* Population genomic screening of all young adults in a health-care system: a cost-effectiveness analysis. *Genet. Med.* **21**, 1958–1968 (2019).
31. Joly, Y., Saulnier, K. M., Osien, G. & Knoppers, B. M. The ethical framing of personalized medicine. *Curr. Opin. Allergy Clin. Immunol.* **14**, 404–408 (2014).

32. Najafzadeh, M., Davis, J. C., Joshi, P. & Marra, C. Barriers for integrating personalized medicine into clinical practice: A qualitative analysis. *Am. J. Med. Genet. Part A* **161**, 758–763 (2013).
33. Collins, F. S. & Varmus, H. A New Initiative on Precision Medicine. *N. Engl. J. Med.* **372**, 793–795 (2015).
34. Smith, G. D. *et. al.* Genetic epidemiology and public health: Hope, hype, and future prospects. *Lancet* **366**, 1484–1498 (2005).
35. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nature Reviews Genetics* **11**, 499–511 (2010).
36. Croucha, D. J. M. & Bodmer, W. F. Polygenic inheritance, GWAS, polygenic risk scores, and the search for functional variants. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 18924–18933 (2020).
37. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177–1186 (2017).
38. Inouye, M. *et. al.* Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults: Implications for Primary Prevention. *J. Am. Coll. Cardiol.* **72**, 1883–1893 (2018).
39. Vassy, J. L. & Meigs, J. B. Is Genetic testing useful to predict type 2 diabetes? *Best Practice and Research: Clinical Endocrinology and Metabolism* **26**, 189–201 (2012).
40. Casamassimi, A., Federico, A., Rienzo, M., Esposito, S. & Ciccodicola, A. Transcriptome profiling in human diseases: New advances and perspectives. *International Journal of Molecular Sciences* **18**, (2017).
41. Curtis, C. *et. al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
42. Villani, A. C. *et. al.* Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science (80-.).* **356**, (2017).

-
43. Tomalin, L. E. *et. al.* Short-term transcriptional response to IL-17 receptor-A antagonism in the treatment of psoriasis. *J. Allergy Clin. Immunol.* **145**, 922–932 (2020).
 44. Byron, S. A., Van Keuren-Jensen, K. R., Engelthaler, D. M., Carpten, J. D. & Craig, D. W. Translating RNA sequencing into clinical diagnostics: Opportunities and challenges. *Nature Reviews Genetics* **17**, 257–271 (2016).
 45. Perera, F. P., Jedrychowski, W., Rauh, V. & Whyatt, R. M. Molecular epidemiologic research on the effects of environmental pollutants on the fetus. *Environ. Health Perspect.* **107**, 451–460 (1999).
 46. Küpers, L. K. *et. al.* Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. *Nat. Commun.* **10**, 1–11 (2019).
 47. Vandenberg, L. N. *et. al.* Urinary, circulating, and tissue biomonitoring studies indicate widespread exposure to bisphenol A. *Environmental Health Perspectives* **118**, 1055–1070 (2010).
 48. Khoury, M. J. Public health genomics: The end of the beginning. *Genetics in Medicine* **13**, 206–209 (2011).
 49. Wald, N. J. The definition of screening. *Journal of Medical Screening* **8**, 1 (2001).
 50. Vineis, P. & Schulte, P. A. Scientific and ethical aspects of genetic screening of workers for cancer risk: The case of the n-acetyltransferase phenotype. *J. Clin. Epidemiol.* **48**, 189–197 (1995).

AUTOR

Juan Ramón González
Associate Research Professor
juanr.gonzalez@isglobal.org
<https://orcid.org/0000-0003-3267-2146>

Juan R. Gonzalez es Associate Research Profesor de ISGlobal, donde dirige el Grupo de Bioinformática en Epidemiología Genética (BRGE).

Es coautor de más de 150 artículos científicos y tiene una amplia experiencia en grandes proyectos internacionales de investigación (ATHLETE, HELIX, MEDIRAD y OPERRA).

Como profesor adjunto en el Departamento de Matemáticas de la Universidad Autónoma de Barcelona, sus actividades educativas se centran en proporcionar conferencias de posgrado sobre métodos estadísticos avanzados y bioinformática. Tras obtener su Ph.D. en Bioestadística, supervisado por el Prof. Edsel Peña de la Universidad de Carolina del Sur (EE.UU.), pasó dos años como estudiante de postdoctorado en el Centro de Regulación Genómica (CRG).

Su trabajo se centra en el desarrollo de métodos estadísticos y computacionales para el análisis y la comprensión de diferentes datos ómicos motivado por una amplia gama de preguntas de relevancia biológica y clínica. Estos incluyen genómica (CNV, inversiones, mosaicismos), transcriptómica (RNA-seq y splicing alternativo) y proteómica entre otros. El análisis del exposoma, su caracterización y cómo integrarlo con diferentes capas ómicas es también otra línea activa de investigación de su grupo.

Es autor del libro “Omic association analysis with R and Bioconductor”, que incluye métodos actualizados para explotar datos ómicos.

Líneas de investigación

- Métodos estadísticos para explotar datos genómicos
- Métodos robustos y escalables para el big data en medicina personalizada
- Rol de las inversiones genómicas en las enfermedades complejas

- Nuevos métodos estadísticos y herramientas para analizar los datos del exposoma
- Rol de la elevada desregulación del cromosoma Y en enfermedades complejas

Publicaciones principales

- Ruiz-Arenas C, Caceres A, Lopez-Sanchez M, Tolosana I, Perez-Jurado L, **Gonzalez JR**. scoreInvHap: Inversion genotyping for genome-wide association studies. Plos Genetics. 2019;15(7):e1008203.
- Hernandez-Ferrer C, Wellenius GA, Tamayo I, Basagana X, Sunyer J, Vrijheid M, **Gonzalez JR**. Comprehensive study of the exposome and omic data using rexpoxome Bioconductor packages. Bioinformatics. 2019;pii: btz526. doi: 10.1093/bioinformatics/btz526.
- Hernandez-Ferrer C, **Gonzalez JR**. CTDquerier: a bioconductor R package for Comparative Toxicogenomics Database data extraction, visualization and enrichment of environmental and toxicological studies. Bioinformatics. 2018;15;34(18):3235-3237.
- Caceres A, **Gonzalez JR**. Following the footprints of polymorphic inversions on SNP data: from detection to association tests. Nucleic Acids Res. 2015;43:e53.
- **Gonzalez JR**, Caceres A, Esko T, Cusco I, Puig M, Esnaola M, Reina J, SirouxV, Bouzigon E, Nadif R, Reinmaa E, Milani L, Bustamante M, Jarvis D, Anto JM, Sunyer J, Demenais F, Kogevinas M, Metspalu A, Caceres M, Perez-Jurado LA. A common 16p11.2 inversion underlies the joint susceptibility to asthma and obesity. Am J Hum Genet. 2014;94(3):361-72.

ANÁLISIS DEL GRUPO DE DISCUSIÓN

1. JUSTIFICACIÓN Y OBJETIVOS

En los últimos años están aumentando de forma progresiva las publicaciones que integran análisis de datos masivos para resolver preguntas de investigación en salud. De forma particular, estos análisis presentan un creciente interés en el área de la salud pública. En las secciones previas de este libro, se ha llevado a cabo una revisión sistemática de la literatura, constatando que el tema es amplio y los enfoques diversos. Los artículos encontrados son heterogéneos en objetivos y metodología, pero pueden considerarse una buena representación de la utilización del término “salud pública de precisión” en la literatura científica. De manera más específica en el informe de experto, el foco principal se basa en el empleo de la epidemiología genética en salud pública de precisión (SPP), una de las áreas que más desarrollo ha demostrado en forma de publicaciones. Se ha incluido información detallada de algunos estudios en los que se está haciendo uso de la genómica para sacar conclusiones en grandes poblaciones.

El objetivo de esta sección es analizar cuál es la valoración que hacen los diferentes agentes sanitarios sobre el concepto de salud pública de precisión, sus implicaciones, y como se prevé que permitirá mejorar el conocimiento epidemiológico y la toma de decisiones en salud pública.

2. MÉTODOS Y PARTICIPANTES

Con antelación a la reunión del grupo de discusión, los panelistas recibieron los epígrafes previos que componen este libro, es decir, la introducción, la revisión sistemática de la literatura y el informe de experto. Asimismo, recibieron las preguntas que se formularían en la sesión grupal.

En la sesión telemática del Grupo de Discusión, moderada por Juan del Llano, se presentaron secuencialmente las preguntas y se permitió a los participantes intervenir de forma flexible contestando a las preguntas originales o reaccionando a las intervenciones de los otros miembros del panel.

Toda la sesión fue grabada y transcrita. A continuación, el analista Diego Benavent llevó a cabo una primera depuración del texto y solicitud de aclaraciones a los participantes, que tuvieron la oportunidad de revisar la transcripción final de sus intervenciones para garantizar que en la elaboración de la transcripción no se hubiese alterado el sentido de sus manifestaciones. Las afirmaciones y opiniones expresadas en el texto reflejan, por lo tanto, la posición de los participantes y no las de sus respectivas instituciones, ni las del equipo editorial de este informe, con la excepción de los comentarios finales. En una etapa posterior, Diego Benavent realizó el análisis de las intervenciones, así como la elaboración de la hoja de ruta.

Participantes:

- **Javier Castrodeza.** *Catedrático de Medicina Preventiva y Salud Pública, Universidad de Valladolid. Ex Secretario General de Sanidad, MS.*
- **José Ramón Banegas.** *Catedrático Medicina Preventiva y Salud Pública, Universidad Autónoma de Madrid y CIBERESP.*
- **Francisco Estupiñán.** *Investigador Senior en Data Science, Instituto Aragonés de Ciencias de la Salud (IACS)*
- **Cristóbal Belda-Iniesta.** *Subdirector General de Evaluación e Investigación, Instituto de Salud Carlos III (ISCIII)*
- **Guillermo Sanz.** *Jefe de Hematología Clínica, H Universitario y Politécnico La Fe, Valencia*

- **Mercedes Alfaro.** *Subdirectora General de Información Sanitaria, Ministerio de Sanidad*
- **Salvador Peiró.** *Director Unidad Investigación Servicios Sanitarios, FISABIO, Valencia*

Equipo organizador de la Fundación Gaspar Casal y presente en la sesión:

- Moderador: **Juan E. del Llano Señarís**
- Secretario, análisis del discurso y redacción del informe: **Diego Benavent Nuñez**

3. PREGUNTAS

1. ¿En qué áreas de salud pública considera que pueden tener más impacto las investigaciones realizadas con metodología de datos clínicos a gran escala y nuevas técnicas para su procesamiento?
2. A su juicio, ¿las conclusiones arrojadas hasta el momento por estudios que emplean metodologías de Salud Pública de Precisión son útiles para guiar políticas sanitarias?
3. ¿Considera que las nuevas tecnologías de “precisión” podrán mejorar las actuales estrategias preventivas, el acceso a la atención médica o los resultados en salud en un futuro próximo (5 años)?
4. ¿Qué impacto cree que puede tener la sobreinformación y el exceso de datos genómicos en el sobrediagnóstico y las intervenciones innecesarias?
5. ¿Conoce o ha utilizado alguna vez sistemas de información geográfica en el contexto de investigación? En caso afirmativo ¿Los considera útiles para obtener datos de alta precisión sobre la población?
6. ¿Sirve un estudio bien diseñado para probar si son óptimos los resultados de una intervención de SPP? ¿Por qué tipo de diseño se inclinaría?

7. ¿Qué barreras considera que son las más críticas para el desarrollo de la Salud Pública de Precisión? ¿Cuáles son los riesgos que considera más relevantes?
8. En el sector privado que es el que tiene mayor acceso a nuevas tecnologías ¿Quién liderará los estudios de Salud Pública de Precisión? ¿Cuál será el papel del sector público y el privado en el uso de nuevas metodologías de investigación?
9. El término Salud Pública de Precisión ha sido criticado por presentar dos concepciones diferentes; una, está más relacionada con el uso de nuevas tecnologías para mejorar procesos tradicionales de salud pública, mientras que la otra está más relacionada con la genómica. ¿En qué sentido entiende que debería desarrollarse el concepto de Salud Pública de Precisión?

4. TRANSCRIPCIONES

LEYENDA DE HABLANTES:

MODERADOR

CASTRODEZA: José Javier Castrodeza Sanz (Universidad de Valladolid)

BANEGAS: José Ramón Banegas Banegas (Universidad Autónoma de Madrid)

BELDA: Cristóbal Belda Iniesta (Instituto de Salud Carlos III)

ESTUPIÑÁN: Francisco ('Paco') Estupiñán Romero (Instituto Aragonés de Ciencias de la Salud)

PEIRÓ: Salvador Peiró Moreno (FISABIO)

ALFARO: Mercedes Alfaro Latorre (Ministerio de Sanidad)

SANZ: Guillermo Sanz Santillana (Hematología La Fe)

BENAVENT: Diego Benavent (Hospital La Paz/ Savana)

CÓDIGOS:

Subrayado Amarillo: Términos en duda.

[Inaudible]: Término/s ininteligible.

Marcas de tiempo: Comienzo y Fin del Audio y cada 10 minutos.

CÓDIGO HABLANTE	TRANSCRIPCIÓN
[00:28:20]	<i>[Inicio de la transcripción]</i>
MODERADOR	Vamos a dar, entonces, comienzo a la reunión. Parecemos británicos, son las cuatro en punto. Muchas gracias a todos por vuestra participación. Creo que tenemos a un grupo de panelistas de lujo para el asunto que nos compete, que es la salud pública de precisión. Contamos con JAVIER CASTRODEZA, que todos le conocéis porque fue Secretario General del Ministerio de Sanidad. Es catedrático, actualmente, de Medicina Preventiva y Salud Pública de la Universidad de Valladolid. JOSÉ RAMÓN BANEAGAS también es catedrático de Medicina Preventiva y Salud Pública de la Universidad Autónoma de Madrid. PACO ESTUPIÑÁN es un investigador muy potente en Ciencia de Datos. Trabaja en el Instituto Aragonés de Ciencias de la Salud que hacen investigación de altísima calidad en datos de vida real y en <i>big data</i>. Gracias, Paco. CRISTÓBAL BELDA, es oncólogo de formación y actualmente es el subdirector general de Evaluación e Investigación del Instituto de Salud Carlos III. GUILLERMO SANZ, el jefe de Hematología en el Hospital Universitario y Politécnico de La Fe, de Valencia. MERCEDES ALFARO, subdirectora general de Información Sanitaria del Ministerio de Sanidad. Y SALVADOR PEIRÓ es investigador en la unidad de investigación de servicios sanitarios de FISABIO en Valencia. DIEGO Benavent, es la persona que me queda por presentaros. Diego es también colega, es reumatólogo, FEA de Reumatología del hospital universitario de La Paz. Trabaja <i>part time</i> en La Paz porque la otra parte de su jornada la dedica a Savana. Es una persona que ya ha colaborado con nosotros en otros proyectos y también, como Paco, muy conocedor de todo el mundo de la ciencia de datos. Además, es el autor, como habéis podido ver, de la revisión sistemática de la literatura que hemos hecho en uno de los dos documentos que os hemos remitido. El segundo documento lo mandó nuestro analista para este estudio, que fue Juan Ramón González, de IS Global, que dirige la unidad de Epidemiología Genética en esta macro organización de Salud Pública que está ubicada en Barcelona. Muy bien, pues vamos a empezar. La primera pregunta es que nos digáis, y abrimos, como siempre, una ronda en la cual hablamos todos y luego ya empezamos a discutir. Yo procuraré daros, quitaros las palabras y llevar una conducción de la reunión razonablemente ordenada que luego nos facilite el análisis del discurso.

(continúa en la pág. siguiente)

	Bien, la primera pregunta la tenéis todos. Es en qué áreas de salud pública consideráis que pueden tener más impacto las investigaciones realizadas con metodología de datos clínicos a gran escala. Es lo que se viene llamando <i>big data</i>, y las nuevas técnicas que hay para su procesamiento. Abrimos fuego, si queréis, con Paco, que es al primero que le veo a mi izquierda en la pantalla. Paco, <i>avanti</i>.
ESTUPIÑÁN	Sí, pues hablando en genérico, a nivel de uso de tecnologías de información o <i>big data</i> para caracterizar o evaluar las intervenciones de salud pública, creo que tiene toda la relevancia. Tengo la precaución, la misma que decía Salvador al principio, de la terminología, ¿no?, porque nos enfrentamos a lo mejor a estos temas desde marcos conceptuales ligeramente distintos. Y el término “salud pública” abarca muchas cosas. Y normalmente, desde donde lo abordamos es la asistencia de poblaciones. Siempre para hacer intervenciones a nivel poblacional necesitas conocer las características de la población, y eso implica segmentar la población y clasificarla. Eso se basa en un conocimiento de esa población. Cuanto mejor detalle tengas o más fino sea el grano o el trazo, mejor. Claro, eso es la predisposición inicial, otra cosa es la capacidad de traducir operativamente ese conocimiento que ganas con la población en intervenciones que sean precisas o que vayan encaminadas a uno de esos grupos característicos o con un comportamiento especialmente diferente. Y ahí es donde se vuelve mucho más complicada esta cuestión, o de ver la utilidad o de abordarla, porque para la toma de decisiones en Salud Pública al final las características de la población es una de las cuestiones a tener en cuenta, pero no es la única. En el caso particular de la genética o de la influencia de la genética en esta ganancia de información, en la caracterización de las poblaciones, a nivel de las personas, de los individuos, creo que todavía está por definir. Hay una sucesión de prioridades que todavía tenemos que abordar que están muy lejos de eso, y ese abordaje a nivel poblacional y del individuo todavía tiene, en mi opinión, que demostrar su utilidad, más allá de su utilidad marginal. Otra cosa es cuando vamos a la integración de ese tipo de tecnología y de conocimiento a nivel, por ejemplo, de las variantes de organismos patógenos del tipo que sean o de la sensibilidad de determinadas cepas o variantes de esos organismos a determinados tratamientos. Y ahí los vemos en los ejemplos más claros en lo que estamos viviendo, o en las plataformas tecnológicas que se nos abren ahora, como la de la vacuna, como una nueva plataforma tecnológica de la vacuna, donde su utilidad es máxima y está a la orden del día.

(continúa en la pág. siguiente)

MODERADOR	Perfecto, pues yo creo que va a recoger muy bien tus comentarios José Ramón, porque en el preámbulo, cuando estábamos esperando la llegada de todos, significaba que en el tema de vacunas es donde ve más espacio para la salud pública de precisión. Adelante, José Ramón, gracias.
BANEGAS	La Salud Pública de Precisión surge por la excesiva variabilidad de procesos de salud entre poblaciones; y consiste en administrar la intervención, o tecnología, adecuada a la población correcta (específica) en el momento apropiado. Uno podría preguntarse si hay que priorizar las investigaciones y tecnologías de salud en poblaciones por su impacto potencial. Si fuera así, habría que tener en cuenta no solo la magnitud del problema y su vulnerabilidad, sino también su trascendencia sociosanitaria, que incluye la percepción de la población y de sus representantes políticos, y los costes asociados. Por eso, en general, las áreas de salud pública de mayor impacto potencial serían aquellas con problemas de mayor frecuencia, o con mayor trascendencia, o más vulnerables, o de menor coste o mayor eficiencia de salud pública. No obstante, lo más experimentado, o quizás con mayor porvenir, podrían ser una serie de áreas específicas, como las 10 que siguen: (i) Metodología y Estadística (análisis bayesianos, linkage y síntesis de datos, análisis de áreas geográficas pequeñas); (ii) Brotes Infecciosos (toxi-infecciones alimentarias clásicas); (iii) Enfermedades Infecciosas Emergentes y “Previsibles” (e.g., SARS-Cov-2; Influenza y sus variantes). En la Covid-19 (u otras epidemias emergentes), la estrategia one-size-fits-all no es sostenible, y el uso de SPP es más factible (búsqueda de casos, rastreo de contactos, cuarentena de contactos... con salud digital, health apps...); y Vacunación Administrada como SPP (por ejemplo, vacunas A-Z o JJ a grupos vulnerables no trombo-génicos); (iv) Gestión Sanitaria (por ejemplo, formación específica de nuevos tipos de gestores) y Medicina del Trabajo (tipos de exposiciones laborales); (v) Envejecimiento (por fragilidad, discapacidad...); (vi) Enfermedades No Comunicables (Cardiovasculares, Cáncer, Demencias –Alzheimer y Vasculares); (vii) Farmacología clínica (fármacos a la carta); (viii) Salud Pública Fetal y Neonatal (trastornos de la nutrición en vida temprana: “todos comemos pero unos menos y peor que otros”); (ix) Sociología y Economía: Determinantes Sociales [N-E, Constructos comprehensivos], Psico-Sociales y Socio-Económicos de la Salud; (x) Biotecnología e Ingeniería de SP, etc.

(continúa en la pág. siguiente)

[00:40:40]	
MODERADOR	Muy bien, José Ramón, muchas gracias. Javier Castrodeza, cuando quieras, gracias.
CASTRODEZA	Buenas tardes a todos. Es una satisfacción conocer y reconocer a gente con la que hemos compartido cuestiones. En mi opinión, no me acaba de convencer el concepto de salud pública de precisión. Empezaré diciendo que tengo muchas dudas al respecto; en España, hay una costumbre de que cualquier palabra que nos gusta, rápidamente la adaptamos. No quiero ser el crítico ni el discordante de vuestras opiniones, pero pienso que la salud pública, su gestión, su actuación y su toma de decisiones, muchas veces está alejado del término de precisión. Por lo tanto, conjugar estos dos términos ya me parece difícil. Yo he tenido la fortuna de vivir operativamente la salud pública; quizá tengo una visión un poco alejada de algunos de vosotros, porque hay una gran distancia entre lo que se pone en el papel, lo que se dice que se va a hacer, lo que luego se quiere hacer, lo que la gente quiere hacer y luego lo que, de alguna forma, la rutina de trabajo. A mí me gustaría que de este grupo de trabajo saliera una idea de alguien que definiera qué es la salud pública de precisión. Si lo logramos, yo creo que ya daríamos por bueno este proceso, porque hablar de salud pública de precisión a mí me produce el siguiente interrogante: ¿De qué estamos hablando? ¿De la epidemiología de precisión? ¿De la epigenética de precisión? ¿De la epigenómica de precisión? ¿O bien estamos hablando de la seguridad alimentaria de precisión? ¿O estamos hablando de los sistemas asistenciales, de todo el proceso asistencial de primaria, de la asistencia sanitaria de precisión? ¿O estamos hablando de la precisión ambiental para la toma de decisiones por su repercusión en salud pública? O como alguien comentaba antes, ¿del espacio sociosanitario de precisión, que genera una información enorme? ¿O estamos hablando de tener un <i>big data</i> que aporte un gran valor? Pero no de la salud pública, porque en la salud pública hay más interlocución, ¿no? Por lo tanto, mi primera duda es: ¿es oportuno un concepto de salud pública de precisión? Yo lo veo difícil, no lo sé definir. A lo mejor aprendo de todos vosotros, y llegamos a una definición precisa. Sí que me interesa trasladar el concepto de precisión como <i>big data</i>, porque es verdad que nuestro sistema sanitario genera tal nivel de información que está infrautilizado. Probablemente dentro de no demasiado tiempo, en los ensayos clínicos ya no sea necesario ni hacer fase I, II, III o IV. Bastará con explotar los <i>big data</i> para saber lo que tiene que pasar.

(continúa en la pág. siguiente)

	Entonces, por ubicar una respuesta a la pregunta que tú formulabas, ¿qué espacios se entienden como los más aprovechables para el tema de la precisión en el ámbito de la salud pública? - yo soy de los que cree que hay que ir paso a paso. Primero desarrollemos el conocimiento y las conclusiones de todo el mundo farmacéutico, que es donde ha nacido todo el tema de la medicina de precisión, con los tratamientos específicos del mundo oncológico. Seguramente que Cristóbal cuando intervenga nos va a dar mucha más luz en ese sentido. Para mí yo esperaré un poco para ver de lo que somos capaces, y a partir de ahí, sabemos que somos conscientes que tenemos un gran <i>big data</i> que tenemos que saber explotar y si me apuráis saber vender y utilizar en beneficio del propio sistema, porque esta es otra de las cosas que implica. Mercedes sabe mucho de esto: que manejamos datos y datos y datos... y tal y luego el sistema no aprovecha nada de la información que tiene que podría llegar incluso a la utopía de poderse autofinanciar.
MODERADOR	Muchas gracias, Javier. De eso se trata, de generar controversia. Muy atractivo lo que has dicho. A ver, que no se olvide que la pregunta, el tractor es con metodología de datos clínicos a gran escala, es decir, DVR y <i>big data</i>, como se ha apuntado. Cuáles son las áreas de salud pública donde pueden tener más impacto la salud pública de precisión. A ver, en pantalla... Salvador Peiró tiene ahora la palabra.
PEIRÓ	A ver, intentaré poner orden. Creo que la salud pública tiene tres grandes áreas. Una es la protección de salud, otra es la prevención (primaria, secundaria, etcétera), y otra es la promoción de salud. En la parte de protección de salud estamos hablando de control de aguas, alimentos, saneamiento, calidad del aire, higiene alimentaria, contaminación, etc., con ese enorme volumen de normas y prohibiciones (en salud pública nos gusta prohibir u obligar). En este terreno de la protección de salud mediante normas y vigilancia, la medicina de precisión, entendida como aplicar una intervención selectiva a un grupo de personas que hemos identificado por características especiales, tiene entre poco y ningún sentido. Ni en la vivienda, ni en el agua potable, ni las residuales, la higiene de alimentos, etc. En estos aspectos hay poco o ningún espacio para la salud pública de precisión. No sé si habrá algún aspecto que podría encajar, pero esta parte de la salud pública está fuera del área de "precisión". En las partes de prevención puede tener cabida en algunos aspectos, y en la parte de desarrollar actuaciones de promoción sobre poblaciones específicas podría también tener cabida.

(continúa en la pág. siguiente)

En segundo lugar, parece que cuando hablamos de salud pública de precisión estamos hablando de aplicar intervenciones específicas, pero en cierta forma poblacionales, a subgrupos de población, en un momento determinado y en un lugar determinado, aprovechando dos tecnologías relativamente nuevas: el *big data* y la genómica. Son tecnologías que salud pública va incorporando paulatinamente. En la Comunidad Valenciana, y desde hace tiempo, con tuberculosis y con otras cosas, la relación entre salud pública y genómica de patógenos es enorme. Iñaki Comas en el IBV-CSIC, Fernando González Candelas en FISABIO, ... y la verdad es que la aportación de la genómica a la salud pública de las enfermedades infecciosas está siendo formidable. En salud pública la genómica de patógenos tiene una importancia enorme, no tengo ninguna duda de esto. Y cada vez tendrá más papel. Y no solo en la COVID. En los brotes alimentarios, en las cadenas de hepatitis C, en tuberculosis, ... una importancia enorme. En tercer lugar, no deberíamos confundir la salud pública con la epidemiología. La epidemiología es una disciplina científica. La salud pública incluye muchas cosas. En el fondo acaba siendo una toma de decisiones en muchos campos. A veces la epidemiología nos sugiere unas cosas y la interpretación de los comportamientos poblacionales, que también están incluidos en la salud pública u otros aspectos, hacen que no se tomen las decisiones que serían razonables desde un punto de vista estrictamente científico. Esto en la epidemia, en la pandemia, no hemos dejado de verlo casi diariamente. Se toman decisiones bastante distintas a las que dictaría el uso exclusivo de la evidencia científica. La salud pública es política. No política en el sentido de políticos, pero son políticas sobre la población para conseguir determinadas cosas. La epidemiología es una ciencia, y la política funciona absorbiendo información de muchas ciencias, de muchos sitios, de mucha intuición también, y de muchas cosas. Por esto, digamos, creo que no debemos confundir ambos parámetros. La epidemiología es una disciplina científica y no solo patrimonio de la salud pública poblacional. La epidemiología está en la base de los ensayos clínicos, de la mayor parte de ciencias clínicas y es parte fundamental de la salud pública pero no es toda la salud pública.

[00:50:35]

(continúa en la pág. siguiente)

PEIRÓ

En el mismo sentido, tampoco creo que debamos confundir medicina de precisión y salud pública de precisión. La oncohematología es el ejemplo más claro de la medicina de precisión; también las enfermedades raras lo están siendo. La medicina de precisión está ahí aportando muchísimo y la salud pública funciona de otra forma. En salud pública, en general, tenemos que hacer las cuentas de otra forma porque no enfrentamos a los pacientes de uno en uno, como hacen los clínicos, sino que los enfrentamos normalmente en dos grandes grupos, experimental y control, o en dos grandes poblaciones sobre las que tenemos que actuar o no, y tenemos que acabar viendo beneficios poblacionales.

Y digo esto porque, lo comentaba al principio, los de salud pública tenemos en la cabeza siempre aquellas pirámides de Geoffrey Rose y aquel juego de que actuando en poblaciones de bajo riesgo se pueden conseguir grandes beneficios poblacionales que a veces no consigues actuando sobre las poblaciones de alto riesgo. Mientras que la medicina de precisión o las tecnologías de precisión buscan sobre todo poblaciones de alto riesgo de padecer algo, en salud pública podemos y debemos mirar al conjunto de la población, incluso la de bajo riesgo.

Pero también coincido con que tenemos que incorporar toda una estrategia, una tecnología a la salud pública. Y también coincido en que puede ser una oportunidad importante para incorporar mucho más el *big data* a cosas que hacemos en salud pública, los aspectos de precisión ... y para incorporar mucho más todavía la tecnología genómica a lo que estamos haciendo en este momento.

Y no me preocupa el nombre. La “medicina basada en la evidencia” es una marca comercial y se vende mucho mejor que cuando Feinstein la llamaba *clinimetría*. Si queremos hablar de salud pública de precisión, a mí no me preocupa lo más mínimo, pero en el fondo lo que tenemos que hacer es aplicar las tecnologías que nos mejoran la salud pública, sobre todo en los aspectos de prevención a la salud pública. Y esto es lo que importa en el fondo.

MODERADOR

Gracias, Salvador, estupenda tu intervención también. Mercedes, cuando quieras.

(continúa en la pág. siguiente)

ALFARO

Yo la verdad es que creo que habéis comentado ya prácticamente todo lo que yo había reflexionado leyendo los documentos que habéis mandado, y pensando en esta reunión. ¿Cómo veo yo que puede favorecer el *big data*, el desarrollo, tanto de la vigilancia de la salud como de las actuaciones de salud pública? Pues hombre, yo me pongo siempre en cosas muy concretas; por ejemplo: sistemas que permitan tratar la información de una manera mucho más rápida y mucho más inteligente para poder predecir o anticipar posibles peligros, posibles amenazas. Es decir, ahora mismo con la pandemia, todos lo hemos vivido, se están haciendo numerosas iniciativas de simulación a la hora de establecer medidas de protección a la población. Por lo menos nosotros en el Ministerio estamos haciendo algunas simulaciones, fundamentalmente las está haciendo el centro de control de alertas, pero no estamos haciéndolas solo nosotros. Se están haciendo simulaciones que ayuden a saber cómo actuar, e incluso algunas están saliendo en los medios. Y están hechas por empresas no estrictamente sanitarias sino que trabajan en *big data* en general. ¿Para qué veo yo que también pueden emplearse las tecnologías basadas en grandes bases de datos y en datos recogidos de fuentes muy diversas? Pues para el rastreo de datos, para el rastreo de casos, por ejemplo. No hemos tenido muy buena experiencia con Radar COVID en ningún país, pues porque posiblemente se hizo, como todo en estos tiempos, muy rápido, y se necesita de una voluntariedad del usuario, de una comprensión de la aplicación que no sé si la hemos llegado a entender en este país, o porque le falta posiblemente sensibilidad para identificar contactos verdaderos; con que estuvieras a metro y medio de una persona más de quince minutos ya te marcaba como que habías estado al lado de un caso aunque hubiera una pared por el medio y fuera el vecino de al lado, con el que no te ves nunca. Pero en el rastreo de casos y la identificación de contactos yo sí que pienso que puede tener bastante utilidad: la trazabilidad precisamente de estas enfermedades, fundamentalmente de las transmisibles, también el rastreo de actividades como el ejercicio físico o la alimentación, que puedan estar marcando tendencias específicas relacionadas con el estado de salud sobre las que hay que actuar, y que hay que conocer bien. Fundamentalmente el tener muchos datos te permite crear patrones que hacen posible, de alguna manera, caracterizar mejor a las poblaciones, y actuar para grupos más específicos. Habéis dicho, y estoy de acuerdo, que la salud pública por definición actúa sobre poblaciones o sobre grandes grupos y no sobre individuos. Es ahí, sobre la persona, donde estaríamos

(continúa en la pág. siguiente)

	hablando más de la medicina de precisión. Pero también se pueden tomar decisiones en salud pública o fomentar actuaciones o hábitos o comportamientos de una manera más cercana a lo que determinados grupos pueden necesitar o a lo que determinados grupos están dispuestos a recibir. Y ahí también pues nos puede ayudar la información que nos dan los <i>big data</i> .
MODERADOR	Muchas gracias, Mercedes. Cristóbal, cuando quieras.
BELDA	Desde mi perspectiva de oncólogo médico, creo que debemos intentar establecer hasta dónde nos han llevado las aplicaciones actuales del entorno Big Data y si las predicciones de su utilidad son realistas. En este sentido, querría destacar que las utilidades finales en la práctica clínica deberían exponerse al mismo rigor científico al que se han expuesto otras tecnologías con aplicación clínica antes de su incorporación a la rutina asistencial.
[01:00:05]	
BELDA	Junto con ello, la pregunta final, a mi modo de ver, es si más allá del conocimiento que nos pueda generar este entorno matemático, su aplicación nos ha permitido tomar decisiones diferentes en una situación en la que no teníamos otra opción. En cualquier caso, hasta ahora no existían un despliegue que permitiera ordenar esta tecnología y su imbricación con el Sistema Nacional de Salud, principalmente desde el punto de vista de generación de conocimiento que, posteriormente, pudiera ser empleado con fines asistencial. Así pues, desde el ISCIII hemos creado la primera infraestructura que será capaz de generar evidencias científicas para su posterior aplicación: el programa IMPaCT. Este programa está configurado por tres actuaciones que establecen las bases iniciales de obtención y gestión de información para su posterior exploración clínica. Por un lado, hemos incluido una cohorte de 200.000 personas, construida sobre base censal, multipropósito y con una amplia representatividad de toda la geografía española. Esta cohorte generará la sobre comportamiento, datos de ambiente, datos de contaminación, datos de enfermedades junto con, y ahí entra la segunda actuación del programa, datos genómicos. Para que cohorte y genómica puedan generar algoritmos es necesario que exista una aproximación matemática que permita buscar esas relaciones: esa es la tercera actuación de IMPaCT, la ciencia de datos en sí misma. Idealmente, el programa deberá ser capaz de generar capacidad

(continúa en la pág. siguiente)

de predicción, como ha dicho Mercedes, en oncología, cardiología, enfermedades infecciosas, etc. En este sentido, creo que nos debemos plantear si ¿realmente la ciencia de datos nos permite establecer la dimensión predictiva, que en política de salud sobre poblaciones pueda tener impacto? Todo ello visto desde una perspectiva científica, tal y como es la orientación del programa IMPaCT. En cuanto a sus aplicaciones, más allá de lo que he expuesto anteriormente, para muy intuitivo, una primera orientación hacia enfermedades infecciosas. Sin embargo, habrá que explorar si realmente la toma de decisiones basada en este tipo de herramientas es mejor, peor o igual que la toma de decisiones basada en la metodología convencional. En cuanto al resto de áreas de conocimiento, tal y como exponía el Prof. Castrodeza en su intervención, quizá en oncología podría ser igualmente intuitivo. Aquí deberíamos avanzar algo más en cuanto al uso de la terminología “precisión” para evitar caer en el uso de dos marcadores cuando, quizá, estamos buscando una aproximación algo más avanzada con un mayor número de variables. En resumen, todavía no me parece sencillo definir las áreas que más se van a beneficiar más allá del beneficio intuitivo que muchos creemos que va a generarse. Aquí es donde el planteamiento del rigor científico debe ser capaz de identificar esos beneficios a través de un método robusto, similar al empleado con otras tecnologías previamente a su incorporación a la práctica clínica. Así, en estos momentos, debemos seguir promoviendo estructuras que como país nos van a permitir obtener la información necesaria para exponer al método científico la utilidad de este tipo de aproximaciones. De momento es el programa IMPaCT, mañana serán un montón de cosas que seguramente se irán haciendo, pero de momento es eso: cohorte, genómica y análisis masivo de datos utilizando capacidades de super computación, y a partir de ahí evidentemente la onco-hematología, la toxicidad asociada a la falta de eficacia de los fármacos y las enfermedades infecciosas entre otras áreas de conocimiento.

MODERADOR

Muchísimas gracias, Cristóbal. Guillermo, cuando quieras.

SANZ

Vamos a ver. La primera cosa que quería comentar es que aparte de trabajar como hematólogo ahora estoy dedicando mucho tiempo a dirigir como *co-chair* dos proyectos europeos del *Innovative Medicines Initiative* dentro del programa *Big Data for Better Outcomes* y que, en nuestro caso, está dedicado a las neoplasias hematológicas. Lo que pretendemos es mejorar los resultados de nuestros pacientes, acelerar el acceso a fármacos

(continúa en la pág. siguiente)

más eficaces y menos tóxicos, es decir, medicina de precisión, y también reducir el coste de los ensayos clínicos de las compañías farmacéuticas y, al final, que el coste de los fármacos sea menor y podamos, de alguna manera, garantizar la sostenibilidad del sistema sanitario público.

Esto, por un lado. Además, quería hacer un breve comentario relativo a lo que habéis comentado de las vacunas y el COVID. Y es que creo que no hemos dado, y probablemente porque no hemos dispuesto de esa enorme cantidad de datos que en este momento podíamos haber usado, de un análisis de cuál es la incidencia, por ejemplo, de trombosis en los pacientes con COVID. Si nosotros estamos hablando de que con las vacunas hemos tenido, por ejemplo, con la vacuna de Janssen, seis casos en siete millones, y los que hacemos clínica estamos viendo, y me quedo corto, que más de un 5% de los pacientes que ingresan en el hospital por COVID desarrollan trombosis y/o embolia pulmonar imaginaros lo que supone esto. Es decir, 3 millones de pacientes en España contagiados; 300.000 ingresados y de ellos un 5% desarrollan trombosis. Evidentemente la relación beneficio/riesgo, no solamente en términos de supervivencia sino en términos de desarrollo de trombosis, es muy superior con las vacunas. Creo que esta cuestión no se ha comunicado bien, lo que es una pena, porque podríamos haber utilizado esta información para estar ahora disponiendo de todas las vacunas existentes para todos los grupos de población.

Refiriéndome a la pregunta concreta que estáis haciendo: ¿en qué tipo de enfermedades podría tener más influencia la salud pública de precisión?, yo soy muy optimista, y creo que en todas. Tanto en las de alta como en las de baja incidencia.

Evidentemente, por las razones que Javier y algún otro de los ponentes, creo que Cristóbal, han comentado. En enfermedades raras, como el cáncer, pero que, a su vez, es una de las tres principales causas de muerte en los países avanzados, estamos obteniendo resultados genómicos desde hace varios años. Y con estos resultados genómicos estamos siendo capaces de acercarnos, de alguna manera, a conseguir medicina de precisión. Pero esto yo creo que es trasladable a cualquier enfermedad, incluyendo enfermedades de un tremendo impacto, como pueden ser la diabetes, la hipertensión o la cardiopatía isquémica. En nuestro hospital, tenemos un grupo que está dirigido por el doctor Bernardo Valdivieso, que precisamente ha conseguido utilizar todos los datos de nuestros registros electrónicos, unirlos y construir “chorizos” de datos con más de 250.000 variables. Con

(continúa en la pág. siguiente)

	estos datos se están haciendo trabajos en determinadas poblaciones que están ofreciendo resultados sorprendentes. Por ejemplo, en diabetes nos estamos encontrando con que los pacientes más pluripatológicos son los que mejor están controlados y que los pacientes más jóvenes, que probablemente se podrían beneficiar más de ese control, no están siendo tan bien tratados. Y disponemos de herramientas para saber no ya en qué centro de salud un paciente no está siendo, digamos, controlado de la glucemia con la frecuencia debida, sino incluso conocerlo a nivel de un facultativo en particular. Es decir, de alguna manera esto lo que supone es que tú puedes lanzar un mensaje a ese centro de salud o a ese facultativo en particular y decirle quién es este, este y este enfermo que tendría que estar mejor controlado y cartas a esos pacientes para que acudieran a hacerse más controles de glucemia.
[01:10:47]	
SANZ	Esto, por un lado. En el campo, por ejemplo, de la hematología, otra cosa que yo creo que va a tener un impacto enorme en los próximos años es el descubrimiento de lo que hemos llamado hematopoyesis clonal de significado indeterminado. ¿Qué quiere decir esto? Que, por encima de los 60 años, cerca de un 15% de la población tiene mutaciones en sus células de la sangre sin tener ninguna enfermedad hematológica. Y dices, bueno, estas personas tienen más riesgo de contraer después una neoplasia hematológica, pero lo que ha sido muy llamativo en estos estudios es que estos pacientes tienen un riesgo relativo cerca de cuatro veces mayor que la población general de morir a consecuencia de una enfermedad cardiovascular, fundamentalmente infarto de miocardio o accidentes cerebrovasculares. Y ya se ha visto que hay dos de los genes más frecuentes que nos encontramos en esta hematopoyesis clonal de significado incierto, que favorecen la arterioesclerosis. Tanto en modelos animales, inicialmente, como ahora en humanos. Y de hecho se están creando consultas multidisciplinarias en los hospitales, incluyendo evidentemente a los cardiólogos, que están tratando de reducir otros factores de riesgo cardiovascular sobre los que podemos actuar. Evidentemente, si uno tiene unos determinados genes alterados, eso de momento, y digo de momento, no lo vamos a poder cambiar. Pero, sin embargo, sí que podemos hacer un seguimiento de esos pacientes y reforzar las medidas de prevención de enfermedad cardiovascular en ellos. Y seguro, seguro

(continúa en la pág. siguiente)

que estos pacientes van a tener una mayor adherencia al tratamiento sabiendo que están en un grupo de mayor riesgo. Y os recuerdo que este riesgo relativo que he comentado es probablemente superior al que tienen otros factores de riesgo como el tabaquismo.

Es decir, estamos, digamos, en un principio. Yo creo que vamos a poder utilizar información genómica, por un lado, pero por otro lado no nos va a hacer falta. Yo creo que una de las principales ventajas que tiene el sistema de salud pública en España es que tenemos acceso a los datos del 90% de la población. Si somos capaces de unir toda la información no solo de un hospital, o de una Comunidad Autónoma, utilizando los sistemas que todos tenemos, como Alumbra o SIA en el caso de la Comunidad Valenciana, y somos capaces de poner todos nuestros datos en el mismo sitio podremos, probablemente, y yo creo que a lo mejor Diego puede hablar de esto, llegar a saber cuándo va a acudir a la puerta de urgencias por un problema vascular un determinado paciente que sobrevive a un infarto o saber, por ejemplo, cuándo va a desarrollar una nueva crisis hipertensiva un paciente con hipertensión arterial.

Todo esto es factible, y además una ventaja del uso de *Big Data* y algoritmos basados en inteligencia artificial es que no parten de ninguna hipótesis preestablecida. Hasta ahora en medicina y en ciencia en general, lo que hemos hecho es partir de una hipótesis y tratar de demostrarla. Con estos sistemas no es necesario. Es decir, tú introduces todo el volumen de datos en el mismo sitio y automáticamente lo que te generan son algoritmos que te pueden decir “oye, pues mira, este paciente, o este tipo de paciente, o de población, va a tener más riesgo de tener esto, esto o lo otro”. Por ello, creo de verdad que desde el punto de vista de la salud pública tenemos que ir hacia este nuevo modelo de salud pública de precisión. Javier decía que cómo definirla. Disponemos de artículos publicados en revistas de muy elevado impacto, como *The New England Journal of Medicine*, que están tratando de alguna manera de definir que es la salud pública de precisión. Fundamentalmente, en estos casos basándose en la aportación adicional que puede tener la genómica a todos los niveles. Estoy hablando de mutaciones, epigenómica, transcriptoma o metabolómica, lo que queramos. Pero yo creo que, en algunos casos, no necesitamos esa información genómica tan precisa que, en este momento todavía es cara, aunque se haya abaratado mucho en los últimos años, en la que sí que podemos realmente obtener resultados en salud. Que es, al final, lo que todos queremos. Es decir, el poder aplicar tanto a poblaciones específicas, ya sean amplias o pequeñas, modelos que les puedan de alguna manera garantizar eso.

(continúa en la pág. siguiente)

MODERADOR

Gracias, Guillermo. La verdad es que todas las intervenciones han sido enormemente interesantes, muy ricas en contenidos. Si tuviera que elegir una palabra que sintetizara el común denominador sería “predicción”. Quizá es el fuerte de la SPP, la capacidad con la metodología que tenemos y el procesamiento con súper computadoras los datos de las bases públicas se podrían modelizar y predecir muchísima actividad relacionada con la atención sanitaria y con los problemas de salud de la población. A vuestro juicio, las conclusiones arrojadas hasta el momento por estudios que emplean metodologías de salud pública de precisión, ¿son útiles para guiar políticas sanitarias? Si os parece, vamos a empezar con Javier, que tiene esta experiencia personal tan rica, para que con la sinceridad que le caracteriza nos conteste.

CASTRODEZA

Yo puedo parecer un poco pragmático. Porque aquí las intervenciones divergen del modelo de Ross, del poblacional orientado hacia grupos de riesgo o hacia los datos menores. No es que quiera ser, digamos, el discrepante de esto, pero a mí no me gusta utilizar términos, digamos, de moda, para que califiquemos la salud pública cuando no sabemos lo que estamos diciendo; si estamos hablando de ciencia, sí; si estamos hablando de otras cosas, sí, si estamos hablando del *big data*, sí. Pero la salud pública es una cosa para mí, muchas veces, muy compleja. Saltando del mundo académico al mundo operativo. Porque claro, en el mundo académico el papel, digamos, lo soporta todo. Luego ya la realidad es muy distinta.

En relación con la pregunta. Las conclusiones arrojadas hasta el momento por estudios que emplean metodología de salud pública de precisión, ¿son útiles para guiar políticas sanitarias? Pues no lo sé. Debieran ser útiles para guiar políticas sanitarias. Si no, partiendo de que no soy muy partidario del término “salud pública de precisión”, podría decir “con precisión”. Mirad lo que estoy diciendo, estoy cambiando la preposición: “con precisión”. Entonces ya me voy encajando más cómodamente en este término. Con precisión, porque la salud pública se nutre de muchas disciplinas, de muchas ciencias en su toma de decisión. Yo he procurado en mi ejercicio de la política sanitaria intentar siempre tomar decisiones amparadas en datos. En datos lo más precisos posibles, de los que tienes que aceptar que son precisos. Porque la génesis del dato también tiene su propia naturaleza. También hemos hablado de grandes masas de datos, pero a veces también pensar que pequeñas masas de datos también son muy precisas y nos permiten tomar decisiones totalmente precisas.

(continúa en la pág. siguiente)

	Por lo tanto, yo creo que debiera de guiar la política, lo que no siempre es posible. Tampoco es aplicable porque los que realizan las decisiones tampoco están dispuestos a aplicarla. Hay más peso de la política que de la parte sanitaria. Por lo tanto, termino, ¿debiera guiar las políticas sanitarias? Sí, debiera guiarlas.
[01:21:13]	
MODERADOR	Gracias, Javier. Mercedes.
ALFARO	Pues no es porque haya sido mi jefe Javier, pero suscribo lo que dice. A ver, yo creo que todavía hacen falta bastantes estudios epidemiológicos y ensayos clínicos para objetivar que <i>big data</i>, que al fin y al cabo es una metodología de análisis avanzado de datos, que nos está dando una información que nos permite tomar decisiones en políticas sanitarias con mucha certidumbre o con mucha más certidumbre de la que tenemos ahora. Es decir, se ha descubierto y se ha identificado creo que es más del 90% del mapa genético humano, se ha visto su relación con patogenicidad y se ha visto cómo pueden influir en la misma factores medioambientales, pero se desconoce todavía realmente cómo se produce esta relación, al menos hasta donde yo he leído. Entonces vuelvo a hablar del valor predictivo: su valor predictivo, los expertos dicen, yo no lo soy, que es todavía bajo. Entonces, yo estoy de acuerdo; debería servir y servirá seguro, seguro, para tomar decisiones con muchas menos dudas, con mucha más certidumbre. Hoy por hoy yo creo que todavía estamos en los albores de esa época, pero bueno. Aquí hay doctores de esta santa madre iglesia que igual conocen bastante más este tema.
MODERADOR	Pues Cristóbal, por ejemplo.
BELDA	Cualquier aspecto de lo que conceptualmente subyace a la “salud pública de precisión” exige un aumento en la robustez científica de sus planteamientos, más allá de que intuitivamente nos guste a todos el propio concepto. En este sentido, y en política sanitaria todos tenéis mucho más conocimiento del que yo pueda ofrecer, exigen un análisis tremendamente sosegado de muchísimas fuentes de información y estudios, así como una perspectiva política que, quizá, va más allá de cualquier capacidad de computación. Por tanto, creo que todavía nos queda un recorrido por hacer antes de incorporar este tipo de planteamientos más allá de que debemos comenzar a ser conscientes de que todo ello debería formar parte de cualquier sistema de salud pública de

(continúa en la pág. siguiente)

	altas prestaciones. Lo que pasa es que probablemente todavía necesitamos un poquito más de, si me permitís la expresión, un poquito más de perspectiva hacia cuál es el impacto que realmente este tipo de tecnologías pueden tener sobre las políticas sanitarias, que además de la ciencia incluyen muchas otras muchas cosas tal y como habéis comentado.
MODERADOR	Fenomenal. Gracias. Guillermo.
SANZ	Bueno, vamos a ver. Yo, si tengo que ser sincero, hasta la fecha creo que la aplicación de los datos que hemos obtenido, digamos, la evidencia, todavía no ha sido de la calidad suficiente como para poder ser aplicados a políticas sanitarias. Pero también pienso que el motivo de esto, y esto se puede ver perfectamente en la bibliografía que ha preparado Diego, es que la <i>n</i> de los estudios disponibles es relativamente pequeña. Si realmente en vez de esto estuviéramos hablando de millones de datos y casos, e insisto, España es probablemente el país europeo que podría liderar este tipo de estudios precisamente porque en Alemania, por ejemplo, hay mucha gente que está en compañías aseguradoras, por no hablar de Estados Unidos, en la que compartir datos entre distintos hospitales y compañías aseguradoras es realmente complicado. Aquí somos mucho más voluntaristas, por un lado, y por otro que tenemos vocación y por lo tanto ponemos por delante la parte pública que la parte del beneficio económico. Y por eso precisamente, es decir, si somos capaces de reunir esos enormes volúmenes de datos, no solo vamos a obtener información que sí que va a ser aplicable en política sanitaria, sino que, además, como comentaba Javier en su primera intervención, nos van a ofrecer el beneficio económico necesario para mantenerlos en el tiempo. Las compañías farmacéuticas se pueden beneficiar enormemente de esto. Enormemente. Es más, en muchos de los proyectos europeos en <i>Big Data</i>, y HARMONY y HARMONY PLUS no son una excepción, participan un elevado número de grandes compañías farmacéuticas con integrantes de sus equipos de global. Y es clarísimo, esto es un <i>win-win</i>. Es decir, si nosotros somos capaces, por un lado, de obtener datos de enorme calidad y con un suficiente número, estos datos de alguna manera, y no me gusta la palabra, se pueden “vender”, ofreciéndoselos a la industria farmacéutica que puede estar interesada en ellos. Quiero decir, hasta ahora no, pero yo de nuevo pienso, de forma optimista, que en un futuro cercano sí que realmente el uso del <i>big data</i> va a resultar en una medicina de o con precisión que nos pueda beneficiar a todos, tanto a nivel individual como colectivo.

(continúa en la pág. siguiente)

MODERADOR	Gracias, Guillermo. Paco.
ESTUPIÑÁN	A ver si me ordeno, porque al contrario que bueno, que Salvador, yo tengo una cabeza completamente dispersa, entonces a mí me cuesta centrar el tiro. Que deberían influir en política sanitaria, estoy seguro de que debería influir en política sanitaria. Eso no me crea ninguna duda en algún momento determinado. Me gusta pensar en algún ejemplo de influencia que haya tenido ya, y se me vienen a la cabeza algunos ejemplos que pueden resultar entre comillas ridículos o fuera de lo que estamos hablando, pero cuando antes hablaba... cuando antes hablaba, bueno, de protección de la salud, de la obligatoriedad de llevar el cinturón, que es muy obvia, ¿no? Es una medida de salud pública trasladada a la política que ha tenido un impacto real en la disminución de muertes por accidente de tráfico, que no ha requerido de la explotación de datos del sistema sanitario. Ha requerido de la explotación de datos del análisis de causas-raíz de los accidentes de tráfico y de las consecuencias de los accidentes de tráfico, pero además que ha requerido una recopilación de información, una recopilación de datos y el análisis de ese conjunto de datos. Y en ese sentido hay algunas cosas que en nuestra historia han supuesto mejoras críticas de la salud poblacional.
[01:30:20]	
ESTUPIÑÁN	Cosas como la monitorización de la potabilización de las aguas o de la calidad de las aguas, que es una monitorización constante en España, en realidad, pues los sistemas de monitorización y potabilización que tenemos, de potabilización y depuración que tenemos, esa monitorización de la calidad de las aguas está automatizada mediante sensor. Y eso son <i>streamings</i> o corrientes continuas de datos, una monitorización de ciclos rápidos de datos sobre esa calidad. Eso no sería una política sanitaria, pero me voy a la prohibición del hábito tabáquico o del tabaco. O a cosas como la reducción del contenido de sal de los productos de bollería, o de azúcar de los productos de bollería, que tienen en la generación de la política, por supuesto, influyen muchísimas otras cosas. Pero de inicio tienen un comienzo de una relación o una asociación donde de alguna manera hemos comprobado el impacto, y ahí hemos hecho la asunción de causalidad y hemos demostrado la relevancia de esa potencial intervención. O la reducción de la velocidad del tráfico alrededor de las capitales, que se ha propuesto en algunas historias, no ya midiendo el impacto de los accidentes de tráfico, que tiene una derivada del

(continúa en la pág. siguiente)

impacto accidentes de tráfico, sino la externalidad positiva del impacto en enfermedades cardiovasculares y respiratorias, ¿no? Por el impacto de la contaminación ambiental. Ese tipo de cosas son ejemplos que yo creo que podríamos poner como ejemplos de, digamos, aproximaciones más rudimentarias, no de las aproximaciones más de medicina de precisión de las que estamos hablando. Pero sí son ejemplos reales de salud pública. Yo creo que ahí sí que podemos tomarlos como algunas de las direcciones a seguir. Ya ni hablamos de lo que hemos comentado ya. La susceptibilidad de ciertas cepas de tuberculosis a cierta medicación que tenemos disponible... o sea, hoy no se entiende el tratamiento de la tuberculosis si no se tiene en cuenta esa susceptibilidad a la medicación, igual que no se entiende el tratamiento de otras enfermedades. En el ajuste... por ejemplo, lo que comentábamos de los cánceres hematológicos o cualquier otro tipo de cáncer. Bueno, podemos hablar de más ejemplos, pero yo creo que estos son los principales.

Por alusiones, y no entra dentro de la pregunta, pero completamente en desacuerdo con Guillermo en... por una parte, en la venta de los datos a las farmacéuticas, o sea, es una cosa que se me han puesto los pelos como escarpas. Y esa aproximación, que la entiendo, creo que sé por dónde viene, "esa aproximación de la coctelera de datos de donde generamos asociaciones que puedan resultar o no relevantes". A mí desde la experiencia del manejo de los datos, esa aproximación me parece interesante para generarse una hipótesis, pero me parece muy complicada de sostener desde el punto de vista de que muchas de esas asociaciones que vas a encontrar son asociaciones completamente espurias. Te puede llevar mucho valorar que esa asociación sea relevante o que sea una asociación causal y te puede llevar otra vida adicional el probar que siendo una asociación relevante y causal puedas intervenir sobre ella. Así que yo lo veo con mucha, mucha, muchísima cautela.

MODERADOR

Gracias. José Ramón.

BANEGAS

Voy a ser breve puesto que estoy de acuerdo con la mayor parte de los ponentes que me han precedido. Algunas, aunque escasas, propuestas ya se han utilizado y han obtenido algunos resultados beneficiosos. La salud pública de precisión se conozca por este nombre u otro, ya se practicaba desde hace mucho tiempo; un cierto *boom* actual podría verse como una oportunidad de impulso de la salud pública con motivo del auge de

(continúa en la pág. siguiente)

nuevas tecnologías, *big data*, y la escasa repercusión en términos de salud pública de lo que son estrategias (por ejemplo, terapéuticas) centradas en (grupos de) individuos concretos con mal pronóstico o difíciles de manejar. En mi opinión, la salud pública de precisión es necesaria porque no funciona (suficientemente) la Medicina de Precisión. Probablemente por el escaso todavía término de desarrollo es una *big data* con una B mayúscula y con una d minúscula, es decir, tenemos más cantidad, aunque sea relativamente escasa también, como decía Guillermo, de datos que de datos de calidad.

Realmente, la utilidad clínica (o poblacional) de un biomarcador (por ejemplo, digitales con reloj-sensor multifunción de parámetros cardiovasculares y de estilos de vida) tiene que pasar por una serie necesariamente lenta de fases de desarrollo: prueba de concepto, validación prospectiva (estudios epidemiológicos), valor adicional (parámetros de curvas ROC), utilidad clínica (Índices Netos de Reclassificación), resultados clínicos (Randomized Clinical Trials), y análisis coste-efectividad. Pero los resultados relativamente tempranos alientan el esfuerzo adicional.

La medicina de precisión, incluso en terrenos como algunos tumores malignos, puede y suele medirse, entre otros indicadores, en términos de micro-vidas (ganancia de unos pocos meses de vida) y calidad de vida relacionada con la salud; y esto está muy bien. Pero la realidad en términos de salud pública solo sería claramente exitosa en términos de ganancia clara de años de vida saludable.

De todos modos, los determinantes poblacionales (por ejemplo, sociales) de la salud pública de precisión no son necesariamente (aunque tampoco lo contrario) determinantes del riesgo individual; por ejemplo, el colesterol sérico puede discriminar riesgo entre individuos, pero podría no hacerlo entre poblaciones en riesgo similar por colesterolemias parecidas. Otro ejemplo serían las políticas poblacionales contra el tabaquismo (por ejemplo, fiscalidad, publicidad, etc) según países, zonas o grupos de riesgo (adolescentes, etc).

Y en el futuro (o ya en el presente) un ejemplo estaría en las vacunaciones, por ejemplo, tipos de diferentes de vacunas anti-SARS Cov-2 no solo en función de la edad, sino por grupos de riesgo (como de riesgo trombótico), es decir, ir en pos de “vacunas de precisión”, en individuos, pero también entre poblaciones dentro de un país o regiones más o menos envejecidas o con más factores de riesgo de agravamiento de esta u otras infecciones.

(continúa en la pág. siguiente)

MODERADOR	Estupendo, gracias, José Ramón. Salvador.
PEIRÓ	En la Comunidad Valenciana tenemos bastante experiencia de uso de la genómica de patógenos para la toma de decisiones. Nos ha servido para abordar los brotes continuos de legionela que hemos tenido durante mucho tiempo y saber cuáles eran las torres de origen; sirvió para identificar el origen del brote de hepatitis C del llamado “caso Maeso”; sirve para el manejo de las tuberculosis (abordaje de resistencias) y para el seguimiento, porque informa del trazado entre contagios. Y se toman muchas decisiones en torno a eso. Pero, apunto: los brotes de <i>legionella</i> no se acabaron por la genómica sino porque se incentivó la sustitución de las torres de refrigeración financiando el 80% de los gastos del cambio. Las políticas y las intervenciones de salud pública no se basaron sólo en la localización y trazado de la contaminación entre torres de refrigeración. El que la genómica sea útil y sea la solución para controlar un brote son dos cosas distintas. Cambiar las torres era lo importante para solucionar el problema y la intervención –la política– de salud pública era subvencionar ese cambio a las empresas, a los hoteles, etc. En la toma de decisiones basada en datos, el principal cambio de tratamiento en la pandemia, el dejar de usar hidroxiclороquina, se basó en un estudio con bases de datos de mundo real que ... ¡era falso!. Se lo habían inventado y fue retirado. Pero este estudio produjo que dejáramos de usar un fármaco que estábamos empleando en muchísimos pacientes. Nuestro grupo de investigación se dedica a eso. Trabaja fundamentalmente en fármaco-epidemiología usando bases de datos y haciendo tres tipos de cosas: 1) Describir, que muchas veces ayuda; a veces simplemente con ver los datos te da muchas ideas y muchas ayudas; 2) Buscar factores asociados o, visto al revés, predecir que funciona. Los patrones de asociación/predicción están ahí. En inteligencia artificial para el reconocimiento de imágenes, de patrones, en anatomía patológica, en imagen, en ojos, en muchas cosas, y son importantísimos. No es causal, pero reconoce un patrón y eso es utilísimo. 3) Y hacemos una tercera cosa, que es la inferencia causal. Y aquí todo se complica. La inferencia causal tiene dos aspectos.

(continúa en la pág. siguiente)

	Uno es buscar causas, claro. Pero el que trabajamos, sobre todo, es buscar si una intervención, si un medicamento, una tecnología, etc., funciona o no funciona, es efectiva o no, para obtener un resultado. Y esto no depende de la predicción. Cuando estabais hablando tenía en la cabeza varios estudios y, entre ellos, uno de Holland de 2015 ... Un BMJ, creo, en el que se había hecho secuenciación y paneles de predicción cardiovascular, genéticos, etc. Se siguió a la gente, se les dieron consejos y prácticamente no hubo ningún cambio en el comportamiento. Y si no hay cambios en los comportamientos en función de esta información y de las políticas que se ponen en marcha, es poco útil. Y digo esto porque no es lo mismo que se pueda detectar un problema (aunque tiene importancia) a que tengamos una solución. No es automático. Me pasa lo mismo con el ejemplo que ha puesto antes Merche. El de la <i>app</i> de rastreo. Podría haber funcionado, pero no ha funcionado. Puede ser un tema de implementación o cualquier otra cosa. Pero para poner en marcha las políticas, para tomar decisiones, nos hace falta tener evidencia no solo de que se pueda hacer, no solo de que hay un problema, no solo de hacer las cosas, sino que funcionan para cambiar, para mejorar. Y esto también es un tema muy importante y requiere su propio estudio. Puede ser con bases de datos de mundo real, puede ser de muchas formas de hacerlo, pero requiere su propio estudio. Pero no hay cambios automáticos.
[01:40:59]	
PEIRÓ	El hecho de que encontremos que unos diabéticos están mejor o peor controlados no quiere decir que enviándoles un mensaje a sus médicos estos diabéticos pasen a estar mejor controlados. Hay que demostrar esto también. Hay que demostrar que la intervención que tenemos es efectiva para cambiar estos comportamientos. O sea, estamos hartos de ver intervenciones bien intencionadas que no funcionan en salud pública. Y creo que de estas hemos hecho muchas. Entonces, las piezas son diferentes. Ya es útil detectar los problemas. Detectarlos y encontrar que son los más pobres, los más viejos o los más jóvenes o los tal o los cual, los que más los padecen. Esto ya es útil. Ya tiene una utilidad porque permite diseñar soluciones. Pero no quiere decir que automáticamente esas soluciones funcionen.

(continúa en la pág. siguiente)

MODERADOR

Gracias. ¿Queréis alguien añadir algo más, o pasamos a la siguiente pregunta, si os parece? Empezamos con Javier, ¿considera que las nuevas tecnologías de precisión podrán mejorar las actuales estrategias preventivas, el acceso a la atención médica o los resultados en salud en un futuro próximo?

Nos ha faltado, en la pregunta anterior, poner los años. ¿No? Todos decíamos “esto lo vamos a ver, esto lo vamos a ver”. Si os parece, vamos a definir futuro próximo como en los siguientes cinco años. Adelante. Javier.

CASTRODEZA

Sí. Como respuesta a la pregunta, por ser breve y conciso, yo creo que estamos en un camino donde las nuevas tecnologías van a ser incuestionables. Como dice José Ramón, la medicina ha sido poco precisa y la salud pública va a ser muy precisa. Bueno, pues este es un planteamiento muy respetable y luego, José Ramón, ya hablaremos un día más de todo esto. Pero yo, como respuesta a la pregunta, estoy convencido de que las nuevas tecnologías van a suponer una mejora de las estrategias preventivas, eso está claro. Una pieza fundamental de la salud pública es la prevención, está claro. Van a suponer un acceso a la atención médica y seguro que vamos a mejorar los resultados en salud en los próximos cinco años. Yo creo que, a lo mejor de forma global, porque incluso nos podemos encontrar en escenarios donde las tasas de mortalidad global nos den alguna que otra sorpresa. Este año ya, el año 2020, nos ha dado una sorpresa desagradable. Yo creo que ya no solamente las tecnologías de precisión, sino toda la inteligencia artificial nos va a sorprender, porque va a acortar los tiempos en las decisiones o en la propuesta de soluciones. Lo que nosotros tardamos en revisar literatura científica hoy día en inteligencia artificial nos hace un algoritmo que en pocos minutos o en pocas horas es lo que podría ser capaz un investigador en toda su vida dedicada a la revisión de un área temática. Por lo tanto, vamos a asistir a cosas que parecen sacadas de ciencia ficción. Y digo la sanidad como podía decir el mundo sociosanitario. Se van a perder puestos de trabajo porque la tecnología sanitaria los va a sustituir; se va a mejorar la atención de muchas personas que no estamos preparados para atenderlos porque ya no forman parte de nuestra cultura, y la tecnología sanitaria nos va a ayudar a atenderlos. Vamos a asistir no solamente a avances en el diagnóstico; sino avances en la especificidad del tratamiento, avances en los productos sanitarios...no olvidemos que grandes logros de salud pública no se logran por vacunas y por medicamentos, sino por

(continúa en la pág. siguiente)

	logros de productos sanitarios, lo que se llaman los <i>medical devices</i> , que van a contribuir a la salud con una aportación que ni pensamos cómo va a ser. Hoy día, desde el mundo laboral hasta el mundo de las fuerzas armadas están diseñando productos específicos para sus campos que van a hacer tareas, que van a ayudar indirectamente a mejorar la salud y las garantías de seguridad de las tareas. Por tanto, sin duda que para mí la tecnología de forma global y quitada la palabra “precisión” va a suponer un reto que ya me gustaría a mí vivir otra vida para verlos.
MODERADOR	Gracias, Javier.
MODERADOR	Pues Paco, venga, esto de la D grande y la d pequeña, que ha sido... la verdad es una genialidad del amigo José Ramón, te lo pregunto a ti porque claro, cuando estamos preguntando nuevas tecnologías de precisión estamos hablando de <i>big data</i> y estamos hablando de la super computación y el procesamiento. Adelante. La pregunta 3.
ESTUPIÑÁN	Hombre, a ver. Hablábamos de los próximos cinco años, ¿no?
MODERADOR	Sí, exacto, de aquí a cinco años.
ESTUPIÑÁN	Se combinan varias cosas. Se combinan, por una parte, las condiciones iniciales que nosotros tenemos con la velocidad y la aceleración a la que se están dando las innovaciones tecnológicas, ¿vale? Hasta tal punto que en muchas ocasiones nos estamos quedando atrás, porque cuando hablamos de tecnología siempre hablamos de instrumentos, que tienen su aplicación... como todas las cosas. Una indicación, la aplicación dentro de indicación, fuera de indicación y luego la capacidad para interpretar los resultados. Y eso requiere de una capacitación. No solo la capacitación técnica de la aplicación de la tecnología, sino la capacitación para decidir cuándo usarla, en qué condiciones, en qué circunstancias y cómo interpretar aquello que obtienes de su uso. Dicho lo cual, yo creo que estamos... y en nuestra vida creo que siempre vamos a estar ahí. Estamos viendo, probablemente de alguna de estas tecnologías, el principio de lo que llaman el <i>hype</i> , de la fascinación tecnológica. Pero porque se está dando ese desarrollo, porque estamos viendo ese desarrollo en directo y ese desarrollo es muy acelerado. Respecto a la D grande y d chica. Aquí, el tema que tengo es que mientras más conoces de un tema, más desconoces, más información tienes de lo que no

(continúa en la pág. siguiente)

	conoces. Y en este sentido, todas estas cosas son un gradiente, o sea... Cuando definimos cualquier cosa de precisión, ¿cuál es la magnitud de la precisión? O sea, si yo estoy viendo la población como toda la población de España y divido la población por grupos de edad ya tengo un nivel de precisión mayor que cuando miro en toda la población de España. Cuando soy capaz de segmentar no por grupos de edad sino por la edad de forma continua ya estoy viendo mayor precisión. Cuando en vez de solo considerar la edad estoy considerando 25 otros factores, ya tengo un mayor nivel de precisión. Que, ¿hasta qué nivel de precisión nos resulta útil? Eso es lo que estamos dirimiendo todo el tiempo, o sea, Voro hablaba hace un momento de la utilidad de simplemente describir o caracterizar, completamente de acuerdo. Es súper útil. ¿Hasta qué nivel? Pues depende. Depende para qué. Y ahí estamos. En eso es tan relevante la evaluación de las cosas que hacemos, la evaluación de las intervenciones y la evaluación del impacto de las medidas que tomamos en base a esa información. Y en eso estamos. Ya veo a Voro pidiendo la palabra...
[01:50:06]	
MODERADOR	Bueno, bueno... esto se está animando. Cristóbal.
PEIRÓ	O déjame que entre yo. Bueno, lo que queráis.
MODERADOR	Venga. Dale, dale.
PEIRÓ	Hemos desarrollado –diría que desde 2005 hasta ahora, sobre todo hasta 2012– unas inversiones enormes en España en sistemas de información sanitarios. Enormes. Y estas inversiones se han hecho básicamente para hacer lo que veníamos haciendo en papel. Hemos cubierto muy bien los sistemas transaccionales, pagamos las facturas a las farmacias sin problemas y otras cosas de estas. Y tenemos toda esta información. Hacemos lo que veníamos haciendo: almacenamos información, podemos pedir las pruebas diagnósticas y otras y almacenarlas. Pero hemos hecho muy poca inversión en la explotación de estos datos y en la capacidad de inteligencia para explotarlos. Y esto es uno de los problemas que tenemos ahora. La capacidad del Sistema Nacional de Salud, en su conjunto, para explotar la información que tiene el propio SNS, es pequeña. Hay cosas. BIFAP es un aspecto importante y yo creo que están trabajando con las bases de datos de farmacia. SIDIAP está

(continúa en la pág. siguiente)

también haciendo estas cosas. Nosotros con VID en Valencia, Paco en Aragón, etc. Trabajamos estas cosas y hay grupos que trabajan estas cosas más específicamente. Pero necesitamos incrementar enormemente la capacidad de explotación de los datos, y esto es la inteligencia del sistema.

Esto requiere mejoras en los sistemas de información, en su facilidad de explotación. Creo que se aprovecharán los fondos *Next Generation* para producir un salto importante en estos sistemas y modernizarlos. El Sistema Nacional de Salud tiene que pensar qué sistemas de información quiere tener dentro de cinco o diez años aprovechando unas inversiones que en este momento se pueden hacer. Y me parece que es muy importante.

Y eso no pasa solo por almacenar datos. Eso pasa también por generar la capacidad, la inteligencia para explotar y analizar esos datos en muchos sentidos. En mucha diversidad. Y eso requiere también incorporar este aspecto a los grupos científicos. Y hablo también de las universidades, porque hay grupos que trabajan en análisis de texto, grupos que trabajan en análisis de píxeles y mucho grupo por ahí que está en esto. Y los necesitamos, y necesitamos aunar esfuerzos.

El mundo está cambiando. En este momento en mi grupo de investigación hay más matemáticos, informáticos e ingenieros que gente que venga del sector sanitario. No es que podamos vivir sin sector sanitario en este campo, pero estamos cambiando en esto. También harán falta nuevas profesiones en los hospitales y en el propio sistema sanitario. Yo diría que en mi Conselleria hay licenciados en derecho (sobre todo) y, también, sanitarios, médicos y enfermeras. Y farmacéuticos. Y esto debe cambiar. Necesitamos informáticos. Claro que tenemos informáticos, pero necesitamos otro tipo de informático que no sale exclusivamente de las licenciaturas informáticas.

Necesitamos cambiar esto. Necesitamos tener un sistema muchísimo más inteligente en la explotación de datos que lo que tenemos actualmente. Más allá de las inversiones en las mejoras de infraestructuras de explotación, de ingesta de datos, de extracción de datos, de análisis... Y además diría que es un "ahora o nunca". Nunca, o mucho más tarde, porque la verdad, si no aprovechamos los fondos *Next Generation* ahora para hacer este esfuerzo después será muy complicado.

Y, otra cosa ahora que está Merche. Si el Ministerio es capaz de liderar esto lo pagaremos una vez. Y si el Ministerio no es capaz de liderar esto lo pagaremos 17 veces.

(continúa en la pág. siguiente)

MODERADOR	Gracias, Salvador. Quizá por alusión y por seguir el hilo conductor, Mercedes.
ALFARO	Pues sí. Pues Voro, tienes toda la razón. Realmente, además, las tecnologías permiten que se trabaje de una manera, como plantean los tecnólogos, digamos, federada. Y yo creo que es una idea muy buena, porque todos somos muy celosos de nuestra propia producción de datos, y estamos dispuestos a ponerla a disposición del resto, siempre que sea recíproco. Pero nos gusta tener cierto control, y cierto dominio sobre los datos. Ahí tenemos una serie de claves, que una de ellas es que los datos sean de calidad. Antes lo habéis mencionado. Yo sí que creo en el poder de las tecnologías de manera que pueda hacerse conversión de texto libre, datos estructurados... estoy segura, pero hoy por hoy los datos estructurados es lo que nos está garantizando que lo que se está cruzando entre sí de verdad tiene sentido y que los resultados que tenemos son inteligibles y son comparables. Entonces, tenemos que ponernos de acuerdo en el Sistema Nacional de Salud en tener unos sistemas de normalización importantes, de la información básica, y unos lagos de datos que formen un espacio nacional de datos donde todos, de manera federada, podamos trabajar, y podamos usar y utilizar esa información para beneficio, fundamentalmente, de todos. Y, además, tanto de la vigilancia como de la asistencia como de la investigación. Los datos, ahora mismo, son la materia prima fundamental, lo sabemos, estoy diciendo una obviedad, para avanzar en todos los aspectos de la sanidad. Dicho esto, cuando se habla de si en los próximos años van a poder mejorar las actuales estrategias, creo que hay dos cosas que son importantes. Una es que tienen que demostrar, y tenemos que poder evaluar que la efectividad que está prometiendo los grandes volúmenes de datos, los <i>big data</i>, se nos devuelve en forma de diagnósticos más tempranos, de reducción de tratamientos ineficaces, de reducción de complicaciones, de reducción de efectos secundarios. Y que, además, han sustituido otros costes sanitarios. O sea, para mí son esas las dos claves. Una, que nos den resultados de ese tipo. Estamos hablando de salud pública y yo aquí estoy hablando un poco más del paciente, pero apliquémoslo a la salud pública, también sirve. Nos sirve por ejemplo para ser más efectivos. Pero, además, debemos asegurarnos de que no estamos añadiendo costes a los que ya tenemos. Es decir, estas tecnologías tienen que sustituir otras cosas que estamos haciendo, porque si no seguiremos encareciendo el proceso y, además, acabaremos teniendo incluso resultados contradictorios.

(continúa en la pág. siguiente)

	Entonces, en los próximos cinco años yo creo que tenemos que ser capaces de montar todo este entramado y empezar... hay sitios ya donde Paco, FISABIO, el Centro de Computación, el Carlos III... ya están sacándole rendimiento a grandes bases de datos. Pero para llegar de verdad a trabajar coordinadamente creo que estos cinco años son los que nos faltan para organizar bien todo eso y empezar ya a tener algunos resultados. Pero tenemos que empezar la casa por los cimientos. Ahora mismo está punto de empezar a correr un borrador de estrategia de salud digital del Sistema Nacional de Salud, que empezará a circular para recibir sugerencias para enriquecerla y que tiene que ser un poco la que nos dé el marco y las reglas del juego para que vayamos todos a una, que no vaya el gaitero por una calle y el tamborilero por la otra tocando la misma pieza, que es que es lo que nos pasa muchas veces.
BANEGAS	A mí me ha gustado mucho la reflexión de Salvador y lo que ha comentado Mercedes. Realmente, bueno, yo primero decir, muy brevemente también, que yo respondo sí a esta pregunta; aunque más como meta abierta a la esperanza, que como predicción “científica”, como tantas proyecciones en ciencias de la vida (incluso a corto plazo, en unos pocos años), pues no sabemos cómo y cuánto van a evolucionar la tecnología y las poblaciones objetivo de las mismas. Parece claro que la mejora de la precisión, en el sentido de información de mayor detalle, más fina, más específica, más directamente aplicable a situaciones concretas, a cada segmento o área de variabilidad, pueda dirigir más efectivamente diagnósticos (de salud de cada comunidad, predictivos, de cribado o confirmación) o intervenciones (preventivas o terapéuticas), en definitiva, de tecnologías médicas, como fármacos, aparatos o programas, aplicadas selectivamente en las poblaciones o en sus “diferentes” partes integrantes. Por ejemplo, los <i>Big Data</i>, aunque actualmente más con la B que con la d en mayúscula, y no absolutamente necesarios en la SPP, al integrar información de varias y variadas fuentes poblacionales (registros sanitarios, administrativos, censos, tests de laboratorio, <i>wearables</i>, health apps, tests genéticos, etc), producen datos más informativos (no solo por mayor precisión estadística) y, además, son una información obtenida de forma más rápida. Pero es necesaria una formación de calidad y específica en su manejo y explotación, con profesional que sepa hacer y hablar con sus colegas científicos (básicos o aplicados) u otros usuarios de la

(continúa en la pág. siguiente)

	información, y con financiación también específica y finalista de dicha formación (que en sí misma es también formalmente evaluable –investigación de evaluación). Los score genéticos (poblacionales) aportan (parece) menos, aunque su evaluación es investigación abierta al futuro y a múltiples problemas interrelacionados que abren nuevos campos y preguntas.
[02:00:04]	
MODERADOR	Gracias. Cristóbal, ahora sí.
BELDA	Estoy completamente de acuerdo con la Dra. Alfaro. En cualquier caso, es necesario poner cierto sentido a todo a través del elemento tractor que subyace a las competencias del Ministerio de Sanidad. Es más, el más reciente decreto de estructura del actual Ministerio de Sanidad suponía un mensaje muy claro en este sentido. Y quizá el momento para poner este orden es ahora. Así pues, se debería delimitar cuál es el papel de los distintos actores que terminan por componer nuestro SNS. Y en esto, me vais a permitir, retomar la posición disruptiva de Javier. Yo no tengo nada claro que tenga que haber equipos multidisciplinares. Yo creo que tiene que haber equipos que configuren las herramientas (ahí sí que pueden ser multidisciplinares), que sean capaces de testar si esas herramientas permiten mejorar los resultados en salud. Los clínicos lo que necesitamos es una herramienta que utilizar y que funcione mejor que lo que haríamos sin disponer de ella. Y a lo mejor nos da igual qué es lo que hay en la caja negra que termina alimentando esa herramienta. Por lo tanto, independientemente de que sea necesario tener cierta formación, lo que necesitamos en una consulta o en un quirófano son herramientas que te permitan funcionar mejor que sin ellas. ¿Y cómo funcionan esas herramientas? Seguro que es súper importante, pero es muy difícil, muy, muy difícil, cuando alguien ha estado manejando fármacos de altísima complejidad como puede ser la quimioterapia, o haciendo técnicas de altísima dificultad, como puede ser la hemodinámica estructural sobre válvulas cardíacas, que además de ese nivel de cualificación le tengamos que exigir un nivel de conocimiento proporcional a lo que subyace a estas herramientas. Yo creo que al final sí que es importante la utilidad en casos de uso concreto, definido, para que ese hemodinamista o ese oncólogo o ese neurólogo que se encuentra delante de un paciente pueda tomar una serie de decisiones sobre las personas.

(continúa en la pág. siguiente)

MODERADOR	Guillermo, a ver tú como clínico si coincides, o quizá...
SANZ	No, sí que coincido. Quiero decir, la respuesta global a esta pregunta es que sí. Pero aquí voy a empezar haciendo de franco-tirador. ¿Tenemos ahora resultados en salud? Es decir, hablando sinceramente. Yo creo que no. Y las herramientas que nos aporta el <i>Big Data</i>, por ejemplo, nos permiten con datos de vida real el conocer la eficacia de distintos fármacos que se pueden utilizar en la misma enfermedad. Y valorando eso y el coste que tienen esos fármacos podemos establecer un criterio de prioridad en el uso de esos medicamentos. Y esto luego se puede negociar, evidentemente, con la industria farmacéutica, porque tienes los datos. Yo recuerdo que, cuando empezamos el proyecto de Harmony, hablé con Francesco Pignatti, <i>Head of Oncology, Heematology, and Diagnostics</i> de la Agencia Europea del Medicamento, que me reconocía una cosa que es dramática, que es... Decía: "Cuando nosotros tenemos que aprobar un nuevo fármaco, prácticamente nos tenemos que fiar de los datos que nos está proporcionando la compañía farmacéutica que puede y que quiere aprobar ese medicamento. Porque cuando nosotros nos vamos a buscar cuál es la evidencia de los tratamientos estándar, nos encontramos con a lo mejor cien artículos, cada uno de su padre y de su madre, que no nos permiten, porque es una población tremendamente heterogénea, saber cuál es el estándar de tratamiento para esa determinada población. Con lo cual, al final, si te parece que el ensayo está muy bien lo apruebas, o si no te parece bien, lo deniegas. Pero este es el problema de no disponer de datos de vida real, de <i>Big Data</i> en vida real. Estos sistemas, ya digo, nos permitirían efectivamente eso. Obtener resultados en salud de verdad. Por otro lado, volviendo a lo que habéis comentado sobre el tema de la no explotación de los datos que tenemos, que ha empezado Voro a tocar este tema, yo creo que esto es crítico. Es decir, no disponemos de los recursos humanos ni suficientes en término de número ni tampoco en término de formación como para poder hacerlo. Es decir, mientras no se reconozca dentro del Sistema Nacional de Salud que no solo tiene que haber médicos, farmacéuticos; que necesitamos biólogos, que no están ahora incluidos. Que necesitamos ingenieros de datos, que tampoco están incluidos, estaremos en una situación en la que no podremos utilizar y explotar esos datos. Es más, lo que tendríamos que hacer, y parece mentira que no lo estemos haciendo, porque es una cosa obvia, es que en las facultades de medicina se debería

(continúa en la pág. siguiente)

	empezar a formar a los estudiantes en medicina en el tema de conocimientos informáticos para poder manejar estos datos. Un problema que hay, que es evidente es que, cuando un médico que no tiene formación en este tema se junta con un ingeniero de datos, están hablando en lenguajes totalmente diferentes. Y no se entienden. No se entienden. Por lo tanto, tendremos que, desde el punto de vista de la formación de nuestros profesionales, hacer un esfuerzo para que estos profesionales tengan ese tipo de conocimientos que nos permitan en un futuro avanzar más rápido en el conocimiento de estas nuevas técnicas que van a suponer, yo creo, un cambio fundamental en nuestro modo de trabajar. Otra cosa que también habéis comentado es el tema de cómo unificar esos datos. Quiero decir, esto ya está más que claro. Es decir, hay un sistema como el OMOP, que es un lenguaje común europeo, cuyo software además es gratuito, que permite establecer toda la información que disponemos en las distintas diecisiete comunidades autónomas y en los distintos hospitales en un lenguaje común que nos pueda permitir analizar esos datos. Es cierto, esto supone un esfuerzo. Y esto, y yo lo siento mucho, no creo que nuestro Sistema Nacional de Salud esté, en este momento, en condiciones de poder financiarlo. Con lo cual vamos a necesitar ayuda. Y la ayuda tiene que venir de donde en este momento existen los fondos necesarios, de aquellos que, de alguna manera, están gobernando, porque hay que decirlo así, nuestro mundo. Y esto es la industria farmacéutica.
[02:10:30]	
MODERADOR	Gracias, Guillermo. Vamos con la siguiente, que es... afortunadamente ya van a ser más acotadas, con un alcance menor y, por lo tanto, más rápidas en su desarrollo. José Ramón, ¿qué impacto crees que puede tener la sobreinformación y el exceso de datos genómicos en el sobrediagnóstico y en las intervenciones sanitarias?
BANEGAS	Tengo que reconocer, realmente, que conceptualmente o teóricamente el impacto puede ser muy grande, pero alguien, de entrada, vendrá finalmente a poner esto en orden (criterios), entre los datos suministrados por los actuales “cazadores” de genes o de “ómicas”, como ocurrió con los cazadores de microbios (Henle y Koch) en el siglo XIX y con los cazadores de factores de riesgo ambientales (A.B. Hill) en el siglo xx.

(continúa en la pág. siguiente)

	En la práctica, podría y debería ser menor si se aplicara un lógico principio de precaución, con traslación eficaz y razonablemente segura de los conocimientos. Ayudaría disponer de grupos de consenso (tipo las antiguas Conferencias de Desarrollo de Consenso en Estados Unidos, cuya traslación a España en el terreno del colesterol e hipertensión arterial vivimos a finales de los años 1980s, o las más recientes de los CDC/NCI e Instituto de Medicina 2011), con elaboración multidisciplinar de Síntesis de Evidencia (MBE y SPBE) y desarrollo de Guías en medicina genómica y en Salud Pública (por ejemplo, cribado poblacional...).
MODERADOR	Gracias. Paco.
ESTUPIÑÁN	A ver, desde el punto de vista del potencial impacto, yo creo que el potencial impacto es muy grande. Una de las cosas que percibimos y que no nos pasa desapercibida es una corriente social, y es como nuestra sociedad, que es una sociedad más o menos acomodada, donde el basal de salud o de salubridad que tenemos es bastante alto, donde las preocupaciones por la salud, son más preocupaciones en términos de bienestar que en términos de asegurarse de que no te vas a morir mañana o de que vas a morirte con 45 años, lo que vemos es una tendencia social hacia la medicalización o la patologización de estados que son normales, de la evolución normal biológica y de la biografía del individuo, ¿no? Entonces, en ese sentido yo creo que más que la ganancia de información, el fallo o la incapacidad de interpretación de esa información nos lleva a convertir en patológicas cosas que son normales o a tener una atención desmedida a la intervención sobre la salud, en términos de bienestar. Y eso es transversal, o sea, eso... Más allá de eso, pues potencialmente se puede ir al sobrediagnóstico y se puede ir a la sobreintervención. Pero como decía, como acaba de decir José Ramón, pues ya hay quien trata de ponerle barreras y quien trata de, con sentido común, intentar que eso no pase en los términos de asistencia sanitaria o en los planes de salud. Pero más allá de eso, esa corriente social vamos, a mí no me pasa desapercibida. Y tiene derivadas bastante importantes, o sea, en términos de lo que estamos viviendo, por ejemplo, una derivada importante son las corrientes negacionistas como reacción a la exposición a una información que no son capaces de comprender. Como una reacción visceral, comportamental a una percepción de un intento de control desmedido. Ese tipo de cosas tenemos que asumirlas también como reacciones sociales a estos avances.

(continúa en la pág. siguiente)

MODERADOR	Gracias, Paco, muy interesante. Cristóbal.
CRISTÓBAL	La verdad es que yo creo que aquí es muy importante que seamos capaces de diferenciar sobre qué tipo de patologías estamos hablando de sobrediagnóstico. No es lo mismo sobre-diagnosticar una patología banal, en lo que decía Paco hace unos minutos, en las cuales pues realmente a lo mejor lo que estamos haciendo es anticipar el diagnóstico de algo que no es que tenga tratamiento, sino que además su tratamiento es tremendamente sencillo o incluso es que ni siquiera es necesario realizar un tratamiento, que diagnosticar tumores en situaciones precoces de tumores o la patología coronaria. Permitidme utilizar el ejemplo de dos tumores muy concretos, como pueden ser el cáncer de páncreas o el sobrediagnóstico del cáncer de próstata. En ambos casos el concepto de sobrediagnóstico en estadios precoces y su utilidad, o falta de la misma, está vinculada a la falta de eficacia de los tratamientos actuales. Este hecho, en el momento en el que esté disponible un tratamiento eficaz para estas fases, tendremos un cambio de perspectiva que nos hará cuestionar las limitaciones de ese mismo sobrediagnóstico. Así, permítanme alcanzar dos conclusiones en esta pregunta: en primer lugar, el sobrediagnóstico puede ser bueno, regular o malo, y esto va a depender mucho, mucho de qué tipo de enfermedad estamos hablando. Y segundo: ese sobrediagnóstico, al final, se tiene que enfrentar a algo elemental, y es si la actuación terapéutica, la de diagnóstico precoz, que hay muchas veces que al sobrediagnóstico lo llamamos diagnóstico precoz, y la rehabilitación posterior puede ser capaz de resolver sobre cadenas de morbilidad que terminan recayendo sobre el propio sistema.
MODERADOR	Gracias, fenomenal. Voro.
PEIRÓ	Coincido bastante con lo que ha dicho José Ramón. El Sistema Nacional de Salud debe tener mecanismos, igual que los tiene ya en farmacia y en alguna otra cosa, para decir qué prestaciones, intervenciones preventivas o de salud pública, tienen sentido o no desde el punto de vista de efectividad para las poblaciones y de coste efectividad. Entonces, hay cribados que no deberemos hacer; hay pruebas que no deberemos usar y hay pruebas que sí. Lo venimos haciendo toda la vida, no necesariamente con la genómica, pero a las cosas hay que darles el valor que tienen. En los países tipo Estados Unidos, y estamos hablando ahora de los paneles de genómica para predicción de resultados, tienen un uso enorme, y probablemente deben marear mucho y crear muchos problemas.

(continúa en la pág. siguiente)

[02:20:21]	
PEIRÓ	Y aquí es donde toca tener esos paneles de consenso, o esos instrumentos, o una agencia tipo Nice (HispaNice), o lo que sea, para decidir sobre estas cosas y tomar las decisiones que tocan. A pesar de eso, nos encontraremos con problemas de sobrediagnóstico no tanto entendido como diagnóstico precoz o estas cosas, sino como un exceso de falsos positivos en muchos terrenos que no generen ningún beneficio a los pacientes y sí cargas para ellos, molestias para ellos y cargas para el sistema sanitario, ¿no? Pero ahí es donde nos toca encajar las piezas. El hecho de disponer de tecnologías o técnicas no implica que las usemos mal. Claro que las podemos usar mal, y esto lo podemos ver todos los días con muchas de las tecnologías técnicas y tratamientos que tenemos disponibles, pero hay que hacerlo lo mejor que podamos. Y eso implica que las aprobaciones que debe dar el Sistema Nacional de Salud tienen que ser aprobaciones no de la tecnología preventiva, no del cribado en sí, sino sobre qué poblaciones tenemos que usarlas y en qué condiciones tenemos que usarlas. No nos vale decir: se aprueba el diagnóstico preventivo genético de algo, sino que tienes que centrar muy bien en qué grupos de padres, de niños, o los que sean y con qué características se va a hacer ese tipo de pruebas o no. Y si en algún caso es toda la población, pues es toda la población. Pero se trata de hacer los encajes lo mejor posible, y ahí es donde las profesionales que entran en el tema, los investigadores que entran en el tema tienen que apoyar una toma de decisiones, ahora sí, basada en la evidencia que hablábamos antes, que permita que el sistema funcione lo mejor posible. Claro que podemos hacer sobrediagnóstico. Lo hacemos ya con muchas cosas. Y sobretratamiento, y lo hacemos también con muchísimas cosas (antibióticos!). Y no hace falta que esté Juan Gervas para recordárnoslo. Pero sí. Lo podemos hacer bien o mal, mejor o peor. Con esto y con cualquier otra cosa.
MODERADOR	Gracias. Guillermo.

(continúa en la pág. siguiente)

SANZ

Bueno, vamos a ver. Lo primero, en el tema que acaba de comentar Voro y algún otro de vosotros, sí que es cierto que hace falta un consenso para, de alguna manera, regular cuáles son las indicaciones en las que se debe hacer este tipo de cribados. Esto es esencial, quiero decir, tenemos un problema, que es que la gente cada vez tiene más acceso a Internet, a las nuevas tecnologías, y muchas veces la información que reciben no es la información correcta. Pero sí que pueden acceder a este tipo de estudios, por ejemplo, genómicos. No les va a valer, a lo mejor, a ellos para nada, pero que sí que pueden suponer una sobrecarga después para el Sistema Nacional de Salud. Esto, por un lado. Por otro, un peligro que yo le veo a todo esto es que siempre vamos a tener que asegurar que estamos cumpliendo el famoso GDPR, el *General Data Protection Regulation* europeo, así como las trasposiciones a las distintas regulaciones de los distintos países europeos, sobre el derecho del paciente a la protección de sus datos. Sí que es cierto, y esto yo creo que expertos en ética lo tienen muy claro, que los datos de los pacientes que están disponibles en las bases de datos de los hospitales son datos que se pueden utilizar incluso sin un consentimiento informado expreso. ¿Por qué? Porque una vez el paciente ha entrado en el centro, en el hospital o en el centro de salud, de alguna manera está dando su permiso para que se usen esos datos. Eso sí, siempre que tú cumplas los criterios de anonimización, es decir, que estos datos cuando se analicen, cuando se usen, no permitan que se pueda identificar a los pacientes que han dado esos datos, y también que el acceso a esas bases de datos gigantes sea restringido y, además, con unos niveles de seguridad lo suficientemente fuertes como para que no se puedan hackear.

Todo esto es fundamental, porque nos podemos encontrar con el problema que se están encontrando ya en algunos países como puede ser Estados Unidos, en los que ese sistema de estudios genómicos para lo que está sirviendo es para que las compañías aseguradoras lo que hagan es aumentar el precio de las pólizas de seguros a unos niveles inalcanzables para una mayoría de la población, porque tienen un determinado riesgo por padecer, por ejemplo, por tener una mutación de BRCA1. Es decir, esto es una cosa con la que hay que tener muchísimo cuidado, y yo creo que es una de las cosas que, de alguna manera, si pretendemos establecer estos sistemas conjuntos de comunicación y de análisis de datos tenemos que garantizar.

MODERADOR

Gracias, Guillermo. Mercedes, cierras tú.

(continúa en la pág. siguiente)

MERCEDES

Prácticamente está todo dicho. Que efectivamente la medicalización es un proceso que ya existe, que es un proceso continuo que se auto alimenta, que crece de forma constante, que se facilita por una situación en que la sociedad desde mi punto de vista va perdiendo su capacidad de control en contra, paradójicamente con lo que estamos proclamando, de que estamos empoderando al paciente, pues yo creo que en ese sentido lo que hay cada vez son unas expectativas más limitadas de que la ciencia encuentra solución para todas las enfermedades y problemas y, entonces, disminuye el nivel de tolerancia, en general, y disminuye pues bueno, ese control un poco sobre tu propia vida y los problemas y de salud que puedas tener en tu vida. Fomentado por la propia sociedad, por los medios de comunicación, por la industria farmacéutica, por la industria tecnológica ahora; incluso por los políticos, por los gestores y por los propios profesionales sanitarios, que son, en el fondo, actores y víctimas de dicho proceso. Entonces claro, sobre todo os habéis referido al tema de los datos, los datos genómicos. Yo creo que ahí hay que tirar de la deontología para empezar. Primero, UNESCO creo que tiene una declaración en la que dice claramente que algo así como que respecto a los datos genómicos “la persona tiene derecho a decidir a no ser informado de los resultados de la investigación”. Lo que pasa que uno, una vez que le han hecho su código genético, pues vas a tener la tentación de preguntar. Pero ¿hasta qué punto es ético crear una incertidumbre y una angustia reactiva a una persona y a su familia si no hay una disponibilidad de dar tratamiento efectivo o de unas medidas preventivas? Lo habéis comentado antes. Yo creo que eso debe ser el núcleo de la discusión fundamental. Y a partir de ahí aprovechar las ventajas que pueda tener, para determinados pacientes y circunstancias, el conocimiento previo de que a los 58 años tienes la posibilidad de desarrollar no sé qué enfermedad sin amargarte la vida desde los 20, ¿no?

CASTRODEZA

Como casi siempre ocurre con los datos en el ámbito sanitario, un exceso de ellos, recogidos sin claros objetivos, provocaran ruido. Por lo tanto, habrá que tener estrategias para hacerlos operativos. Por otra parte, el sobre diagnóstico habrá que contemplarlo para implementar estrategias que ayuden a paliar sus efectos y a minimizar su impacto sobre las personas y los pacientes.

(continúa en la pág. siguiente)

MODERADOR	Estupendo. Muchas gracias. José Ramón, la quinta pregunta. ¿Conoces, o has utilizado alguna vez, sistemas de información geográfica en el contexto de la investigación? En tu caso, epidemiológica, de altísimo nivel; pues en cardiovasculares o en otras crónicas en las cuales la Autónoma es muy fuerte. En caso afirmativo, ¿los consideras útiles para obtener datos de alta precisión sobre la población?
BANEGAS	Pues dos síes. Gracias, Juan. Clásicamente, los indicadores comparativos geográficos de mortalidad infantil y vacunaciones en diversos países están dotados de Fuerza Política y Social (buena propiedad para empezar). Los estudios geográficos para determinar áreas más vulnerables o de alto riesgo, son clásicos. Un ejemplo fue el Estudio de Siete Países de Ancel Keys, el estudio de chocolate y Premios Nobel de Medicina según países (de Franz Messerli), o el primero de John Snow sobre el cólera en Londres (clusters de hogares según compañía suministradora del agua de bebida –hoy con GPS y la prevalencia, la resolución hubiera sido en horas).
[02:30:10]	
BANEGAS	Son claramente útiles para obtener datos de alta precisión (o detallados) en áreas grandes (e.g., el GBD/NCD-RisC internacional), o pequeñas (este último, estudios de áreas pequeñas en investigación de cáncer en España, CNE/López-Abente y Pollán). La Revisión Sistemática que se nos ha ofrecido en este Grupo de Discusión sobre SPP, brinda excelentes ejemplos de la utilidad del GIS (sistema de información geográfica) para guiar la política de Salud Pública. Por ejemplo, Franchini y cols y su validación del algoritmo CARPEDIEM de métricas de predicción de riesgo de insuficiencia cardíaca congestiva en atención primaria de salud; o el de Liu y cols sobre previsión de Covid-19 con 2 días de antelación y en tiempo real. Son de destacar también los trabajos del grupo de Salvador Peiró en España. Es oportuno mencionar que el código postal es más importante que el código genético, el predominio explicativo de la Geografía sobre la Genética. En EE.UU., más de 1/3 de las muertes por 5 enfermedades en los Estados son potencialmente prevenibles, con el objetivo de llevar a los estados con mayor mortalidad al nivel de los que la tienen menor.

(continúa en la pág. siguiente)

MODERADOR	Fenomenal, muchísimas gracias. Salvador, venga.
PEIRÓ	Sí, nosotros hemos hecho muchos estudios de estos. En buena parte con el grupo de Paco, ¿no? Tanto de epidemiología de enfermedades como de utilización de servicios por áreas geográficas pequeñas y similares. También hemos hecho estudios, ya de base individual, de agregación de enfermedades por sitios y cosas de este estilo. En Valencia los SIP, las tarjetas, están todas georreferenciadas y esto permite hacer algunos ejercicios que, normalmente, no dan para publicación. Los hacemos para Conselleria. De todas formas, otra vez aquí tenemos un choque con la idea de precisión. En los estudios de base geográfica, los que llamamos “ecológicos”, lo que estamos haciendo es atribuir a toda la población que reside en un área los valores medios de esa población, y si tuviéramos que intervenir ahí, tendríamos que desagregar por los subgrupos que hay dentro de esa población. Son trabajos que requieren una gran explotación de datos, que funcionan bien y que, en este sentido de que tienen datos de base real, son interesantes. Pero ahí la unidad de análisis no son los individuos, sino las zonas geográficas que se están analizando y a las que estás aplicando valores medios. Es como cuando dices que la India está muy mal en algo: No quiere decir que todos los indios estén muy mal en ese algo. Y esto es un tema importante porque diría que lo que se busca con el término “precisión” es justamente no ir al bloque completo, sino ir a los grupos específicos, ¿no?
MODERADOR	Gracias. Paco. Por alusión.

(continúa en la pág. siguiente)

ESTUPIÑÁN	Bueno, es que esto es el núcleo de nuestra actividad... una gran parte de nuestra actividad es este tipo de análisis, o sea, que el valor para nosotros es obvio. Dentro justo de lo que acaba de decir Voro, que la precisión aquí es muy poco precisa, y lo tenemos que decir. Por ampliar un poco la información de Voro, tenemos mucho que aprender en otras disciplinas, ¿vale? Porque hay profesionales de otras disciplinas haciendo esto... tenemos los compañeros geógrafos y los compañeros sociólogos, y otras disciplinas, que tienen este conocimiento complementario, que tienen en el ADN esta aproximación de estudios geográficos de las poblaciones ligadas a territorio para cualquier otro tipo de distribución de determinantes de salud. En general los consideramos (los determinantes de la salud). Y hay estudios muy interesantes de llevar esta aproximación a unidades más pequeñas, a unidades de desagregación más pequeña. La propia sección censal, que es una cosa que a nosotros nos cuesta bastante trabajar, porque la distribución territorial de las áreas administrativas de sanidad en ocasiones difiere de la distribución territorial de las áreas censales, y tenemos conflictos ahí que son difíciles de manejar en ocasiones. Pero incluso más abajo, llevarlo a la propia organización vertical de las comunidades de vecinos. No solo pensar en la territorialidad, sino si en la finca en la que vive ese vecino tiene ascensor o no y cómo influye eso en el mantenimiento de la autonomía de una persona mayor, que tiene que bajar cinco pisos de escalera o puede coger el ascensor e ir a comprar. Ese tipo de aproximaciones en unidad más pequeña que sigue siendo una unidad geográfica yo creo que son interesantes y de las cuales nos podemos beneficiar. Dicho lo cual, nosotros estamos haciendo estos estudios desde siempre, y el atlas de variaciones es obviamente un producto de esta aproximación, y lleva aquí muchos años. Muchos más años de los que yo estoy involucrado en la cosa, o sea, que tampoco puedo hablar mucho más de eso.
MODERADOR	Fenomenal, has puesto un botón de muestra absolutamente reseñable. Cristóbal.
BELDA	Buenas tardes. Me vais a disculpar, yo no he utilizado jamás un sistema de localización geográfica entre otras cosas porque todos mis modelos animales los he tenido siempre bastante bien localizaditos. En fin...
MODERADOR	Gracias. Guillermo.

(continúa en la pág. siguiente)

SANZ	Bueno, vamos a ver. Es fundamental tener en cuenta la territorialidad, porque el impacto que tiene en tu salud el lugar en el que vives es más que obvio. Quiero decir, independientemente de que los problemas de salud sean distintos en un país o en una región o en otra, el acceso a los sistemas sanitarios es diferente. Y esto, esta inequidad en este acceso, desgraciadamente también se ha dado en España. Se ha dado en España, y alguien decía antes lo del código postal, que es más importante muchas veces que la genómica. Se da en el acceso a fármacos y especialmente, por ejemplo, a los fármacos de alto impacto económico. Tenemos un sistema autonómico, el que nos hemos dotado, y en este momento una persona con una enfermedad oncohematológica puede tener acceso a un fármaco en una comunidad y en otra no. Y, de hecho, estamos viviendo ahora situaciones en las que un enfermo de determinada comunidad va a otra porque sabe que en esa otra le pueden dar un fármaco que en la suya no está financiado. Es decir, que...
PEIRÓ	Guillermo, yo me compré la casa en un distrito donde creo que podré llegar a los 80 conservando los meniscos, conservando la próstata y sin que me hayan cambiado ni la rodilla ni la cadera. Hay sitios donde es imposible llegar a los 70 con el útero puesto, o con las amígdalas o el menisco. El código postal es fundamental. Poblacionalmente (que no individualmente) es un determinante de la utilización de servicios sanitarios mayor que la propia morbilidad.
SANZ	Pues no comento más, eso es lo que quería decir.
MODERADOR	Gracias. Fenomenal. Mercedes, vamos contigo.
ALFARO	En mi caso, la casi única experiencia de geolocalización real es reciente, es con el COVID. No sé si sabéis que se montó un sistema de información que se llama SERLAB, donde todos los laboratorios públicos y privados mandan diariamente, a través de las consejerías, todos los resultados de pruebas diagnósticas COVID, PCR y antígenos, que llegan al Ministerio con un decalaje de 36- 48 horas desde que se han realizado. Tanto las positivas como las negativas tienen que ir identificadas por el código de tarjeta sanitaria, bien autonómico o bien nacional o, en su defecto, por el DNI o el pasaporte. Nosotros lo primero que hacemos cuando recibimos todos esos registros es cruzarlos con la base de datos de tarjeta sanitaria en la que ya está incorporado incluso

(continúa en la pág. siguiente)

	MUFACE privado. Falta solo por incorporarse ISFAS y MUGEJU privado, pero todos los demás ya están incorporados en la base de datos de población protegida financiada públicamente, porque las mutualidades en un 99% están financiadas con impuestos. Y, entonces, nosotros lo que hacemos es un cruce, para el que hubo que cumplir una serie de requisitos “duros” que nos pedía de la Agencia de Protección de Datos, con una única finalidad, que es la de geolocalizar a la persona en su sitio de residencia, y poder indicar dónde se concentran los casos día a día, y cómo van evolucionando los casos a lo largo de los días por código postal.
[02:40:49]	
ALFARO	Esto, además, amplía la información a las Comunidades Autónomas, nosotros estábamos viendo en el mes de marzo que las comunidades estaban montando unos sistemas de georrepresentación importantes, porque en estos casos es importantísimo saber por dónde va extendiéndose la pandemia, pero evidentemente los estaban montando dentro de sus límites. Y por fuera de su frontera no conocían si los municipios limítrofes también tenían una incidencia determinada o si iban creciendo los casos de una manera determinada, ¿no? El SERLAB lo que les proporciona a las Comunidades Autónomas es poder ver el mapa entero de España con todos los distritos postales suyos y del resto de Comunidades Autónomas. Y tiene datos desde el principio de la pandemia, al principio con peor calidad porque había menos laboratorios informantes, pero desde más o menos mediados de abril con bastante exhaustividad en el registro. Y yo creo que está siendo una de las herramientas más útiles. Y es muy curioso los patrones geográficos. En las tres olas anteriores, cuando el cauce del Ebro comenzaba a ponerse negro, en el mapa de calor, bueno, era para echarse a temblar. En esta cuarta ola no se está poniendo tan negro el cauce del Ebro. Entonces bueno, son herramientas absolutamente útiles, por ejemplo, en este caso práctico
CASTRODEZA	En mi caso, hemos trabajado en métodos de triaje, como métodos de interpolación geoestadístico en temas de gripe. Los datos que se obtienen por estos métodos son ponderaciones y/o estimaciones.

(continúa en la pág. siguiente)

MODERADOR	Perfecto, magníficos ejemplos los que nos ponéis, muchas gracias. Pasamos a la pregunta 6, que arranca Voro: ¿Sirve un estudio bien diseñado para probar si son óptimos los resultados de una intervención de salud pública de precisión? ¿Por qué tipo de diseño te inclinarías?
PEIRÓ	A ver, la primera pregunta es un rotundo sí. Elaboración de intervenciones de cualquier tipo de salud pública incluidas, si el estudio está bien hecho pues claro que sirve para probar si la intervención es efectiva o no, al menos en el contexto y en la forma que se ha hecho. Por otro lado, las intervenciones de salud pública suelen ser complejas. La misma intervención en un entorno distinto puede no funcionar ... o funcionar de modo distinto. Los diseños ... yo soy un ecléctico. Cada pregunta de investigación tiene un diseño que es idóneo a esa pregunta (y no siempre es el posible). En evaluación de intervenciones, normalmente nosotros trabajamos con diseños de cohortes y análisis de diferencias en diferencias, variables instrumentales, regresiones discontinuas, ... Otras veces intentamos construir una cohorte que asemeje el <i>target trial</i> (lo más parecida posible al ensayo clínico que nos habría gustado hacer).
MODERADOR	Vale.
PEIRÓ	Los diseños a veces son más sólidos y otras veces son menos sólidos. Depende mucho del volumen y calidad de las fuentes de datos.
MODERADOR	Fenomenal. Guillermo.
SANZ	No, no tengo nada más que añadir. Yo sinceramente, respondería a la primera pregunta que sí, obviamente, pero la segunda, como dice Voro, depende un poco del estudio que quiera realizar el tipo de análisis que tienes que hacer. La gran ventaja, por ejemplo, que tiene todo el tema de <i>Big Data</i> , de inteligencia artificial es que tú dispones de esas herramientas, y esas herramientas además son gratuitas y te pueden permitir esos análisis en un periodo muy reducido de tiempo. Pero vamos, depende un poco de...
MODERADOR	De la pregunta de investigación. Fenomenal. Gracias. Paco.

(continúa en la pág. siguiente)

ESTUPIÑÁN	Yo creo que está todo dicho. O sea, siempre es así, ¿no? Siempre el diseño depende de la pregunta de investigación. Si acaso, como es a tener en cuenta en estos estudios la consideración del tiempo, o sea, que sean estudios longitudinales o de seguimiento de las poblaciones. Siempre que sea posible el tener o tratar de formular la posibilidad de tener un grupo de control, algún tipo de control siempre que sea posible, porque al final lo que vas a ver son diferencias entre una población que recibe una intervención y... el tener en cuenta los riesgos competitivos, y en el establecimiento de estos controles, que no siempre pueden ser, o sea, estamos hablando de que no son ensayos clínicos, de que son como muchos ensayos pragmáticos. En el establecimiento de estos controles, ser muy exigente en términos del control de los sesgos; de los sesgos de información, de los sesgos de medida y la conclusión. Pero, en general, en el trabajo que tenemos.
MODERADOR	Fenomenal. Mercedes.
PEIRÓ	En lo de tener en cuenta o no los riesgos competitivos hay muchísima discusión. Nosotros normalmente los valoramos, pero los incorporamos como análisis secundarios, de sensibilidad, para que no nos digan que no lo hemos tenido en cuenta, pero la mayoría de las veces no les hacemos demasiado caso.
ESTUPIÑÁN	Exacto, en lo que se traduce al final son en extensos, a veces extensísimos análisis de sensibilidad. Sí, sí. Pero está ahí, en el abordaje, y está siempre en la parte de atrás.
ALFARO	Qué sinceridad, me encanta. Yo nada que añadir. Al contrario. Y el sí rotundo que ha dicho Voro. Es que yo creo que una intervención, si no se prueba con un estudio bien diseñado realmente pues no deberíamos adoptarla.
CASTRODEZA	Cualquier estudio bien diseñado dará respuesta apropiada en base a sus fortalezas y debilidades.

(continúa en la pág. siguiente)

BANEGAS	Sí creo que sirve. Los tipos de diseños posibles son, en principio, todos, pero fundamentalmente los estudios prospectivos (sugere- ntes y con potencial utilidad clínica en salud pública y SPP –grupos más vulnerables dentro de un país o región...) y los ensayos clínicos (corroboradores). Los estudios epidemiológicos proporcionan prevalencia de genotipos (trasversales), riesgos asociados a variantes genéticas (longitudinales), e interacciones gen-ambiente y gen-gen. Los estudios de evaluación clínica permiten cuantificar el valor añadido de la genómica en la reclasificación de sujetos compara- do con factores de riesgo tradicionales (e.g., Net Reclassification Index). Los estudios de ciencias sociales y comportamentales, evalúan la influencia de perfiles genómicos sobre conductas familiares y poblacionales. Los estudios de ciencias de la comuni- cación y educación utilizan información genómica para mejorar la salud. La investigación de servicios de salud y vigilancia de salud pública, valora el impacto del genoma sobre resultados de salud, desigualdades y aspectos económicos. En todo caso, hay que modernizar (y adaptar) los métodos de vigilancia, informática y epidemiología (e.g., uso de randomiza- ción mendeliana con variables genéticas instrumentales, machi- ne-learning, propensities scores, etc).
MODERADOR	Vale. Arranca, por ejemplo, Guillermo. ¿Qué barreras consideras que son las más críticas para el desarrollo de la salud pública de precisión? ¿Y cuáles son los riesgos que consideras más relevantes?
SANZ	Aquí lo primero es pensar que nos ceñimos a España. El primer riesgo que veo es el tema de las 17 comunidades autónomas. Este es el primero. El segundo que se puedan poner de acuerdo en un modelo común, que se puedan poner de acuerdo en el tema de acceso y explotación de datos. Por otro lado, el tema de la financiación de los recursos. Yo creo que la gran ventaja que tenemos ahora, y yo creo que en esta sesión ha quedado de manifiesto, es que la pandemia por COVID ha demostrado para qué sirve esto. Antes ni los políticos ni la opinión pública tenían claro cuál era el papel que podría jugar en salud el uso de estas grandes bases de datos que permitieran desarrollar esa salud pública de precisión de la que estamos hablando. Entonces, yo diría que los dos más importantes son: tratar de ponerse de acuerdo y, por otro lado, el tema de la financiación.

(continúa en la pág. siguiente)

MODERADOR	Gracias. Fenomenal. Mercedes.
ALFARO	Riesgos, bueno, pues primero creo que es necesario que haya unas buenas hipótesis antes de trabajar en salud pública de precisión. A mí me da la impresión de que a veces, y no solo en salud pública, sino cuando se trabaja con grandes volúmenes de datos, parece que la gente va un poco como de pesca, es decir, hay tantos datos que vamos a ver lo que sale. Y nos van a salir, seguramente, unos resultados que no esperábamos y que los vamos a encontrar y no saber cómo interpretar. A mí eso me parece un riesgo verdaderamente importante, que lleguemos a modelos casuales mucho más que causales.
[02:50:14]	
ALFARO	En salud pública se tiene en general bastante cuidado en utilizar fuentes que están diseñadas para la investigación ya específicamente o bien fuentes como las historias clínicas, primarias para otros fines cuyos datos tienen un uso secundario en investigación. Pruebas de laboratorio, censos de población, registros de enfermedades, a las que el <i>big data</i> lo que hace es incorporar otras fuentes de información como son las <i>apps</i> , los <i>wearables</i> , las redes sociales... todas estas cosas, que da la sensación de que los datos que hay en estas plataformas a veces se parecen un poco a las “muestras por conveniencia”, y, por lo tanto, que pueden tener un número importante de sesgos de selección y de información. Y, contrariamente, con grandes cantidades de datos puede que aparezcan asociaciones que se deben más al azar que a ninguna otra causa. Son los principales riesgos que le veo al <i>big data</i> en general. Entonces vuelvo a lo de antes. El estudio tiene que estar bien sustentado metodológicamente; sea cual sea el origen de los datos tiene que plantearse con un método conocido, tiene que cumplir unos criterios técnicos adecuados, pero los riesgos fundamentalmente que le veo son estos. Y luego, por supuesto, que no cree tampoco más brechas, es decir, que los países más ricos no se lancen a hacer unos avances importantísimos en análisis de datos y en modelos predictivos y en anticipación de problemas, y que los de renta más baja se queden atrás. Es otra cosa que preocupa también de toda la transformación digital, es otro riesgo que subyace. No es solo del <i>big data</i> y del <i>big data</i> en salud pública.
MODERADOR	Gracias. Paco.

(continúa en la pág. siguiente)

ESTUPIÑÁN

A ver, yo creo que para mí una de las barreras más importantes ya está mencionada... lo mencionaba Voro al principio. En cuanto al comentario que hizo sobre... que se ha hecho mucha inversión en la digitalización de los sistemas de información o digitalización de la atención sanitaria en términos de hacer lo que se hacía antes en papel, transformarlo en el plano digital. Pero tenemos que crecer en la capacidad de explotación. Yo creo que una de las barreras es, y lo hemos comentado, la capacitación; el tener personal cualificado y capacitado para poder hacer ese análisis. Y no en el sentido que hablaba antes en el comentario que hacía Cristóbal, que es completamente pertinente. O sea, al final no creo que la idea sea pedirle al profesional sanitario, que es el usuario final, que conozca en detalle todo el proceso, el procedimiento, pero sí que tiene que tener una cierta idea... o sea, igual que no se le pide a un profesional sanitario que, en general, que conozca de forma detallada la generación de una tomografía axial computarizada, o una resonancia magnética, por más que lo damos en física médica, pero no es un conocimiento que requiramos ni nos vamos a poner a arreglar el aparato, pero sí que tenemos que tener una cierta idea de cómo funciona.

Dicho lo cual, esto lleva a una barrera que es la de siempre. Es un tema de recursos, pero también no solo de recursos, sino de que, si pudiéramos ahora mismo a todo el sistema educativo, de formación, a producir este tipo de profesionales, lo mismo no cumplíamos los requerimientos que teníamos en los próximos cinco años. O sea, eso tenemos que ser realistas en ese sentido. O sea, nuestra capacidad de progreso tecnológico por una de las cosas que va a venir limitada va a ser por la eficacia de ese perfil. Que son perfiles técnicos que son extraños, que son perfiles híbridos. Luego está la inversión en recursos, en recursos materiales, en recursos tecnológicos y en eso. Pero... y esto lleva a la transformación cultural. Al final no sabes qué es primero, si la transformación cultural o la venta de la utilidad que permite incorporar los perfiles, que te permite hacer la transformación cultural. Es como... es complicado, a veces una cosa se produce por la otra o la otra por la una, pero bueno. Esa es la principal barrera que veo.

En cuanto a riesgos, yo sí que tengo en el fondo de mi cabeza siempre el riesgo inherente a cualquier tipo de decisión de salud pública, cualquier tipo de intervención en salud pública, que son las cosas que a veces ni siquiera puedes medir o que no piensas en el diseño, que son las externalidades. Externalidades hay de todo tipo, positivas o negativas, y hay algunas externalidades que no son previsibles de ninguna manera. Entonces, siempre como

(continúa en la pág. siguiente)

	expresión vulgar, con la mosca detrás de la oreja, de que toda intervención que hagamos de salud pública tiene que ir de alguna manera acompañado de un plan de evaluación, de un plan de medición. Para tener en consideración estas potenciales, externalidades negativas.
MODERADOR	Genial, muchísimas gracias. Cristóbal.
BELDA	Me inclinaría por modelos de ensayos clínicos pragmáticos. En estos momentos estamos cambiando todos nuestros modelos de investigación con lo que debemos ser cautos pues corremos el riesgo de reducir el rigor con el que se termina explorando la fiabilidad científica de cualquier tipo de intervención. Por tanto, mi aproximación sería una aproximación evidentemente pragmática, pero intentando también mantener un nivel de rigor que sea proporcional al nivel de rigor que le hemos exigido históricamente a todas las actuaciones clínicas que nos han ido acompañando en los últimos años.
MODERADOR	Pues ahora, la siguiente, Cristóbal. Las barreras que son más críticas para el desarrollo de la SPP y los riesgos que consideraban...
BELDA	La barrera más importante es estructural. Y esa barrera estructural es algo que como hemos venido diciendo a lo largo de las últimas dos horas, está muy vinculado a la capacidad de generar cohesión alrededor de las actuaciones del Ministerio de Sanidad. Una vez que se resuelva esa barrera, la siguiente barrera es muy cultural, y es muy cultural en la medida en que probablemente, en la mayoría de las situaciones, no nos hemos todavía llegado a plantear que vamos a tener cierto tipo de herramientas que nos ayuden a la toma de decisiones. Por lo tanto, de forma muy, muy resumida, una de las principales barreras va a ser estructural en la medida en que tenemos que desplegar sistemas que nos permitan hacer uso inteligente de los datos, y en segundo lugar sería la barrera cultural. Es decir, que seamos capaces de incorporar este tipo de tecnologías como herramientas convencionales en la toma de decisiones, tanto de políticas de salud pública como en el propio ejercicio operativo de la salud pública.
[03:00:45]	
MODERADOR	Gracias. Voro, venga, cierras tú

(continúa en la pág. siguiente)

PEIRÓ

En general estoy de acuerdo en todo lo que se ha dicho. La limitación básica es que tenemos que mejorar nuestras capacidades, hablo del Sistema Nacional de Salud y salud pública, de explotación de los datos. Tenemos que mejorar mucho con esto. Es una mezcla de capacidad investigadora, de capacidad del sistema para explotar estos datos en su práctica diaria, en su gestión.

Hay problemas debajo. España arrastra un problema histórico, y es que la propia política de salud pública y de no salud pública funciona, ha funcionado un poco siempre por “ocurrencias” (que no ha querido nunca que se evaluaran). Y me estoy acordando mientras hablo de esto de aquella ministra (no recuerdo el nombre) que debió pasar por Reino Unido o algún sitio de estos, vio que cortaban los medicamentos por unidades y decidió que iba a poner esto en España. Y se estaba cargando, probablemente, el mejor sistema de trazado de medicamentos que hay en el mundo, porque en España podemos saber hasta qué lote se lleva cada paciente. Y además pensaba que tres paracetamoles son más baratos que veinte, cuando en el precio de un medicamento la cantidad es lo último que importa. El precio del medicamento no depende del corte. De hecho, es más caro dispensar los paracetamoles de uno en uno que de veinte en veinte.

Hay una tendencia a no querer evaluar las políticas o las intervenciones que se ponen en marcha. Probablemente porque la confrontación política en España es muy dura y juega también en este frente. Y nadie quiere ver si algo le ha salido mal cuando lo ha puesto en marcha. Y tampoco nadie quiere que se evalúen las cosas *ex ante*: si alguien dice “voy a ver cómo funciona esto, porque igual no funciona bien” probablemente es la precondition para que eso no se ponga ya en marcha. Debemos buscar una mejor relación con el mundo de la política para hacer políticas. Y estoy diferenciando todo el rato entre *politics* y *policy*, que son dos palabras que no dicen lo mismo.

El tema de las 17 comunidades autónomas. Hay países muy pequeños, y estoy pensando en los nórdicos, más pequeños que muchas de nuestras CCAA, que tienen un enorme potencial de explotación de datos. De los mejores del mundo en el caso de los nórdicos. El tamaño importa, pero no es lo más importante.

MODERADOR

Finlandia.

(continúa en la pág. siguiente)

PEIRÓ

Sí, y Dinamarca, y Noruega, y la propia Suecia. Su capacidad de hacer estudios con bases de datos es brutal, y trabajan muy bien. No estamos hablando de tamaño, estamos hablando de generar esa capacidad. Y a mí en muchos casos no me molestaría que alguna comunidad tirara adelante y fuera marcando y abriendo caminos.

Prefiero una política más, digamos, liberal, ... en España hay experiencias muy importantes, con este tipo de datos. Una es lo que está desarrollando el grupo de Ximo Dopazo en Andalucía, incorporando la genómica a la historia clínica y la explotación de la genómica. Esos datos son uno de los capitales más importantes que tenemos y se está derrochando en muchísimas comunidades. O se lo están quedando las empresas a las que les encargamos que nos hagan estos análisis, o se quedan en el ordenador de un jefe de servicio, de un servicio y luego se borra. Que lo haga una comunidad mientras otras no lo están haciendo no es motivo para frenarlos a ellos sino para aprender de ellos e incorporar sus avances.

La otra experiencia importante en España también es la de Eurobioimaging en la que participa, entre otras, la Comunidad Valenciana con el Biobanco de imágenes que reúne prácticamente todos los datos de imagen en un gran repositorio centralizado con el que está trabajando en este momento toda Europa.

Nosotros muy poco, pero el resto de Europa bastante. Son cosas donde no vale el “tenemos que hacerlo todos y mientras no lo hagamos todos, no lo hace ninguno”. Al contrario, son para aprovechar muy bien las experiencias que vamos teniendo. Y me parece importante porque además me parecen dos ejemplos fantásticos que nadan entre la investigación y la clínica directa. Que mejoran ambas cosas a la vez.

Y este es otro punto importante. En muchos terrenos, y este es uno, la investigación y la gestión diaria y la clínica directa y la salud pública se van a diferenciar cada vez menos. No vamos a saber muy bien dónde acaba la gestión del día a día y dónde empieza la investigación. Y esto es importante tenerlo en cuenta porque también tenemos que combinar estas perspectivas. Los centros de investigación tenemos que dar mucho más soporte a la gestión del día a día y a las decisiones del día a día. Y esto nos lleva a otro ritmo, porque los tiempos de la investigación son unos tiempos y los tiempos de la gestión no tienen nada que ver con ... cuando te piden algo en realidad querrían haber tomado la decisión ayer, y tú dices “dadme un año” ... Y hay que empezar a encajar las piezas para que podamos dar respuestas y estudios que no van a publicación pero que son capaces de dar soporte al día a día de lo que necesitan nuestros sistemas de salud.

(continúa en la pág. siguiente)

BANEGAS	Algunas de las principales barreras, bajo el esquema de Khoury de realizar las promesas de la Genómica y Big Data en SP [Desarrollo de una Agenda de Investigación Translacional Robusta], son: – Falta de datos sobre magnitudes de riesgos, especialmente para efectos conjuntos de genes y ambiente. – Actualmente, débil poder discriminatorio de la genómica comparado con otros factores de riesgo. No está claro el balance de beneficios/daños de usar genómica en prevención y cuidados clínicos. – El cambio conductual es difícil, y no es claro si la información genética es importante. – Falta de preparación para manejar información comunicativa y educativa. – La tecnología genómica es cara cuando se aplica a poblaciones, y con beneficios y daños desconocidos. Algunos riesgos: actuar no tanto basados en la (escasa) evidencia sino en la pasión y fascinación ideológica y humanitaria de la SP (y SPP). Sería una SP de “imprecisión”. Posible falta de uso de la Medicina Compasiva (no solo fármacos) y, ante todo, de Ética (incluyendo confidencialidad y consentimiento informado). Otro riesgo sería el de tecno-optimismo acrítico (Horton, 2018).
MODERADOR	Fenomenal. Muchas gracias. Cristóbal. La pregunta ocho. El sector privado, que es el que tiene mayor acceso a nuevas tecnologías, ¿quién piensas que liderará los estudios de salud pública de precisión? ¿Y cuál será el papel del sector público y el privado en el uso de las nuevas metodologías de investigación que ya se están probando?
BELDA	El sector sanitario privado en estos momentos está terminando un proceso de concentración que le ha hecho adquirir un tamaño en el que algunas de las empresas del sector son más grandes que los sistemas sanitarios de muchas Comunidades Autónomas. Es decir, estamos hablando de un sector sanitario privado, en el que tenemos compañías como Quirón, como Vithas, como HM, que tienen tamaños mayores que muchas Comunidades Autónomas, y que ven y que atienden pacientes haciendo un enorme esfuerzo de integración de sus estructuras de información clínica.

(continúa en la pág. siguiente)

	Eso, ¿dónde nos lleva? A que estas empresas probablemente están en unas condiciones de poder competir directamente en la exploración de la eficacia de ciertos tipos de perfiles clínicos en mejores condiciones que en las que está el sector público. Claro, son perfiles poblacionales que no tienen por qué corresponderse con la totalidad de la población, aunque hay que tener en cuenta que en Madrid y Barcelona el 50% de sus ciudadanos tienen un seguro de asistencia privado, y el 30% de ellos no consume sanidad pública nunca. Por tanto, yo creo que aquí, más allá de ver quién puede ser el que realice un tipo de estudios u otro tipo de estudios, creo que es muy importante el ejercicio de interoperabilidad. Porque si los sistemas de análisis masivos de datos no incluyen algunos de los sectores poblacionales que en estos momentos están dentro de otro sistema sanitario que es completamente distinto al sistema sanitario público, probablemente nos vamos a perder una buena parte de la información.
[03:10:49]	
BELDA	Teniendo en cuenta que es posible que algunas compañías puedan implantar sistemas de precisión clínica antes incluso que el propio sistema sanitario, con tamaños poblacionales gigantescos. Y en esto hay que tener en cuenta también otra cosa, y es que 1.300.000 funcionarios están dentro de ese sistema sanitario. En cualquier caso, el sector privado está siguiendo un camino para la unificación de toda su información clínica a través de un modelo de inversión que es muy eficiente.
MODERADOR	Gracias, Cristóbal. Guillermo, te quedaste con ganas antes de hablar del <i>partenariado</i> público-privado, y te quedaste con ganas de desarrollar un poco el papel que ves, porque Cristóbal ha mencionado a los prestadores de servicios que, efectivamente, el sector privado está creciendo de una manera muy marcada en determinados lugares de España, pero hay otro sector que es el farmacéutico, la biomedicina, que también juega un rol muy importante. Adelante, Guillermo.
SANZ	Bueno. En esto, varias cosas. Bueno, un comentario sobre lo que ha dicho Cristóbal. Yo, de todas maneras, veo muy difícil que Quiron, Vithas y los demás compartan entre ellos los datos, lo veo difícil porque están compitiendo por el mismo sector de población. Volviendo al tema de la industria farmacéutica, lo primero que me gustaría decir es que lo que tenemos que hacer es

(continúa en la pág. siguiente)

romper el molde de que la investigación se realiza únicamente en las instituciones académicas. No nos debemos engañar. Es decir, la industria farmacéutica tiene una capacidad de investigación tremenda. Yo hace poco estaba en Cambridge, Massachusetts y justo enfrente del MIT, uno de los centros más importante en ciencia médica e investigación, hay un edificio de Novartis de doce plantas, el más importante de investigación de esta compañía en el mundo. Uno de sus objetivos es captar a los investigadores de más calidad, digamos, de la parte pública. La industria farmacéutica desde siempre ha sido considerada como el lado oscuro, es decir, hablando de La Guerra de las Galaxias, serían los Sith y Darth Vader. Y esto, aunque puede ser verdad y, el ejemplo del tema de Astrazeneca y el número de dosis de vacuna COVID que han dejado de enviar a la Unión Europea es uno de ellos, de alguna manera nos pone en una situación de ventaja. En los proyectos internacionales en que participo y con un enorme contacto con representantes del más alto nivel de las principales industrias farmacéuticas, ellos tienen asumido el papel de malos de la película. Y eso es bueno para las instituciones públicas.

Es bueno porque nos permite partir de una situación de ventaja. ¿Quién debe liderar esto? Lo debemos liderar desde el sector público, es decir, hay que hacer un esfuerzo, y lo hemos comentado antes, en financiación, en todas estas cosas. Pero a lo que se tiene que tender, de alguna manera, es a un modelo común, lo que llaman los ingleses un *Public Private Partnership* (PPP), una cooperación público-privada que nos permita, de alguna manera, poder hacer lo que en este momento no estamos pudiendo hacer, que es explotar los datos. Es decir, si ahora yo necesito tres ingenieros de datos en mi hospital y en este momento esa figura no está contemplada por el Sistema Nacional de Salud y no puedo contratarles y, sin embargo, puedo establecer algún tipo de convenio con una industria farmacéutica que me financie esos tres ingenieros, ello me va a permitir analizar los datos de mi centro. Evidentemente esto sería una gran ventaja. Esto, como he comentado antes, es absolutamente “vendible” a la industria farmacéutica. Además, lo que la industria tiene que hacer, y es parte de su papel, es aportar sus datos. Es decir, no es simplemente decir “bueno, pues vamos a poder ofrecer los datos de un conjunto inmenso de población”, sino también incorporar a este sistema los datos de los ensayos clínicos de las compañías farmacéuticas. Esto lo estamos consiguiendo en HARMONY, y en HARMONY la industria farmacéutica está introduciendo en nuestra base de datos los datos de ensayos clínicos.

(continúa en la pág. siguiente)

	¿Por qué? Pues porque esos ensayos clínicos tienen datos de enorme calidad. Cuando, por ejemplo, hablamos, y a todos se nos llena la boca, del tema de empoderar al paciente, de que sea el paciente el que tiene que tomar decisiones; el que tenemos que disponer de los famosos <i>Patient Reported Outcomes</i>, nos encontramos que esta información prácticamente no existe en nuestras historias clínicas. Sí que es verdad que en los últimos años algunas instituciones estamos tratando de dar ese salto y empezar a incluir este tipo de información en nuestras bases de datos internas. Pero esto en los ensayos clínicos de la industria farmacéutica está prácticamente en todos los de los últimos años, y esto es porque las agencias reguladoras, a una de las cosas a las que se está dando más valor para la aprobación de nuevos fármacos, entre otras cosas porque los pacientes están incorporándose a los comités de valoración de los nuevos medicamentos, es al tema de calidad de vida y a qué es lo que piensa el paciente que le puede beneficiar de ese tratamiento. Con lo cual, yo creo que aquí tenemos que darle una vuelta y tratar de ver cuál sería el mejor modelo, como dicen los ingleses, el <i>business model</i>, que se podría utilizar para que las dos partes obtengamos un beneficio. Que lo necesitamos ambos.
MODERADOR	Gracias, Guillermo. Voro, ¿cómo conseguimos esas sinergias que ha empezado a mencionar Guillermo?
PEIRÓ	¿Cómo? Complicado. O quizás no tan complicado. En este campo tenemos espacio para el sector privado sanitario, que estaba comentando Cristóbal. Lo hay. Hay espacio para la industria farmacéutica (y no solo como financiadora de estudios), como parte integrante del desarrollo de los estudios. Y hay un tercer espacio para la industria tecnológica e innovadora, tanto las grandes tecnológicas como estas PYME innovadoras que tenemos por muchas partes, muy revolucionarias, que quieren ser grandes. Y todo el mundo tiene que formar aquí alianzas. Gracias fundamentalmente a los incentivos puestos en marcha por el Carlos III la investigación ha desarrollado un montón de alianzas en estos años que se han traducido en nombres como los CIBER, como las plataformas, como las RETICs o, simplemente, como los propios Institutos de investigación que integran universidades y otros centros y son un poco la base organizativa (el ladrillo) de estas alianzas. Y eso que los Institutos se pierden aquí un poco en la medida que tienen dimensión de centro, digamos de hospital y su primaria, cuando las grandes bases de datos tienen economías de escala enormes que no debe limitarse a la población de un departamento.

(continúa en la pág. siguiente)

[03:20:27]	
PEIRÓ	Un instituto, analizando las 130.000 o 150.000 personas que tiene el departamento ... cuando le cuesta el mismo esfuerzo que analizar a cinco o a treinta millones de personas. En este terreno creo que tenemos que encajar mucho, estimular que estas piezas se junten y se junten para hacer una investigación mayor. Y que también tenemos que estimular que el propio sistema, los de salud pública y otras, estiren de las empresas innovadoras para encajar las piezas que necesitan para desarrollar todo esto. En cierta forma sorprende ver como a veces queremos que nos solventen problemas de bases de datos que teníamos hace diez años, cuando deberíamos estar pensando en qué queremos tener en 2025 y en 2030. Y, de hecho, las empresas ofrecen cosas que no tienen nada que ver con lo que les están pidiendo, que superan mucho lo que les están pidiendo. Creo que hay que encajar todas estas piezas. Para nosotros el núcleo central es la capacidad investigadora que tenemos en este momento, que, la verdad, ha mejorado mucho en los últimos diez, quince años gracias a esta política del Carlos III de estimular la colaboración y las redes de muchas formas, y de establecer como centros de excelencia en base a los institutos y alguna otra cosa. Pero creo que hay que tirar todavía más de allí y establecer que los grupos que encajan en esas piezas formen alianzas, porque la investigación con datos, las economías de escala son enormes. Si hay algún campo donde investigar el trocito de cada uno no tiene ningún sentido es este.
MODERADOR	Paco. Si quieres, estira un poco el llamado que hiciste a la necesidad de talento, de reclutar talento y a las dificultades que hay, y si realmente este <i>partenariado</i> público privado puede servir para conseguir más objetivos que si va uno solo.
ESTUPIÑÁN	Vale, a ver. La idea del <i>partenariado</i> público privado es interesante. El cómo establecerlo, el cómo se materializa. Siempre, el diablo está en los detalles. El cómo establecerlo y en quién guía esa relación y los términos en los que se da esa relación al final es lo relevante. Yo creo que lo que no deberíamos tratar de hacer es parchear estas barreras o necesidades estructurales de las que estábamos hablando antes del sistema público con esa intrusión. Sobre todo, porque necesitas esa capacitación para interpretar que lo que estás recibiendo en esa relación es adecuado. O sea, es apropiado, es acorde a tus intereses y es en la

(continúa en la pág. siguiente)

medida de tus expectativas, ¿no? Necesitas tener ese conocimiento de las cosas. Ese conocimiento tiene que ser previo, en mi opinión. Yo sí que veo el cómo se produce esa interrelación y cómo se pueden dar esta situación de beneficio mutuo, lo veo por vías... Voro comentaba el fomento de la capacidad en investigación a través de estos programas o el refuerzo de las instituciones de investigación. A mí me resulta muy interesante, muy interesante la señal que se está dando desde las instituciones españolas pero también desde las instituciones europeas, de la necesidad de apertura y de reproducibilidad de las cosas que hacemos en términos científicos y de evaluación, o sea, todo el movimiento de ciencia abierta tanto en los términos de acceso a datos disponibles públicamente como en los términos de acceso a los últimos desarrollos de software, de programas disponibles públicamente, o acceso a plataformas tecnológicas, que nunca lo hubiéramos pensado, pero plataformas físicas, tecnológicas, como el movimiento *maker* que está detrás de poder hacer objetos físicos, no solo de *software* o de entelequias o de programas, ¿no? Y que no están sujetos a una licencia, o están sujetos a una licencia abierta que te permite modificarlos, utilizarlos para otras cosas que no fueron definidas, en principio. Yo creo que por ahí está el camino, ¿no? El sistema, además, se está dotando a nivel de sistema sanitario, se está dotando de instrumentos como los instrumentos de compra pública de innovación para tratar de decir qué es lo que quieren en términos de “yo tengo este problema, ofrecedme las soluciones tecnológicas para solucionarlo”. Y yo lo que quiero adquirir es esa solución tecnológica. Y esa solución no tiene por qué estar en el mercado, no tiene ni siquiera por qué existir. Hay un cierto margen de riesgo ahí, pero estás señalando hacia dónde quieres ir con esa cosa.

Y de alguna manera, un poco más ágil que lo que conocemos de los concursos públicos de adquisición de programas, o de lo que sea, que existían. Yo creo que, en ese sentido, en el sentido de la ciencia abierta, en el sentido de otros mecanismos de compra, y en el sentido del fomento de las instituciones y de la interlocución de estas instituciones públicas, de investigación, con los privados, ¿no?, es por donde van las cosas. Destacar también el tema de interoperabilidad. O sea, es necesario al final, yo creo que porque... yo creo que llegaremos al camino de la interoperabilidad por medio del camino de la ciencia abierta. Va a ser... Yo creo que va a ser la única manera de hacerlo, porque en ese camino es donde se unen todos los incentivos. Pensad que muchísimos datos de salud, de estilos de vida, están en nuestro teléfono móvil, o están en el dispositivo que llevamos en la

(continúa en la pág. siguiente)

	muñeca, en el reloj inteligente; o está en nuestra báscula inteligente, o esta mañana en la nevera, que sabe nuestros hábitos, qué es lo que comemos, mejor que nosotros. O está en los propios supermercados, en las tarjetas de fidelización. Claro, son fuentes de información a las que en la actualidad no tenemos acceso, y si tuviéramos acceso a lo mejor tendríamos problemas después, por cuestiones de interoperabilidad. Yo creo que la solución a esos problemas de interoperabilidad vendrá porque compartamos esos datos, nos enfrentemos a tener que solucionar esos problemas y ganemos la utilidad de haberlos usado. Entonces los incentivos se suceden de esa manera. ¿No? Hasta que no demuestras utilidad, no tienes el interés real en usarlo más que de forma experimental o exploratoria. No sé, yo creo que la oportunidad está ahí. Yo creo que tenemos para todos... hay, hay trabajo para todos, por decirlo así, y hay muchas oportunidades para la alianza y la mejora continua entre todos.
MODERADOR	Gracias, Paco. Mercedes, desde la perspectiva de la autoridad sanitaria, la capacidad de respuesta, ¿va a mejorar? ¿Tenéis herramientas y elementos como para que la toma de decisiones pueda ser un poco más rápida?
ALFARO	Bueno...
MODERADOR	¿Y puede ayudar a esto contar con un marco legal y legislativo? Porque si no el <i>partenariado</i> público privado y el rol que pueda jugar en esto quedaría muy cojo. Por ejemplo, con la dinamización de los fondos que van a venir ahora del plan de resiliencia.
ALFARO	Sí. A ver, aquí hay varias cosas. Algunas las habéis mencionado, otras perdonadme, pero tengo que poner el punto prosaico. A ver. Vamos a hablar primero de proveedores privados, porque estamos hablando de un intercambio de datos que habitualmente se utilizan en primera instancia y como uso primario para la clínica, que es lo que se está recogiendo fundamentalmente a través de la historia clínica y de los demás documentos interoperables en el SNS. Ahí, es verdad que hay un millón trescientos y pico mil funcionarios, y luego habrá un número pequeño de personas que solo tienen un seguro privado, en resumen, unos y otros utilizan únicamente un proveedor privado. Sin embargo, por ejemplo, respecto a las camas hospitalarias que hay en España, aunque el 40% de los hospitales son privados, suelen ser más pequeños y hay más numéricamente, el 85% de las camas son financiadas por el sector público.

(continúa en la pág. siguiente)

[03:30:09]**ALFARO**

Y existen, además, convenios muy importantes, sustitutorios en muchos casos, como es la Fundación Jiménez Díaz o el hospital de San Juan del Alfarache, en Sevilla, y como estos otros hospitales sustitutorios que trabajan para la pública con un área de población asignada. Hay otros casos en los que la pública les tiene concertados procedimientos, listas de espera. Entonces ahí el intercambio de información tiene que existir y, además, tiene que existir de una manera normalizada. Y hoy en día no está completamente desarrollado ese sistema y evidentemente hay que hacerlo. Pero luego también está la persona que teniendo derecho a una asistencia sanitaria pública, que es la inmensa mayoría, e incluyo a MUFACE tanto los que toman opción pública como, de alguna manera, a los de opción privada, por el tipo de financiación que tiene su asistencia, que acudiendo a servicios privados o a servicios públicos, en un momento determinado quieren ir a la inversa y quieren que su información sea interoperable y quieren que se pueda fusionar, y quieren que esté disponible para su correcto tratamiento. Y eso, evidentemente, es un fin muy legítimo. Y es un fin en el que hay que trabajar sin ninguna duda.

Pero habrá que trabajar también con unas reglas del juego porque, bueno, imaginaos, os voy a poner un ejemplo muy típico, y se puede dar en los dos sentidos, aunque más en uno que en otro. Yo soy una persona que tengo una póliza de seguro privado y en la pública me hago una prueba diagnóstica cara. Esa prueba diagnóstica cara está a disposición mía y, por supuesto, si me voy al proveedor privado está a su disposición, pero a mi ese proveedor privado no me ha bajado la póliza o me ha cobrado menos por mi asistencia, y sin embargo ha tenido un insumo menos o ha resuelto el problema en menos visitas. Y además es un insumo que yo he pagado por otro lado, con mis impuestos. Quiero decir, se necesitan unas mínimas reglas del juego a la hora de intercambiar insumos, vamos a ser prosaicos, entre proveedores, unos con una cuenta de resultados, muy legítimo, y otros con un sistema de redistribución de renta a través de la sanidad, como es la pública. Entonces son dos mundos que se tienen que hablar, pero que para eso tienen que establecerse reglas del juego. Lo mismo para la utilización de datos para la investigación, luz y taquígrafos. Yo creo que hay que compartirlos, pero tiene que ser también con unas reglas del juego. Es decir, yo cuando era más joven y estaba en un hospital y los laboratorios te pagaban un viaje a un congreso a mí siempre me pareció que los laboratorios

(continúa en la pág. siguiente)

	debían dar donaciones a los hospitales que salieran en el tablón de anuncios y que dijeran: tal laboratorio ha dado tanto, tanto y cuánto. Y ser la dirección médica del hospital, los departamentos, los jefes de servicio, los que decidieran a qué congreso se iba y quiénes iban con esas donaciones. O sea, a mí la colaboración público-privada, siendo defensora a ultranza de lo público, y me conocéis todos, creo que tampoco hay que demonizarla, pero que tiene que tener unas reglas del juego muy claras. Para la asistencia y para la investigación.
MODERADOR	Ya. Y la respuesta que es una demanda muy mayoritaria por parte de la administración pública, y que se ha visto, por ejemplo, en la primera ola de la COVID, que cuando a los gestores y a los decisores se les permite actuar bajo un criterio de urgencia, se han buscado la vida bastante bien y han reaccionado con una grandísima flexibilidad... Claro, la innovación tecnológica también requiere, una respuesta más ágil. Lo que viene, la digitalización es una ola tremenda, que nos tiene que coger a todos preparados. No sé si a la administración pública la va a coger preparada o no.
ALFARO	La autoridad sanitaria depende de administraciones públicas, como todos los demás ministerios. Y es cierto que la administración española pues necesita una vuelta, un giro, un giro muy importante que no lo resuelve solo la digitalización. Es decir...
MODERADOR	Ya, ya, ya.
ALFARO	Si digitalizas determinados procesos sin haberlos modernizado y sin haberlos repensado, pues digitalizas lo que tienes, punto, digitalizas, si tienes un caos, un caos; si tienes una cosa mediocre, digitalizas lo mediocre. O sea, que la digitalización no es una varita mágica. Sí te enseña muchas cosas que tú no habías pensado que se pueden hacer para mejorar tus procesos, pero realmente hay un planteamiento, un replanteamiento de esos procesos, y uno de ellos son las contrataciones públicas, y otro de ellos es la relación con los administrados o con el paciente o con el usuario o con la familia del usuario. En esta pandemia se ha visto. Entonces, es que yo creo... hombre, yo tengo la esperanza, de que en el Ministerio de Función Pública ahora mismo hay un equipo que procede en gran parte del Ministerio de Sanidad, y ha vivido la pandemia en primera persona. Fue el equipo que trajo Illa, el Subsecretario, el Secretario General

(continúa en la pág. siguiente)

	Técnico... O sea, una serie de gente que conocen las necesidades de la sanidad Y ahí tengo yo un rayito de esperanza de que comprendan desde que tiene que haber expertos en semántica en la plantilla, o ingenieros, hasta que hay que hacer una colaboración público privada bien reglada, pasando por agilizar la compra pública innovadora que es un rollo tártaro en el que da muchísima pereza meterse, en resumen, que hay que modernizar la administración española y en concreto la sanitaria.
BANEGAS	Para liderar hay que querer ser líder: ¿quién quiere ser líder realmente? Inicialmente, podría pensarse que el sector público, cuyo cometido es obligado por definición. Pero sectores público y privado deberían ser más complementarios que competitivos, lo mismo que la SP añade el componente social a la genómica de precisión. El sector privado añadiría la más temprana adquisición de algunas tecnologías novedosas a evaluar localmente, quizás conjuntamente con el sector público, la educación sanitaria (p.ej., dudas sobre vacunas); la más efectiva y eficiente labor investigadora complementaria (y ética) al manejar a una parte no desdeñable de los mismos pacientes del sector público (por ejemplo, en chequeos racionales y segmentados por grupos de riesgo); y la imperiosa necesidad de colaboración en momentos de crisis como el de la pandemia Covid-19. Debería haber acuerdos de mínimos entre ambos sectores para su cooperación, basados en normas éticas, y no cada sector operar con sus códigos por separado.
CASTRODEZA	Los debieran liderar los más capaces. Y como siempre debiera haber una colaboración público-privada estrecha.
MODERADOR	Perfecto. Muchísimas gracias. Última pregunta. El término salud pública de precisión ha sido criticado por presentar dos concepciones diferentes. Una está más relacionada con el uso de nuevas tecnologías para mejorar procesos tradicionales de salud pública, como hemos ido viendo a lo largo de la tarde, mientras que la otra está más relacionada con la genómica. ¿En qué sentido entendéis que debería desarrollarse el concepto de salud pública de precisión? Y con esto finalizamos. Arrancamos contigo, Guillermo.

(continúa en la pág. siguiente)

SANZ	Bueno, yo creo que son inseparables. Ni la genómica por sí sola vale para nada, si no tiene en cuenta pues factores económico-sociales, geográficos, de exposición a agentes potencialmente tóxicos y demás. O a factores de riesgo que de alguna manera ya conocemos. Es decir, que en ese sentido yo creo que no son excluyentes. Yo creo que hablar de salud pública de precisión es hablar de la interrelación en nuevas tecnologías, con las tecnologías tradicionales, de la aplicación de estas nuevas herramientas que nos dan las nuevas tecnologías en el ámbito de la salud, y en los casos en los que sea posible también el poder incluir los datos genómicos, que no va a ser posible en todos los casos. Pero que yo creo que el camino que lleva la ciencia es que cada vez vamos a tener información genética en un número mayor de enfermedades, insisto, tanto de alta como de baja incidencia.
MODERADOR	Sí, sí. Gracias. Me encanta que haya calado en el estamento clínico el pensamiento salubrista, que se tengan en consideración los determinantes de la salud y que los manejeis con la pericia que lo acabas de hacer, Guillermo, gracias. Además, ponía a lo que tú apuntas antes Voro el ejemplo de Joaquín Dopazo en la Junta de Andalucía, que está integrando datos genómicos en la historia clínica digital. O sea, que se puede. Paco.
SANZ	No, perdona, y una última cosa que ha comentado también Voro Peiró, que es: la imagen es fundamental.
MODERADOR	Sí, por supuesto, claro.
SANZ	En este momento los radiólogos se están comiendo a un montón de las especialidades tradicionales. Es decir, el ámbito de actuación del área de la imagen cada vez es mayor.
MODERADOR	Sin duda.
SANZ	En técnicas, por ejemplo, cardíacas, lo que se te ocurra. Entonces, la disponibilidad, la incorporación de las técnicas de imagen a esas bases generales yo creo que también va a ser fundamental. Imagen, genómica y todo lo demás.
[03:40:00]	
MODERADOR	Sí, sí, además en Valencia tenéis a Lluís Martí Bonmatí que es un crack en esta materia, sí, sí. Fantástico. Paco.

(continúa en la pág. siguiente)

ESTUPIÑÁN	Sí, yo creo que la crítica, como hemos podido ver en este propio parlamento y discusión sigue vigente, ¿no? Entre otras cosas, porque esa definición de salud pública de precisión...
MODERADOR	Chirría, ¿no?
ESTUPIÑÁN	Todavía nos chirría a todos. En ese sentido, tal cual yo lo veo, lo veo igual que Guillermo. Al final lo que guía la aproximación de salud pública es el conocimiento de los determinantes de salud y la genética es un determinante más, con un peso específico, dependiendo; en cada situación tiene el suyo relativo, y dentro de ese sentido es mucho más amplio el ámbito de la salud pública que el de la genética. Por eso yo creo que la salud pública de precisión es, al final, todas esas tecnologías que te ayudan a ganar información relevante sobre aquello de lo que estás estudiando. Sigo pensando que también la distinción que se ha hecho aquí por parte de Salvador es pertinente. Es la diferencia entre lo que sería salud pública y epidemiología puramente. Y yo creo que en ese sentido estamos hablando más de epidemiología, en estas ocasiones. Nos vamos más hacia el estudio de las cosas que hacia la acción sobre las cosas.
MODERADOR	Correcto.
ESTUPIÑÁN	Entonces, yo creo que en la progresión a lo mejor llegaremos a un momento en que, o bien salud pública de precisión tiende a ser un término paraguas en el que luego tengamos que entrar a especificar, o bien que ese término acabe como tantos otros en desuso, y evolucione hacia algo que sea más específico y en lo que todos podamos coincidir.
MODERADOR	Genial, gracias. Cristóbal.
BELDA	Bueno, yo creo que la salud pública de precisión debería desarrollarse en un contexto de un sistema de salud pública de altas prestaciones proporcional al nivel de calidad de los servicios asistencial del más alto nivel que tenemos en España. Bajo el liderazgo y la coordinación del Ministerio de Sanidad, y en coordinación con la forma en la que cada Comunidad Autónoma, desde sus competencias, plantee su salud pública de cercanía, debemos ir hacia un sistema de altas prestaciones. O vamos hacia ese concepto o al final seguiremos repitiendo las mismas cosas, tal y como comentaba la doctora Alfaro, pero gastándonos mucho más dinero, al fin y al cabo.

(continúa en la pág. siguiente)

MODERADOR	Gracias. Pues Mercedes.
ALFARO	Ahora, ya lo tengo. Nada, vamos a ver, pues lo dicho. ¿Cómo veo yo que puede ser mucho más precisa la salud pública? Pues desde el punto de vista epidemiológico, el poder llegar a conclusiones mucho más certeras, mucho más seguras y que permitan aplicar de una manera mucho más eficaz pues las estrategias y los programas de salud pública, sin olvidar en ningún momento algo que se ha comentado, no sé si ha sido Guillermo o Voro. Que, si ayuda a penetrar los grupos o colectivos a los que hay que influir, bienvenido será, porque muchas veces el problema es que es tan generalista... “paseen ustedes, no coman tal...” Que, realmente, cala poco. Cala cuando ya la persona se ha puesto enferma o cuando se ve en un apuro. Entonces, yo creo que la precisión tiene que ir para... en ese sentido. Siempre y cuando redunde en eso. En unas estrategias mucho más eficaces y unos programas mucho más aceptados.
CASTRODEZA	No lo tengo claro. Para mí es necesario avanzar en determinar lo que entendemos por SPP. Como casi siempre la SP lo abarca todo.
BANEGAS	Como no sabemos cómo y cuánto puede avanzar el conocimiento, hay que apostar por líneas variadas. Yo apuesto por el término más amplio de SPP-multidisciplinar, que obviamente incluye al primero (genómica -toda enfermedad es 100% genética y 100% ambiental). Es decir, Integración Robusta de Conocimiento de ciencias básicas, clínicas y poblacionales (Khoury). El mayor impacto de salud vendría de la incorporación de datos socioeconómicos a la genómica (p.ej., pobreza, educación, vivienda, desigualdades). Pero los estudios tipo GWAS (o los EWAS) en sí mismos podrían tener valor y mayor validez interna (si tienen una finalidad explicativa) al no considerar otros factores (posibles confusores), aunque menor relevancia de salud pública (explican poco –p.ej. en presión arterial solo explican el 2% o 1/20 de la varianza que se esperaba explicar –posible mejora: medir el efecto acumulado de muchas variantes de bajo riesgo [Score Genético]). Los GWAS sirven también para indagar interacciones genético-ambientales para tratamientos específicos y mayor adherencia en estilos de vida sanos.

(continúa en la pág. siguiente)

	Dadas las evidencias disponibles, aún incipientes, y su escaso grado de desarrollo en España, parece razonable impulsar el movimiento SPP, empezando por el principio de lo ya existente hacia la confección e implementación de una Estrategia Nacional.
MODERADOR	Pues muchísimas gracias a todos. La verdad es que me vais a disculpar porque, vamos, nos hemos ido de hora, son muchas preguntas, pero realmente ha sido, al menos para mí, reconozco que enormemente enriquecedor el debate. Se han dicho cosas muy interesantes, y ahora tenemos el reto de darles forma, y de retroalimentarnos. Gracias por vuestras contribuciones, que desde la Fundación os agradecemos mucho. Gracias, y buenas tardes a todos.
	[Fin de la transcripción] [03:46:40]

5. RESULTADOS DEL ANÁLISIS FOCAL

ALCANCE DEL IMPACTO DE LA SALUD PÚBLICA DE PRECISIÓN

Son varias las áreas dentro de la salud pública que están avanzando gracias al uso de grandes datos y nuevas metodologías de investigación. Existe un cierto consenso entre los panelistas en que la infectología y la oncohematología son las disciplinas en las cuales la “salud pública de precisión” (SPP) puede alcanzar un mayor impacto. Respecto a la primera, se señala la relevancia de la genómica de patógenos, es decir, el uso de la genómica para rastrear agentes infecciosos. Gracias a ello, se puede obtener información epidemiológica del microorganismo causante, así como de su perfil de resistencias. Esto permite ayudar a la trazabilidad de enfermedades transmisibles y a la selección de antibióticos de forma más eficiente. Si bien la pandemia de COVID-19 ha evidenciado aún más esta utilidad de la genómica, ya existían previamente trabajos a este respecto, como se remarca en el grupo de discusión; se pueden leer algunos de estos ejemplos en la revisión sistemática de la literatura. En relación con la oncohematología, si bien pudiera ser un área con cierto solapamiento con la medicina de precisión, hay casos de uso de prevención y promoción de la salud que probablemente han de considerarse dentro del espectro de la SPP. Un ejemplo que se proporciona es la reciente evidencia de que los pacientes con hematopoyesis clonal de significado incierto presentan un significativo aumento

de riesgo cardiovascular. Por tanto, esta población podría establecerse como diana para intervenciones de salud cardiovascular. Así, en algunos centros se está comenzando a llevar a cabo un seguimiento estrecho de esos pacientes y reforzar las medidas de prevención de enfermedad cardiovascular en los individuos a riesgo.

En todo caso, los panelistas concuerdan en que será la capacidad de modelizar y predecir la que aporte mayores avances en las diferentes esferas de la salud pública. Así, los modelos predictivos son una gran promesa en el avance del conocimiento, tanto para prevenir la enfermedad, como para promover y proteger la salud de la población. Será fundamental la correcta evaluación y validación de estos modelos para poder implementarlos en las diversas áreas y hacer uso de ellos con suficiente confianza para extraer nuevas conclusiones en investigación.

En respuesta a la segunda pregunta- si las conclusiones arrojadas hasta el momento por estudios que emplean metodologías de Salud Pública de Precisión son útiles para guiar políticas sanitarias- los panelistas remarcaban la aún escasa evidencia de calidad para la toma de decisiones en salud pública. Se comenta que los estudios incluidos en la revisión sistemática tienen un tamaño muestral discreto y se señala que apenas hay estudios realizados bajo metodologías entendidas como SPP que aporten conclusiones distintas (y a la vez robustas) que las de estudios realizados con metodología tradicional. Por tanto, hacen falta más trabajos de calidad para objetivar que el *Big Data* y el complejo análisis de dichos datos puede aportar conocimiento sólido que permita tomar decisiones en políticas sanitarias con certidumbre. No obstante, existe en el grupo un claro consenso de que los estudios bajo metodologías de SPP sí que deberían guiar las políticas sanitarias, desde el momento en que alcancen un adecuado rigor metodológico que permita obtener conclusiones válidas. Para ello, no solo hacen falta muchos datos, si no datos de calidad; ésta debe estar presente desde la extracción al análisis de estos, para permitir su adecuada interpretación y la subsiguiente generación de conocimiento.

Se señalan algunos ejemplos en los que parece pertinente el diseño de estudios con metodologías de SPP; entre ellos se propone la monitorización de la calidad de las aguas, la reducción de la sal de los productos de bollería, el control del tráfico en las ciudades o la detección de brotes de enfermedades infecciosas.

Todas estas intervenciones tienen como denominador común el potencial uso de grandes datos para poder extraer conclusiones en poblaciones concretas, con mayor granularidad que otros estudios que emplean metodologías tradicionales. Para poder alcanzar su implementación basada en evidencia, primero se deberá demostrar la relación o asociación que permita llevar a cabo una asunción de causalidad. En una etapa siguiente, se deberá demostrar la factibilidad y efectividad de la intervención propuesta. Las decisiones de salud pública han de tomarse basadas en datos que permitan aceptar que son adecuadas; sin embargo, el avance en el conocimiento epidemiológico o la detección de áreas de mejora de salud pública no será suficiente. Para alcanzar el verdadero potencial de la SPP, habrá de demostrarse que las intervenciones derivadas de dichos hallazgos son efectivas y eficientes.

Respecto al potencial de estas nuevas metodologías de investigación, en general los panelistas mostraron acuerdo en que mejorarán los resultados en salud, incluyendo avances en la optimización de estrategias preventivas y en el acceso a la atención médica. Algunas de las áreas donde se esperan importantes mejoras son el diagnóstico de enfermedades, la especificidad del tratamiento o el desarrollo de productos y dispositivos sanitarios. Se indican algunos pasos ineludibles para dirigirse hacia dichos avances, como definir la magnitud de la precisión que se pretende alcanzar, aumentar la capacidad del Sistema Nacional de Salud (SNS) de explotación de los datos y delimitar el papel de los distintos actores participantes en los proyectos de SPP.

Para poder desarrollar estudios que mejoren los actuales resultados, serán necesarios avances estructurales –tanto tecnológicos como de recursos humanos–, que permitan una adecuada explotación e interpretación de los datos. Así, se señala la capacitación de los profesionales en estas nuevas metodologías como una condición necesaria para el avance en la implementación de estas tecnologías. Dicha capacitación será necesaria tanto en la propia aplicación de la tecnología, como en la decisión sobre en qué circunstancias usarla y cómo interpretar lo que se obtiene de su uso. Uno de los pasos que se remarca fundamental para alcanzar ese impacto positivo es la inversión en formación. Es evidente la escasa existencia de perfiles profesionales mixtos, con formación transversal en ciencias de la salud y ámbito técnico. A pesar de haber aumentado en los últimos años, se estima que la demanda de este tipo de perfiles

estará muy por encima de la actual capacidad de formación. Asimismo, se señala la necesidad de introducir perfiles técnicos en estudios de salud pública. Una formación académica flexible es el camino para adaptarse a estas nuevas necesidades, permitiendo disponer de profesionales formados en salud digital desde distintas ramas de formación universitaria.

Dada la tendencia a la medicalización de la salud en los últimos años, los panelistas contemplan un importante impacto potencial de la implementación de intervenciones (incluyendo decisiones terapéuticas) sin un claro beneficio-riesgo favorable. Esto sería la consecuencia lógica en el caso hipotético de hacer un manejo e interpretación incorrectos de los datos, generando por tanto información sesgada. Dicha situación puede llevar a convertir en patológicas situaciones que no lo son o a prestar una atención desmedida a ciertas intervenciones sobre la salud. Sin embargo, se muestra confianza en la aparición de guías que permitan una adecuada diseminación de la información, así como en la instauración de medidas de barrera que establezcan una correcta toma de decisiones basadas en evidencia. El SNS tiene la responsabilidad de proporcionar mecanismos para indicar qué prestaciones o intervenciones preventivas tienen sentido desde el punto de vista de efectividad y de coste-efectividad para las poblaciones.

El impacto de las intervenciones hay que valorarlo de forma individualizada en distintas áreas. Así, el diagnóstico precoz puede significar un avance fundamental en ciertas patologías (por ejemplo, cáncer de páncreas), mientras que podría suponer una preocupación innecesaria en patologías médicamente menos relevantes. Por ello, el razonamiento clínico realizado desde una perspectiva de salud pública es decisivo para encauzar los distintos problemas de salud de la población. Haciendo una analogía del dicho de medicina de precisión, se pretende abordar a la población correcta, en el momento correcto, con la intervención correcta. Esto es, la SPP trata de encontrar el equilibrio, para el cual es tan pernicioso una intervención necesaria no realizada como una intervención realizada cuando no era necesaria.

SISTEMAS DE INFORMACIÓN GEOGRÁFICA

En cuanto a la cuestión relativa al uso de sistemas de información geográfica, se ponen de manifiesto una amplia variedad de respuestas, sin duda influenciada

por los distintos perfiles y actividades profesionales de los panelistas. Los sistemas de información geográfica son un conjunto de herramientas que permiten la conjunción de datos relacionados con el espacio físico con programas informáticos o software. Con ellos, es posible integrar, analizar y representar la información geográfica sobre ciertos elementos asociada a un territorio. En la discusión se señalan algunos proyectos a nivel internacional, como el *Global Burden of Diseases Risk*, o nacional, como estudios del Centro Nacional de Epidemiología sobre cáncer en territorios pequeños.

A pesar de la variedad en las respuestas en relación con la utilización práctica de sistemas de información geográfica, los panelistas se encuentran en general abiertos a su uso y alineados en cuanto a su potencial. Según se reportó en la revisión sistemática de la literatura, existen publicaciones incipientes que evidencian que esta tecnología puede ser una modalidad útil para guiar las políticas de salud pública y obtener datos de precisión sobre los riesgos de la población, incluyendo análisis de comorbilidades, variación en tasas de incidencia, eficacia de programas de tratamiento o análisis de factores socioeconómicos asociados al riesgo de enfermedad. En el debate se mencionan algunos de estos artículos, que abordan distintos objetivos como la valoración del riesgo de insuficiencia cardíaca en la población (algoritmo CARPEDIEM) o la predicción a corto plazo de la evolución de la pandemia por COVID-19. Sale a debate cierta frase utilizada entre expertos en salud pública, que dice que “*el código postal es más importante que el código genético*”. Asimismo, se comenta un caso de uso en España desarrollado durante la pandemia de COVID-19, que consiste en un sistema de información llamado SERLAB, donde todos los laboratorios públicos y privados envían los resultados de las pruebas diagnósticas de infección por COVID-19 realizadas cada día. Esto está permitiendo conocer en tiempo real la evolución diaria de cada región.

VALIDACIÓN, DISEÑO, BARRERAS Y RIESGOS DE LOS ESTUDIOS DE SPP

La primera parte de la sexta pregunta –si un estudio bien diseñado sirve para probar si son óptimos los resultados de una intervención de SPP– concitó una respuesta positiva unánime entre los participantes; coincidieron en que un estudio bien diseñado es la forma de validar los resultados de una intervención de SPP.

Frente a la cuestión relativa al tipo de diseño, las respuestas fueron en general laxas en dirección a que el diseño específico vendrá determinado por la pregunta de investigación concreta. No obstante, parece haber cierto consenso en algunas características de los estudios, como el diseño longitudinal, la existencia de un grupo control, y la valoración de los riesgos competitivos. Entre los diseños particulares que se mencionan, caben destacar la aproximación diferencias en diferencias, presente en alguno de los artículos de la revisión sistemática, o el diseño “target trial”. Además, se insiste en la necesidad de modernizar y adaptar los métodos de vigilancia, la informática y la epidemiología para contribuir a una mayor calidad en los estudios.

En cuanto a las barreras existentes para implementar y desarrollar tecnologías de SPP, los panelistas señalan importantes barreras de tipo estructural. La dificultad en la interoperabilidad entre sistemas informáticos sin duda repercute en la capacidad de explotación de los datos. Esta falta de interoperabilidad puede venir condicionada por la escasa coordinación entre el eje público-privado, así como por la territorialidad en comunidades autónomas que manejan sistemas operativos distintos. También se señala una barrera cultural, con una dificultad en la adopción de la tecnología ligada a una falta de formación de profesionales cualificados y capacitados para el uso de metodologías de investigación con grandes datos en salud. Se insiste en que esta escasa existencia de perfiles profesionales mixtos podrá amortiguarse si aumenta la flexibilidad en los entornos académicos en busca de una mejor adaptación a las necesidades de una sociedad digital.

Respecto a los riesgos, se señala que los análisis de grandes datos pueden asociar sesgos de información y sesgos de selección. Se señala la necesidad prioritaria de establecer buenas hipótesis e investigarlas mediante una metodología adecuada, para poder llegar así a modelos causales y no caer en conclusiones erróneas por hallazgos casuales, lo que es fundamental para poder extraer conclusiones en la dirección correcta. También se señalan las externalidades como un potencial riesgo, en situaciones en que las intervenciones puedan afectar a terceros de manera no previsible sin que esto sea claramente medible ni reflejado en los costes. Por ello, toda intervención de salud pública tiene que ir acompañada de un plan de evaluación. Por otro lado, se hace referencia a un potencial incremento de las diferencias entre países ricos y pobres,

en que se hagan avances en análisis de datos y en modelos predictivos que queden circunscritos a algunos países.

PAPEL DE LOS SECTORES PÚBLICO Y PRIVADO

Si hay una palabra que puede resumir las respuestas a la pregunta de cuál será el papel del sector público y privado en el uso de nuevas metodologías de investigación, ésta es *colaboración*. Así, en general los panelistas no apuestan claramente por un liderazgo del sector público en el desarrollo de la SPP, sino que insisten en la colaboración público-privada. Las alianzas entre ambos sectores son claves en el desarrollo de tecnologías emergentes que puedan realizar estudios en el marco de la SPP. Esa colaboración deberá verse reflejada en la interoperabilidad de los sistemas y el fomento de la investigación común. Con ello, no solo se hará posible la realización de investigaciones, si no se atenuarán potenciales sesgos que se podrían facilitar en caso de no existir regulación. Los sistemas de análisis masivos de datos deben incluir parte representativa de los distintos sectores poblacionales que pueden estar en el sistema sanitario público o el privado para no perder una buena parte de la información. De este modo, se concede importancia al papel regulador del sistema público en la implementación y acreditación de estas nuevas tecnologías. Asimismo, se señala que desde las instituciones nacionales e internacionales se está planteando la necesidad de apertura y de reproducibilidad de las investigaciones en términos científicos y de evaluación, con un movimiento de ciencia abierta que permita el acceso a datos disponibles públicamente, a los avances en software, o al acceso a plataformas tecnológicas.

LA IMPORTANCIA DE UN NOMBRE

Desde los primeros momentos de la discusión del panel, los participantes ponen de manifiesto posiciones variadas en cuanto a la interpretación del término de SPP. Se mantiene divergencia de opiniones en relación con la nomenclatura y su extensión, así como cierto escepticismo para el término por parte de algunos panelistas a lo largo del debate.

En todo caso, los participantes aceptan, en general, que el concepto comprende la interrelación de nuevas tecnologías y salud pública. En ese sentido, se hace hincapié en las diferencias entre salud pública y epidemiología. Mientras que la epidemiología es una ciencia que se circunscribe al estudio

de la distribución y los determinantes de estados relacionados con la salud, la salud pública comprende las actividades organizadas por las Administraciones Públicas. Dicha distinción es relevante dado que implica que algunos hallazgos notables a nivel científico puedan no verse trasladados a intervenciones prácticas. Respecto a la distinción entre la parte genómica y el resto de las áreas que se contemplan en SPP, varios panelistas señalan que van ligados en su manejo. Así, se menciona que SPP se entendería como la interrelación de todas las nuevas tecnologías que permitan mejorar la salud pública; en los casos en que sea posible, incluirá los datos genómicos, que al fin y al cabo constituyen un determinante más de salud. La revisión sistemática de la literatura confirma la percepción de los panelistas de una falta de homogeneidad en uso del término “salud pública de precisión” en estudios de investigación.

Con todo esto, parece poder concluirse que el consenso del grupo es que, más allá de la terminología concreta, el objetivo común es alcanzar mayor concreción y granularidad en los grupos en los que hay que influir, redundando en estrategias y políticas de salud pública más efectivas y coste-efectivas.

HOJA DE RUTA

A lo largo de todo el libro se ha puesto de manifiesto la falta de consenso terminológico para el concepto de “salud pública de precisión” (SPP). Por ello, el primer objetivo de esta hoja de ruta es esbozar el significado y la extensión que comprende el término a partir de las aportaciones de las distintas secciones del libro, incluyendo la deliberación de los panelistas. De una manera sencilla, podría resumirse que la SPP emplea el uso de grandes cantidades de datos de alta calidad para dirigir intervenciones poblacionales con mayor eficacia y eficiencia. Es este sentido, su principal motor son los determinantes de salud, de forma similar a lo que ocurre en los estudios de salud pública realizados con metodologías tradicionales. Gracias al uso de nuevas tecnologías, se pretende perfilar con mayor granularidad y precisión distintos grupos de población para comprender mejor las enfermedades y así poder optimizar las acciones de salud pública.

Uno de los debates abiertos a lo largo del libro versa sobre si la SPP debe únicamente centrarse en la genómica o por el contrario beneficiarse además de otras tecnologías y fuentes de datos. Tanto por la evidencia recopilada en la revisión sistemática y el informe de experto, como por la posición general de los panelistas, se puede concluir que el término en cuestión va más allá del uso de la genómica. Si bien el uso de datos genómicos es una parte importante en muchos estudios de SPP, otras tecnologías que aportan información relativa

al ambiente, al comportamiento o al estilo de vida, aportan sin dudas grandes beneficios en el conocimiento. En el informe de experto queda patente como las “-ómicas” pertenecientes a la ciencia básica, como la genómica, transcritómica o la epigenómica pueden proporcionar una gran cantidad de biomarcadores moleculares de un individuo. Sin embargo, es importante remarcar también el impacto que puede tener toda la información granular de una sociedad progresivamente más digitalizada; desde la historia clínica electrónica a los dispositivos móviles, son numerosas las fuentes que prometen proporcionar información precisa para perfilar a los individuos y permitir un mejor cuidado de la población. Con estos datos de gran magnitud, y siempre que se preserve su calidad, parece plausible llegar a conclusiones más certeras y en menor tiempo que lo que se ha conseguido previamente.

A pesar de las numerosas críticas que recibió el término en su aparición, principalmente en relación con la falta de concreción de su significado, su utilización progresiva en estudios de investigación está permitiendo delimitar de forma más precisa su verdadero significado. Por lo tanto, la primera conclusión va en dirección de establecer un concepto inclusivo de SPP, el cual va claramente más allá del estudio de la genómica, e incluye diferentes tecnologías que permiten una evaluación detallada de la interacción entre los determinantes de salud.

Entre las distintas áreas en las que la SPP ha mostrado desarrollo hasta el momento, la prevención de enfermedades transmisibles es probablemente la de mayor impacto. Más aún tras la experiencia adquirida durante la pandemia COVID-19, es prioritario el uso de la genómica en la detección precoz e investigación de brotes de enfermedades transmisibles. La mayor disponibilidad y menor coste de las técnicas genómicas en microbiología, permite la obtención de información precisa sobre agentes infecciosos, evitando el uso de métodos lentos y costosos. Además, la vigilancia epidemiológica se podrá beneficiar del uso de grandes cantidades de datos de distintas fuentes (por ejemplo, de dispositivos electrónicos, sistemas de información geográfica o de la propia historia clínica) para la detección de brotes, acelerando la comunicación de enfermedades de declaración obligatoria.

Con el enfoque en la fuente de obtención de los datos (más que en el área de aplicación), es relevante destacar el uso de biomarcadores moleculares, para ser utilizados en distintas áreas consideradas en el espectro de la SPP.

Aunque muchos de estos biomarcadores no son modificables, sí que pueden ser utilizados para identificar factores modificables asociados a las enfermedades. Así, los biomarcadores genéticos permitirían estratificar la población mediante el planteamiento de *scores* genéticos, contruidos a partir de la suma de variantes genéticas, para indicar qué individuos están en mayor riesgo. De este modo, en caso de existir una estrategia preventiva para la condición en cuestión, habría la posibilidad de ponerla en conocimiento del individuo, en una búsqueda de mejorar la adherencia a conductas que mejoren el pronóstico o la probabilidad de desarrollar la condición.

Por tanto, parece evidente que el desarrollo de la SPP vendrá estrechamente ligado a la implementación de metodologías que empleen *Big Data* y procesamiento de datos mediante tecnologías comprendidas en el espectro de la inteligencia artificial. Aunque se espera que su uso permita un avance significativo en la obtención e interpretación de los datos de salud pública, queda patente que hacen falta más estudios de calidad que demuestren la efectividad de potenciales nuevas intervenciones. Para poder llevar a cabo estos estudios, es imperativo plantearse hipótesis de investigación concretas y aplicar la metodología adecuada a los estudios que las contrasten. En ese sentido, se debe tener presente que la magnitud de los datos no es directamente proporcional al número de preguntas de investigación que estos pueden abordar; el hecho de disponer de una gran cantidad de datos proporcionará mayor precisión y granularidad de información, pero se deberá definir a priori la pregunta de investigación de forma concreta y precisar las variables a estudio, para no perder el foco de las hipótesis a contrastar.

Los datos son un gran activo, pero sin un adecuado abordaje metodológico no podrá obtenerse toda la información que albergan ni mejorar el conocimiento actual. Para mantener el rigor metodológico, será necesario preservar la calidad del dato desde el primer momento de su recogida y a lo largo de todo el análisis. Además, otro aspecto clave para facilitar la aceptación de estos estudios es la transparencia en el reporte de dicha metodología, que permite su mejor comprensión e interpretación. Es necesario adaptar las investigaciones a las recomendaciones en el manejo de datos y su análisis. A este respecto, recientemente han sido publicados los conjuntos de recomendaciones más reconocidos para los ensayos clínicos en inteligencia artificial: Standard Protocol Items:

Recommendations for Interventional Trials-Artificial Intelligence (SPIRIT-AI) y Consolidated Standards of Reporting Trials-Artificial Intelligence (CONSORT-AI). Ambas recomendaciones, centradas en el uso de inteligencia artificial en los ensayos clínicos, recomiendan que los investigadores proporcionen descripciones claras de las intervenciones, el entorno en el que se integrarán, los aspectos relevantes en el manejo de los datos o el análisis de los errores (1, 2).

En la propia conceptualización propuesta de la SPP se pone de manifiesto uno de los principales retos que tendrá que afrontar, que es la unificación de los datos de diferentes fuentes, tanto sanitarias como sociales. Para ello, en primer lugar, será necesario una mejor integración de los datos de salud de la historia clínica electrónica con los sistemas de salud pública, lo cual supone un reto de gran magnitud. Entre las dificultades de este proceso se encuentran las barreras estructurales ligadas a la rigidez y dificultad de interoperabilidad de los sistemas sanitarios de diferentes comunidades autónomas, la falta de claridad en la legislación o las limitaciones de la Administración para afrontar los retos de la revolución digital. Si bien la mayoría de los datos vendrán de organismos públicos, ciertas fuentes serán entidades privadas. Por ejemplo, la integración de los datos de salud móvil (mHealth), así como diversa información relacionada con la exposómica, se verá facilitada en el contexto de una adecuada colaboración público-privada. Por ello, es prioritario establecer un marco normativo que permita la interoperabilidad entre los sistemas, a la vez que garantice la seguridad y confidencialidad de la información. En el momento que la evidencia permita consensuar nuevas intervenciones de salud pública basadas en grandes datos, una escrupulosa consideración de los aspectos éticos será crucial para llevar a cabo dichas intervenciones. Además, será necesario validar el uso de estas nuevas herramientas tecnológicas (como las aplicaciones móviles o los wearables), para facilitar el automanejo de los pacientes y permitir diseñar estrategias personalizadas para determinadas poblaciones.

Otro de los retos para los que debemos prepararnos con el advenimiento de la SPP es el potencial riesgo de intervenciones innecesarias. A diferencia de lo que ocurre en la medicina de precisión, centrada en el tratamiento de pacientes concretos, la SPP requiere un foco de salud pública basado en la prevención de enfermedades en poblaciones. Se deberá mantener cautela al analizar los resultados de los estudios, evitando que se interprete causalidad de asociaciones

que puedan ser espurias, como resultado del contraste de un gran número de hipótesis. Para poder extraer conclusiones adecuadas, los proyectos de investigación deberán diseñarse cuidadosamente desde las etapas iniciales, realizando un proceso robusto de integración de los datos y adhiriéndose a los principios de la medicina basada en la evidencia (3). Requieren especial atención los sesgos intrínsecos de los estudios observacionales; para controlarlos serán necesarios diseños epidemiológicos robustos que garanticen la diversidad y representatividad de la población a estudio (4). Con ello se permitirá ir más allá de descubrimientos epidemiológicos y de la comprensión de los determinantes de salud, para ver ese conocimiento reflejado en intervenciones eficaces en las poblaciones a nivel de prevención, screening, o diagnóstico.

Se pone de manifiesto la relevancia de la inversión en formación y capacitación de los profesionales para poder superar los desafíos mencionados. Las instituciones académicas tienen la llave para facilitar este cambio, mediante una formación más flexible que permita un aprendizaje transversal entre las ciencias de la salud y las ciencias técnicas, tanto a profesionales como a gestores sanitarios. La formación de grado puede ser el punto de partida para incorporar nuevas competencias transversales que permitan adquirir una mejor capacidad de análisis de datos y una mayor comprensión de las distintas áreas de la informática médica. Asimismo, a nivel de formación especializada se requiere el desarrollo de programas específicos que preparen a los profesionales, que incluyan cursos suplementarios para quienes busquen desarrollar una habilidad en particular (de manera análoga, por ejemplo, a un reumatólogo que realiza un curso de ecografía musculoesquelética para complementar su formación). Del mismo modo, será necesario atraer talento técnico al campo de la Salud Digital, y garantizar que reciban capacitación y acreditación que les permita colaborar con profesionales de la salud de manera fluida.

La SPP está dando sus primeros pasos y tiene muchos retos aún por delante para alcanzar todo su potencial. Indudablemente, la pandemia de la COVID-19 ha acelerado el uso de herramientas que emplean grandes datos en distintos niveles de salud pública, a la vez que ha puesto de manifiesto los retos y debilidades a los que debe enfrentarse. Todo ello ha acelerado el desarrollo tanto de las nuevas metodologías de investigación como de salud pública. Así, en los próximos años se espera seguir evolucionando en esa línea para alcanzar un

mejor conocimiento de la etiopatogenia de las enfermedades, así como de su transmisión y diseminación con información geográfica detallada, para poder desarrollar intervenciones más precisas. Es el momento de adaptarse a este nuevo paradigma que promete acercarnos a una asistencia preventiva, participativa y personalizada en salud pública.

REFERENCIAS

1. Rivera SC, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. *BMJ*. 2020 Sep 9;370:m3210. doi: 10.1136/bmj.m3210. PMID: 32907797; PMCID: PMC7490785.
2. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med*. 2020 Sep;26(9):1364-1374. doi: 10.1038/s41591-020-1034-x. Epub 2020 Sep 9. PMID: 32908283; PMCID: PMC7598943.
3. Bayer R, Galea S. Public Health in the Precision-Medicine Era. *N Engl J Med*. 2015 Aug 6;373(6):499-501. doi: 10.1056/NEJMp1506241. PMID: 26244305.
4. Khoury MJ, Iademarco MF, Riley WT. Precision Public Health for the Era of Precision Medicine. *Am J Prev Med*. 2016 Mar;50(3):398-401. doi: 10.1016/j.amepre.2015.08.031. Epub 2015 Nov 4. PMID: 26547538; PMCID: PMC4915347.

APÉNDICE 1

Lista de abreviaturas

Abreviatura	Definición
AI	Inteligencia Artificial
BRFSS	Behavioral Risk Factor Surveillance System
CDC	Centers For Disease Control And Prevention
CIBER	Centro De Investigación Biomédica En Red
CIE-9	Clasificación Internacional de Enfermedades-9
CONSORT-AI	Consolidated Standards Of Reporting Trials-Artificial Intelligence
DDD	Diferencias en Diferencias en Diferencias
DE	Desviaciones Estándar
DNA	Ácido desoxirribonucleico
DPD	Dihidropirimidina Deshidrogenasa
ENCODE	Encyclopedia Of DNA Elements
ETS	Enfermedades De Transmisión Sexual
EWAS	Epigenome-Wide Association
ExAC	Exome Aggregation Consortium
FFQ	Food Frequency Questionnaire
GBD	Carga Global de las Enfermedades
GIS	Sistema De Información Geográfica
GLEAM	Global Epidemic And Mobility

GPS	Global Positioning System
GWAS	Genome-Wide Association Study
HIV	Virus de Inmunodeficiencia Humana
IMC	Índice de Masa Corporal
IU	Intervalo de Incertidumbre
LASSO	Least Absolute Shrinkage and Selection Operator
ML	Machine Learning
NIH	National Institutes of Health
OR	Odds Ratio
pharmaSNPs	Pharmacogenomic Snps
PPP	Public Private Partnership
PRISMA	Proposal To Improve The Publication Of Systematic Reviews And Meta-Analyses
PYME	Pequeña Y Mediana Empresa
RCT	Randomized Controlled Trial
RNA	Ácido ribonucleico
ROC	Receiver Operating Characteristic
RSL	Revisión Sistemática De La Literatura
SPP	Salud Pública De Precisión
S	Sensibilidad
SARI	Infección Respiratoria Agudas Graves
SEIR	Susceptible-Expuesto-Infecioso-Recuperado

SNP	Single Nucleotid Polymorphism
SPIRIT-AI	Recommendations for Interventional Trials-Artificial Intelligence
STEM	Spatiotemporal Epidemiological Modeler
TBC	Tuberculosis
TGCA	Atlas del Genoma Del Cáncer
TPMT	Tiopurina Metiltransferasa
VPN	Valor Predictivo Negativo
VPP	Valor Predictivo Positivo
WES	Whole Exome Sequence Analysis
WOS	Web Of Science

APÉNDICE 2

AUTORES:

Diego Benavent

Reumatólogo en el Hospital Universitario La Paz de Madrid, donde realiza actualmente su doctorado en espondiloartritis axial. Además, colabora como Medical Expert en Savana, una empresa médica dedicada a la investigación con inteligencia artificial. Es profesor de reumatología de la academia AMIR y participa como *fellow* en el equipo de la Outcome Measures Library (OML) de EULAR, una valiosa fuente de información sobre medidas de resultado. Ha sido autor de numerosas publicaciones internacionales y capítulos de libros, tanto en reumatología como en nuevas tecnologías aplicadas a la salud. Fue Vocal de Residentes de la Comisión Nacional de la Especialidad y Representante de los Médicos Jóvenes Europeos (EJD) en la sección de Reumatología de la UEMS.

Es el responsable de la redacción de la introducción, la revisión sistemática de la literatura, el análisis del grupo de discusión y la hoja de ruta.

Juan Ramón González

Associate Research Profesor de ISGlobal, donde dirige el Grupo de Bioinformática en Epidemiología Genética (BRGE). Es coautor de más de 150 artículos científicos y tiene una amplia experiencia en grandes proyectos internacionales de investigación (ATHLETE, HELIX, MEDIRAD y OPERRA). Como profesor adjunto en el Departamento de Matemáticas de la Universidad Autónoma de Barcelona, sus actividades educativas se centran en proporcionar conferencias de posgrado sobre métodos estadísticos avanzados y bioinformática. Tras obtener su Ph.D. en Bioestadística, supervisado por el Prof. Edsel Peña de la Universidad de Carolina del Sur (EE.UU.), pasó dos años como estudiante de postdoctorado en el Centro de Regulación Genómica (CRG). Su trabajo se centra en el desarrollo de métodos estadísticos y computacionales para el análisis y la comprensión de diferentes datos ómicos motivados por una amplia gama de preguntas de relevancia biológica y clínica. Estos incluyen genómica (CNV, inversiones, mosaicismos), transcriptómica (RNA-seq y splicing alternativo) y proteómica entre otros. El análisis del exposoma, su caracterización y cómo integrarlo con diferentes capas ómicas es también otra línea activa de

investigación de su grupo. Es autor del libro “Omic association analysis with R and Bioconductor”, que incluye métodos actualizados para explotar datos ómicos.

Es el responsable de la redacción de Salud Pública de precisión: el papel de la epidemiología genética.

Juan del Llano

Licenciado en Medicina y Cirugía (1981) y Doctor en Medicina (1990) por la Universidad Complutense de Madrid. Especialista en Medicina Preventiva y Salud Pública (MIR 1982-85, Hospital La Paz). M. Sc. Community Health, Usher Institute, University of Edinburgh (1985-86). European Healthcare Leadership Programme, INSEAD (1999-2000). Programa de Alta Dirección de Instituciones Sanitarias, IESE, Universidad de Navarra (2004). Advanced Health Leadership Forum, Universitat Pompeu Fabra, Barcelona y University of California, Berkeley (2005). Salzburg Global Seminar, ediciones 1994 y 2015. Desde 1997 es director de la Fundación Gaspar Casal. Desde 1989 es director Académico y profesor de Salud Pública y Gestión Clínica del Master de Administración y Dirección de Servicios Sanitarios, Fundación Gaspar Casal y Universitat Pompeu Fabra (anterior al 2000, ICADE-UPCO). Presidente de la Asociación Española de Evaluación de Tecnologías Sanitarias, desde 2003. Editor Asociado de Gestión Clínica y Sanitaria, desde 1999. Investigador Asociado del Centro de Investigación en Economía y Salud (CRES), Universitat Pompeu Fabra, desde 2000. Evaluador del Área de Biomedicina, Agencia Nacional de Evaluación y Prospectiva. Profesor Ayudante Doctor, Agencia Nacional de Evaluación, Calidad y Acreditación (PAD: 2012-7582). Ha sido Asesor de la Subdirección General de Atención Hospitalaria del Instituto Nacional de la Salud (1987); vicepresidente Ejecutivo de la Asamblea de Madrid de la Cruz Roja Española (1988); Socio director de Análisis y Diseños Operativos en Salud, S.A. (1989-95) y Gerente Senior de Nuevos Clientes de MSD (1995-97).

Es el responsable de la moderación del grupo de discusión, y ha revisado la revisión sistemática de la literatura, la salud Pública de precisión: el papel de la epidemiología genética, el análisis del grupo de discusión y la hoja de ruta.

Este libro ha sido financiado por:

