



# Registro de datos: Cuántos? Cuáles y por qué?

Dra. CRISTIAN BRUJAN SOLANO

# Objetivos

- Definir los conceptos esenciales que permitan dar perspectiva a la discusión del uso de estos métodos en aplicaciones clínicas.
- Discutir los principales desafíos técnicos y conceptuales del análisis de datos a gran escala.
- Discutir la importancia de la regulación de los datos.



<https://www.ontsi.red.es/ontsi/sites/ontsi/files/Informe%20Big%20Data%20en%20Salud%20Digital.pdf>

[REV. MED. CLIN. CONDES - 2019; 30(1) 83-94]

REVISTA MÉDICA CLÍNICA LAS CONDES  
<https://www.journals.elsevier.com/revista-medica-clinica-las-condes>



## Una guía conceptual para usar y entender Big Data en la investigación clínica

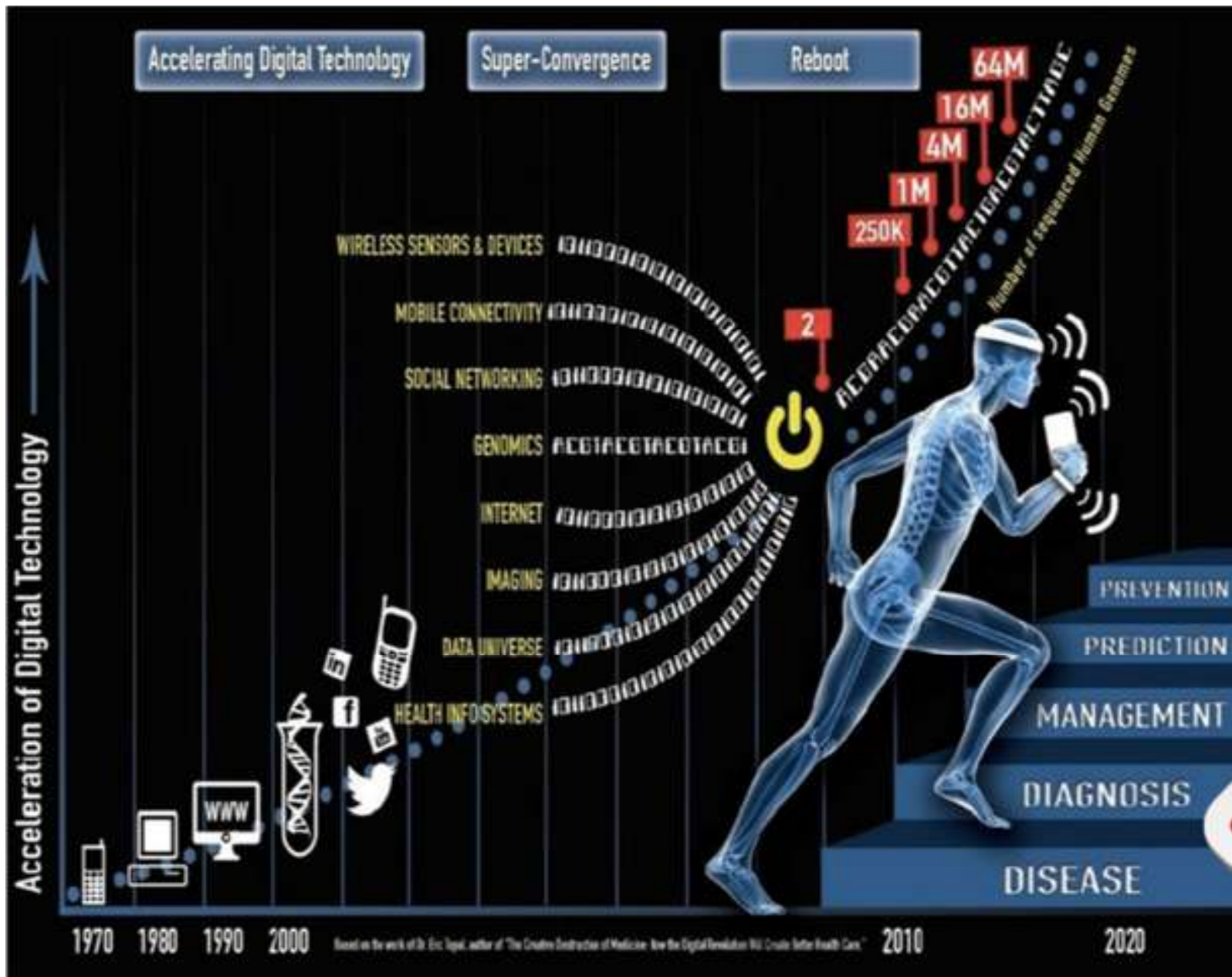
*A conceptual guide to use and understand Big Data in clinical research*

Marcelo E. Andía MD/PhD<sup>a</sup>, Cristóbal Arrieta MS PhD<sup>b</sup>, Carlos A. Sing Long MS PhD<sup>c</sup>

<sup>a</sup> Departamento de Radiología, Escuela de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile.

<sup>b</sup> Centro de Imágenes Biomédicas, Pontificia Universidad Católica de Chile, Santiago, Chile.

<sup>c</sup> Instituto de Ingeniería Matemática y Computacional, Pontificia Universidad Católica de Chile, Santiago, Chile.



# ¿Qué es Big Data?

- Conjunto de datos o combinaciones de conjuntos de datos cuyo tamaño (volumen), complejidad (variabilidad) y velocidad de crecimiento (velocidad) dificultan su captura, gestión, procesamiento o análisis mediante tecnologías y herramientas convencionales, tales como bases de datos relacionales y estadísticas convencionales o paquetes de visualización, dentro del tiempo necesario para que sean útiles.

# Características de Big Data

- Volumen: Cantidad de datos
- Velocidad:
  - De la generación de datos
  - Tiempo de procesamiento de a cuerdo a su uso
- Variedad:
  - Naturaleza de los datos
    - Inamgen radiológica
    - Prueba de laboratorio
    - Antropometría
    - Información cualitativa del expediente



# Características de Big Data

- Veracidad
- Variabilidad: Caducidad de los datos: Remoción o actualización
- Valor: Relevancia en el contexto de análisis

# Cuántos?

- La mayoría de la información procede de los países occidentales, pero esta realidad, poco a poco, va cambiando.
- El 51% de los datos generados actualmente tiene como origen (2012)
  - Estados Unidos → 32%
  - Europa Occidental → 19%
  - China → 13% , para 2020 → 22%
  - India → 4%







# Cuáles son las fuentes de datos a nivel de salud?

## **LOS HISTORIALES O REGISTROS MÉDICOS ELECTRÓNICOS INTEGRADOS.**

- Historia clínica electrónica (EHC).
- Registros electrónicos de salud (EHR).
- Los sistemas de intercambio de información en salud.
- Las aplicaciones de telesalud o telemedicina.
- Servicios de telemedicina.
- Servicios de asistencia remota.

# Transformación de datos en conocimiento

- Big Data permite el análisis no causal de grandes volúmenes de datos que puede estructurar nuevo conocimiento, especialmente en el área de la genómica.
- Mejora del aprovechamiento de la información
- En Big Data, la información no se recoge con una finalidad inmediata.
- Salto en la investigación clínica: la propia lógica descentralizada y distribuida de los sistemas.
  
- Nuevos instrumentos para los profesionales de la salud.
- Promoción del autocuidado de la salud.

# 10 DATOS RELEVANTES SOBRE EL 'BIG DATA' EN 2020

Cada usuario generará

**1,7 MB**

de información nueva por segundo.



Nuestro universo de datos pasará de 4,4 zettabytes a

**44 ZB**

(44 billones de GB).



Habrán más de

**50.000 millones**

de dispositivos *smart* conectados por el mundo, listos para recoger y analizar datos.



Las soluciones tecnológicas *big data* aportarán

**206.000 M€**

a la economía de Europa, lo que supone un incremento del PIB del casi 2%.

El negocio *big data* tendrá un valor de

**9.400 M\$,**

lo que representará el 10% del mercado global de herramientas de gestión de la información.



Se necesitarán

**8 millones de profesionales**

especializados en *big data*.

La cantidad de información

generada por empresas **aumentará 75 veces**, por lo que solo el personal de IT se incrementará 1,5 veces.



Cada persona producirá cada día **más de 1,5 GB** de datos, lo que equivale a:



**1,5 millones** de mensajes



**750 imágenes**



Toda la jornada escuchando música



**4 horas** de videos reproducidos

Cada uno de los **coches autónomos** que circulen en carretera de aquí al 2020 producirá más de

**4.000 GB**

de datos al día.



Actualmente, solo se analiza y utiliza el **0,5% de todos los datos** producidos. El potencial del *big data* está aún por explotar...



# Actuar con prudencia en big data en salud

- Es esencial asegurarse de la calidad de los datos y la información de los modelos
- Asegurar que los procesos sanitarios están adecuados según los recursos disponibles
- Personal capacitado para el uso e interpretación adecuada

# Conclusiones

- La aplicación de Big Data en el entorno de salud, supondrá una mejora de magnitud aun no predecible.
- La calidad de atención en los pacientes ,así como en la prevención, diagnóstico y tratamientos de enfermedades.
- No obstante el verdadero valor del Big Data en salud se conseguirá, si los actores implicados en el proceso se unen, estos son la administradora de salud pública y privada, hospitales, clínicas, universidades y centros de investigación





*Gracias*