Data-Driven Healthcare

Canada's Big Data Consortium
Working Group on Predictive Health

Greg Butler

Data Science Research Centre

and

Centre for Structural and Functional Genomics

and

Computer Science and Software Engineering
Concordia University, Montreal, Canada

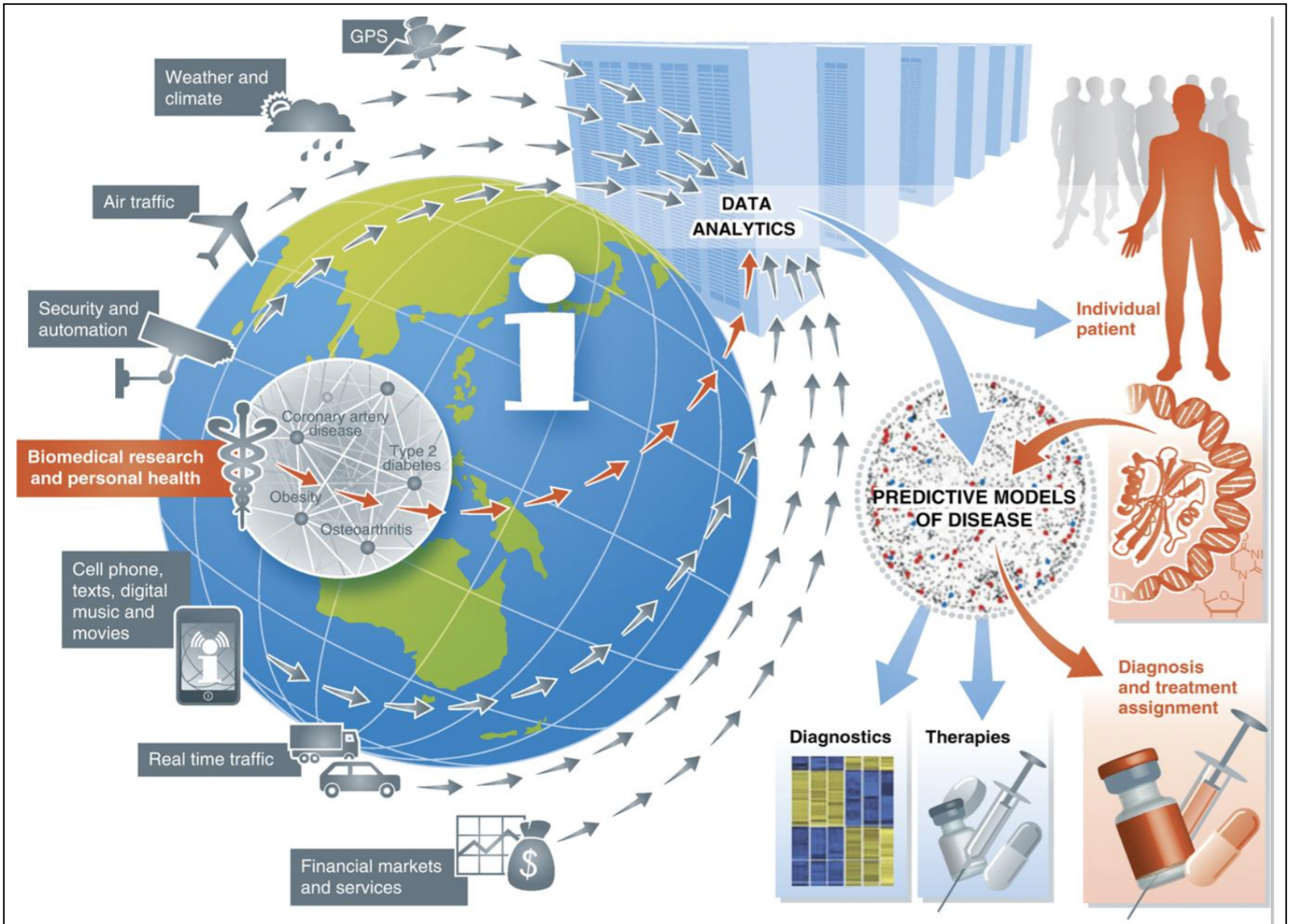
gregb@cs.concordia.ca

Big Data Day — 21 February 2018

Overview

1. Overview in Pictures
2. Big Data
3. Working Group on Predictive Health
4. Big Data in Healthcare and Health Research
5. Working Group Recommendations

Big Data



Eric E. Schadt, The Changing Privacy Landscape in the Era of Big Data, *Molecular Systems Biology* 8, 612 (2012).

Basics of Data-Driven (Clinical) Healthcare

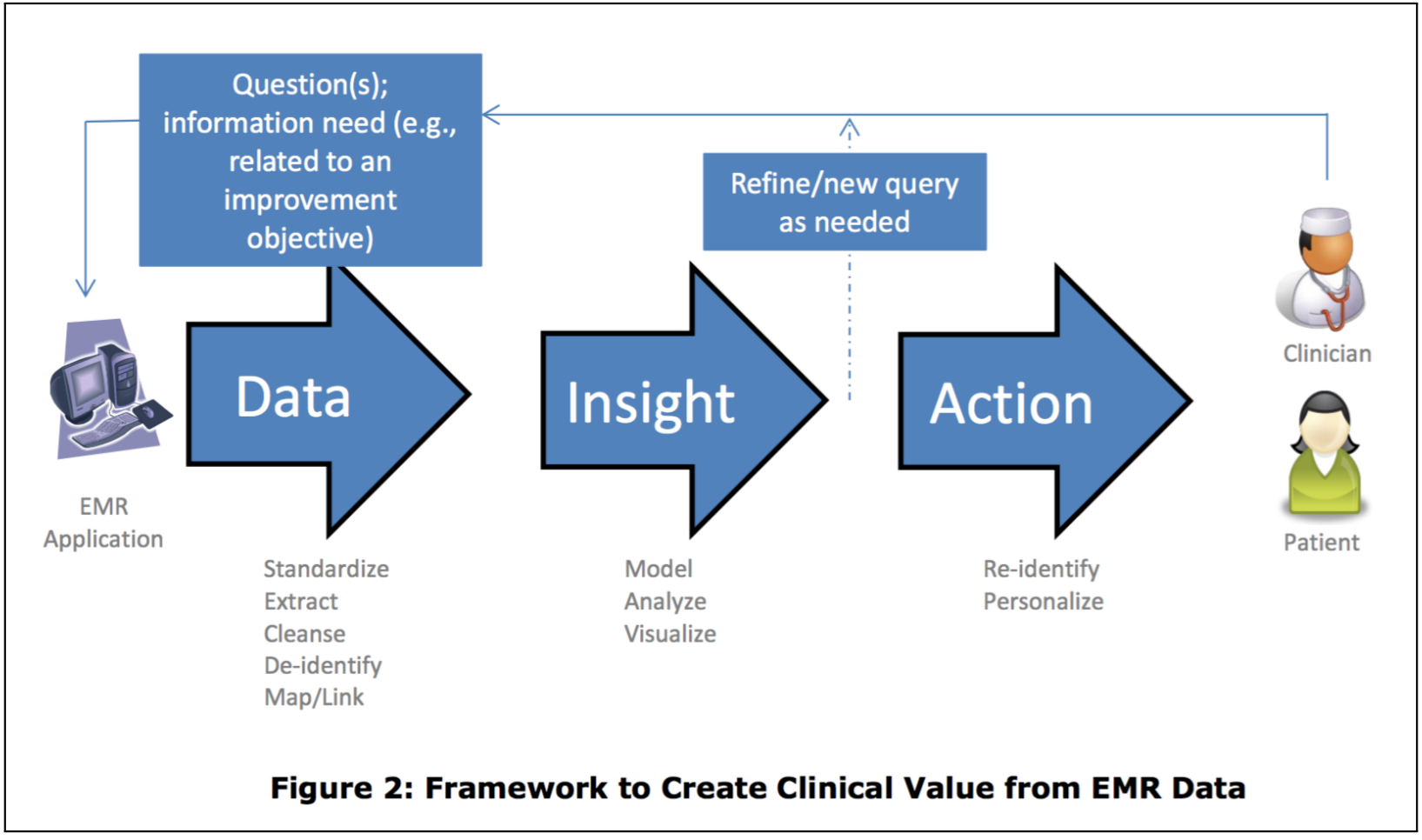
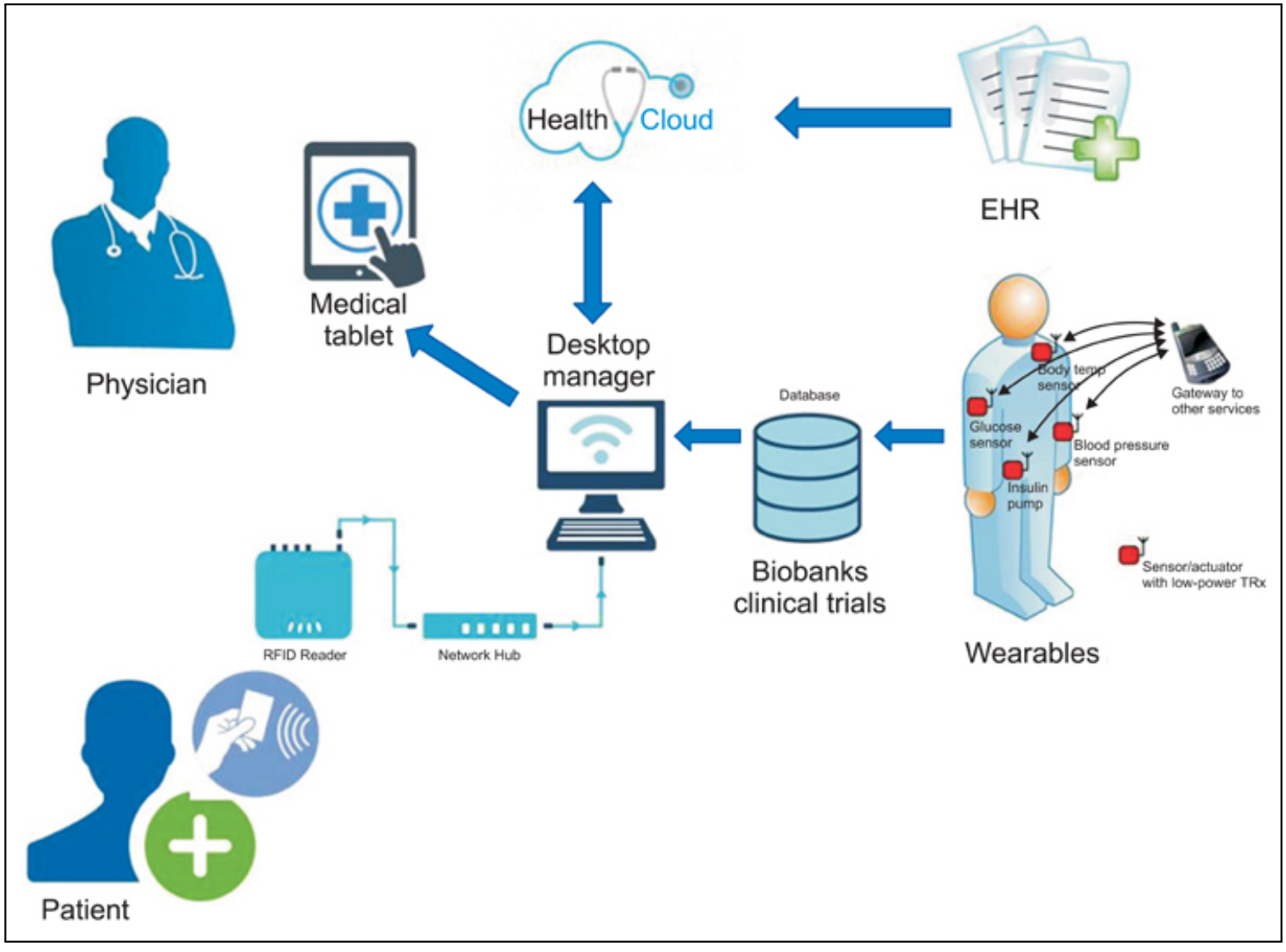


Figure 2: Framework to Create Clinical Value from EMR Data

(Infoway Health Canada 2016)

The Big Data Future



Dimitrov (Health Informatics Research, 2016)

Big Data (http://dsrc.encs.concordia.ca/what-is-bigdata.html)

Big Data

Definition of *"Big"* has changed as we have become more advanced

History

Hollerith Cards 1890 (US population census)

Economic Data 1952 (GDP etc)

Computers 1959 — The First Digital Data Tsunami

World Wide Web 1990's — The Second Digital Data Tsunami

Social Media 1985 — The Third Digital Data Tsunami

Internet of Things 2000 — The Fourth Digital Data Tsunami

Big Science — 1960's onwards

Deep Knowledge — 2011 onwards

A key notion is **actionable data** that is useful in supporting decisions, determining actions, and adding value to an endeavour.

Big Data

The 5 V's

Volume: amount of data
Variety: different types of data
Velocity: rate at which data is generated
Veracity: trustworthiness, level of noise
Value: usefulness of data to a business

Drivers

Transactions
Mobile
Social Media
Internet of Things

MGI Report

McKinsey Global Institute, *Big data: The next frontier for innovation, competition, and productivity*, May 2011.

Canada's Big Data Consortium

Established by Ryerson University in mid 2014
Bring together Govt, Industry, and Universities
Four academic founding partners: Ryerson, SFU, Dal, Concordia

Working Group on Predictive Health

Infoway Health Canada since 2001
- spent $2.1B plus matching funds
- ...and still no wide use of EHR's
- ...and no digitally connected healthcare system

Canada Institute for Health Information
- ...to provide comparable, actionable information

Ontario Health Innovation Council report
- ...alas, reality is promoting SME health instrument makers

St Michael's Hospital data warehouse now online linking
- research, clinical, and patient data

Alberta govt has open data as default

Big Data in Healthcare and Health Research

- Cost Effectiveness
- Quality of Physical, Mental, and Emotional Health
- Data Access: Open, Secure, Privacy Control with End-User
- New Technology, New Data, New Social Norms
  eg, Internet of Things, Gamification, attitudes to privacy

Some Examples of Predictive Health

- Predict need for intervention in real-time for crisis healthcare such as heart-attacks, and strokes;
- Predict state of a disease during disease progression;
- Predict disease epidemics and locate source of outbreak;
- Predict diagnosis of disease;
- Predict lifestyle risk factors;
- Predict environment and nutrition factors for and against good health;
- Predict genetic factors for and against good health.

Data-Driven Healthcare

Working Group Concern

Quality of

- Physical Health
- Mental Health
- Emotional Health

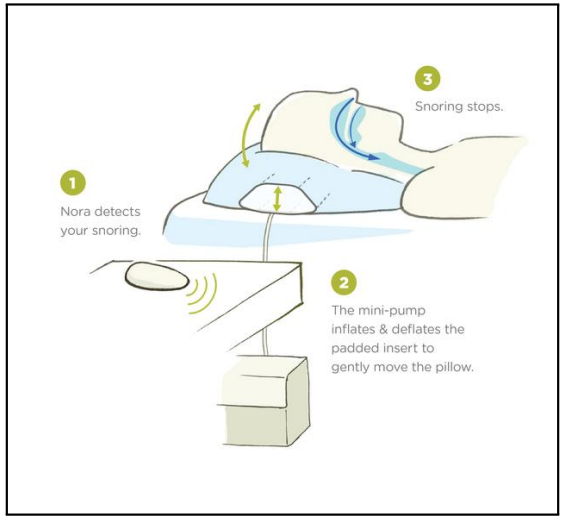
Working Group Concern

Effective Healthcare
(measured as days-of-life $\times$ quality-of-life-per-day)

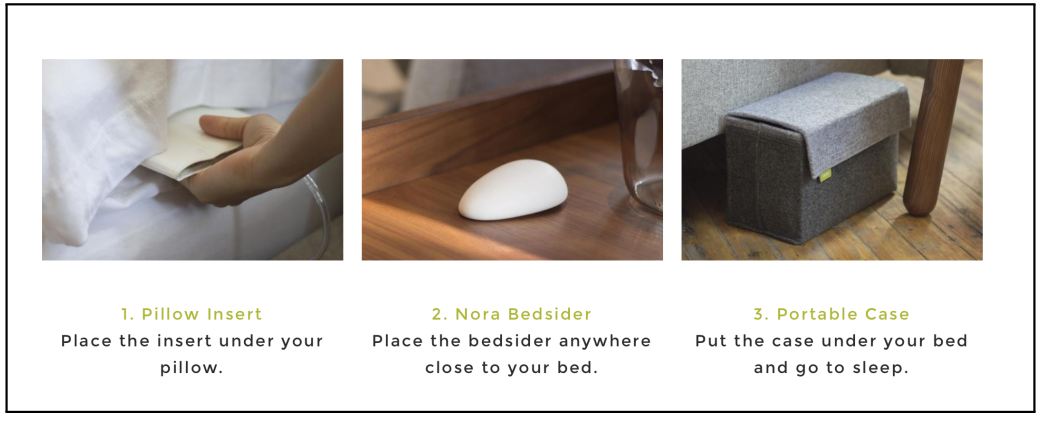
and

Cost-Effective Healthcare
(measured by dollars-per-person-day)

NORA Smart Snoring Solution



http://www.smartnora.com



http://www.smartnora.com

Deep Genomics

`http://www.deepgenomics.com`

"Founded in 2015, Deep Genomics brings together world-leading expertise in machine learning and genome biology. Were inventing a new generation of computational technologies that predict what will happen within a cell when DNA is altered by genetic variation, whether natural or therapeutic."

RightBlueLabs

`http://www.rightbluelabs.com`

Prevention of sports training injury through data mining

- Reduce Illness
- Optimize Recovery
- Maximize Performance

"To deliver an intuitive system that helps identify precursors of burnout. We ultimately aim to reduce attrition and improve performance in sport, the military, government services, and other related industries."

Results

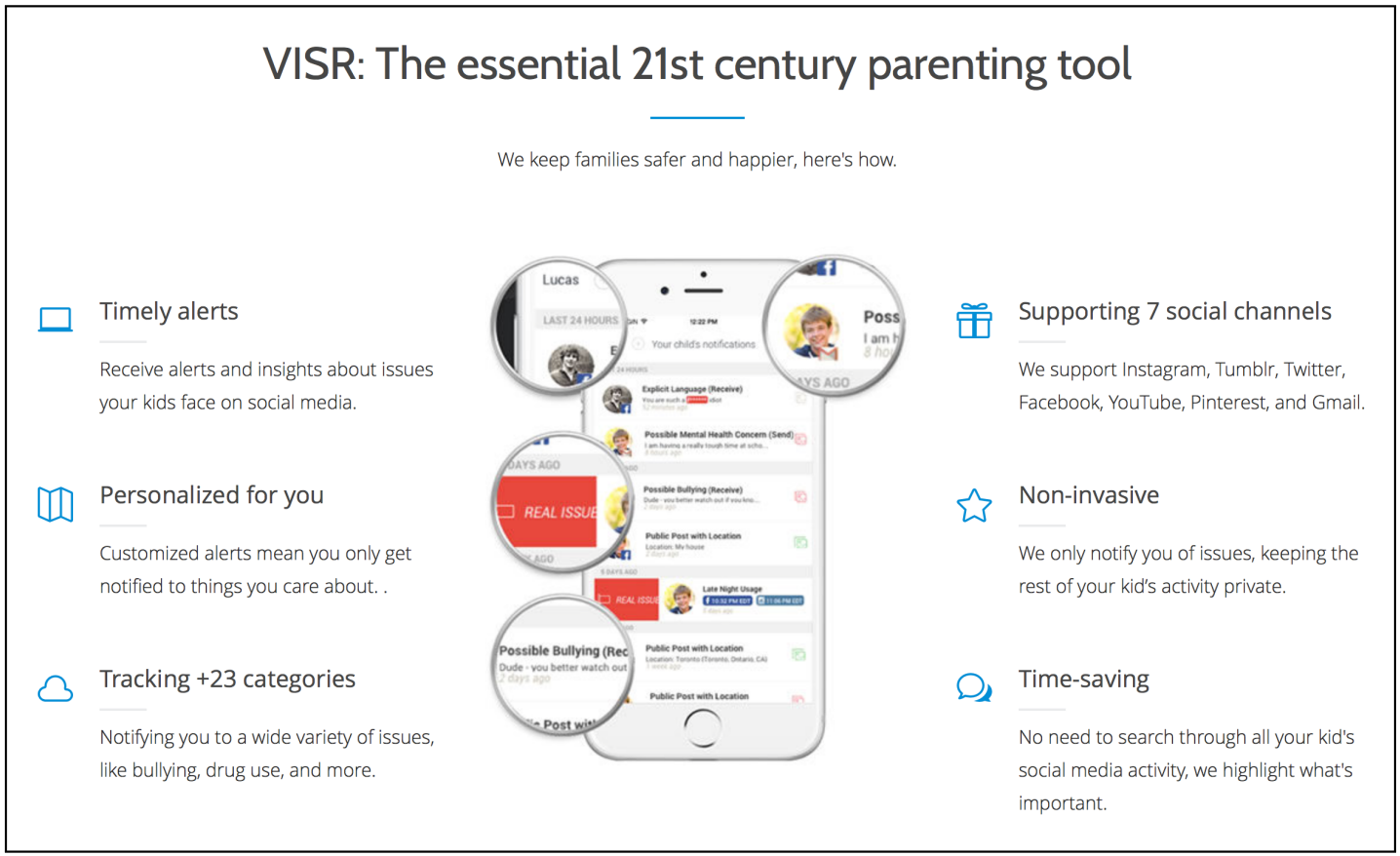
Lowered injury rates by 23%

Swimming, Skating, Hockey, Badminton, Wrestling

VISR

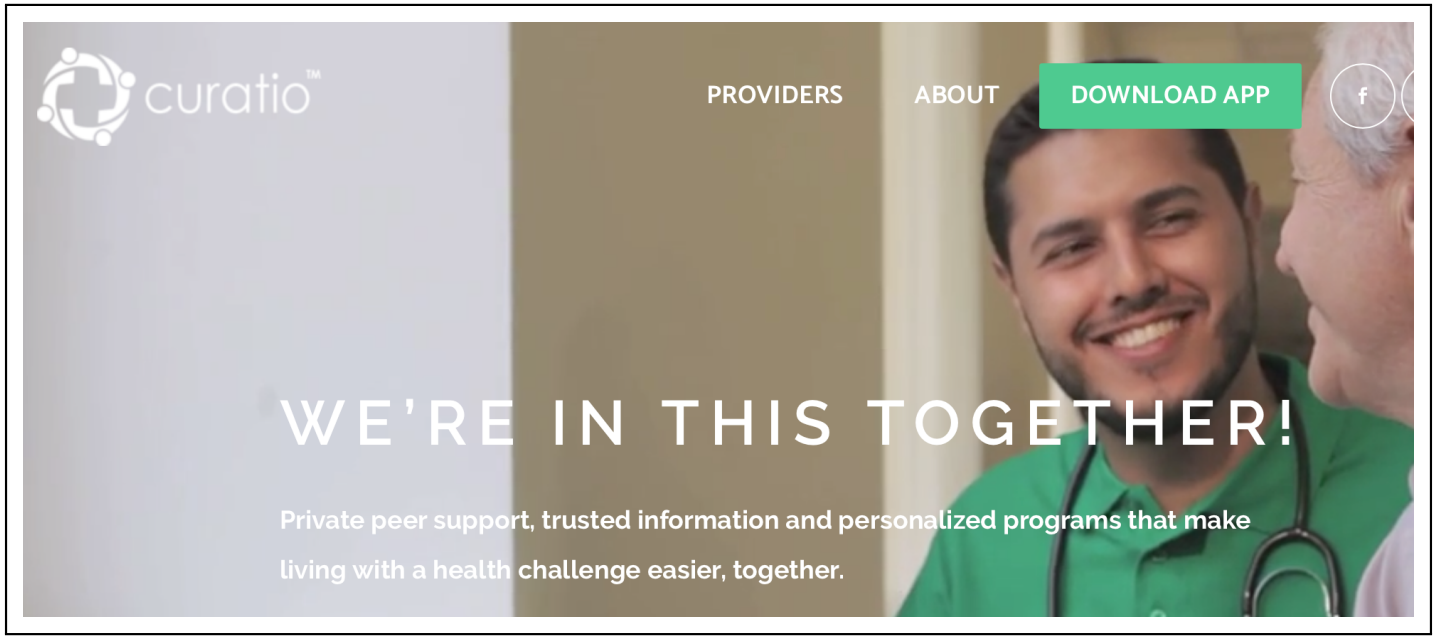
Better mental and emotional health via social media data mining

"On a mission to help families better navigate technology, by notifying parents about safety and wellness issues their kids face on social media"



http://www.visr.co

curatio — curatio.me



https://curatio.me

Curatio connects patients to support

MOBILE HEALTH PLATFORM READY TO ROLL AFTER WINNING OFFERS FROM FOUR DRAGONS

Globe & Mail Report on Business Feb 2017

Lifestyle

Nutrition
Exercise
Sleep

Smoking
Alcohol
Drugs

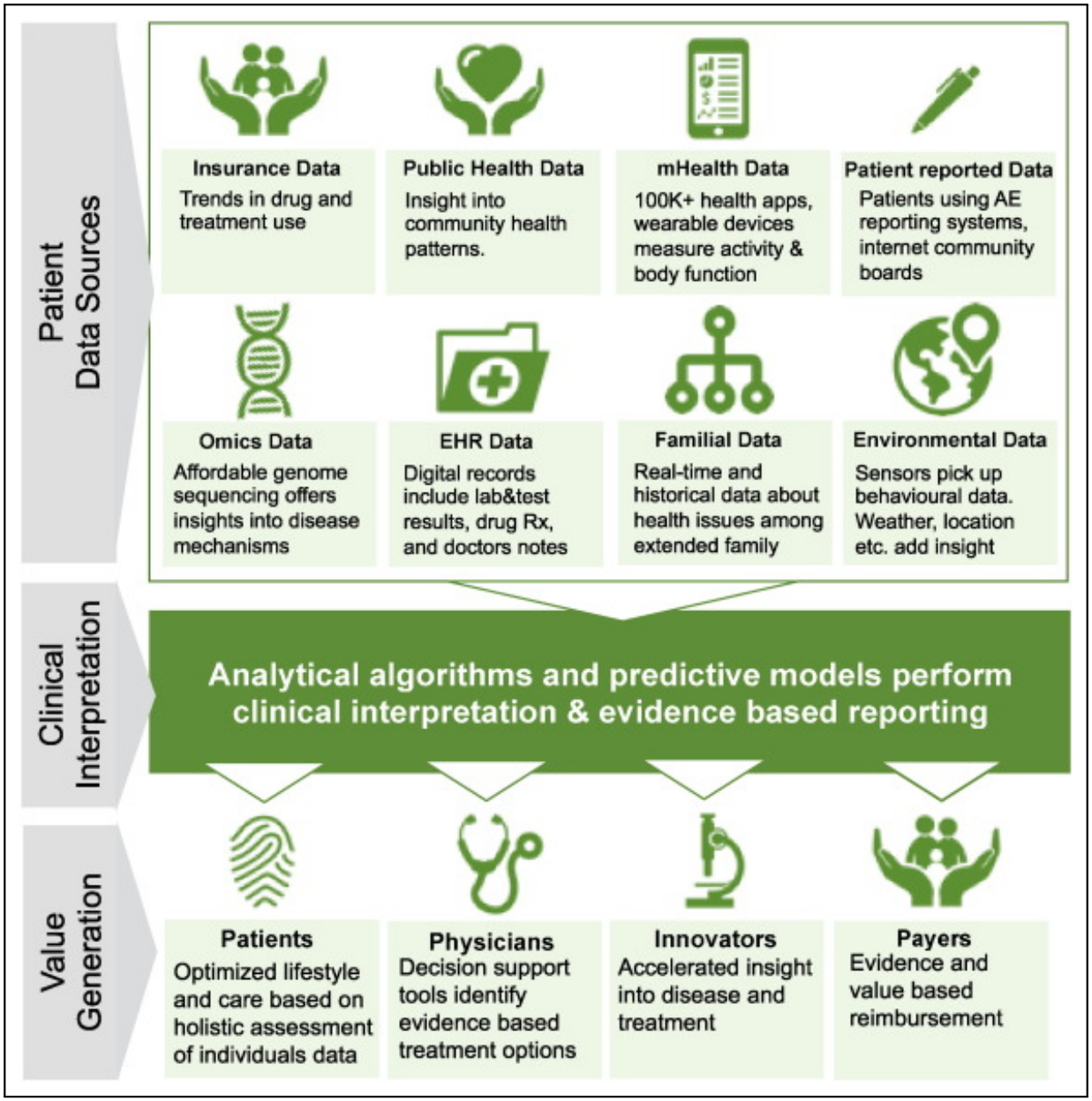
Fast Cars
Guns

Environment

Air Contamination
Land Contamination
Water Contamination

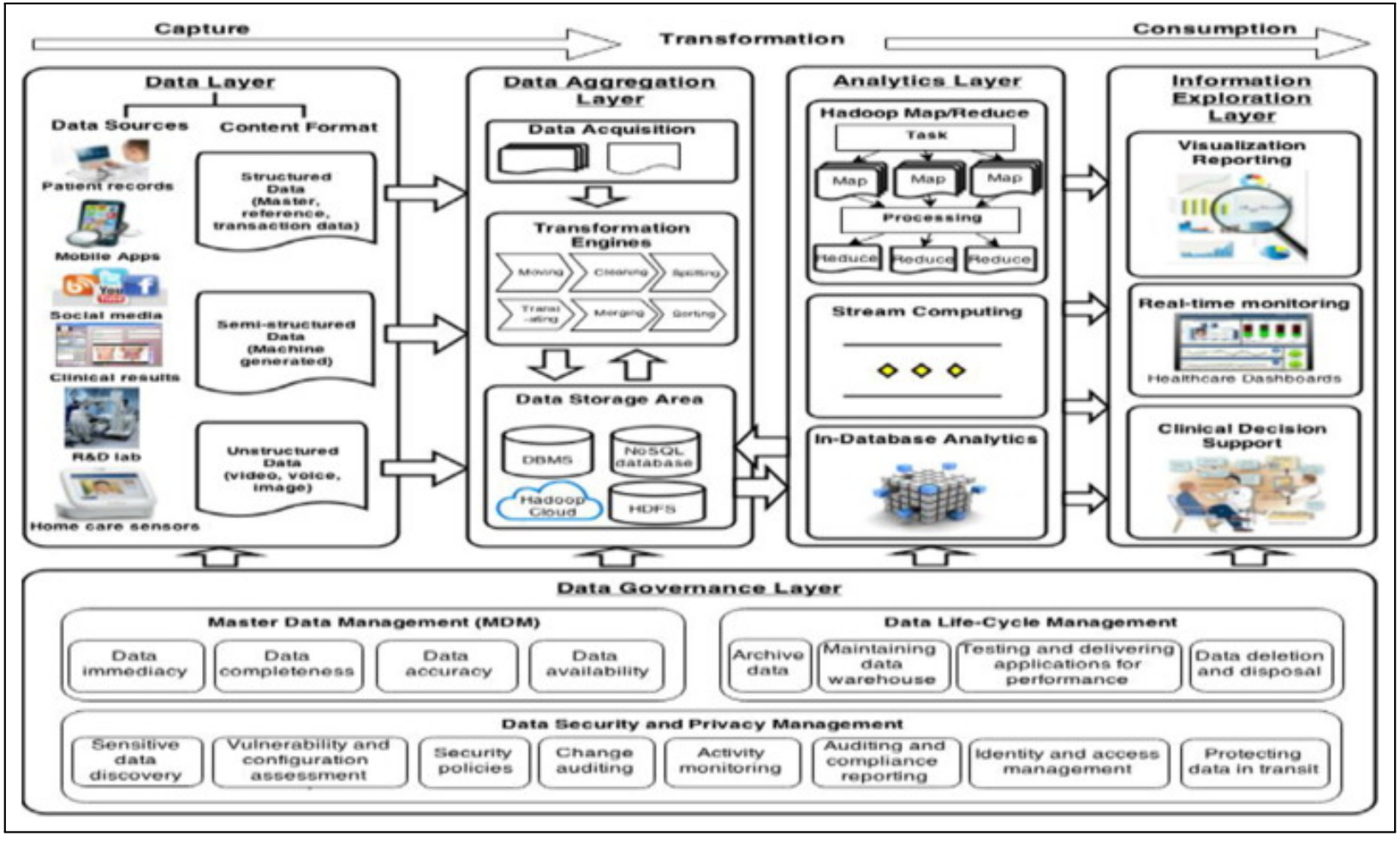
Radiation
Noise
Sunlight

The Data Perspective



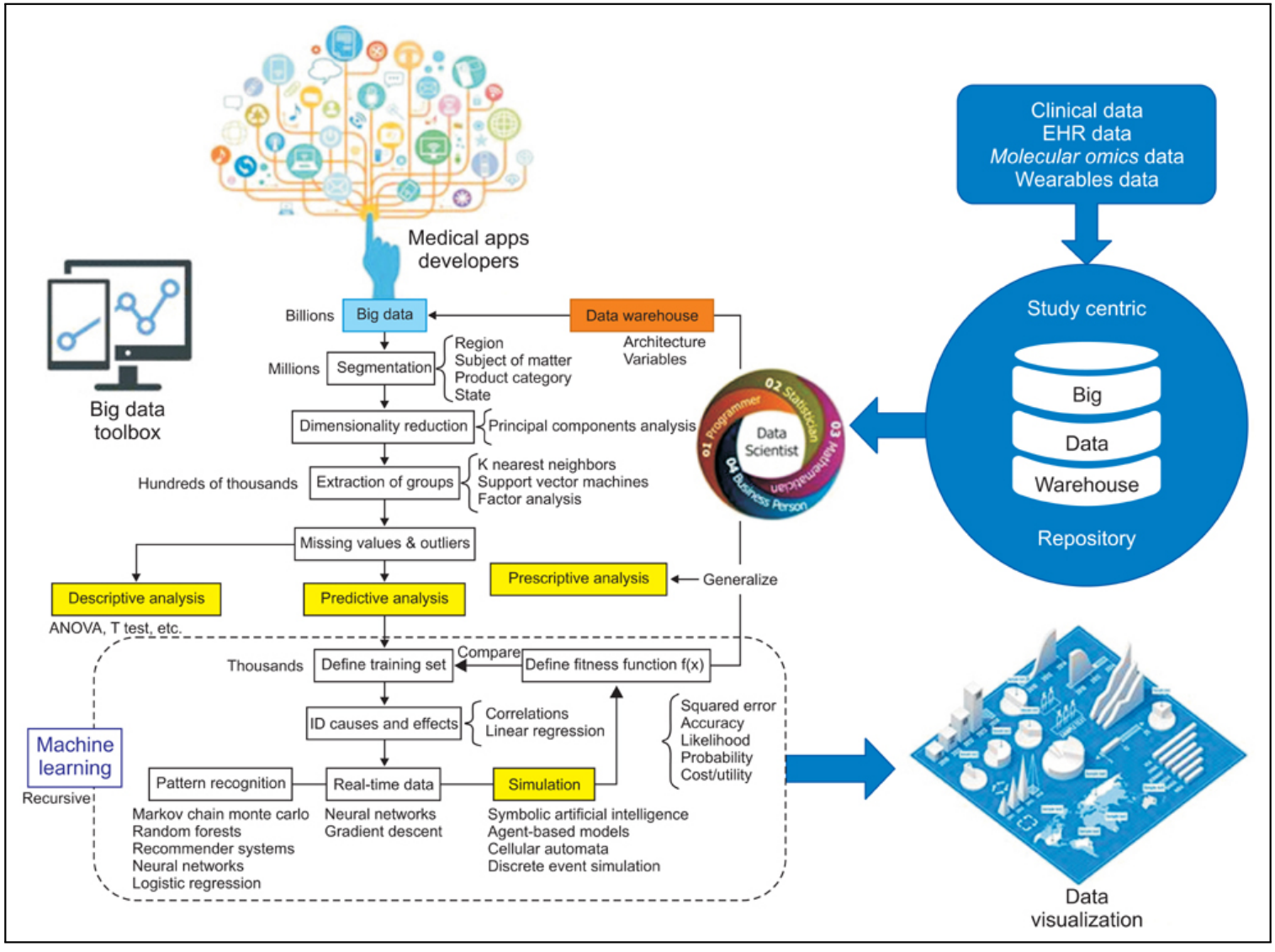
Taglang et al (Gynecologic Oncology, 2016)

The IT Perspective



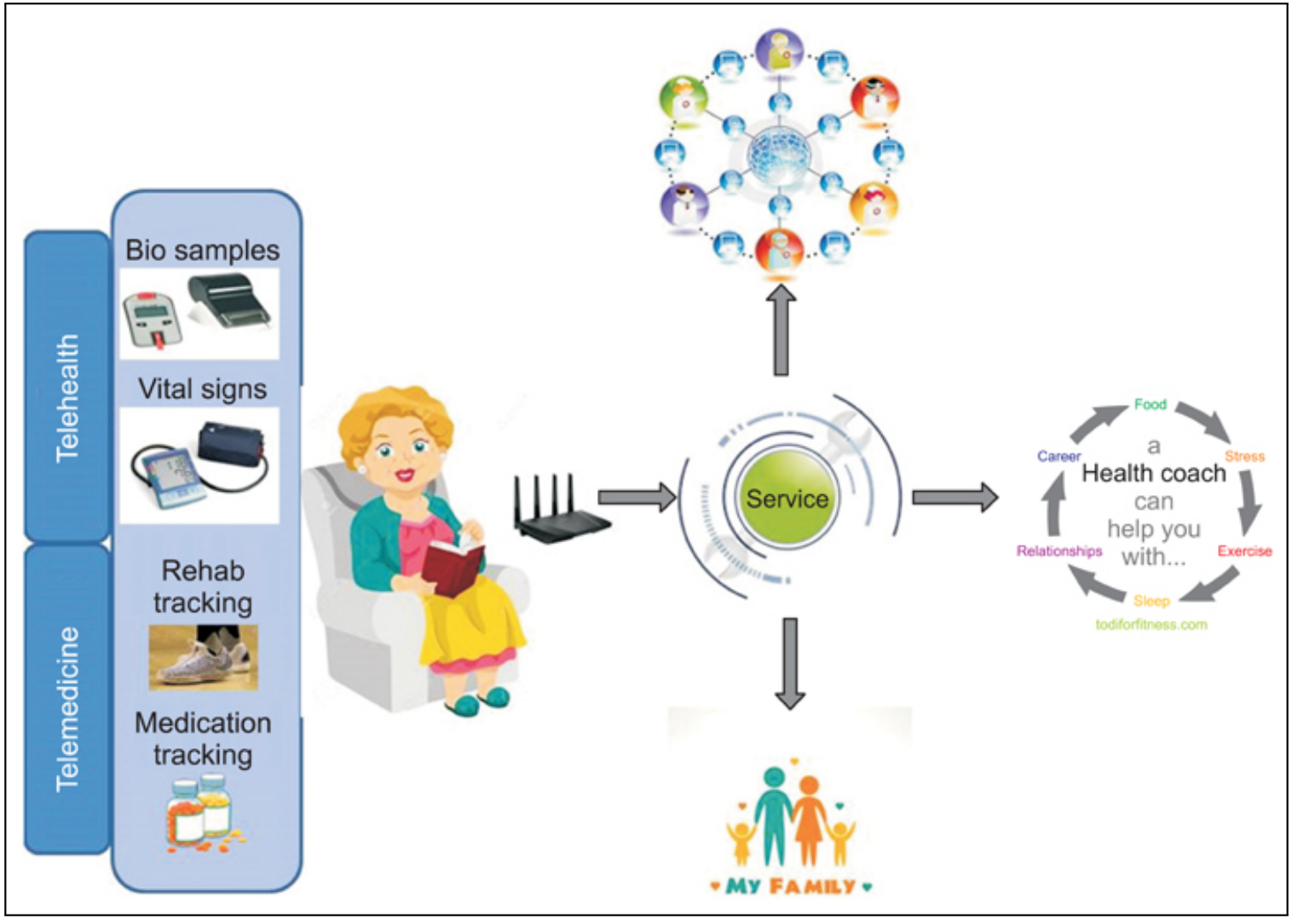
Wang et al (Tech. Forecasting & Social Change, 2016)

The Big Data Analytics Perspective



Dimitrov (Health Informatics Research, 2016)

The Elderly or Remote Patient Perspective



Dimitrov (Health Informatics Research, 2016)

Whitepaper of Working Group — March 2018

This white paper foresees a future with real-time predictive analytics delivered at all points of care by providers — both real and virtual — to better improve patient diagnoses, outcomes, quality of care, management of resources, and scientific discoveries.

In a very short period of time, cloud technology will be universal, with each patient having a holistic digital representation of their entire health history.

Interactions with the health data ecosystem will be at patients' fingertips, and the fingertips of healthcare providers.

By this point, open data and advanced technological infrastructure should be prominent in allowing large-scale analytics and research initiatives to take place, ...
... and for innovations to flourish.

Recommendations of Working Group

Data Access is Key!

Open data by default will be the future.
Data secure in the cloud and on personal smartcards
Privacy and sharing of data is in control of patient!
Anonymised data using *"id servers"* for data integration
Future is linked open data using RDF, ontologies, semantic web

Incubators need Technology + Business + Healthcare

See example at St Michaels Hospital and MARS in Toronto.
Physicians must buy-in ... or ... it's a no-go!

Scale Out beyond Incubators

Need community engagement ...

- for acceptance, use, and data sharing
- to establish agreed terminology for data sharing
- to reap benefits of innovation from Big Data

Closing Remarks

Big Data Important

Great space for innovation!
Great space for SME!
Great space to leverage IT, CS, and people skills!

Big Data (*Extremely*) Important in Health and Life Sciences

Working Group on Predictive Health

My thanks to ...

Eugene Wen, MD, PhD (Manulife), CDO, Chair
Greg Dashwood (Microsoft Canada), Technology, Co-Chair
Abidin Ashok (Ryerson), Coordinator
Ayse Bener, PhD (Ryerson), CS Academic
Greg Butler, PhD (Concordia), CSE Academic
Linda Koechli (Ryerson) Business Academic
Brigid Elmy (Ryerson) Business Academic
Muhammad Mamdani, PharmD (St Michael's Hospital) Director, Healthcare Analytics Research and Training
Laura Morin (Industry Canada)
David Hume (BC Govt) Citizen Engagement
Greg Horne (SAS)
Mike Miller (Ernst & Young)
Young Lee (Deloitte)
Daniel Zikovitz (GE Healthcare)
Jas Klotia (GE Digital)
Nicholas Yee (Manulife)
Paresh Yadav (PHEMI) IT
Daniel Lewis (RightBlueLabs) SME CTO — prevent sports training injury
Ronen Benin (RightBlueLabs) SME CEO
James Schuback (RightBlueLabs) SME Data scientist
Meir Dick (VISR) SME — monitor social media
Robert Reichmann (VISR) SME CEO
Gabriel Musso, PhD (BioSymetrics) SME — big data

A Starting Point for Data-Driven Healthcare

Big Data Analytics in Health White Paper,
Canada Health Infoway
`http://www.infoway-inforoute.ca/`

Thank You!

Questions Please?

References

Infoway Health Canada: Big Data Analytics in Health White Paper, May 2013.
Infoway Health Canada: Clinical Analytics in Primary Care White Paper, February 2016.

McKinsey Global Institute: Big data: The next frontier for innovation, competition, and productivity, June 2011.
McKinsey: The big-data revolution in US health care: Accelerating value and innovation January 2013.

The Catalyst towards an Ontario Health Innovation Strategy, Ontario Health Innovation Council (ohic.ca), December 2014.

R. Fang et al, Computational Health Informatics in the Big Data Age: A Survey, ACM Computing Survey July 2016 http://dl.acm.org/citation.cfm?id=2932707

Y. Wang et al, Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations, Technological Forecasting and Social Change Available online 26 February 2016 http://www.sciencedirect.com/science/article/pii/S0040162516000500

D.V. Dimitrov, Medical Internet of Things and Big Data in Healthcare, Healthc Inform Res. 2016 Jul;22(3):156-163. English. http://synapse.koreamed.org/search.php?where=aview&id=10.4258/hir.2016.22.3.156&code=1088HIR&vmode=FULL

C. Auffray et al, Making sense of big data in health research: Towards an EU action plan, Genome Medicine June 2016 http://genomemedicine.biomedcentral.com/articles/10.1186/s13073-016-0323-y

Guillaume Taglang, David B. Jackson, Use of "big data" in drug discovery and clinical trials, Gynecologic Oncology 141 (2016) 17–23. http://www.sciencedirect.com/science/article/pii/S0090825816300464

Michael KK Leung, Andrew Delong, Babak Alipanahi, Brendan J Frey, Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets, Proceedings of the IEEE 104:1 (2016) 176–197. http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7347331

Privacy and Security

"**Privacy** *refers to an individuals right to control the collection, use, and disclosure of his/her personal health information (PHI) and/or personal information (PI) in a manner that allows health care providers to do their work.*

Security *is about ensuring the information gets to the right person in a secure manner."*

Ontario's Ehealth Blueprint `http://www.ehealthblueprint.com`

Privacy by Design 2009

Seven Foundational Principles

1) being proactive not reactive;
2) having privacy as the default setting;
3) having privacy embedded into design;
4) avoiding the pretence of false dichotomies,
   such as privacy vs. security;
5) providing full life-cycle management of data;
6) ensuring visibility and transparency of data; and
7) being user-centric

Prof. Ann Cavoukian, formerly Information and Privacy Commissioner of Ontario; now Ryerson University. `http://www.privacybydesign.ca`

Canada's Personal Information Protection and Electronic Documents Act (PIPEDA) 2000

Ten Privacy Principles

Accountability: An organization is responsible for personal information under its control and shall designate an individual or individuals who are accountable for the organization's compliance with the following principles.
Identifying Purposes: The purposes for which personal information is collected shall be identified by the organization at or before the time the information is collected.
Consent: The knowledge and consent of the individual are required for the collection, use or disclosure of personal information, except when inappropriate.
Limiting Collection: The collection of personal information shall be limited to that which is necessary for the purposes identified by the organization. Information shall be collected by fair and lawful means.
Limiting Use, Disclosure, and Retention: Personal information shall not be used or disclosed for purposes other than those for which it was collected, except with the consent of the individual or as required by the law. Personal information shall be retained only as long as necessary for fulfilment of those purposes.
Accuracy: Personal information shall be as accurate, complete, and up-to-date as is necessary for the purposes for which it is to be used.
Safeguards: Personal information shall be protected by security safeguards appropriate to the sensitivity of the information.
Openness: An organization shall make readily available to individuals specific information about its policies and practices relating to the management of personal information.
Individual Access: Upon request, an individual shall be informed of the existence, use and disclosure of his or her personal information and shall be given access to that information. An individual shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate.
Challenging Compliance: An individual shall be able to address a challenge concerning compliance with the above principles to the designated individual or individuals for the organization's compliance.

`https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/`

`the-personal-information-protection-and-electronic-documents-act-pipeda/p_principle/`

Data Sharing Policy

Global Alliance for Genomics and Health

Framework for Responsible Sharing of Genomic and Health-Related Data (`genomicsandhealth.org`)

Regulatory and Ethics Working Group

Leadership

Bartha Maria Knoppers,
McGill University

Kazuto Kato,
Osaka University

Focuses on ethics and the legal and social implications of the Global Alliance, including harmonizing policies and standards, and developing forward-looking consent, privacy procedures, and best-practices in data governance and transparency.

Global Alliance
for Genomics & Health