

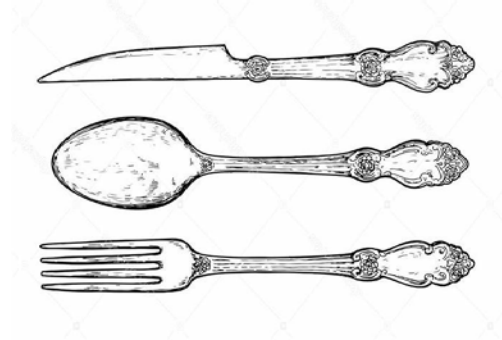
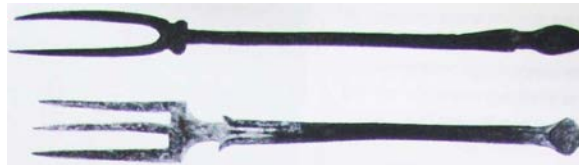
# **Ethics and Regulatory Policy for Utilizing Health Data toward Open Science in Japan**

**Tamami Fukushi, Ph.D.**  
Deputy Manager  
Group for Research and Analysis  
Department of Research Infrastructure  
Japan Agency for Medical Research and Development

# Today's agenda

- Health data and co-evolution of “Open Science” and “Data Science” in 21<sup>st</sup> century
- Ethical, Legal and Social Issues (ELSI) for utilization of health data in the era of open science
- Related Activity in AMED

## Co-evolution of “Open Science” and “Data Science”



## Academic Context



### National Academy of Science and Medicine

*Research funders have seen the **value of openly sharing the results of the research** that they have supported, not just in the form of publications, but **also in the form of the data** that have been produced in the course of the investigation. They have begun to require that applicants prepare **data management plans** as part of their grant proposals.*



.....

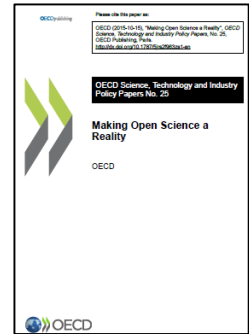
*To evaluate more fully the benefits and challenges of broadening access to the results of scientific research, described as “**open science**,” the National Academies of Sciences, Engineering, and Medicine appointed an expert committee in March 2017.*

# Policy Context



## OECD

*Open science* – There is no formal definition of open science. In this report, the term refers to efforts by researchers, governments, research funding agencies or the scientific community itself to make the primary outputs of publicly funded research results – publications and the *research data* – publicly accessible in digital format with no or minimal restriction as a means for accelerating research; these efforts are in the interest of enhancing transparency and collaboration, and fostering innovation.

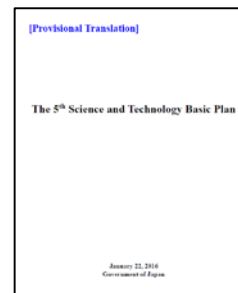


<https://www.oecd-ilibrary.org/docserver/5jrs2f963zs1-en.pdf?expires=1580461110&id=id&accname=guest&checksum=0FD196172D2DFDF43549ED0E627AEEDB>

## JAPAN

*Promotion of open science*

*Open science* is a concept that encompasses open access by making *research data* into *open data*. As open access is promoted, research outcomes will become available to all manner of users, including both those in the academic community, the business world and private citizens. This will result in the accelerated production of knowledge through new collaborations between researchers in different organizations, as well as between specialized fields and across borders, thereby enabling the production of new sources of value.



<https://www8.cao.go.jp/cstp/english/basic/5thbasicplan.pdf>

# Time lines



**Table 1. Abbreviated Timeline Highlighting Recent Transformations in Scientific Practice**

Year	Event
2003	Open Access publishing is boosted with <i>PLOS Biology</i> launch
2004	Google's Dean & Ghemawat publish MapReduce in <i>OSDI</i>
2005	Amazon launches Mechanical Turk
2005	Reproducibility goes critical with Ioannidis's "Why Most Published Research Findings Are False" in <i>PLOS Medicine</i>
2006	Facebook opens account creation to the general public
2006	<i>PLOS One</i> launches with a central goal of facilitating post-publication peer review
2006	Amazon launches Elastic Compute Cloud (EC2) as part of Amazon Web Services (AWS)
2006	Netflix announces the Netflix Prize competition
2007	Twitter becomes independent and "debuts" at South by Southwest (SXSW)
2007	iPhone launch sparks the "smartphone revolution"
2007	iPython changes scientific computing when introduced in Pérez & Granger, <i>Computing in Science and Engineering</i>
2008	GitHub launch makes scientific version control easier
2008	DJ Patil (LinkedIn) and Jeff Hammerbacher (Facebook) coin the phrase "Data Science" to describe their jobs

Table 1 from Bradley Voytek, 2017 <https://doi.org/10.1016/j.neuron.2017.11.015>

2013: G8

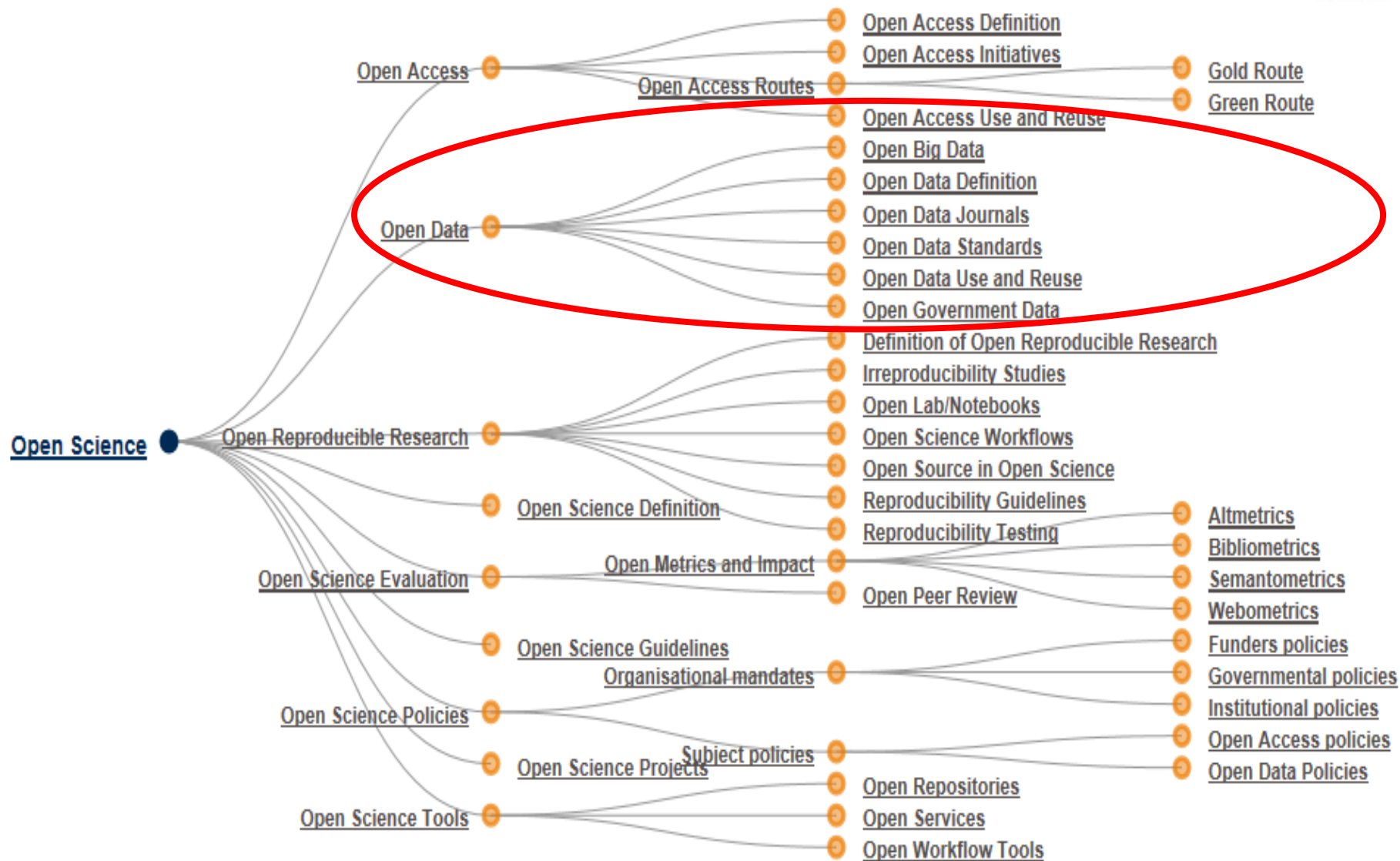
2016: Open Science WG in G7

**FAIR Principle**

G20 supports FAIR Principle

2019: **CARE Principle**

# Possible components included in “Open Science”



# FAIR principles



## Findable

- F1. (Meta)data are assigned a globally unique and persistent identifier
- F2. Data are described with rich metadata (defined by R1 below)
- F3. Metadata clearly and explicitly include the identifier of the data they describe
- F4. (Meta)data are registered or indexed in a searchable resource

## Accessible

- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol
  - A1.1 The protocol is open, free, and universally implementable
  - A1.2 The protocol allows for an authentication and authorisation procedure, where necessary
- A2. Metadata are accessible, even when the data are no longer available

<https://www.go-fair.org/fair-principles/>



## **Interoperable**

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (Meta)data use vocabularies that follow FAIR principles
- I3. (Meta)data include qualified references to other (meta)data

## **Reusable**

- R1. Meta(data) are richly described with a plurality of accurate and relevant attributes
  - R1.1. (Meta)data are released with a clear and accessible data usage license
  - R1.2. (Meta)data are associated with detailed provenance
  - R1.3. (Meta)data meet domain-relevant community standards

# CARE principle for indigenous data



<https://www.gida-global.org/care>

# Reaction to “Open Science” in Japan

# Reaction to “Open Science” by Governmental Bodies



## Cabinet Office

2015

launched Discussion group on open science based on its international trend in Council for Science, Technology and Innovation (CSTI).

2018

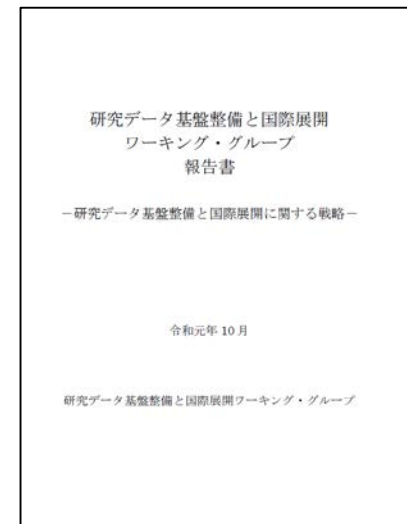
Working group on research data infrastructure and international evolvement was launched in the discussion group.

2019

The working group drafted guideline for storage and publishing research data and developing and managing data repository.



Public-funded research data will be opened based on data policy and data management plan.



<https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/index.html>

### “Contract Guidelines on Utilization of AI and Data”.

#### Outline of Data Section

##### 1.Objectives

Help businesses reduce the costs of implementing a data utilization contract.

##### 2.Targeted readers

A wide range of people involved in concluding contracts, including business operators, associated business and management units, and parties developing systems for data distribution and utilization.

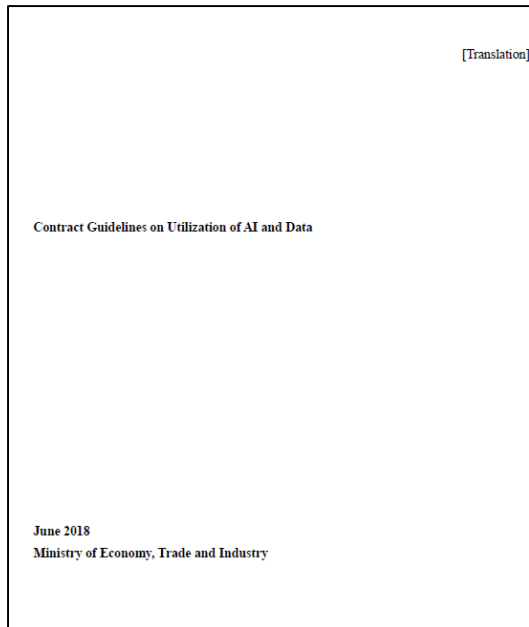
##### 3.Outline

Categorizes data utilization and explains structures, main legal issues and proper contract preparation. processes for each contract type.

- data provision
- data creation
- data sharing (platform type)

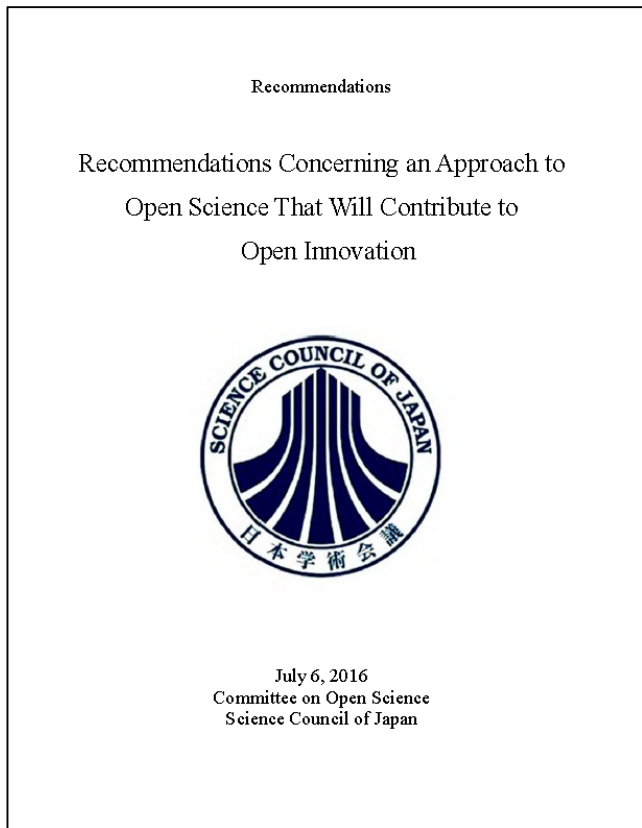
Modified from the contents on

[https://www.meti.go.jp/english/press/2019/0404\\_001.html](https://www.meti.go.jp/english/press/2019/0404_001.html)



2016

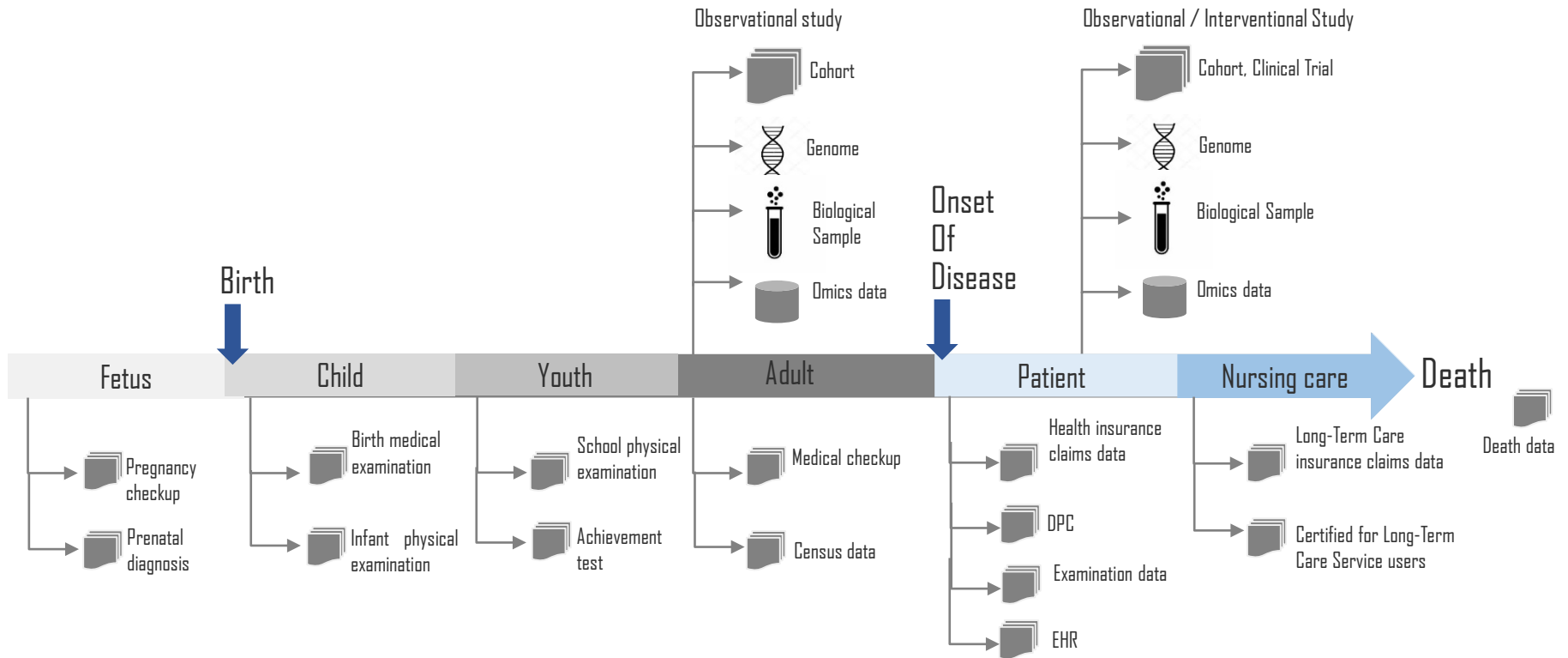
Science Council of Japan published recommendations.



- (1) Establishment of a research data infrastructure that allows the management of and ensures the openness of interdisciplinary research data
- (2) Establishment of data strategies by research communities
- (3) Career design for data producers and data curators

# Health Data Utilization in Japan toward “Open Science”?

# Health data in human life course contains diverse innovation potential.



In Japan, health data is managed by multi, disconnected system rather than consistent one.



**May, 2018**

“Act on Anonymized Medical Data That Are Meant to Contribute to Research and Development in the Medical Field” was enacted.

## Back Ground

- Japanese research and development community needed to re-organize medical data utilization framework in accordance with Amended Act on the Protection of Personal Information.

The Act enable us to

- Utilize both input and outcome of medical data
- Aggregation and connection of distributed medical data

Certified medical data anonymization agencies can anonymize medical data provided by medical institution, and they also provide the anonymized data to end users.

# Reaction by Governmental body



ひと、くらし、みらいのために  
厚生労働省  
Ministry of Health, Labour and Welfare

政策について  
厚生労働省について 統計情報・白書 所管の法令等 申請・募集・情報公開

## 医療・介護データ等の解析基盤に関する有識者会議

回数	開催日	議題等	議事録/議事要旨	資料等	開催案内
第10回	2019年11月15日 (令和元年11月15日)	1 医療保険制度の適正かつ効果的な運営を図るための健康保険法等の一部を改正する法律の施行について(案)(NDB・介護DB・DPCDBの連結解析機) 2 その他	議事録	資料	開催案内
-	-	「医療・介護データ等の解析基盤に関する有識者会議」報告書	-	資料	-
第9回	2018年11月15日 (平成30年11月15日)	レセプト情報・特定健診等情報データベース(NDB)及び介護保険総合データベース(介護DB)の連結について	議事録	資料	開催案内
第8回	2018年10月25日	レセプト情報・特定健診等	議事録	資料	開催案内

関連リンク

May, 2018

Ministry of Health, Labour and Welfare (MHLW) launched Expert meeting on analysis base for medical and nursing data.

[https://www.mhlw.go.jp/stf/shingi/other-hoken\\_553056.html](https://www.mhlw.go.jp/stf/shingi/other-hoken_553056.html)

July, 2019

MHLW launched Expert meeting on the use of assured number of health insurance to promote consolidation of medical information

ひと、くらし、みらいのために  
厚生労働省  
Ministry of Health, Labour and Welfare

政策について  
厚生労働省について 統計情報・白書 所管の法令等 申請・募集・情報公開

## 医療等情報の連結推進に向けた被保険者番号活用の仕組みに関する検討会

回数	開催日	議題等	議事録/議事要旨	資料等	開催案内
-	-	-	-	報告書	-
第3回	2019年9月24日 (令和元年9月24日)	被保険者番号履歴を活用した医療等情報の連結について	議事録 NEW 10月17日	資料	開催案内
第2回	2019年9月4日 (令和元年9月4日)	被保険者番号履歴を活用した医療等情報の連結について	議事録	資料	開催案内
第1回	2019年7月31日 (令和元年7月31日)	(1) 被保険者番号活用の仕組みを巡る状況 (2) 被保険者番号活用の仕組みの基本スキーム (3) 仕組みの具体化に当たっての論点	議事録	資料	開催案内

関連リンク

## 医療等情報の連結推進に向けた被保険者番号活用の仕組みに関する検討会

回数	開催日	議題等	議事録/議事要旨	資料等	開催案内
-	-	-	-	報告書	-
第3回	2019年9月24日 (令和元年9月24日)	被保険者番号履歴を活用した医療等情報の連結について	議事録 NEW 10月17日	資料	開催案内
第2回	2019年9月4日 (令和元年9月4日)	被保険者番号履歴を活用した医療等情報の連結について	議事録	資料	開催案内
第1回	2019年7月31日 (令和元年7月31日)	(1) 被保険者番号活用の仕組みを巡る状況 (2) 被保険者番号活用の仕組みの基本スキーム (3) 仕組みの具体化に当たっての論点	議事録	資料	開催案内

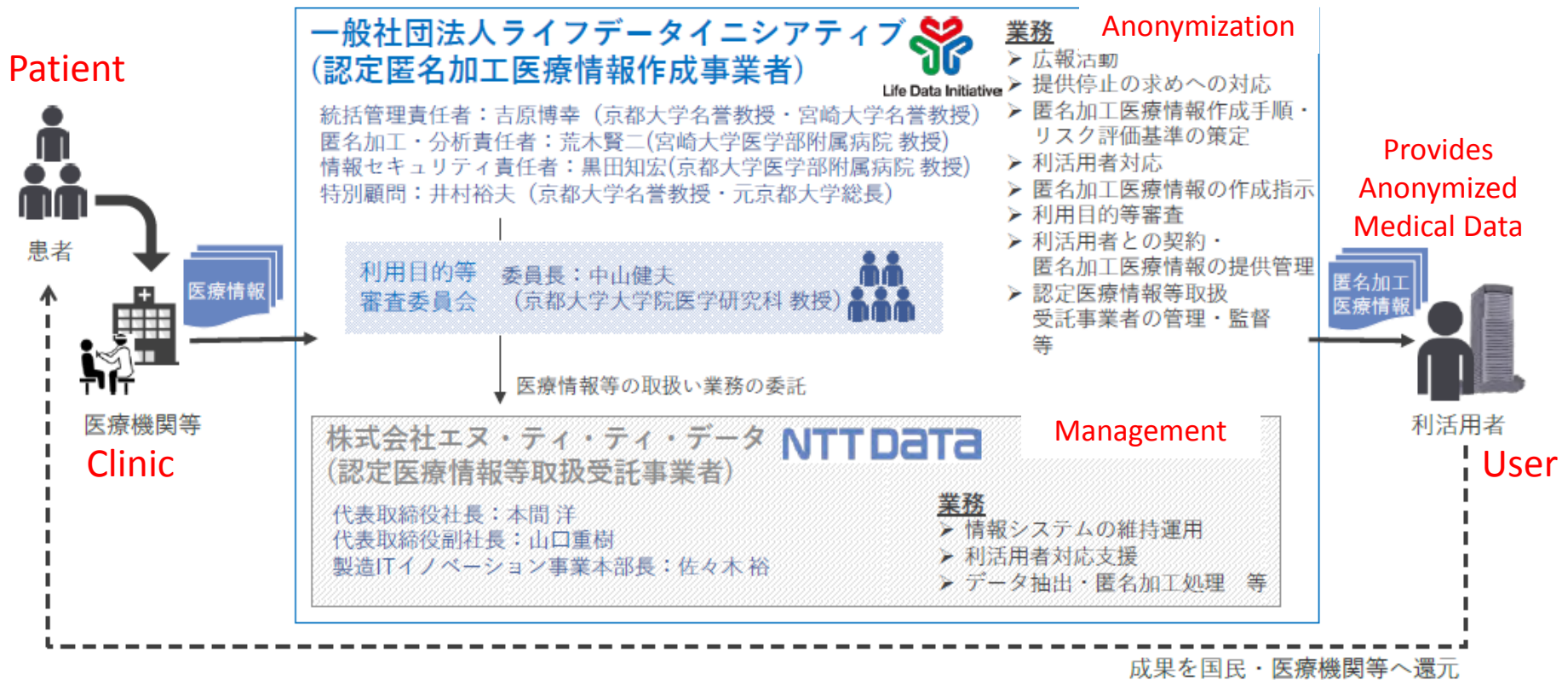
[https://www.mhlw.go.jp/stf/newpage\\_05744.html](https://www.mhlw.go.jp/stf/newpage_05744.html)

December, 2019

Two organizations were certified to anonymize and manage medical data, based on the “Act on Anonymized Medical Data That Are Meant to Contribute to Research and Development in the Medical Field”.



一般社団法人ライフデータイニシアティブ及び株式会社エヌ・ティ・ティ・データの事業実施体制



<https://www8.cao.go.jp/iry/nintei/nintei/jigyousha/pdf/besshi2.pdf>

# Reaction by Academia

November, 2019

Biomedical science-related working groups in Science Council of Japan published recommendation to progress establishment of infrastructure for sustainable life science data management.



- (1) Data sharing policy
- (2) Data base strategy in early stage of research project development
- (3) Integrated data base center with super computer
- (4) Human resource and education
- (5) Resource allocation for data utilization

<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-24-t279-1.pdf>

# Health Data-Related Activity in AMED



# Tohoku Medical Megabank Project



東北メディカル・メガバンク機構  
TOHOKU MEDICAL MEGABANK ORGANIZATION

## Community-Based Cohort :

Recruited either  
at municipal specific health checkups  
at Community Support Centers

Recruited participants: Total: 84,073  
(Miyagi : 52,212, Iwate : 31,861)

## BirThree Cohort :

Recruited at OB&GYN clinic or hospital

Recruited participants: 73,500 people

## Questioners (mail・Web)

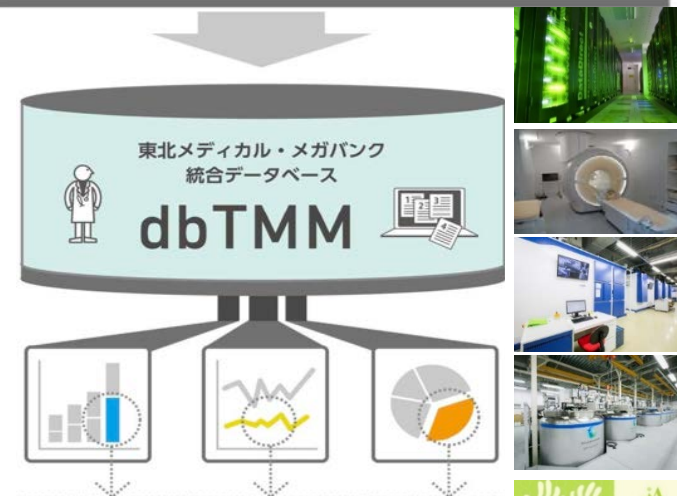
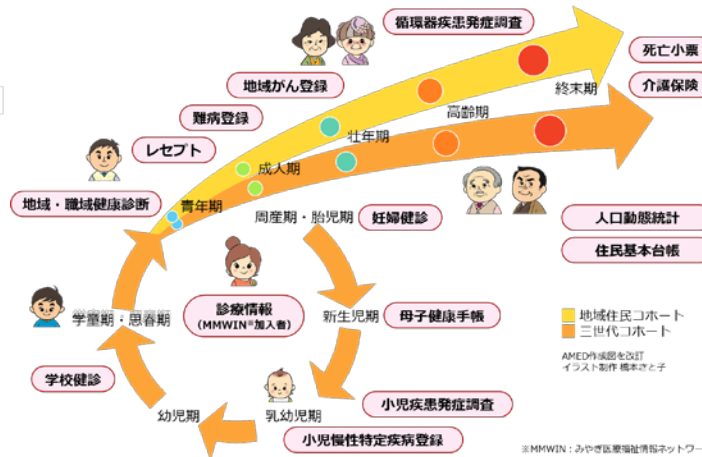
## Medical chart review

## Official data, disease registry

## Health exam (blood・Physiological exam, etc.)



## Collection of data over life course

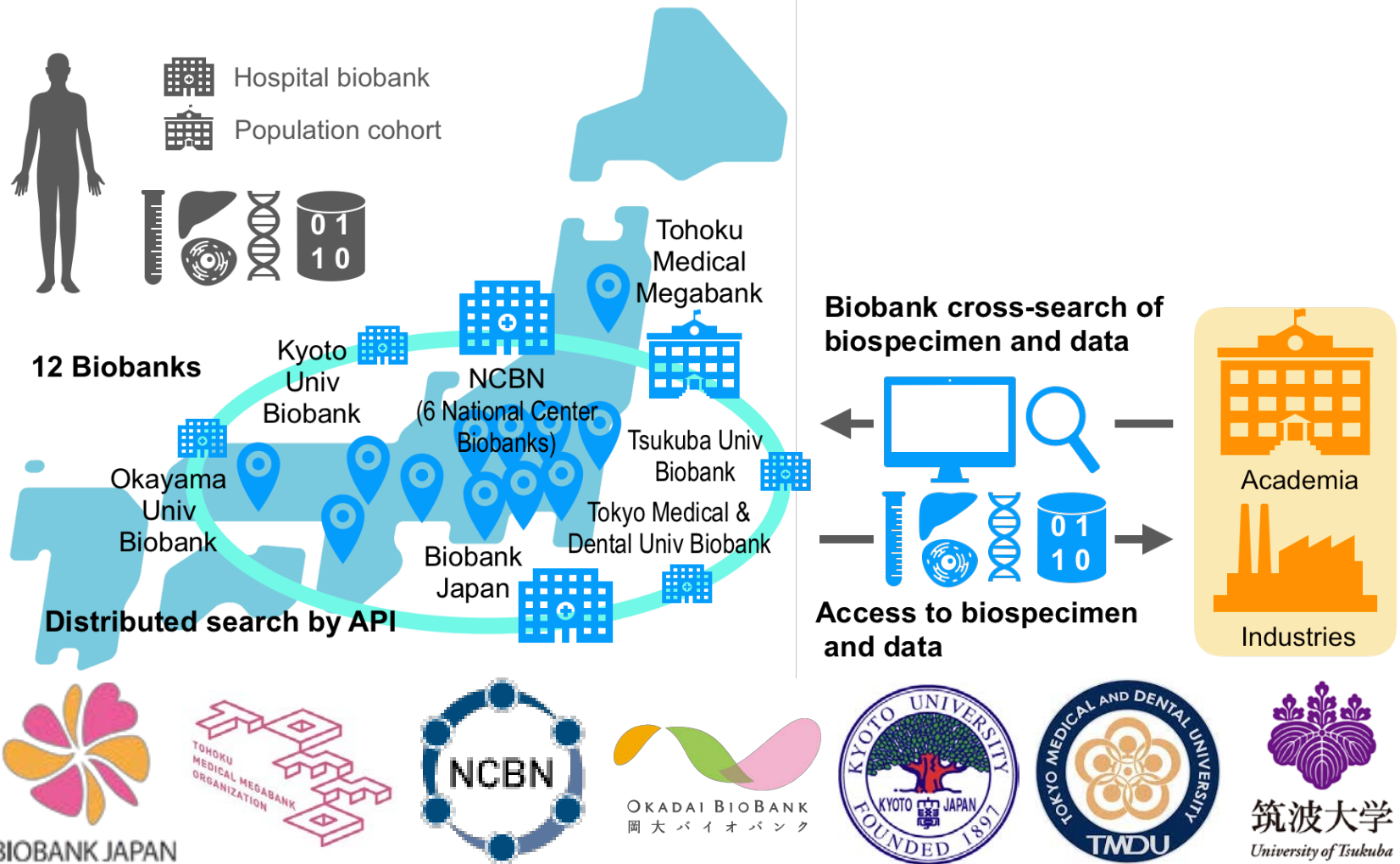


Personalized prevention, and genomic medicine

Provided by AMED Division of Biobank (2019)



# Research and development of biobank network in Japan



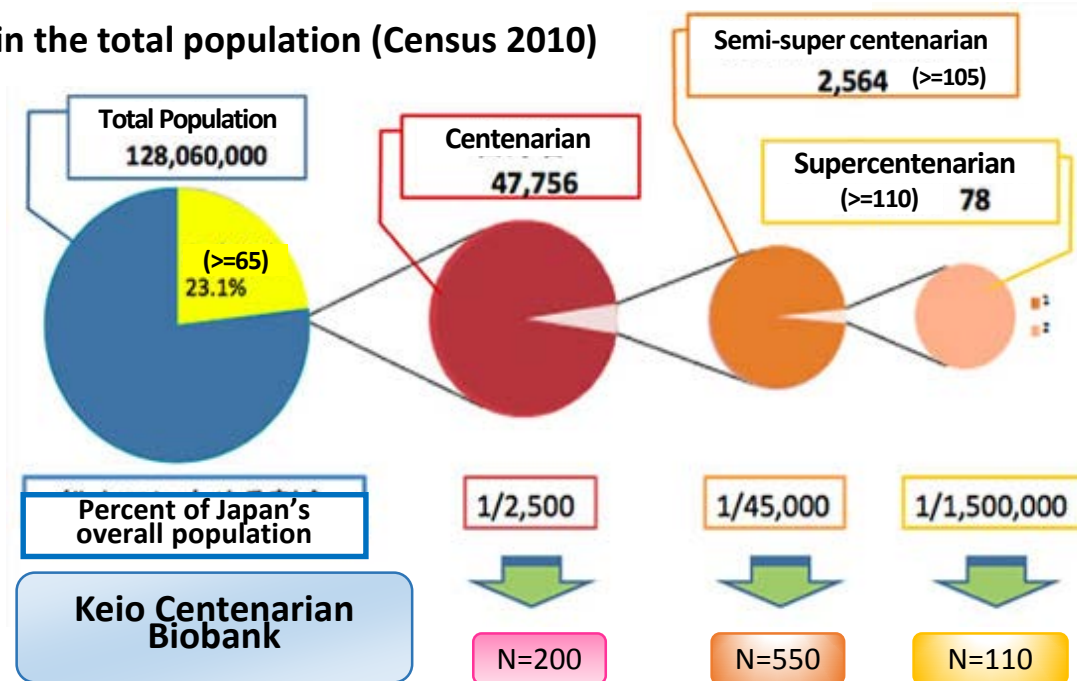
Provided by AMED Division of Biobank (2019)

# The Japanese Semi-Supercentenarian Study (JSS)



Keio University School of Medicine, Center for Supercentenarian Medical Research

## Percent of Centenarian and Supercentenarian in the total population (Census 2010)



## International Collaborative Researches

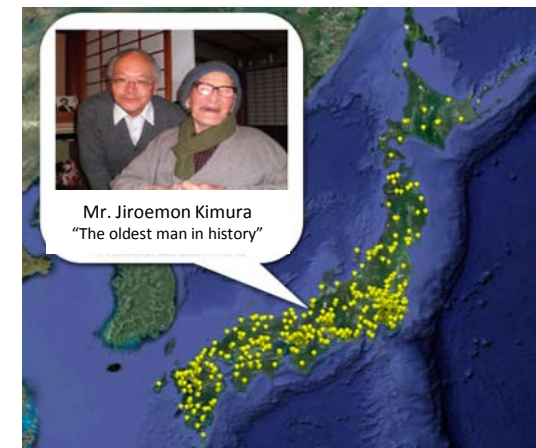
- 5 Country Oldest Old Project (5COOP)  
France, Switzerland, Denmark, Sweden, Japan



- Research with Newcastle University, UK



## Distribution map of Supercentenarian



## Key Outcomes from Cohort Studies

- Length of telomere are kept longer in centenarians.
- Chronic inflammation is related to life expectancy and Activities of Daily Living (ADL) in population  $\geq 100$  years old.
- Both the APOE  $\epsilon 4$  allele and education appear to be associated with cognitive function even in centenarians, but the interaction between the  $\epsilon 4$  allele and education might depend on gender.

Keio University School of Medicine, Center for Supercentenarian Medical Research, 2017

<http://www.hosp.keio.ac.jp/annai/shinryo/supercentenarian/> 24



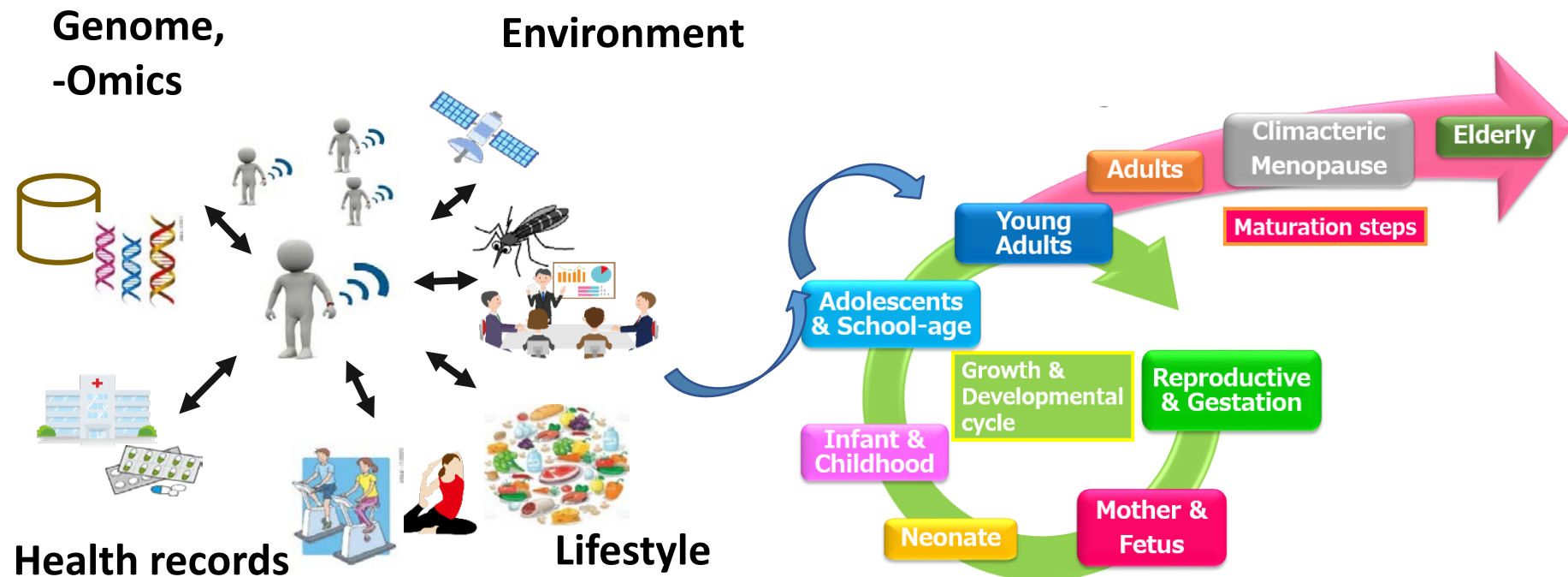
# Internal Discussion Group

## Team “Life Course Data for R&D”

### Goal of the discussion

Draft research & development strategic plan (to be supported by AMED) for

- Identifying the right intervention to the right population at the right timing
- Predicting risks and diseases throughout individual’s life-course
  - data linkage/sharing, multidisciplinary approach, and implementation
  - novel understanding of intervention point through a combined approach of genomic, lifestyle and environmental factors
- Developing effective interventions through a bio-psycho-social approach



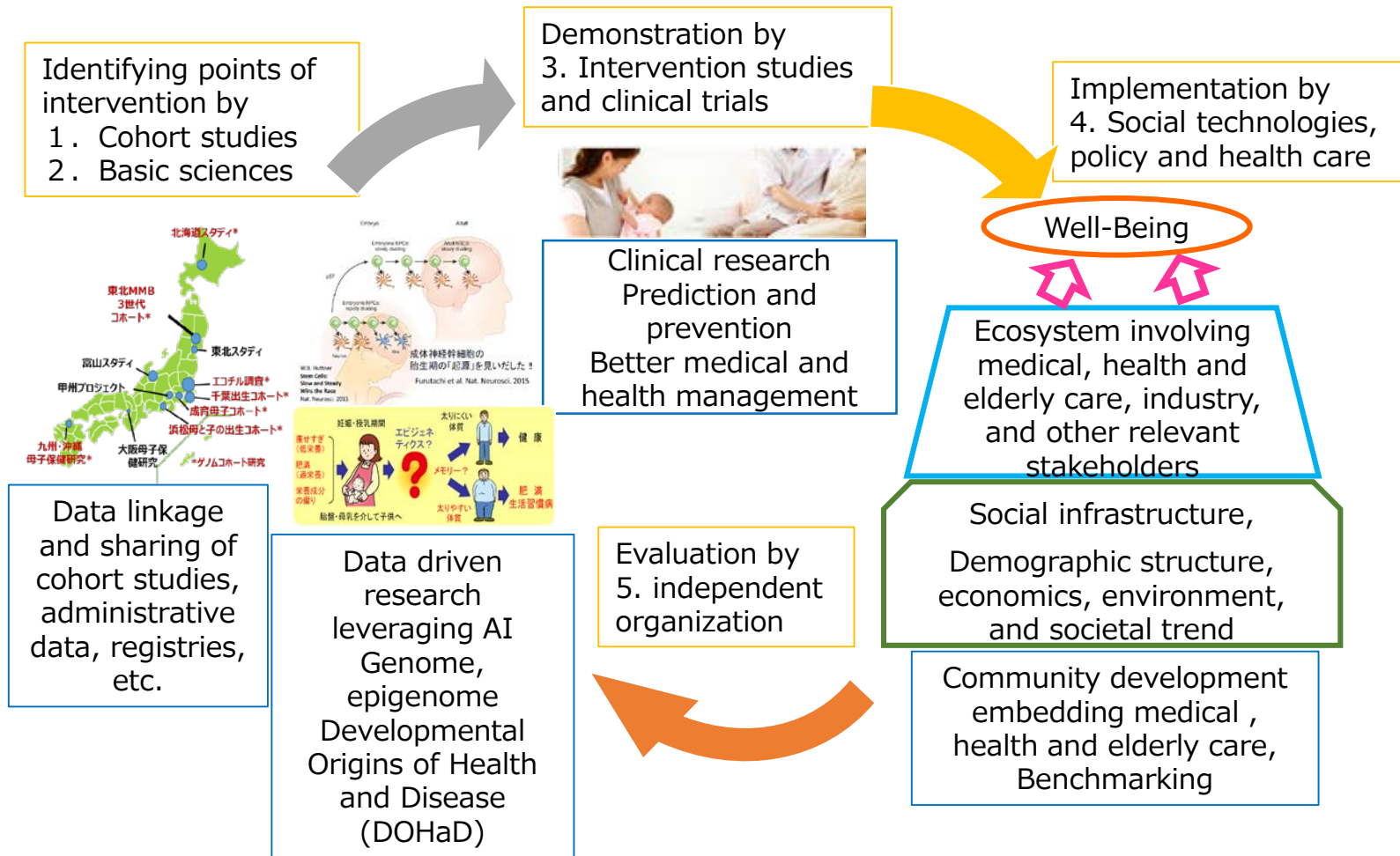
# Life Course Approach for Society of 21th Century

## Novel research strategy for health management and disease promotion



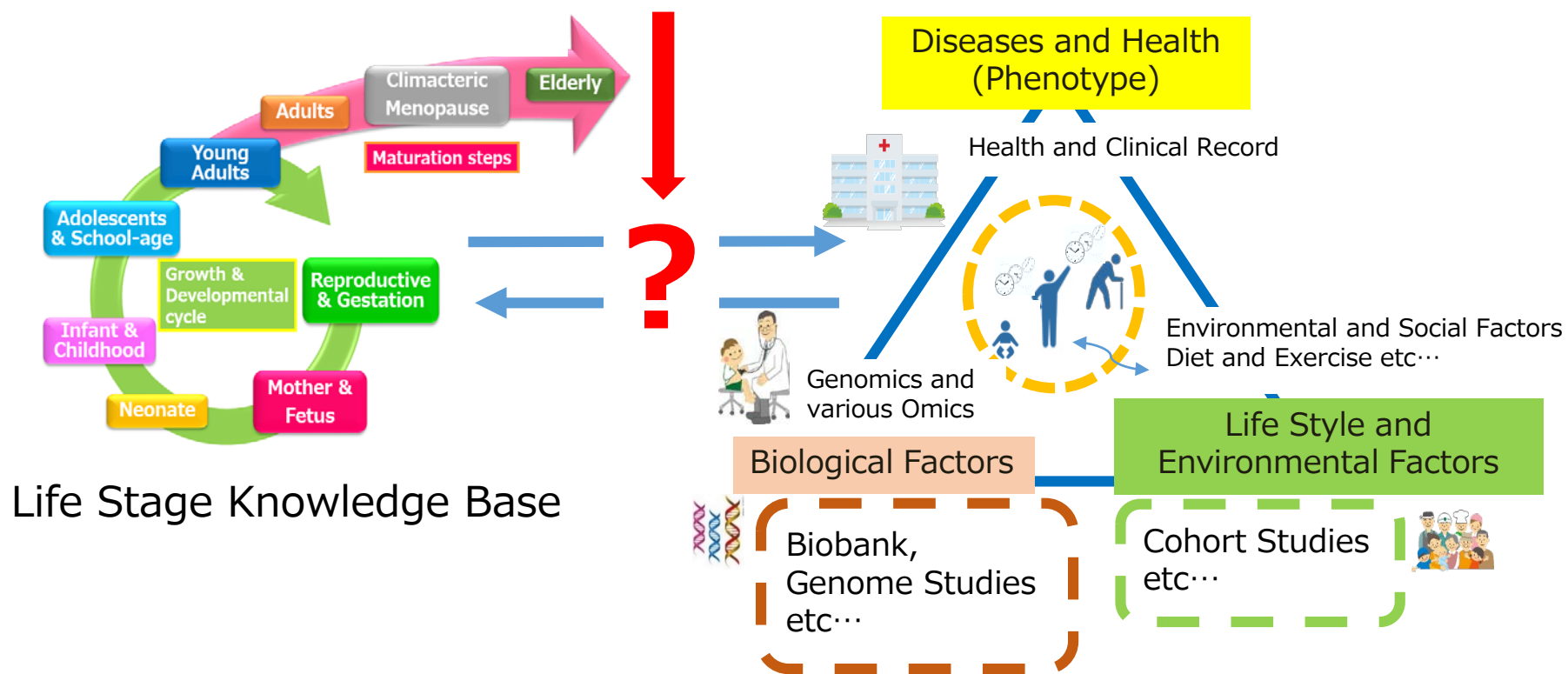
With bio-psycho-social approach, we will make impact on society by support of research and development based on scientific evidence

→ **Preparing for an unprecedented society to reach in 20-30 years.**



Better consent mechanism would be needed for keeping longitudinal study with ongoing health data management.

enables research participants to provide consent as his/her life changes



Life Stage Knowledge Base

# New Ethical Approach for Utilization of Health Data

# Ethical, Legal and Social Issues (ELSI) pointed to facilitate “Open Science” including health data



- Privacy
- Informed Consent
- De-identification
- Inequality – digital divide
- Research integrity

[https://www.ukdataservice.ac.uk/media/604711/big-data-and-data-sharing\\_ethical-issues.pdf](https://www.ukdataservice.ac.uk/media/604711/big-data-and-data-sharing_ethical-issues.pdf)



- Openness within the relevant legal and ethical constraints
- Informed consent
- Transparency
- Quality
- Care for any potential negative consequences
- Unambiguous distribution of responsibilities
- Balance the power across stakeholders
- Access to an ethics review body(ERB)

[https://www.oecd-ilibrary.org/science-and-technology/research-ethics-and-new-forms-of-data-for-social-and-economic-research\\_5jln7vnpxs32-en](https://www.oecd-ilibrary.org/science-and-technology/research-ethics-and-new-forms-of-data-for-social-and-economic-research_5jln7vnpxs32-en)

# Possible Solution -Dynamic Consent-



Dynamic consent is an emerging mechanism which enables study participants to provide consent and facilitates the ongoing management of clinical studies. It is a “personalized, online consent and communication platform”.

## Potential benefit

### For **Participants (including patients)**

- More appropriate, granular and flexible consent options
- Access to better study information,
- Opportunity to increase scientific and medical literacy
- Two-way communication between participants and researchers, building trust

### For **Researchers/Research organizations**

- Better electronic consent records
- Retention of participants in longitudinal studies
- Clearer data sharing frameworks for health information
- Working toward addressing ethical, legal & social issues relevant to clinical studies

#### ✕References

- Regulatory and Ethics Breakout Summary (2019) *Proceedings of National Initiatives Meeting*, London, UK.
- Consent Process for Mobile App Mediated Research Systematic Review (2017) *JMIR Mhealth Uhealth*. 5:e126.
- Dynamic consent: a potential solution to some of the challenge f modern biomedical research (2015) *Eur J Hum Genet*. 23: 1416

Dynamic consent may help to take consent at various phases of clinical research through ongoing communication on line.

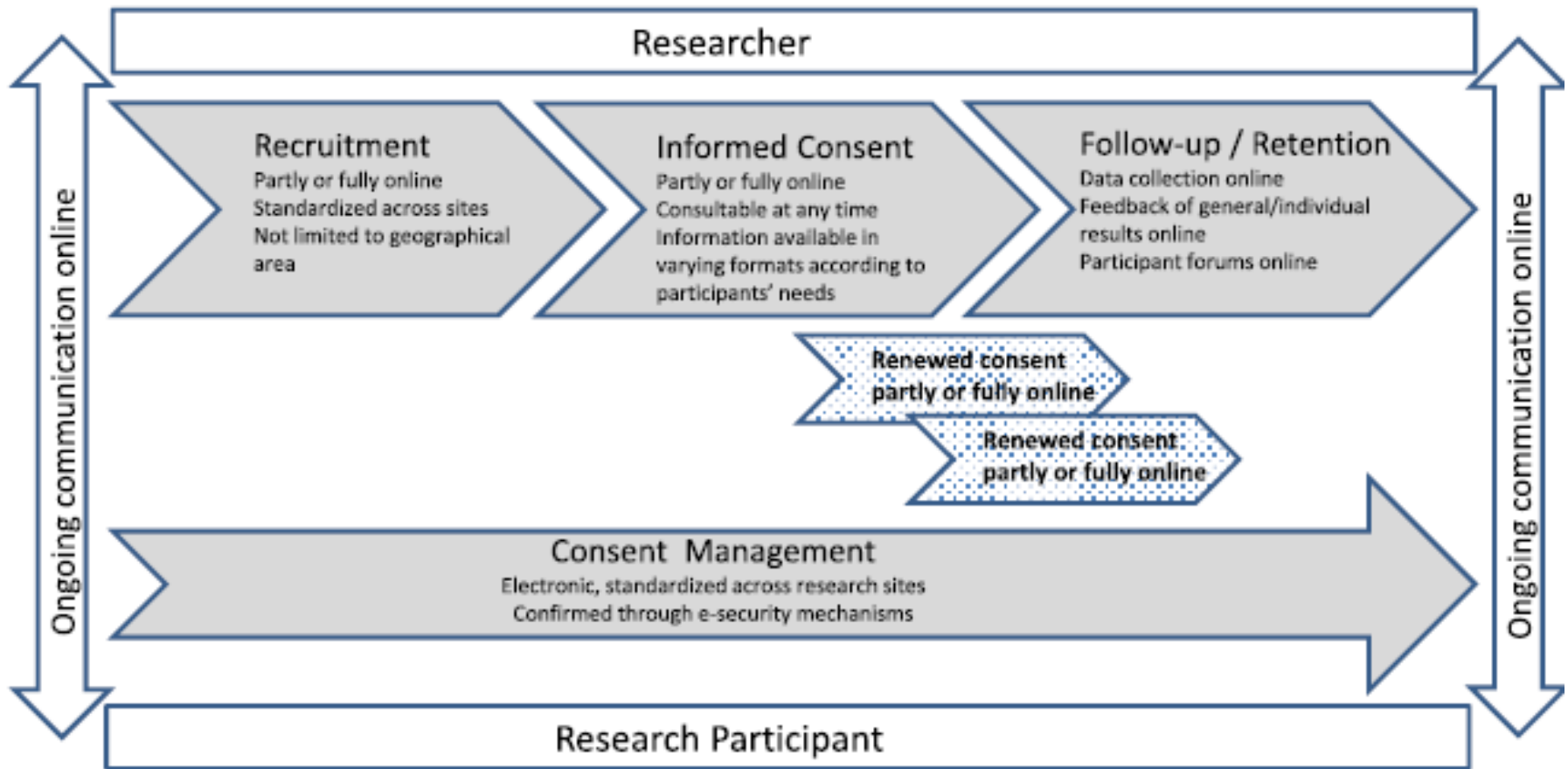


Fig. 1 from Budin-Ljønsne I, et al. BMC Med Ethics 2017, DOI: [10.1186/s12910-016-0162-9](https://doi.org/10.1186/s12910-016-0162-9)

# Pros/Cons of Dynamic consent



## **Pros**

For research participants (Patient and Public)

More appropriate, precise and flexible selection/change of consent

Accessibility to research information

Literacy for science and medicine

Bilateral communication to trust

For researcher and institutions

Certain evidence of consent

Sustainability

Better framework for data sharing in medicine/health field

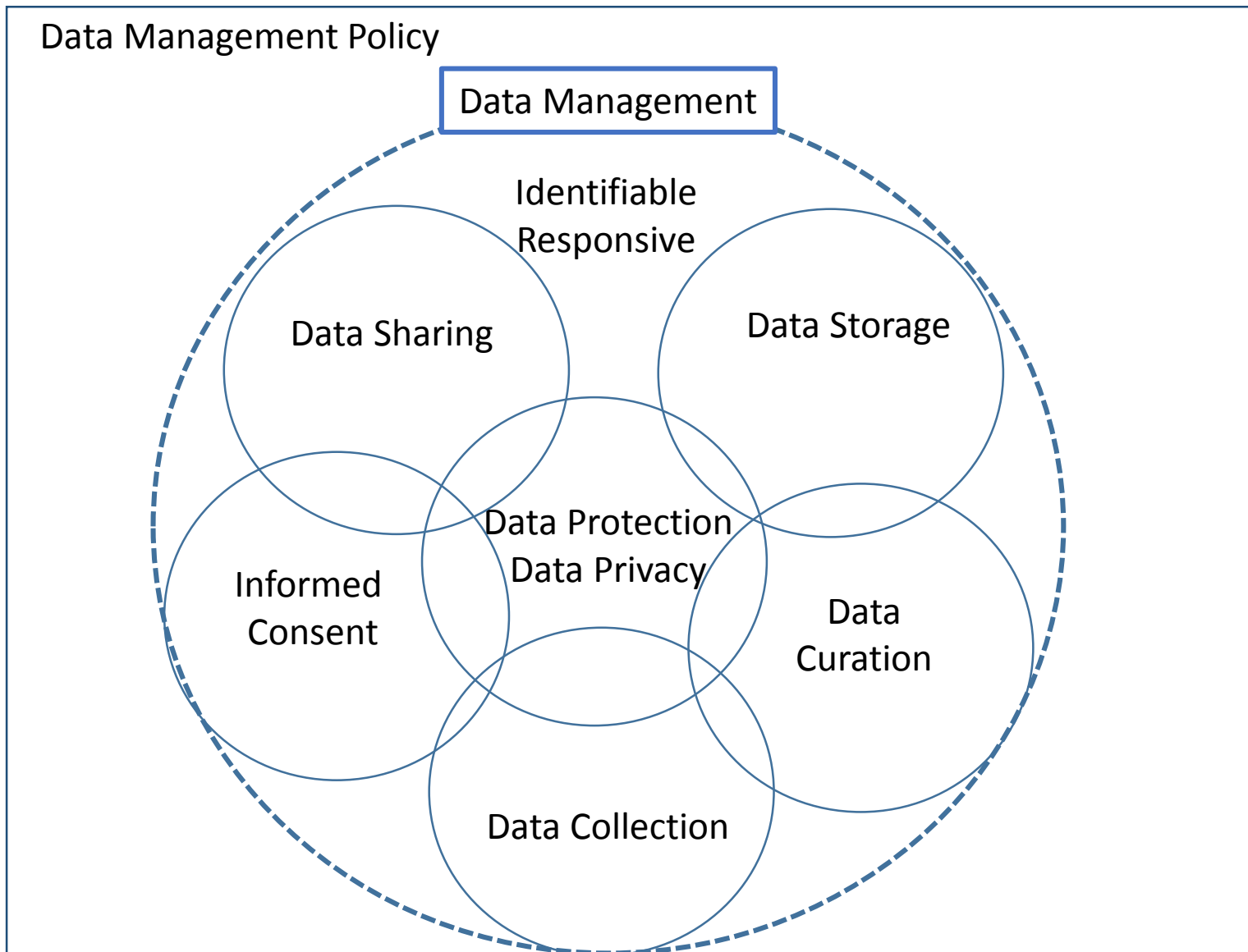
Accountability to ELSI in biomedical research

## **Cons**

- Game change in communication between research participants and researchers
- Usability (easy use for every participant)
- Less opportunity for research participation due to the limited access of internet (especially for Elderly)
- Less opportunity for face-to-face meeting in consent process
- Consent fatigue
- Difficulty to estimate “informedness”



# Data Management is important to survive in the era of “Open Science” as well as to respect/protect research participant in health research.



# Acknowledgement



## AMED

### Team “Life Course Data for R&D”

Makoto Suematsu, MD, Ph.D.

Hiroya Hirose

Megumi Haga

Yuriko Suzuki, MD, Ph.D.

Haruko Tsurumi, MD, Ph.D.

Michiaki Iwata, Ph.D.

Koichi Tomoshige, MD

Shogo Kato, Ph.D.

Takuzo Takayama, Ph.D.

Yoko Izumi, MD (currently in MHLW)

Daisuke Yasumizu, MD (currently in MHLW)

### Department of Research Infrastructure

Osamu Kato

Taisuke Kawasaki

Keiko Katsui, Ph.D.

Madoka Nakamura

**Thank you for your attention.**

**Tamami Fukushi, Ph.D.**  
**tamami-fukushi@amed.go.jp**