April 5, 2022

Asia Pacific Conference
of Pharmaceutical Associations

Mission and Progress of Tohoku Medical Megabank Project

Masayuki Yamamoto

Tohoku University
Tohoku Medical Megabank Organization

Tohoku Medical Megabank Organization

東北大学 東北メディカル・メガバンク機構は、
未来型医療を築いて震災復興に取り組むことを目的に
設置されました。
機構は、東日本大震災の被災地の地域医療再建と
健康支援に取り組みながら、
医療情報とゲノム情報を複合させたバイオバンクを構築します。
このバイオバンクに集まった情報と
その解析結果に基づく新しい医療の創出を通じて、
被災した東北地区への医療人の求心力向上、
産学連携の促進、関連分野の雇用創出、
さらには医療復興を成し遂げたいと考えています。

Established on February 1, 2012

ToMMo = Friends (Tohoku University)

TMM = ToMMo + IMM (Iwate MM)

We wish to deliver the most advanced medicine to the people who suffered from the earthquake and tsunami

We think that the most advanced medicine is personalized medicine and personalized healthcare

http://www.megabank.tohoku.ac.jp/index.php

APAC Asia Partnership Conference of Pharmaceutical Associations

Longitudinal Population Studies Based on Cohorts and Biobanks Are Essential to Promote People's Health

- Longitudinal population studies with intensive follow-up are indispensable approach for the understanding how genetic background, lifestyle (or environment) and ageing provoke pathogenesis of non-communicable diseases
- Longitudinal population studies can be realized through cohorts and biobanks

UK Longitudinal Population Studies Strategy (2014)

APAC
Asia Partnership Conference of Pharmaceutical Associations

TMM's Challenge

The first study in Japan

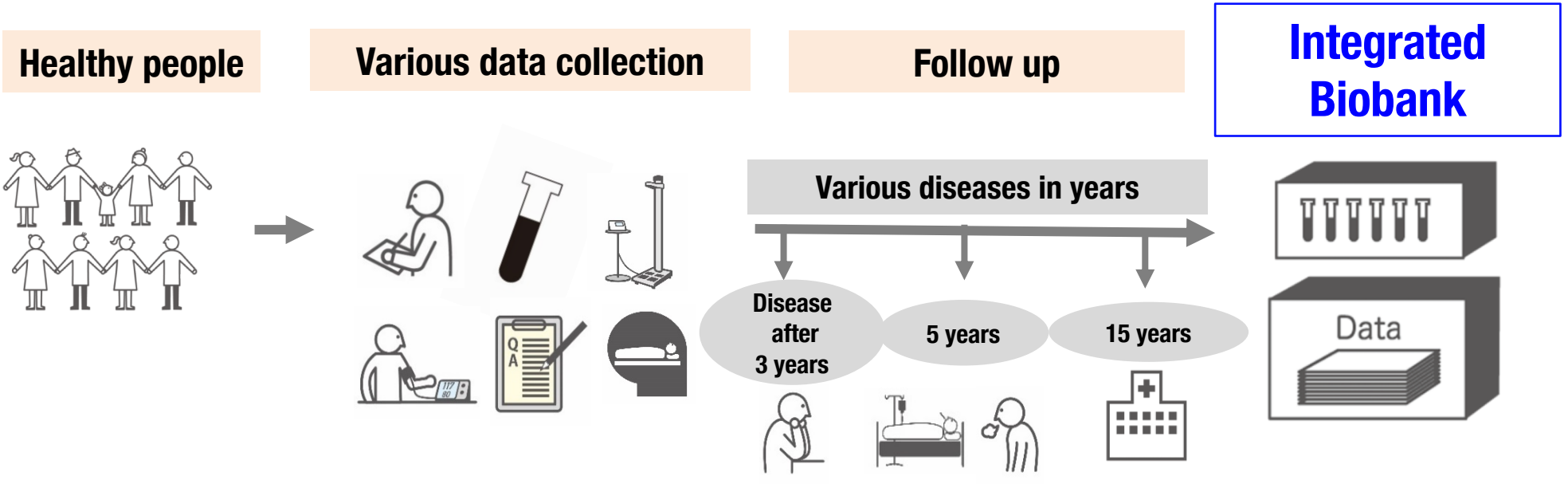
- to establish a large-scale population cohort study with a biobank
- to declare its commitment to providing specimen and data not only to academia but also to industry
- to declare its commitment to the return of genomic results (ROGR) to the participants

What makes TMM unique

- Two strategically established cohort studies: "Comm cohort" and "BirThree cohort"
- High level of consent rate from local residents
- Detailed and longitudinal data collection with comprehensive database
- Genome and omics data generation by sample analysis

Genome Cohorts and Integrated Biobank Leading to Personalized Medicine and Healthcare

- Cohort studies with genome analyses are referred to as Genome Cohorts
- Genome cohort studies are come of age for the personalized healthcare



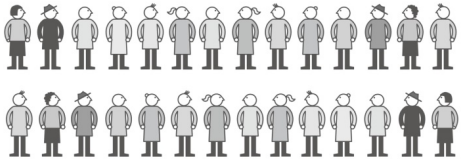
- Prospective cohort study enables us better estimation of the association between genetic effects and diseases or disease markers
- To approach for common diseases that involve an interplay between genetic and environmental events, genome cohort studies are necessary

APAC
Asia Partnership Conference of Pharmaceutical Associations

TMM Strategically Established Two Types of Cohorts
Community Cohort & Birth and Three-Generation Cohort

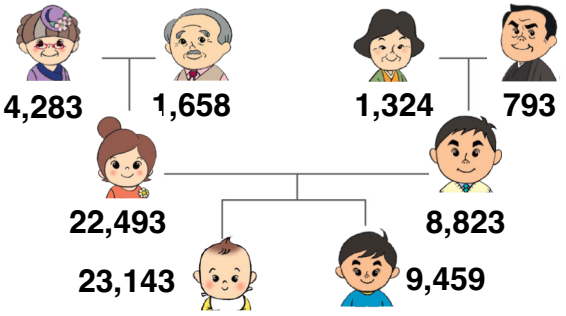
Community-Based Cohort

Recruit 80,000 residents from coastal areas in Miyagi and Iwate prefectures



Birth and Three-Generation Cohort

- Recruit 70,000 people including offspring, parental and grandparental generations
- Request expectant mothers for cooperation in maternity hospitals



We finished recruit of
more than 84,000 participants for CommCohort by March 31, 2016,
and, finished recruit of
more than 73,000 participants for BirThree Cohort by March 31, 2017

Recruited more than 157,000 participants in total

APAC
Asia Partnership Conference of Pharmaceutical Associations

Collected Items of TMM Cohort Studies

Blood test :

Collect 34 ml blood

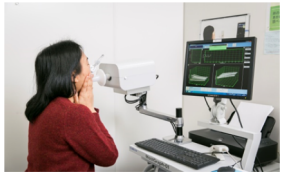
Blood test items

Peripheral Blood Tests
Hemogram
glucose level
HbA1c
GOT
GPT
γGTP
Total cholesterol
HDL cholesterol
neutral fat
urea nitrogen
Cr（eGFR）
uric acid
Serum pepsinogen
Helicobacter pylori
Glycoalbumin
Specific IgE（5 items）
Total IgE
Cystatin C

Other biospecimens; urea, plaque, saliva, breast milk…

Lifestyle information by Questionnaires

- Standard items (Exercise, alcohol consumption, smoking, diet, medical information, personal relationships, items related to women's health, name and address, etc.)
- Disaster related items (Depression, disaster situation , stress)
- Genome related items



Detailed survey at our admission centers

Physiological examinations
Audiometry, body composition, calcaneal bone density, carotid artery echography, home blood pressure, leg extension test, **MRI,** ophthalmological examinations (fundus, axial length, intraocular pressure, retinal tomogram), oral examination, respiratory function, tablet questionnaire, etc.

Follow up survey*

- Public data and incidence registries

- Center-based survey (Second / Third)

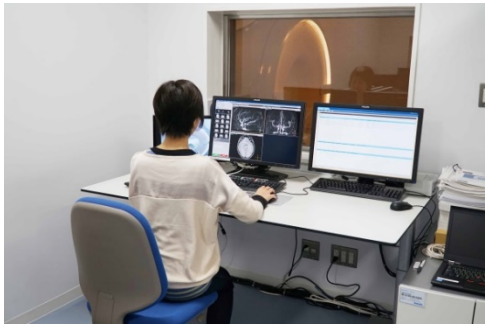
- **Re-Visit Community Support Center every few years** to perform physiological function tests, obtain samples for biobank, etc.
- Conduct add-on cohort surveys with the cooperation of companies etc.

*Follow up survey has been agreed upon by the participants, and being implemented from FY2017 (every few years)

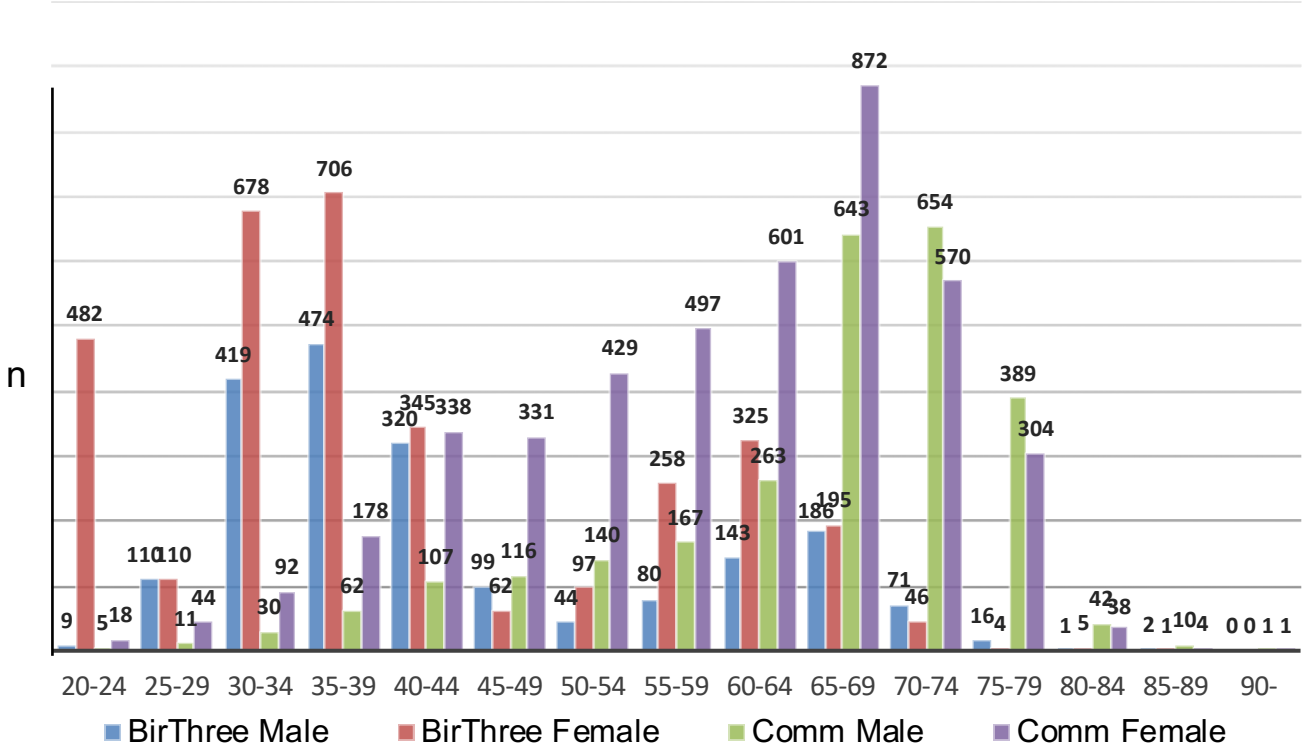
APAC Asia Partnership Conference of Pharmaceutical Associations

MRI & Cognitive / Psychological Data Collection

- The Brain-MRI Project was established in 2014
- Building Japan's largest brain image database from over 12,219 adults
- Follow-up (2nd phase) started from late October 2019



Demographic distribution of Brain-MRI project participants



MRI Project (n=12,219)	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
	The First Stage Jul. 2014-Oct. 2019						The Second Stage Nov. 2019 - the present					

APAC Asia Partnership Conference of Pharmaceutical Associations

Follow-up Data Collection

Mail and web-based questionnaire

- Health outcomes provided from all cohort participants
- Information from resident registry

Medical information

- National Health Insurance subscribers (about 56,000) in TMM CommCohort Study
- Information of public health examinations and medical expenses
- Electrical medical information collected from regional medical welfare network

Public data and registration of disease outcomes

- Infant health checkup information, regional cancer registration data, and demographic statistics etc.

On-site survey (every five years)

- Physiological test and blood sampling etc.
- Add-on research data under collaboration with industries
- Lifelog using mobile devices (coming soon)

ToMMo Community Support Center (CSC)
Health Check / Assessment Center

- Seven centers in Miyagi Prefecture
  (plus, four assessments centers in Iwate Prefecture)
- Detailed health examinations
- MRI imaging for volunteers





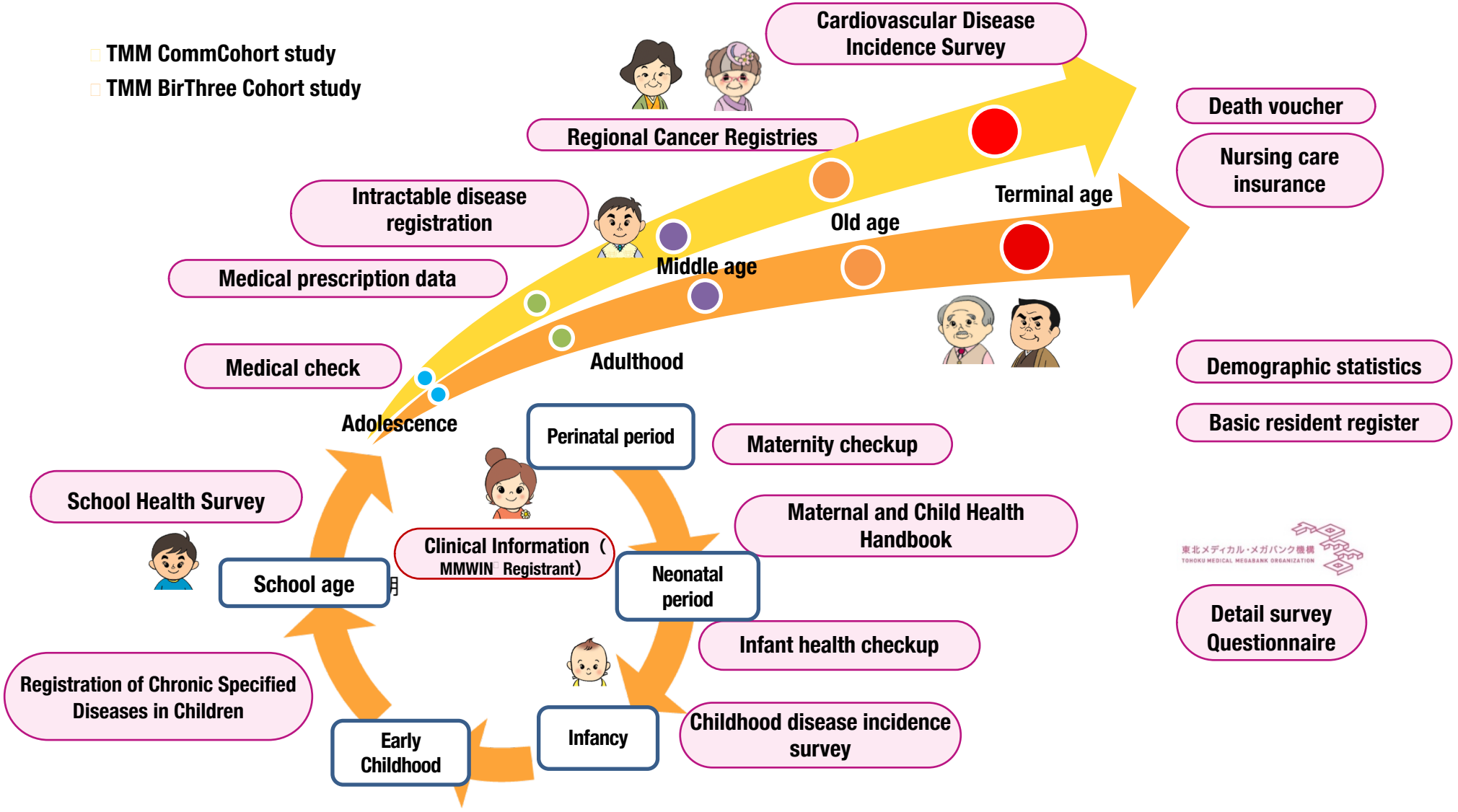


10

Data Accumulation through Life-course

“Linkage” of life course data enables us to conduct detailed longitudinal analysis

- TMM CommCohort study
- TMM BirThree Cohort study



※MMWIN：Miyagi Medical Welfare Information Network

AMED作成図を改訂
イラスト制作 橋本さと子

Progress of ToMMo's Integrated Biobank

We have made up our mind to establish an "Integrated Biobank" which

- consists of biospecimens and data of genome and omics linked to de-identified health and clinical information of cohort study participants
- aims at the development of personalized medicine
- contributes as an infrastructure for medical research in academia and industry,
- is based on the trust of residents in our university's 150 years of work
- in medicine

Biobank

Collects, stores, and distributes biological specimens and related information for advancements of medicine and science

Biobank is beneficial for the society

Large size biobank is good for

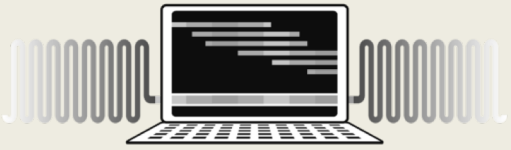
- Efficient use of resources
- Good quality control
- Reasonable use of resources

More than 4 million sample and 12 Pb information storage in total

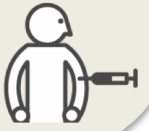
Information Available from ToMMo' Integrated Biobank

Integrated biobank

- To avoid rapid depletion of samples, ToMMo sets up an analytical center that conducts standard analyses of samples
- ToMMo distributes data and information first, and then bio-samples



Blood & Urine



Whole blood, serum, WBCs are stored
→ **metabolome and proteome**

Genomic DNA & RNA

DNA extracted from blood is also stored
→ **genome and transcriptome**

Questionnaire

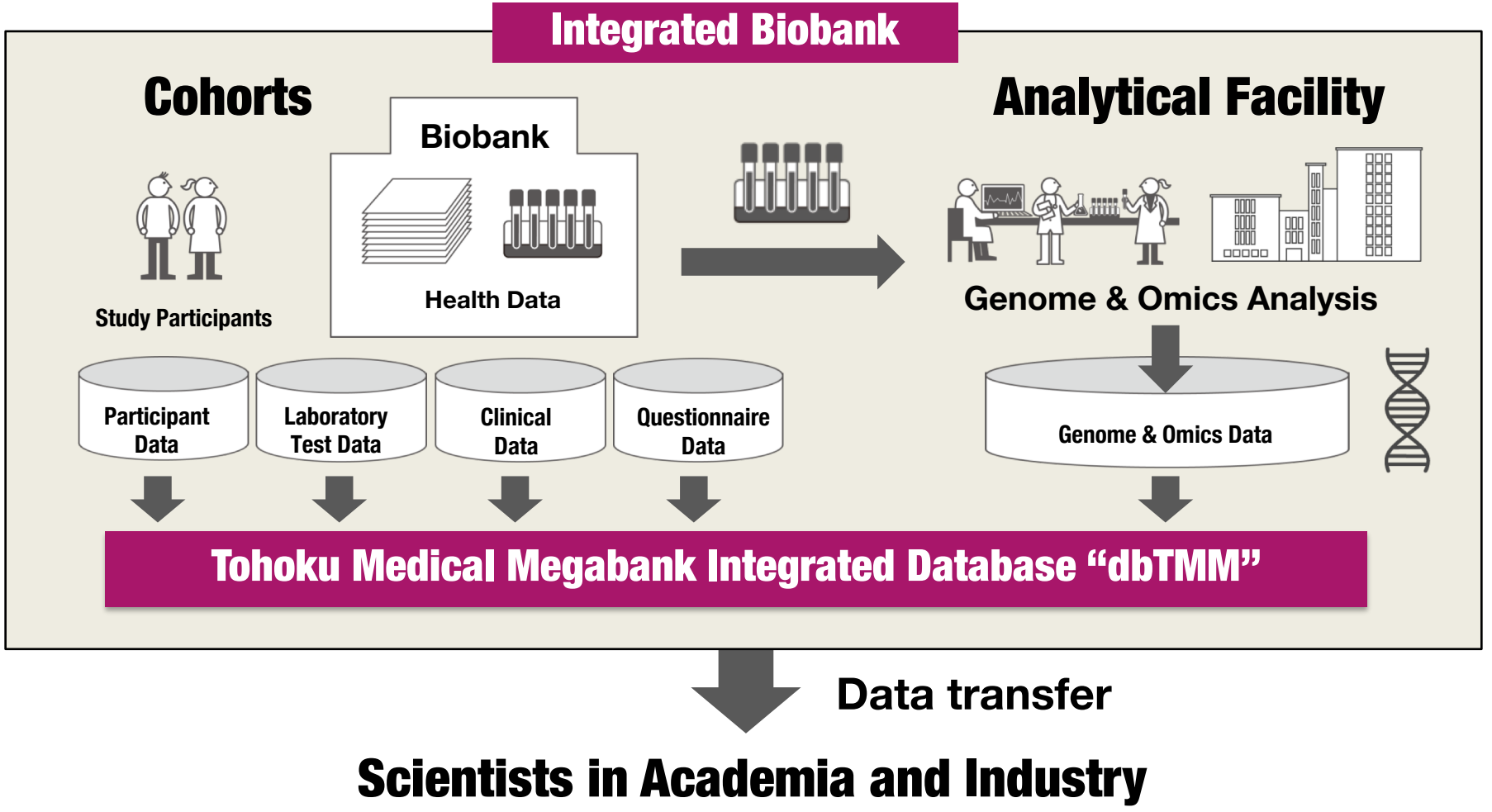
Main part is for lifestyle (including food), psychological condition, experiences of the disaster

+ **MRI** &

More than 10 physiological examinations, and cognitive and psychological assessment

Database Tohoku Medical Megabank (dbTMM)
Helping Users to Identify Sample Information Easily

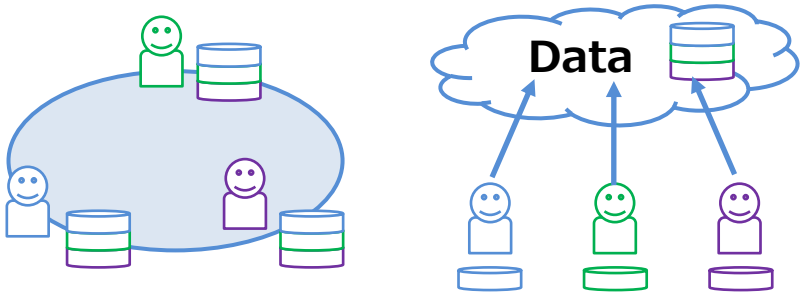
Tohoku Medical Megabank (TMM) is an integrated biobank retaining both biobank and genome / omics analytical facilities



ToMMo Utilizes Data Visiting

A New Paradigm

From Data Sharing ⇨ To Data Visiting



- Remote access to large-scale ToMMo database, dbTMM, from a highly secured remote security area

AMED Supercomputer
Outsourced to ToMMo

- Remote access via VPN line
- Safe sharing of wide variety of data including our whole genome sequence data

Installation of Remote Security Area

- In operation： 28 locations
- In preparation：8 locations

Shared space in Nihonbashi Mitsui Tower (near Tokyo station)

Shiga University of Medical Science
Osaka University
Iwate Medical University
Tohoku Medical and Pharma University
Tohoku University
Okinawa
Kyushu University
Nagasaki University
Nagoya univ
Chukyo area1
RIKEN AIP
Juntendo University
Nihonbashi Branch
Tokyo Medical and Dental University
DBCLS
Tokyo University
National Cancer Center
National Center for Child Health and Development
National Center of Neurology and Psychiatry
etc.

Current Status of Genome and Omics Studies in TMM and Future Directions

TMM has been working on
Genome, Metabolome, Transcriptome, Proteome, Epigenome and Metagenome analyses

Open Access

Japanese Multi Omics Reference Panel

jMorp

iMethyl Database
(Iwate Medical University)

iMETHYL
integrative DNA methylation database

Genome data:
- Whole genome sequencing for approx. 14,000 cohort participants
- **Open access**: frequency data of all SNVs (app. 76M) on autosomal, X chromosome, mitochondria and INDEL allele

Omics data:
- Plasma metabolome data from approx. 40,000 cohort participants
- **Open access**: subject number, mean, SD, CV, and detection rate range in proteins

Epigenome data
- whole genome DNA methylation database from 102 Japanese monocytes and CD4+ T cells
- **Open access**: specific demethylations at the gene region, distribution of methylation levels of each CpG site, distribution of expression levels of each gene, and information of each SNV

WGS of Japanese Population Is Important

Human genome contains many variations and polymorphisms that determine or contribute to the disease susceptibility of individuals AND

- **Japanese genome is quite different from that of Europeans and other ethnics**
- **In order to develop genome medicine and personalized healthcare, we need to determine Japanese genome structure precisely !**

WGS panel of population cohort participants will strongly support the clinical sequence for diagnosis and treatment of diseases

- **Clinical sequence operation requires large scale reference panel of healthy people**

WGS of patients alone will never tell us the responsible variants

Healthy individuals occasionally retain loss-of-function mutations as well as suppression mutations that contribute to the maintenance of health and Identification of these mutations leads to innovative drug development

TMM Has Been Conducting Whole Genome Sequence Analysis of General Japanese Populations

2015 1090

2016 2049

2017 3554

2019 4773

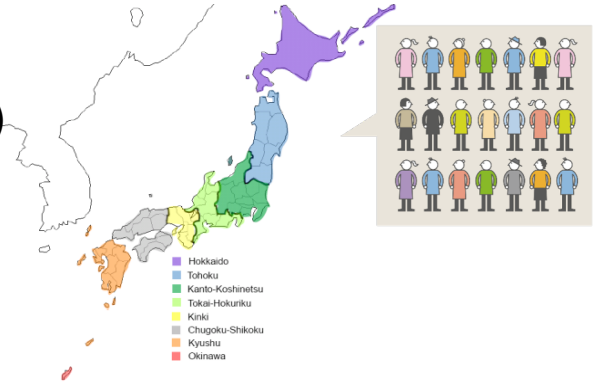
2020 8380 !

2021 14,129 !

TMM has accomplished @ December 2021

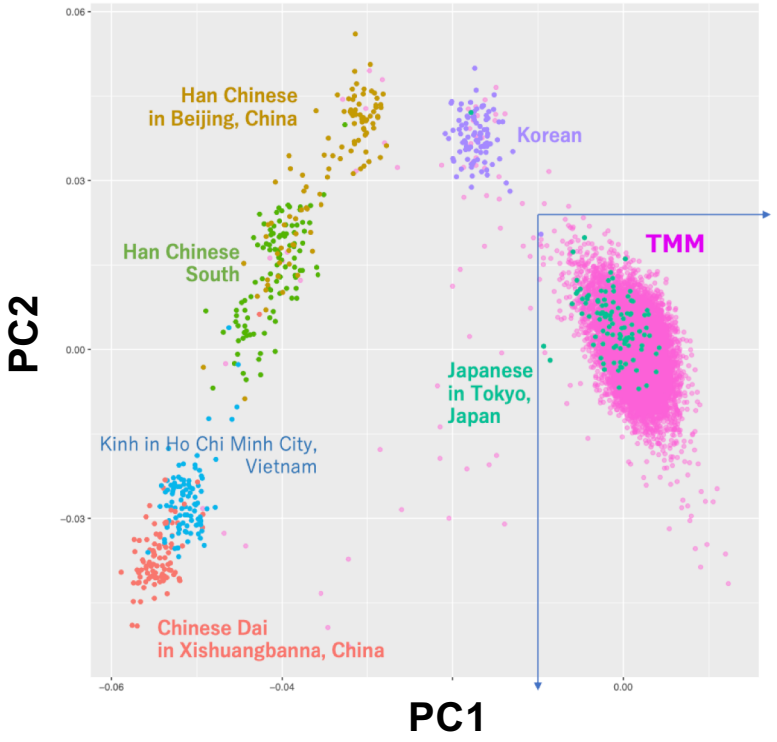
- 14,129 WGS
- JG2.1 (3 deep & long-read sequence analyses)
- Completed Japonica Array analysis > 120K

14,129 people



WGS Reference Panel

- Information about the position and frequency of SNVs



ToMMo Whole Genome Variation Database

Position and frequency of sequence variation

Genome structure of Japanese

ToMMo's whole genome sequence database will facilitate clinical sequence studies

Drug development

ToMMo's integrated data will be of important for segmented drug development for specific group of people

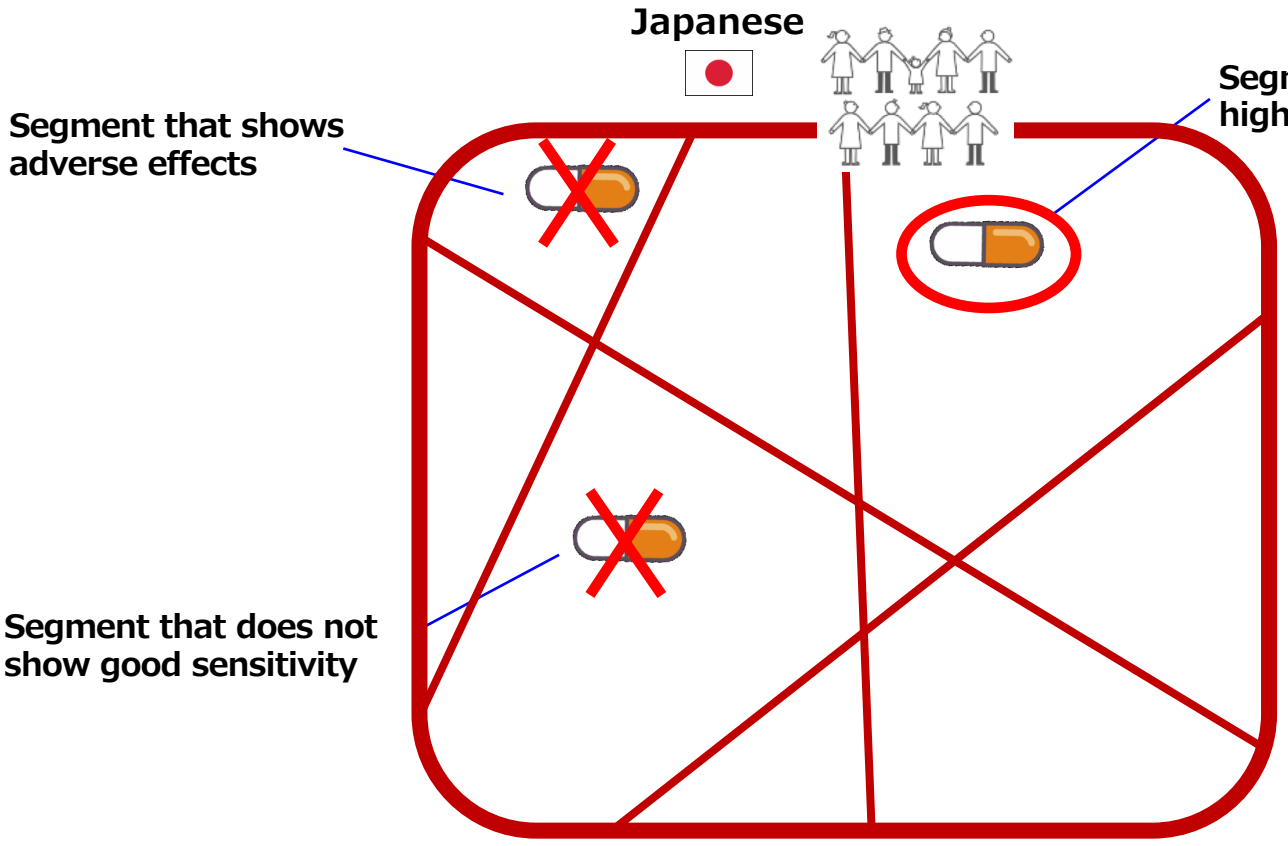
Personalized healthcare and ethnic array

Generate a special array that enable efficient imputation of Japanese genome

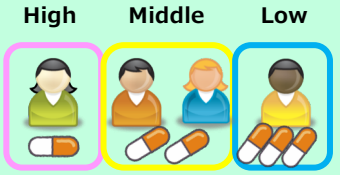
Segmented Drug Development Utilizing Genome Information

Identifying segments in Japanese by using genome information

- Precise identification of segments can by accomplished by the use of genome and omics information
- Low molecular weight chemicals can be target drugs by the use of segment information

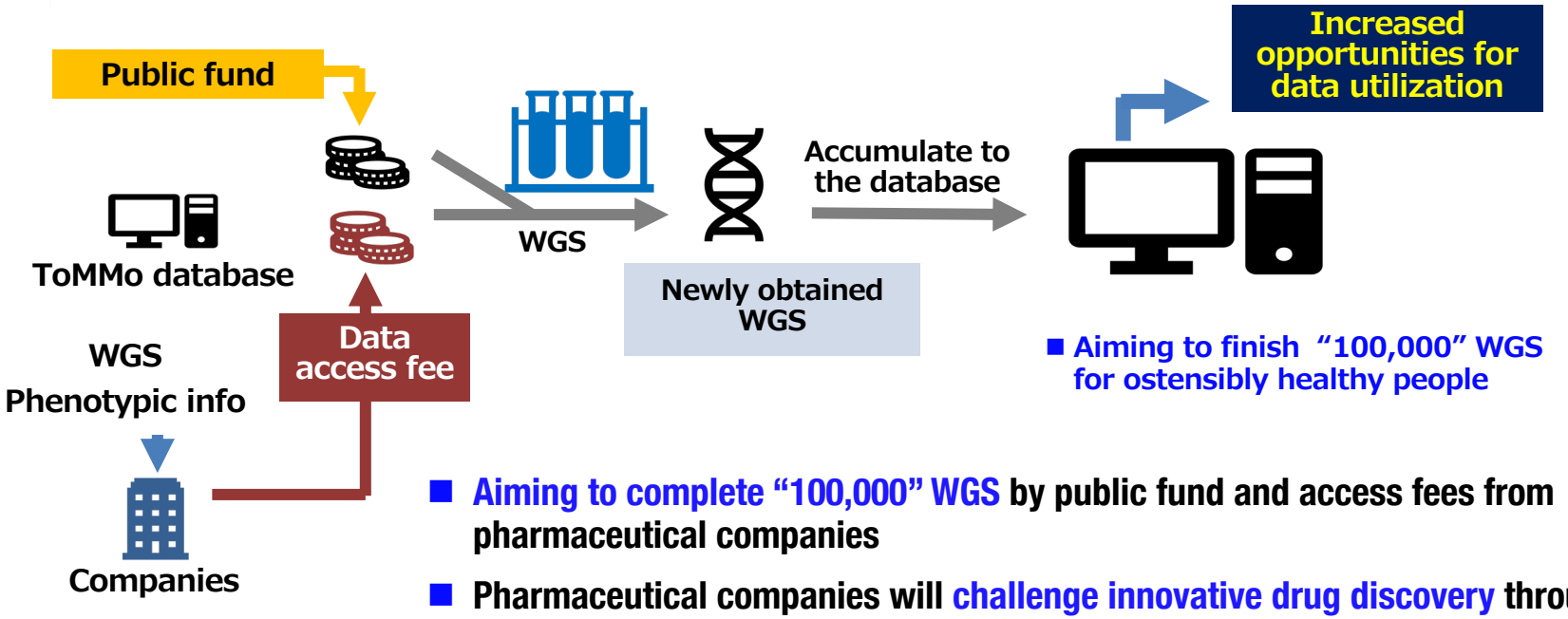


Personalized Drug Development



- Selection of optimized dosage
- Optimized clinical trials
- Use of LCLs for Cohorts in a dish

Consortium for Integrated WGS Analyses



- **Aiming to complete "100,000" WGS by public fund and access fees from pharmaceutical companies**
- **Pharmaceutical companies will challenge innovative drug discovery through integrated analyses of WGS / big data and ToMMo cohort data**
- **ToMMo will build large scale public genome database of Japanese after the priority period for consortium participants**

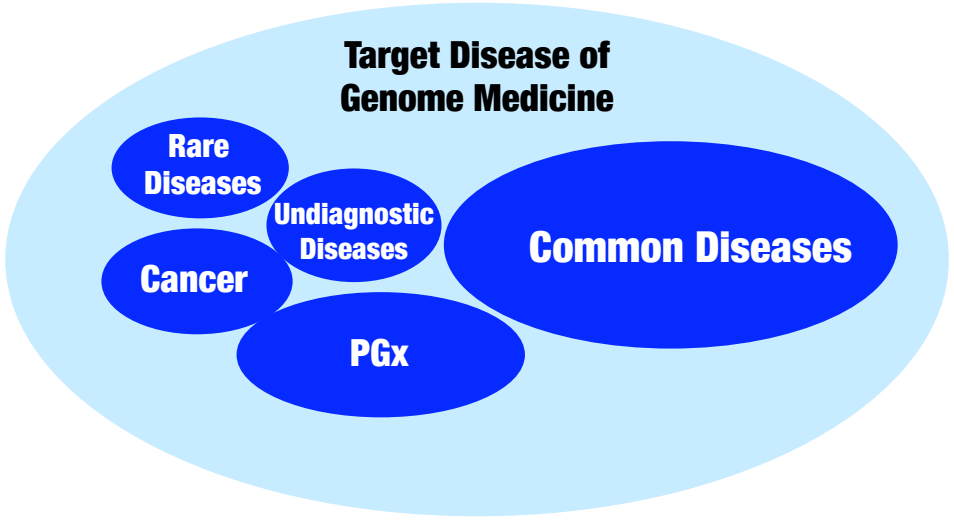
Risk Assessment of Common Diseases Is Important for Personalized Healthcare

NATURE MEDICINE | VOL 24 | OCTOBER 2018 | 1483 |

editorial

GWAS to the people

Thanks to improvements in data collection and analysis, some polygenic risk scores that predict disease risk are approaching the same predictive accuracy offered by tests for monogenic mutations. The time to think about how best to incorporate polygenic tests in the clinic is now.



Polygenic Risk Score (PRS)

- WGS has been touted as the next great leap in healthcare, but it is a costly process
- By comparison, the array chips used in PRS generation are commonly under $100, if not much less, making this approach far more affordable

Japonica array is the key-technology for personalized healthcare

Japonica Array NEO (2019)

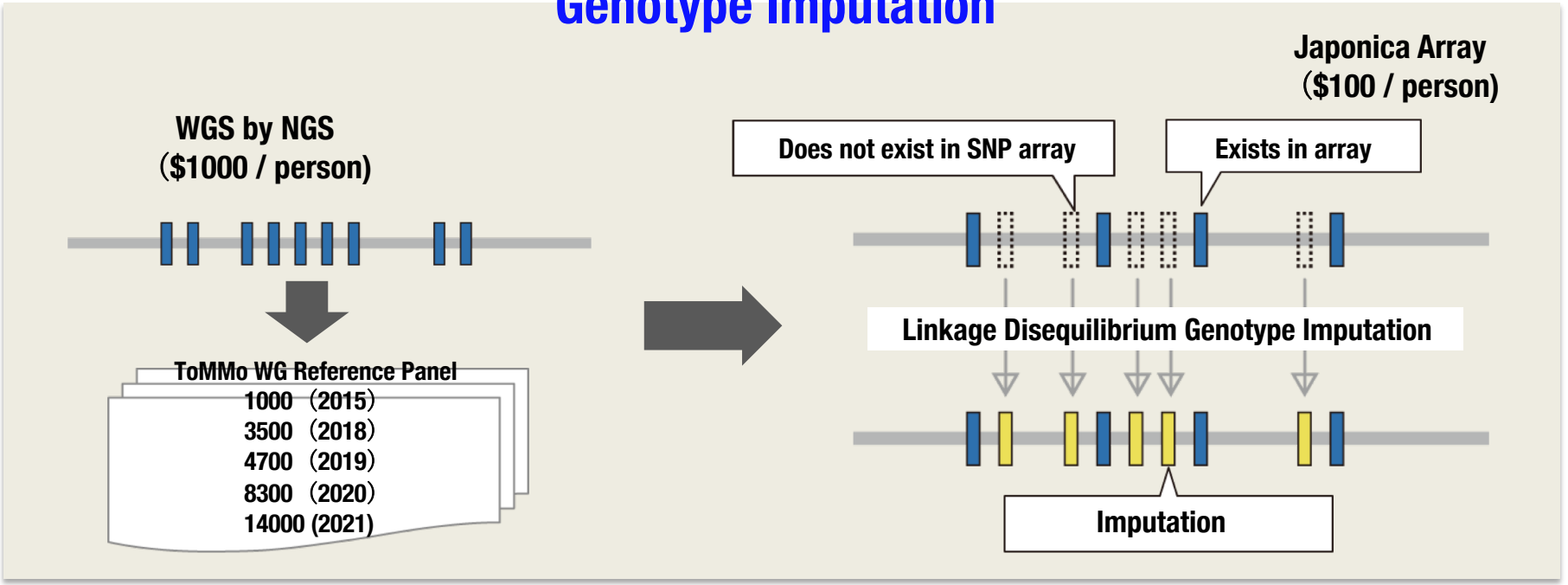
DNA Array Highly Adopted for Japanese Population

Japonica Array® is designed for studies in Japan,
and will contribute to personalized healthcare and medicine

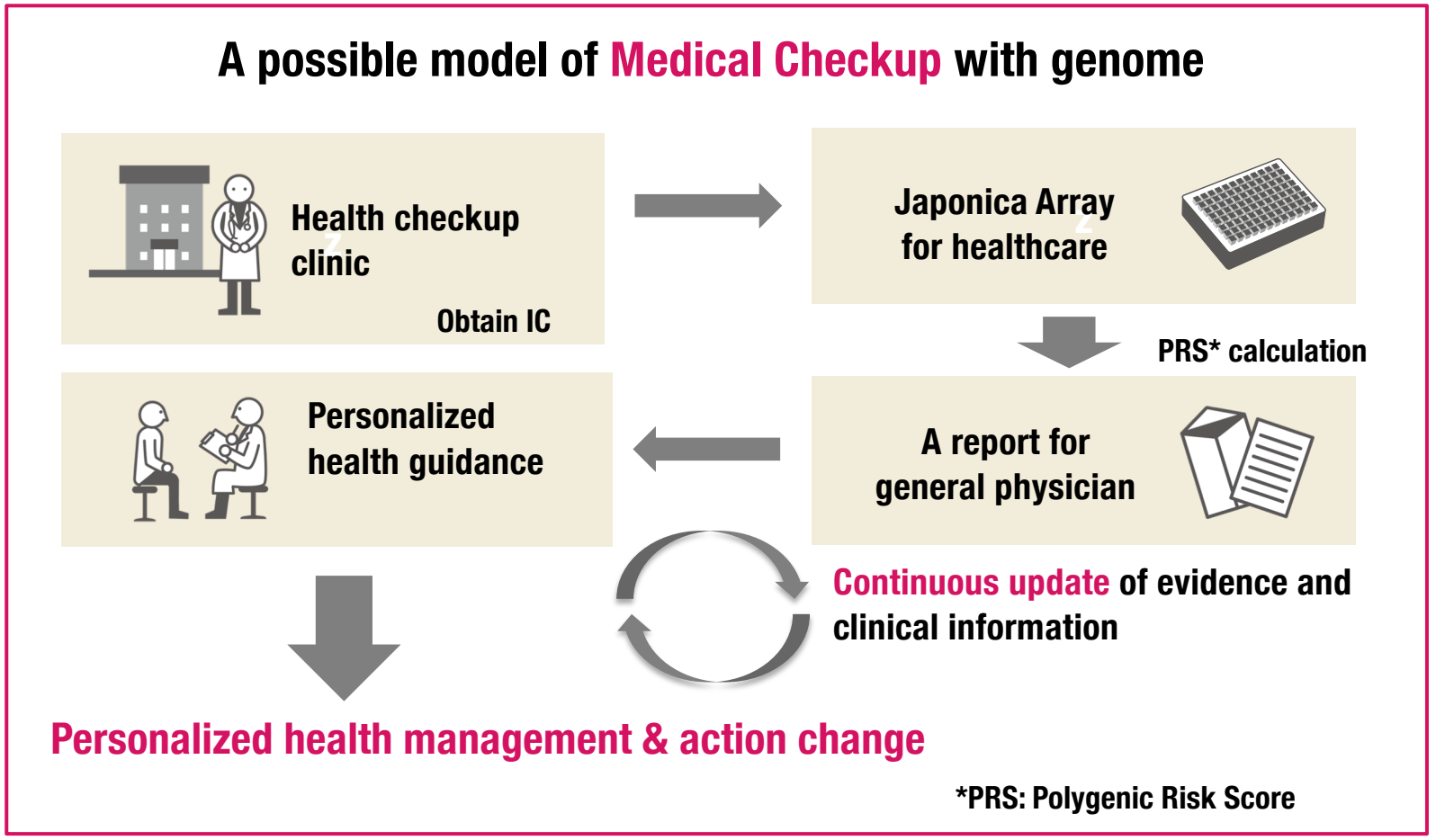
- Based on Japanese WGS data (3.5KJPNv2)
- Japonica Array is designed to minimize the number of probes but to maximize the capacity of genotype imputation for Japanese
- Inexpensive: providing the low cost WGS information will realize mega-size survey of genes responsible for common diseases

Japonica Array® was started marketing ver.1 in 2014, ver2. in 2017, and NEO in 2019

Genotype Imputation



In Near Future …



SUSTAINABLE DEVELOPMENT GOALS
2030年に向けて
世界が合意した
「持続可能な開発目標」です

3 すべての人に健康と福祉を

Our model

Personalized healthcare =

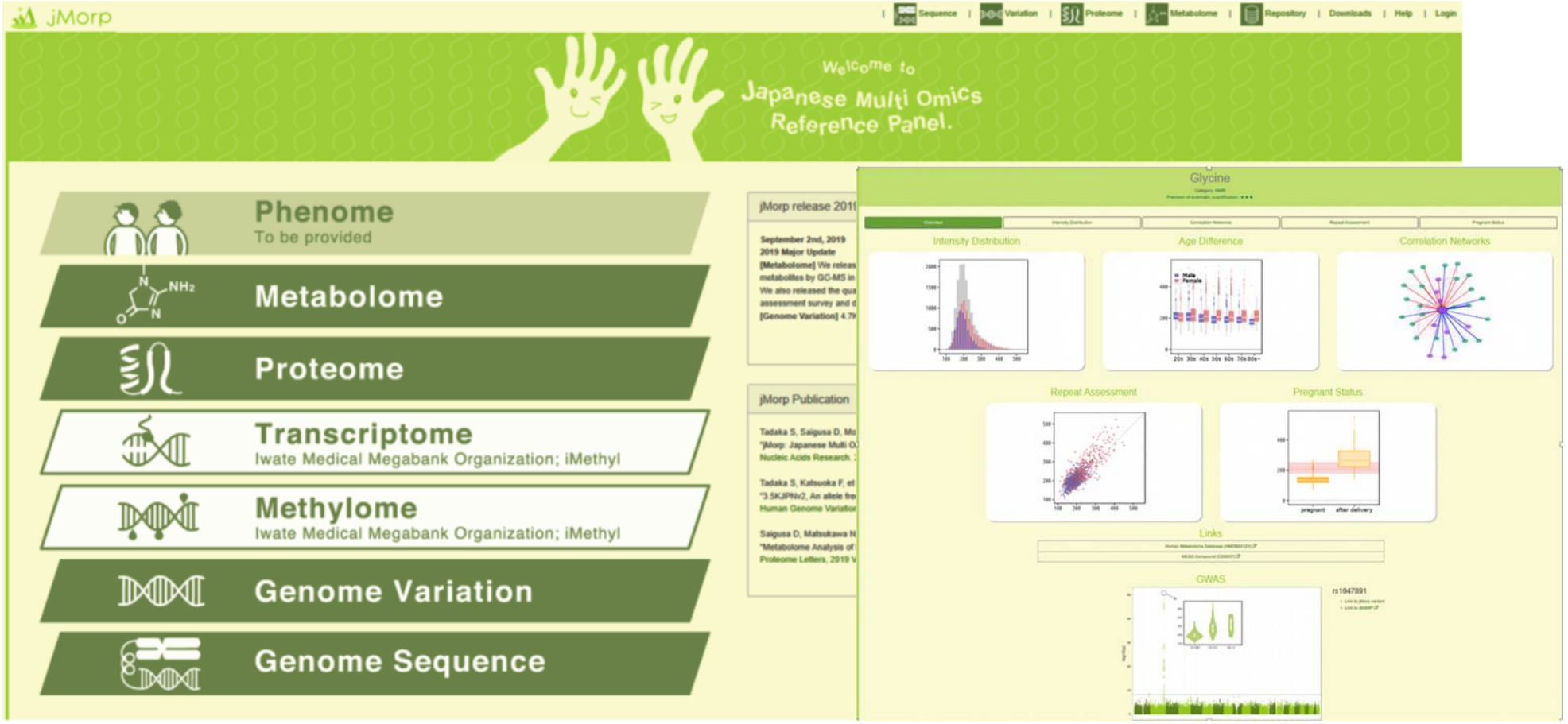
Ethnic specific SNP array + Genome reference panel + Return PRS by general physicians

jMorp (Japanese Multi Omics Reference Panel) Contains Various Big Data Including WGS

jMorp

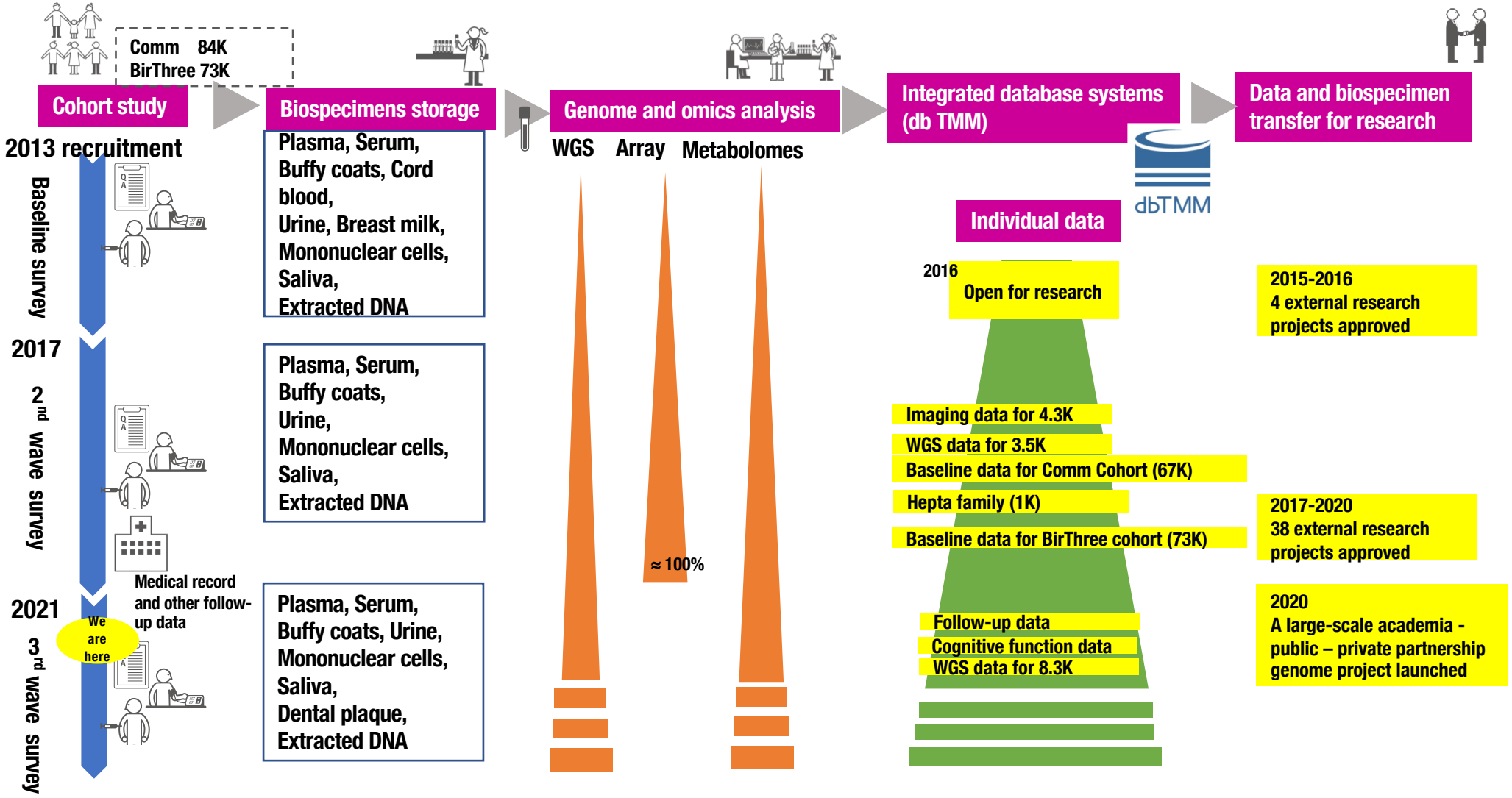
https://jmorp.megabank.tohoku.ac.jp

- ToMMo integrated database enables to generate health-science big-open-data
- Information in the integrated database will be open to research laboratories in Japan
- ToMMo integrated data will be of important for new drug development for specific group of people



https://jmorp.megabank.tohoku.ac.jp/

Our progress in ten years



Summary

- We have established Tohoku Medical Megabank to realize personalized healthcare (PHC) and personalized medicine
- In the Tohoku Medical Megabank Project, we have designed and are operating of two types of cohorts strategically ; community cohort and birth-and-three-generation cohort, which in collaboration will realize cutting edge accomplishments in Longitudinal Population Studies
- We have established an integrated biobank and are conducting genome-omics analyses for the establishment of genome medicine
- Tohoku Medical Megabank is contributing to the development of data visiting system and integrated database dbTMM

http://www.megabank.tohoku.ac.jp/index.php