

## ゲノム・データ基盤の構築に向けた取組について

令和3年3月16日

ゲノム医療協議会



## 目次

1. はじめに	… 1
2. 現在の主な取組	… 1
2.1 全ゲノム解析等の実施	… 1
(1) 「全ゲノム解析等実行計画」の推進	… 1
(2) コントロール群の構築	… 2
(3) 官民共同による全ゲノム解析の実施	… 2
2.2 ゲノム・データ基盤の構築及びデータ利活用推進のための環境整備	… 3
(1) ゲノム解析体制の構築に向けたスーパーコンピュータの連携	… 3
(2) ゲノム情報等の利活用推進のためのデータベース等の構築	… 3
(3) ゲノム医療を進めるための新たなサービス	… 4
3. 令和3年度以降の検討事項	… 5
3.1 全ゲノム解析等の実施	… 5
3.2 多因子性疾患におけるゲノム解析等の検討	… 5
3.3 ゲノム・データ基盤の構築及びデータ利活用推進のための環境整備	… 5
3.4 CANNDs の運営管理のための新規事業の開始	… 5
(参考資料)	
ゲノム医療研究開発イメージ	… 7
課題の整理	… 8
ゲノム・データ基盤の構築に向けた取組について	… 9
ゲノム・データ基盤の構築に向けた取組について (工程表)	…10

## 1. はじめに

我が国では、国民が健康な生活及び長寿を享受することのできる社会（健康長寿社会）を形成するため、政府が講ずべき医療分野の研究開発及び健康長寿社会に資する新産業創出等に関する施策を総合的かつ計画的に推進するべく「健康・医療戦略（令和2年3月27日閣議決定）」（以下、「本戦略」という。）を策定。

第2期に当たる本戦略は、2040年頃までを視野に入れ、令和2年度からの5年間を対象とし、令和6年度末までに全体の見直しを行うこととしている。

また、同戦略に即して策定した「医療分野研究開発推進計画（令和2年3月27日健康・医療戦略推進本部決定）」（以下、「本計画」という。）では、国立研究開発法人日本医療研究開発機構（以下、「AMED」という。）が中核的な役割を担う医療分野の研究開発において、ゲノム・データ基盤の整備を推進するとともに、「全ゲノム解析等実行計画」等の実行により得られるデータの利活用を促進することで、ライフステージを俯瞰して遺伝子変異・多型と疾患の発症との関連等から疾患の発症・重症化予防、診断、治療等に資する研究開発を推進し、病態解明を含めたゲノム医療、個別化医療の実現を目指すこととしている。

本戦略に基づき設置された「ゲノム医療協議会」（以下、「本協議会」という。）は、令和元年10月から令和3年3月までの間に5回の協議会を開催し、有識者等から成る構成員等でゲノム医療推進のための取組について協議を行った。

今回の報告では、本協議会において協議した結果の中から、令和6年度末までに世界最高水準の医療の提供に資する基盤を構築し、我が国におけるゲノム医療を更に推進するため、関係府省等が連携して検討し、解決すべき当面の課題及びその取組について取りまとめた。

## 2. 現在の主な取組

現在、本計画に基づき、関係府省が所管する事業等にてゲノム医療推進のための取組を進めているところであるが、世界最高水準の医療の提供に資する基盤構築のためには、これまで以上に府省の垣根を超え、我が国が一丸となって取組を進めて行く必要がある。

本計画を着実に推進し、ゲノム医療における基礎から実用化までの一貫した研究開発を進めることを目的として、令和4年度末を目途に以下の取組を実施する。

特に、「短期的視点で検討、議論していくべき課題」に関しては、関係府省、AMEDが連携、協議の上、医療分野の研究開発関連の調整費等の活用も念頭に入れ、着実に実施する。

### 2.1 全ゲノム解析等の実施

#### （1）「全ゲノム解析等実行計画」の推進

- 「全ゲノム解析等実行計画」に基づき、がん、難病における全ゲノム解析等を着実に実施する。

○全ゲノム解析等の成果を患者に還元する。新たな個別化医療等を実現し、日常診療への導入を目指す。また、全ゲノム解析等の結果を研究・創薬などに活用する。

[現在の解析数（解析中（解析予定）のものを含む）]

（がんの全ゲノム解析）

- ・ 遺伝性のがん等 : 3,000 症例
  - ・ 難治性がん : 750 症例
  - ・ 難治性がん・希少がん（小児がんを含む） : 12,250 症例（予定）
- 合計 : 16,000 症例

（難病の全ゲノム解析）

- ・ 難病 : 1,500 症例
  - ・ 調整費 : 1,800 症例
- 合計 : 3,300 症例

## （2）コントロール群の構築

○がんや難病に関するゲノム医療を進めるために必要となるコントロール群<sup>（※1）</sup>の構築を目的として、国内バイオバンクやコホートが所有する検体について、全ゲノム解析等を実施する。

[現在の解析数（解析中のものを含む）]

- ・ 東北メディカル・メガバンク（TMM） : 8,000 症例
  - ・ ナショナルセンター・バイオバンクネットワーク（NCBN） : 8,000 症例
  - ・ バイオバンク・ジャパン（BBJ） : 500 症例
  - ・ 日本多施設共同コホート研究（J-MICC）等<sup>（※2）</sup> : 2,000 症例
  - ・ 健康長寿社会の実現を目指した大規模認知症コホート研究 : 4,000 症例
- 合計 : 22,500 症例

（※1）一般集団中に0.01%の頻度で存在する遺伝子変異のうち、疾病発症の寄与が高いもの（相対危険度5.0以上）を検出するためには合計2.8万症例のコントロール群が必要

（※2）「神奈川県みらい未病コホート研究」からの提供検体を含む

## （3）官民共同による全ゲノム解析の実施

○創薬研究等に資する10万人規模のゲノム・データ基盤の整備を目的として、官民のマッチングファンドによる全ゲノム解析を実施する。

[現在の解析数（解析中のものを含む）]

- ・ 東北メディカル・メガバンク（TMM） : 20,000 症例<sup>（※）</sup>

（※）コントロール群のTMM（8,000症例）、J-MICC等（2,000症例）を含む

[今後の解析予定数]

- ・東北メディカル・メガバンク（TMM）：80,000 症例<sup>(※)</sup>

(※) 官側が負担する解析症例数：40,000 症例

## 2.2 ゲノム・データ基盤の構築及びデータ利活用推進のための環境整備

### (1) ゲノム解析体制の構築に向けたスーパーコンピュータの連携

○今後の大規模なゲノム解析を、国内のインフラを活用しながら効率的に進めること等を目的として、将来的なクラウドとの併用も視野に入れ、研究機関等が所有するスーパーコンピュータ（以下、「スパコン」という。）等を連携することによる大規模ゲノム解析基盤を構築する。

[連携拠点となる国内研究機関]

- ・東北メディカル・メガバンク（TMM）
- ・国立遺伝学研究所（遺伝研）
- ・国立がん研究センター（国がん）
- ・東京大学医科学研究所（医科研）

○後述する CANNDs<sup>(※)</sup> を通じて利活用を推進するシーケンスデータについては、解析結果の均質化のため、AMED の研究班が指定するプロトコールに基づき、解析を行うことを原則とする。

(※) Controlled shAring of geNome and cliNical Datasets の略

### (2) ゲノム情報等の利活用推進のためのデータベース等の構築

○AMED が支援したゲノム研究の成果等を一元管理し、産業界も含めた利活用を推進するとともに一研究機関での管理によるデータの散逸や流出等へのリスクマネジメントへの対応を目的として、複数機関によるストレージ連携を介したデータベース及び検索システム等を構築する。

[連携拠点となる国内研究機関]

- ・東北メディカル・メガバンク（TMM）
- ・国立遺伝学研究所（遺伝研）
- ・国立がん研究センター（国がん）
- ・東京大学医科学研究所（医科研）
- ・国立国際医療研究センター（NCGM）

○AMED 内にデータ利用に係る会議体<sup>(※)</sup> や事務局等を整備することにより、ワンストップサービスに対応するとともに、厳密なデータ管理体制を構築する。

(※) CANNDs にデータを格納するに当たり公開を免除するデータの判断を担うとともに、アカデミア、企業等からのデータ利用申請について、visiting 解析環境（リモートアクセスにより利便性と安全性を確保した解析環境）の利用の可否を審査する。また、既存のデータベース等と役割分担・連携を図り、利用者の便に供する。

○複数機関によるストレージ連携を介したデータベース等の構築及びデータ利用に係る会議体等の整備を一体的に進めることにより、プライベートクラウドの高度なセキュリティと産官学で幅広く活用できるパブリッククラウドの要素とを兼ね備えた「ハイブリッドクラウド」として整備する。

○産業界も含めた更なるデータ利活用を推進するため、AMED「ゲノム医療実現のためのデータシェアリングポリシー」や公募要領等を改定し、令和3年度以降、以下を行うことを原則とする。

- ・AMEDが支援したゲノム研究から得られたデータについては、AMEDが定める一定要件を満たす場合を除き、AMEDが指定する公的データベース<sup>(※)</sup>を通じて、制限公開すること

(※) 情報・システム研究機構 (ROIS) 国立遺伝学研究所/生命情報・DDBJセンターと科学技術振興機構 (JST)/バイオサイエンスデータベース (NBDC) が共同で運営する Japanese Genotype-phenotype Archive (JGA) 等

- ・学術研究以外の目的での第三者への提供も可能な同意<sup>(※)</sup>を取得すること

(※) 全ゲノム解析等実行計画において作成する統一的な同意説明文書と同様に、新薬開発への活用や将来の追加解析にも耐えうるものとして作成

### (3) ゲノム医療を進めるための新たなサービス

○本計画に基づき、ゲノム医療を進めることを目的として、AMEDが提供する以下のサービスの総称を CANNDs と定義する。

[CANNDs が提供するサービス]

- ・スパコン連携による大規模ゲノム解析基盤(将来的なクラウド併用含む)
- ・ストレージ連携を介したデータベース及び検索システム等
- ・データ利活用に係る審査等のワンストップサービス

○CANNDs については、データ利活用に係る技術検証等のため、令和3年度より試験運用を開始し、令和4年度より民間企業も対象とした本格運用を開始する。

○CANNDs ではシーケンスデータ (FASTQ) や臨床解析<sup>(※)</sup>の結果だけでなく、AMEDが支援する研究から得られたデータについても利活用の対象とする。

(※) 解析 (マッピング、バリエーションコール) 等を含む

### 3. 令和3年度以降の検討事項

令和3年度以降については、「2. 現在の主な取組」の進捗を確認しつつ、令和4年度末を目途に以下事項について、引き続き検討する。

#### 3.1 全ゲノム解析等の実施

○がん、難病等における全ゲノム解析等については、これまでの解析の結果や、国内外の研究動向等を踏まえて症例数の検討を行うとともに、がんや難病等の医療の発展や、個別化医療の推進等、より良い医療の推進のため、医療分野の研究開発関連の調整費等の活用も念頭に入れた全ゲノム解析等の更なる推進を検討

○コントロール群の構築に不足する5,500症例を追加するため、バイオバンク・ジャパン（BBJ）等の検体の全ゲノム解析等を検討

#### 3.2 多因子性疾患におけるゲノム解析等の検討

○ゲノム・オミックス解析等を活用した糖尿病、循環器疾患、認知症等の多因子性疾患<sup>(※)</sup>に関する予防、早期診断、治療最適化に資する研究や、ポリジェニックリスクスコアの開発に資する研究については、今後、優先順位等も含めて、効果的・効率的な実施を検討

(※) 複数の遺伝因子に加え、環境・生活習慣や老化が関わって発症する疾患

#### 3.3 ゲノム・データ基盤の構築及びデータ利活用推進のための環境整備

○国内において不足するデータサイエンティスト等について、大規模ゲノム解析基盤を活用した人材育成

○ゲノム情報等の更なる利活用推進のため、後述するCANNDsと既に構築されているバイオバンク横断検索システムや、AGD<sup>(※)</sup>との連携、統合を検討

(※) Amed Genome group sharing Database の略

#### 3.4 CANNDsの運営管理のための新規事業の開始

○CANNDsの運営管理等を目的として、関係府省、AMEDで協議の上、令和4年度より開始する新たな事業を検討

○新たな事業については、AMEDが令和2年度末に取りまとめる研究開発データ利活用基盤に係る構想との整合性を図り、AMEDが支援するゲノム研究以外の成果も含めた利活用基盤の運営管理と合わせて検討する必要があることから、内閣官房健康・医療戦略室及び厚生労働省が事務局を務める健康・医療データ利活用基盤協議会において検討

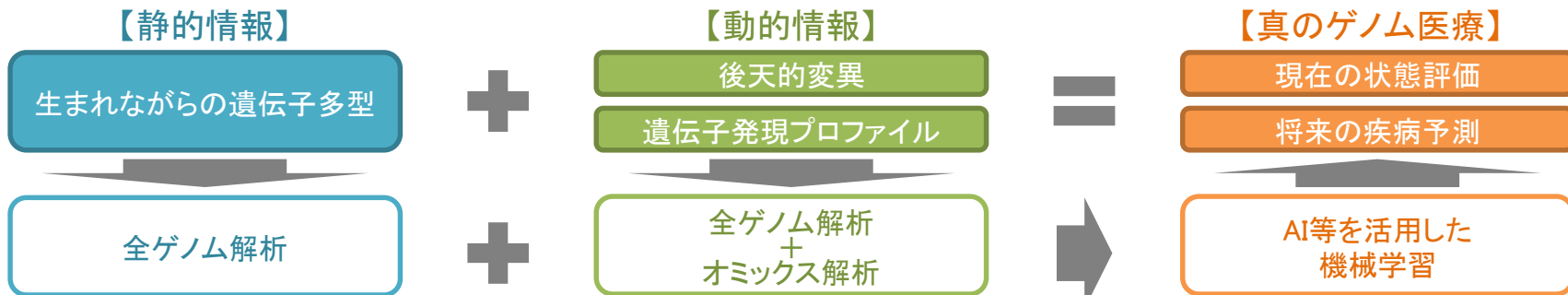


(参考資料)

# ゲノム医療研究開発イメージ

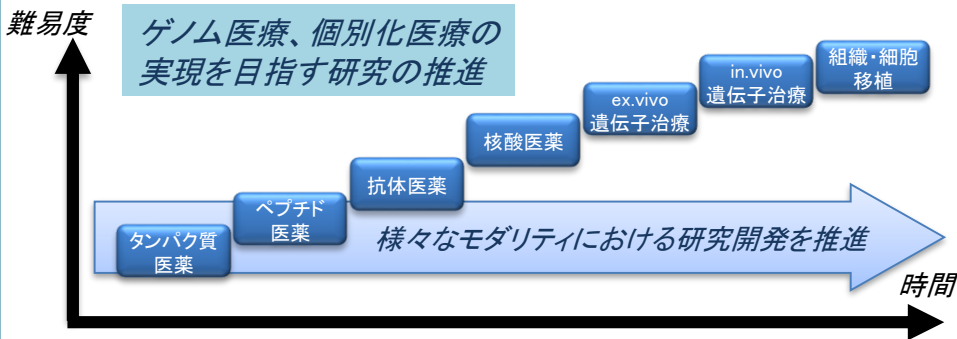
- 以下のような研究開発のコンセプト、枠組みも踏まえ、我が国におけるゲノム医療に係る研究開発が進むべき方向性や、何年後を目途に何を実用化していくのかまで見据え、検討を進める。

## 研究開発のコンセプト

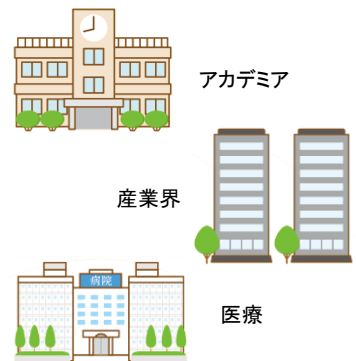
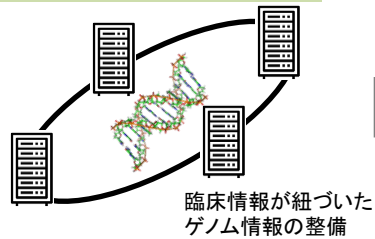


(理化学研究所生命医科学研究センター 寺尾 知可史先生提供資料より作図)

## 研究開発の枠組み



### ゲノム・データ基盤の構築 データシェアリングを推進



### ゲノム医療の実用化に向けて解決すべき課題

- 倫理的、法的、社会的課題への対応
- ゲノムリテラシー向上のための普及啓発等



関係府省において解決のための対応を検討する。

### ゲノム医療の研究開発を推進する人材の教育・育成



データサイエンティスト等の教育・育成

# 課題の整理

- 今後の議論は、課題解決の達成時期等を見定めて、以下のような優先順位をつけて議論を進める。
- 各省、AMEDは、本協議会での議論等を概算要求、調整費の配分方針等に反映させる。

## ただちに実施すべき課題

達成時期: 2020(令和2)年度末

進捗確認: 次回、ゲノム医療協議会にて進捗を確認

- 産業利用も含めた利活用促進のため、AMED「ゲノム医療実現のためのシェアリングポリシー」、公募要領等の改訂

## 短期的視点で検討、議論していくべき課題

達成時期: 2022(令和4)年度末

進捗確認: 2022(令和4)年度までは、ゲノム医療協議会にて毎回進捗を確認

- ゲノム医療を更に進めるための「全ゲノム解析等実行計画」の推進
- がんや難病に関するゲノム医療を進めるために必要となる2.8万症例のコントロール群の構築
- 官民共同による創薬研究等に資する10万人規模の全ゲノム解析の実施
- データサイエンティスト等の教育・育成のための環境整備を含む、ゲノム医療の研究開発推進のための人材育成
- ゲノム・オミックス解析等の活用も含めた多因子性疾患に関するゲノム研究の効果的、効率的な実施
- 3大バイオバンクを含むゲノム・データ基盤の構築及びデータ利活用推進のための体制整備<sup>(※)</sup>
- 同意の法的根拠について整理するとともに、産業界での利活用も可能な同意の要件の検討

※ゲノム医療協議会及び健康・医療データ利活用基盤協議会の双方からの指摘事項。進捗については、健康・医療データ利活用基盤協議会にも報告

## 中期的視点で検討、議論していくべき課題

達成時期: 2024(令和6)年度末

進捗確認: 2022(令和4)年度までは、年度の最終開催(3月目途)にて進捗を確認

(課題については、2022(令和4)年度末に進捗に応じた見直しを実施する)

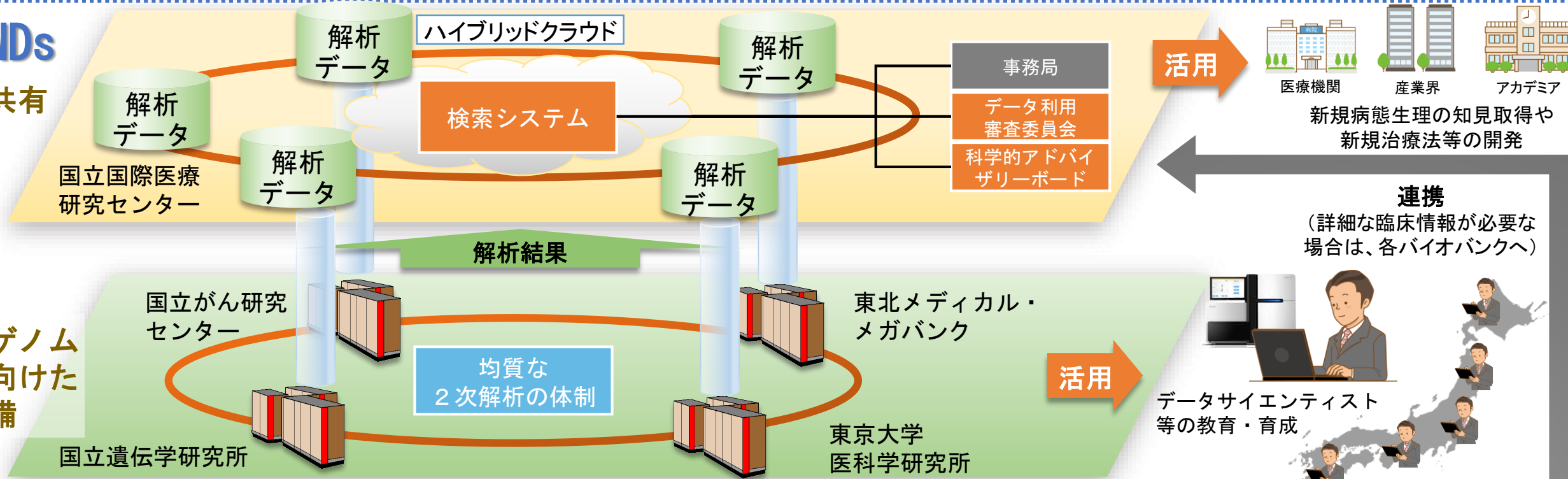
- 全ゲノム解析を活用したがんの新規原因遺伝子等の探索や、さらにオミックス解析も活用した難病等の新規原因遺伝子等の探索による、病態解明、早期診断、創薬等に資する研究の実施
- ゲノム解析等を活用した糖尿病等の多因子性疾患に関する予防、早期診断、治療最適化に資する研究の実施
- 海外とのゲノム研究における連携・協力と流出防止に係る方針の検討
- 倫理的、法的、社会的課題への対応するための法整備についての検討

# ゲノム・データ基盤の構築に向けた取組について

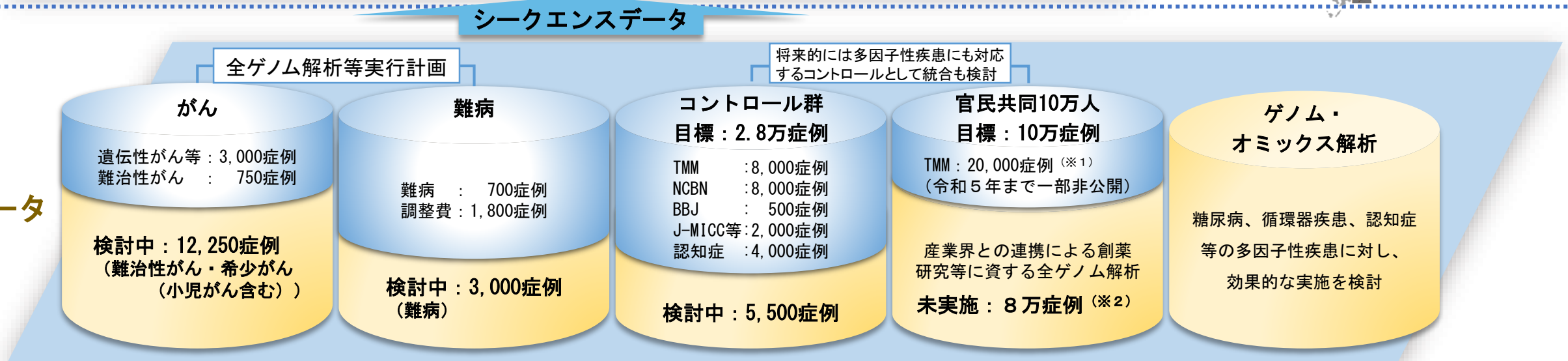
医療分野研究開発推進計画を着実に推進し、ゲノム医療を進めることを目的として、「短期的視点で検討、議論していくべき課題」に対応するため、医療分野の研究開発関連の調整費等の活用も念頭に入れ、関係府省、AMEDとも協議の上、今後、以下の取組を実施する。

## CANNDs データ共有の推進

## 大規模ゲノム解析に向けた基盤整備



## ゲノム解析データ



## 生体試料



# ゲノム・データ基盤の構築に向けた取組について（工程表）

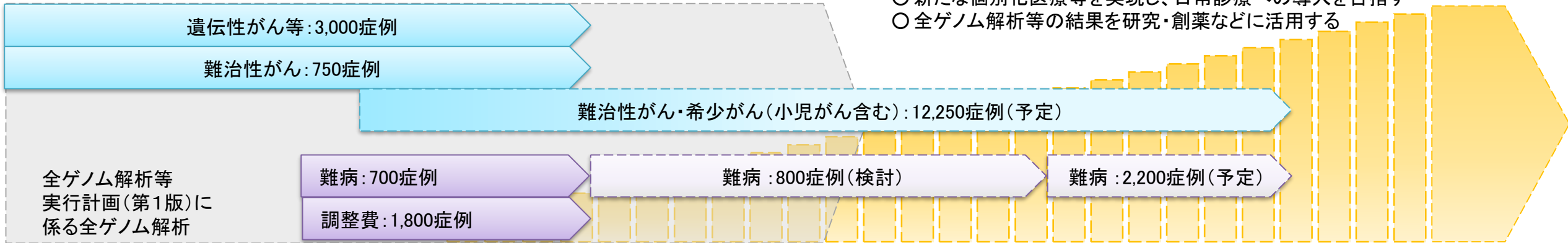
令和2年4月

令和3年4月

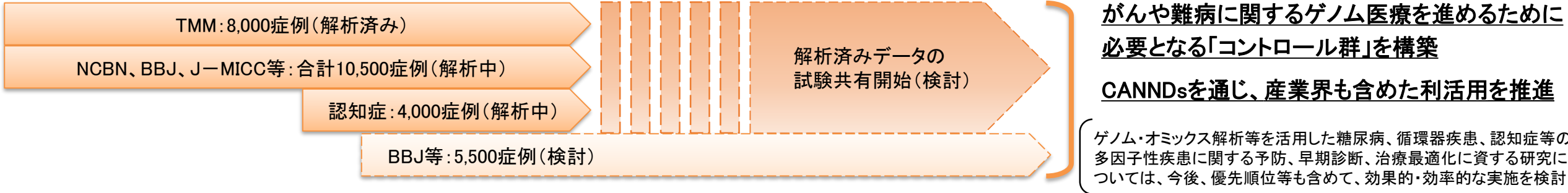
令和4年4月

令和5年4月

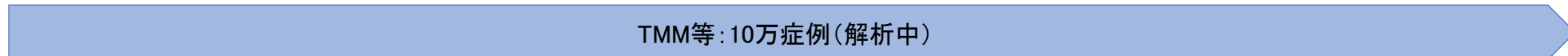
## <全ゲノム解析等実行計画に係るがん、難病の全ゲノム解析>



## <コントロール群の構築>

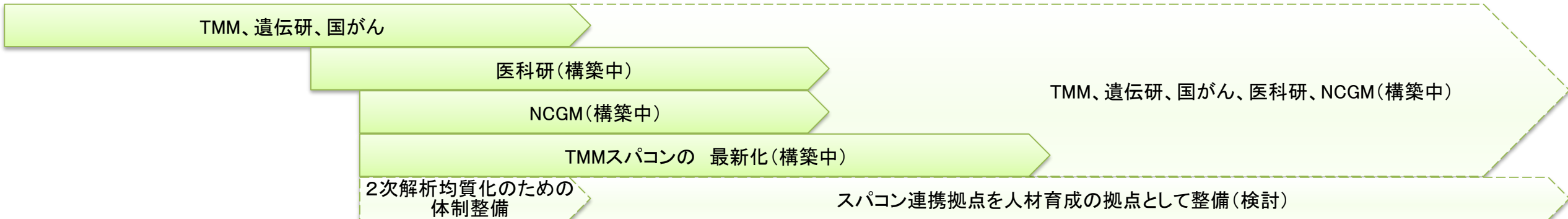


## <官民共同による全ゲノム解析>

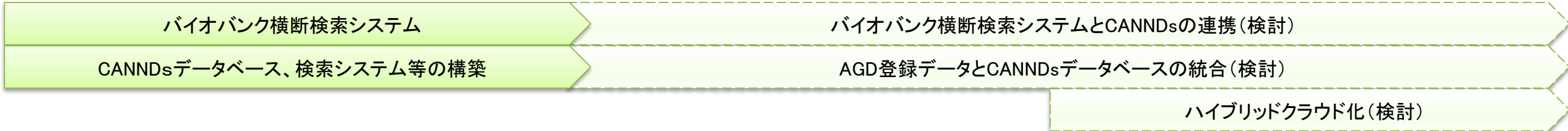


## <ゲノム解析体制の構築に向けたスーパーコンピューターの連携>

大規模データ解析基盤として一体的に構築



## <ゲノムデータ等の利活用推進のためのデータベース等の構築>



全ゲノム解析

環境整備