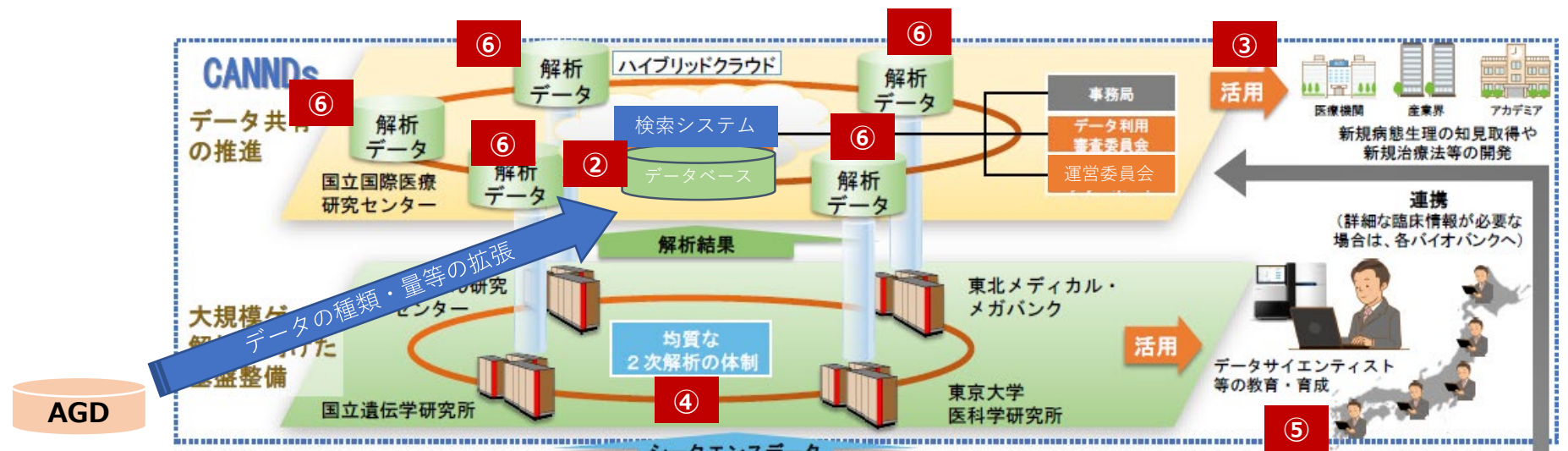


資料 5 : CANNDs実施計画の進捗について (AMED)

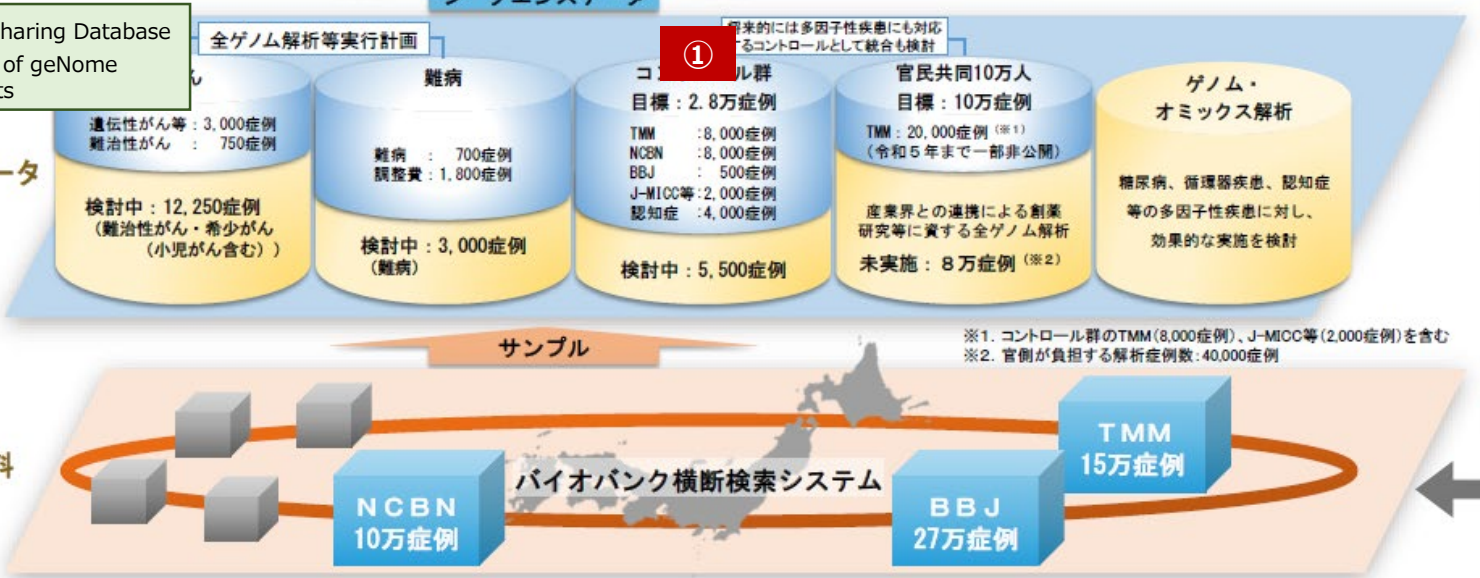
データシェアリングポリシーの改定とCANNDsの機能拡張 (R4年度以降)

- AMEDが支援する研究から得られたデータの利活用を推進するため、ゲノム医療実現のためのデータシェアリングポリシーを改定するとともに、AGDとCANNDsを統合し拡張する。
- AGDについては、格納するデータの種類、量、セキュリティレベルの大幅な変更と、検索機能の大幅な強化 (カタログ付加、外部システム連携等) を計画する。
- CANNDsの検索システムとAGDの検索システムを統合し、一体的に検索できるように拡張する。



AGD: AMED Genome group sharing Database
 CANNDs : Controlled shAring of geNome and cliNical Datasets

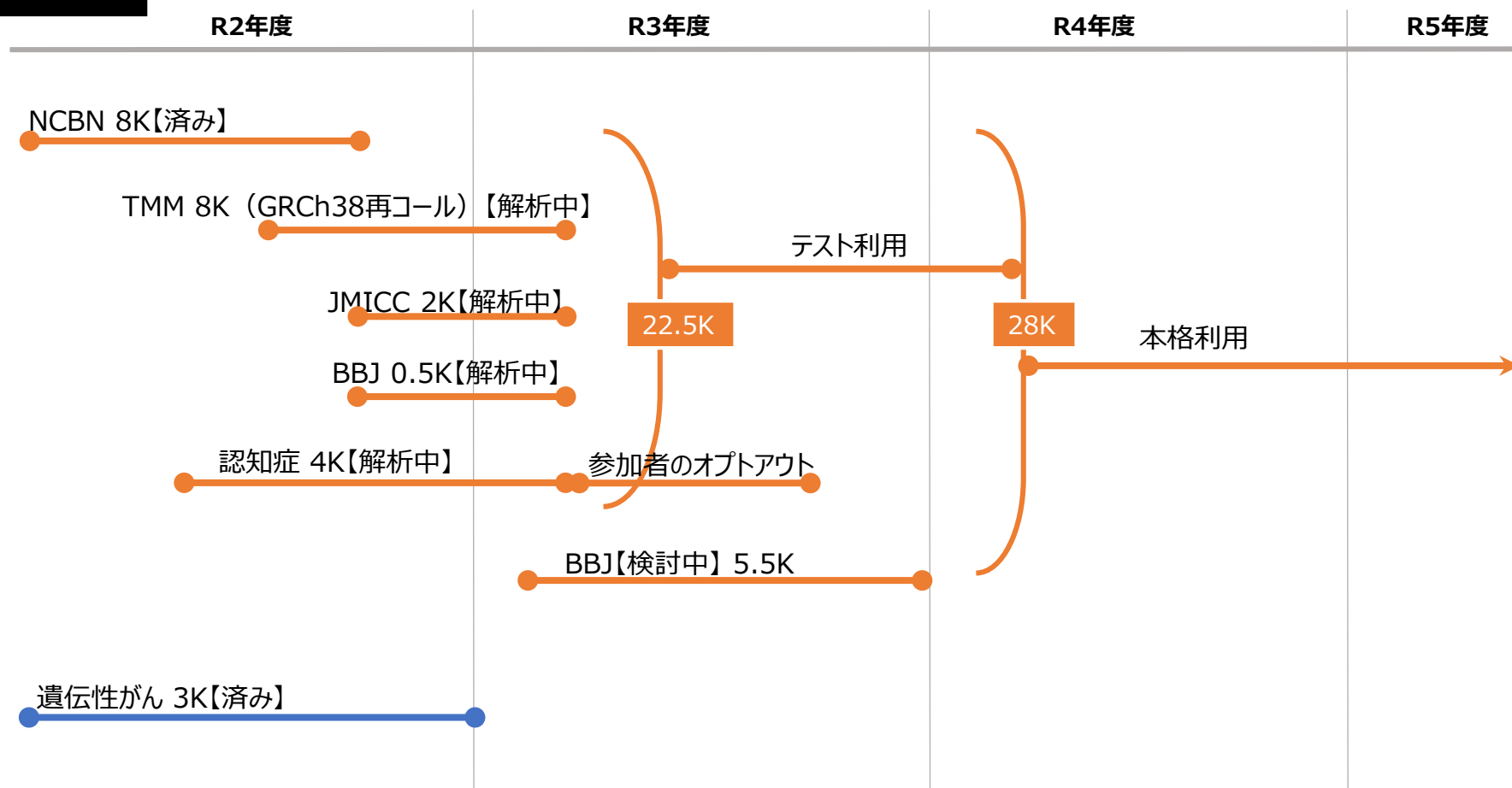
- ①コントロール群
- ②CANNDsの整備
- ③One Stopな利用受付・審査サービス
- ④テスト利用によるVisiting計算環境の検証
- ⑤研究公募準備 (ゲノム解析の均質化を含む)
- ⑥高度セキュリティルームの設置



※1. コントロール群のTMM (8,000症例)、J-MICC等 (2,000症例) を含む
 ※2. 官例が負担する解析症例数: 40,000症例

① コントロール群

工程表

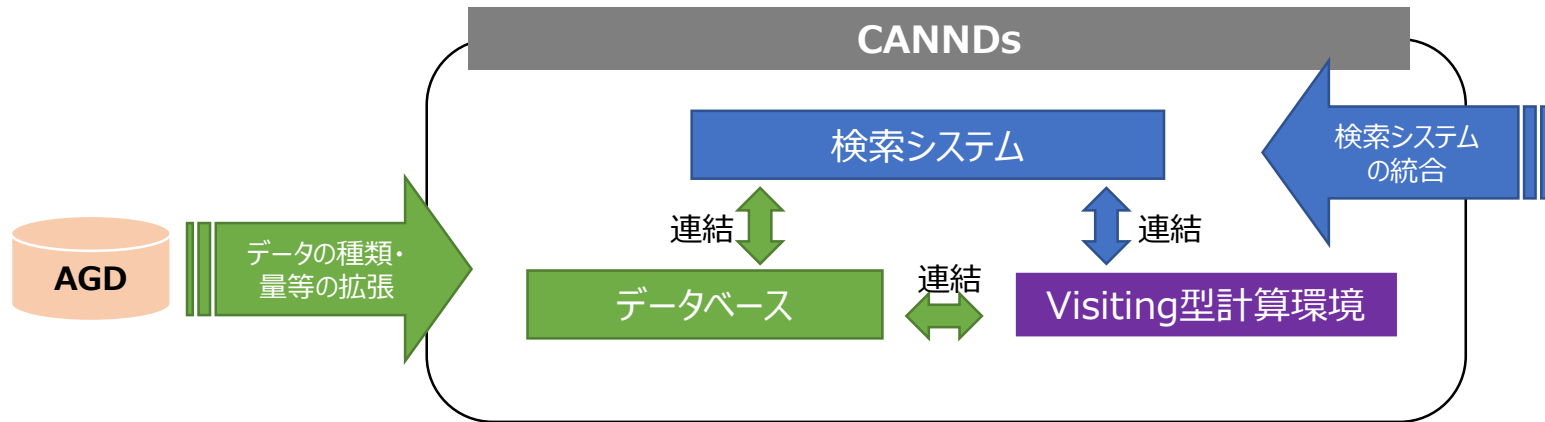


■ コントロール群として解析する検体の条件

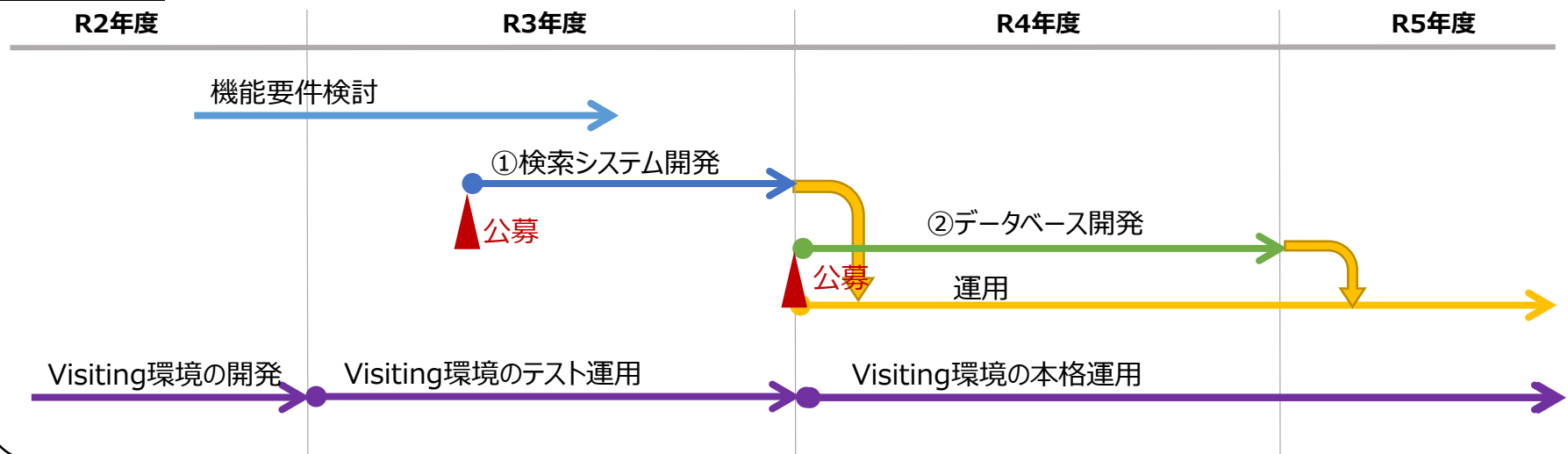
1. 詳細で正確な臨床情報が得られており、将来に亘り追跡できること
2. 将来に亘り検体が保管され、今後必要とされるトランスクリプトーム、プロテオーム、メタボローム等を含むオミックス解析が可能であること
3. 企業の利用を含む幅広い利活用に関する同意が十分に得られていること

② CANNDsの整備

- AMEDが支援する研究から得られたデータの利活用を推進するため、AGDをCANNDsと統合し拡張する。
- 格納するデータの種類や量の大幅な変更に対応するため、AGDの機能を拡張した後継のデータベースシステムを開発する。
- CANNDsの検索システムとAGDの検索システムを統合した検索システムを開発する。



工程表



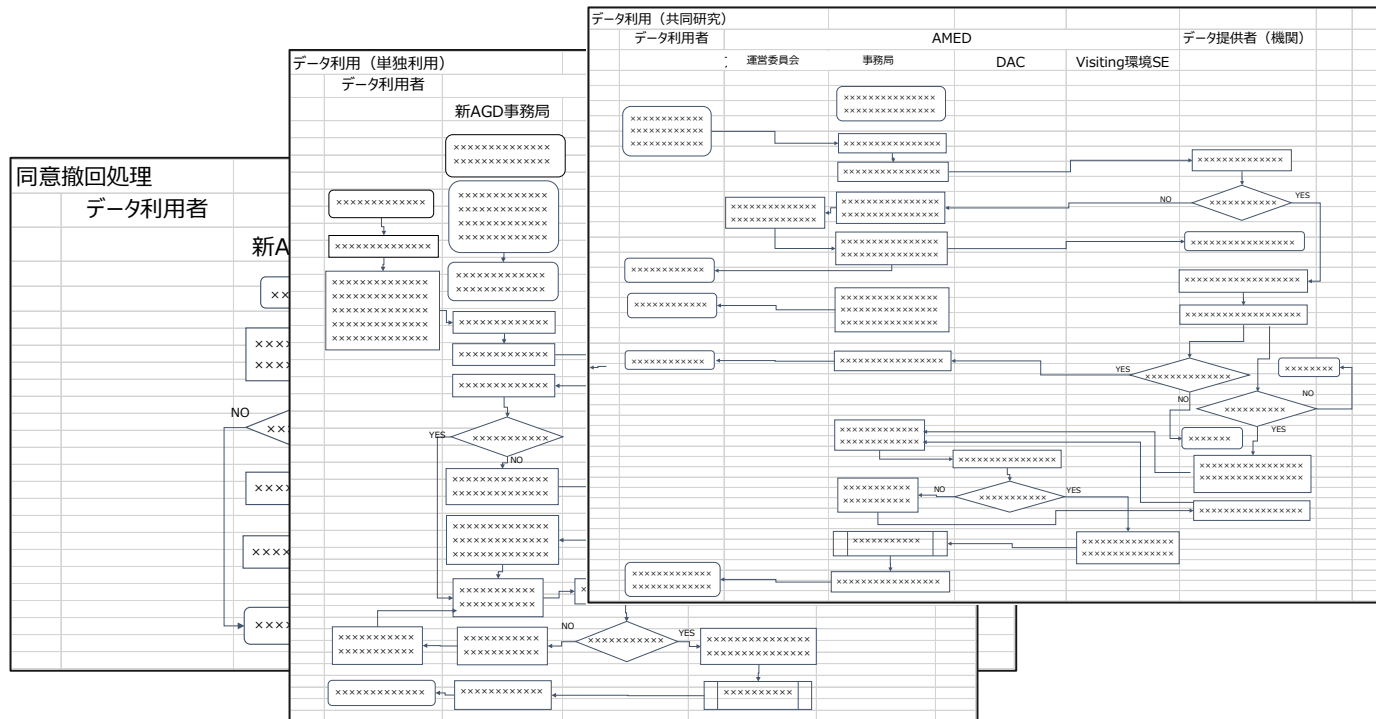
開発計画

	【A】 検索システム	【B】 データベース
機能概要	データベースとVisiting型計算環境に格納されるデータのメタデータとカタログを対象とした検索、および利用者の認証・承認機能。	任意のデータセットについて、当該データおよびメタデータとカタログの登録、保全、提供等を管理する機能、およびこれらの処理を記録し分析可能とする機能。ハードウェアを含む。
開発範囲	ソフトウェアのみ。Visiting型計算環境との連結を含む	ソフトウェアとハードウェア。検索システムおよびVisiting型計算環境との連結を含む
開発期間	R3年度	R4年度（予定）

③ One Stopな利用受付・審査サービス

● 処理フローの作成を進めている。

- i. 処理の全容
- ii. データ登録
- iii. データ利用申請を受け利用許可まで（利用者の単独利用）
- iv. データ利用申請を受け利用許可まで（データ提供者との共同研究利用）
- v. 同意撤回処理
- vi. CANNDs Visiting型計算環境の利用申請を受け利用許可まで



④ テスト利用によるVisiting計算環境の検証

<目的>

- Visiting型解析環境の動作検証を行い、R4年度に開始する本格利用に備える。

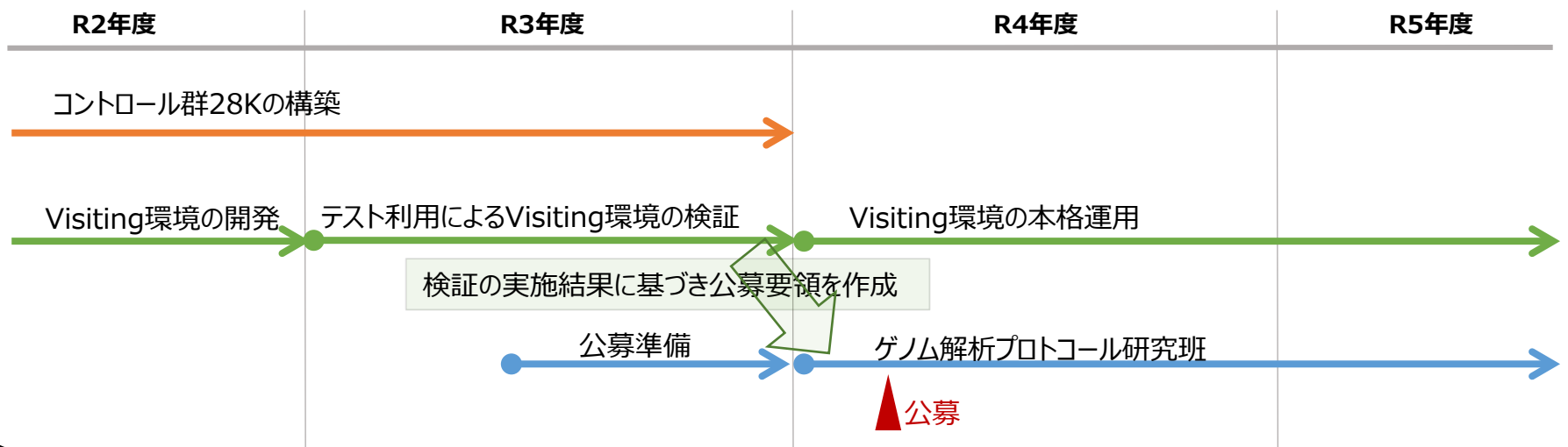
<実施内容>

- コントロール群の構築に従事した研究者による、Visiting型計算環境の動作検証を実施する。
 - 利用者認証の動作検証
 - ユーザーインターフェースの動作検証
 - 仮想マシン上のユーザ利用領域の動作検証
 - データアクセスの動作検証
 - 計算速度の検証
 - ネットワーク転送速度の検証
 - セキュリティの検証
 - 処理フローのテスト実施による検証
- 民間企業による試行的利用を通じて仕様改良のためのフィードバックを得る

⑤ 研究公募準備（ゲノム解析の均質化を含む）

- CANNDs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ルに基づき解析を行うことを推進するとともに、データサイエンティスト等の教育・育成を目的とした、ゲノム解析プロトコル研究班（若手研究者対象）の公募を検討する。

工程表



⑥ 高度セキュリティルームの設置

- CANNDs Visiting型計算環境の利活用においては、Genomics Englandと同様に、**自席からのアクセスを可能**とする計画である。コントロール群のゲノム情報と基本臨床情報（性別、年齢、出生地区/居住地区、疾患名）の利用もこれに準じる予定。
- 一方、がんや難病の全ゲノムデータと詳細な臨床情報等、**機微性が一段と高いデータ**をCANNDsで利活用するためには、一層高度なセキュリティが求められる。それに対応できるアクセス環境を整備するため、CANNDsのデータベースを担う連携拠点（東北メディカル・メガバンク機構、国立遺伝学研究所、国立がん研究センター、東京大学医科学研究所、国立国際医療研究センター）およびAMEDに**遠隔セキュリティルームを設置**する。

