

# 「全ゲノム解析等に係る厚生労働科学研究班」からの報告

令和4年3月2日

全ゲノム解析を基盤としたがんゲノム医療の実装に向けた患者還元、解析・データセンター、ELSI体制構築についての研究班

中釜 齊 (研究班長、国立がん研究センター・理事長)

## 患者還元WG

河野 隆志 (WG長、国立がん研究センター研究所・分野長)

織田 克利 (東京大学医学部附属病院・教授)

## 解析・データセンターWG

井元 清哉 (WG長、東京大学医科学研究所・教授)

白石 友一 (国立がん研究センター研究所・分野長)

## ELSI WG

横野 恵 (WG長、早稲田大学社会科学総合学術院・准教授)

田代 志門 (東北大学大学院文学研究科・准教授)

## 準備室WG

青木 一教 (WG長、国立がん研究センター・研究所・副所長)

徳永 勝士 (国立国際医療研究センター研究所・プロジェクト長)

## 目次

患者還元WGからの報告	P 3
解析・データセンターWGからの報告	P20
ELSI WG からの報告	P37

# 患者還元WG

○班長：河野隆志（国立がん研究センター）

大項目	項目		第3回	第4回	第5回	第6回	第7回	第8回
患者還元	ICT/AI技術	1 eコンセント等による参加者の利便性向上			○	●	☆	☆
	対象疾患、患者数	2 対象疾患および対象ステージ詳細	○	●			☆	☆
	医療機関要件	3 EP、全ゲノム解析体制の要件など	○	○	●		☆	☆
検体の保存・利活用	保管、管理ルール	4 採取、処理、移送、保管、管理SOP	○	○	●		☆	☆
シーケンス	受託要件	5 NGS台数、WGS実績等	○	●			☆	☆
	技術的要件	6 WGSデプス、RNA seq範囲	○	○	●		☆	☆
	QC	7 QC方法、タイミング等	○	○	●		☆	☆
データ利活用		8 アカデミア、産業界の役割について				●	☆	☆
		9 データ共有ルール（データシェアリングP）				●	☆	☆
		10 データ利活用審査委員会				●	☆	☆
解析・データセンター	ゲノム解析	11 ゲノムデータベース構築	○	○	●		☆	☆
		12 統一パイプライン	○	●			☆	☆
		13 高度な横断的解析（AI活用含む）			○	○	☆	☆
	臨床情報等の活用	14 臨床情報DB構築（API自動収集）	○	○	○	○	☆	☆
		15 レポート作成システム			○	○	☆	☆
	データ共有システム	16 研究支援システム			○	○	☆	☆
	集中管理システム	17 集中管理システム			○	○	☆	☆
	情報管理・システム構築に関して	18 情報管理			○	○	☆	☆
		19 システム開発や環境構築			○	○	☆	☆
人材育成	20 バイオインフォマティクソン等について			○	○	☆	☆	
ELSI	ICF	21 統一ICF挿入文作成	○	○	●		☆	☆
	ガイダンス	22 IC手法、2次所見・結果還元在り方等	○	●			☆	☆
	患者・市民視点	23 PPIの推進			○	○	☆	☆
	人材育成	24 遺伝カウンセリングなど			○	●	☆	☆

※ ○ 協議継続、●方針決定、☆AMED研究班の経験等を踏まえ必要な検討・修正

# 1. eコンセント等による参加者の利便性向上

# 1. ICT/AI技術：ICプロセスへのICT技術の導入(eコンセント等による参加者の利便性向上)

令和3年6月から施行された新統合指針においてもeコンセント(電子的IC: 電磁的方法によるインフォームド・コンセント)が明記され、実施する際の留意事項についてもガイダンスで示されている。 <https://www.mhlw.go.jp/content/000769923.pdf>

**適切なツールや通信環境を整えた上での導入を検討する。その際には、以下の利点・留意点を理解して行う必要がある。**  
なお、導入においては、説明・同意取得・記録保存の電子化 (e-consent) だけでなく、患者が主体となって情報管理が行えるオンラインプラットフォームの整備を合わせて検討すべきである。

## eコンセントの利点

- ・ 効率の向上  
ICFの内容の修正やトラッキング、一元的管理が容易
- ・ 参加者の理解の向上  
個々のペースに合わせて説明コンテンツを利用できる  
同意内容の振り返り・確認が容易になる  
家族等とも共有しやすい
- ・ 参加率の向上
- ・ リンク接触利便性の向上

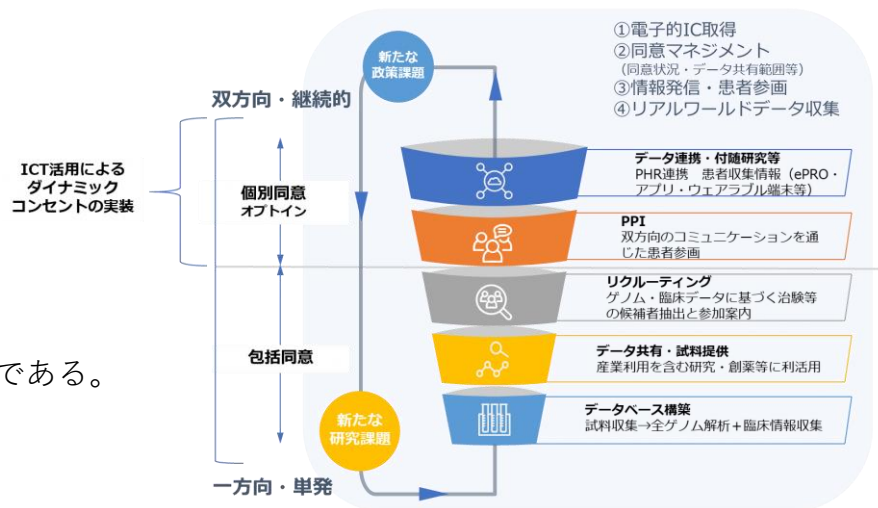
## 留意すべき点

- ・ 不慣れな参加者においては、内容の理解や意思表示が困難である。
- ・ 質問の機会が失われやすく、丁寧な対応が失われやすい。
- ・ 代諾者が必要な場合の対応や要件の確認が困難である。

## 医療現場導入に際しての検討すべき点

- ・ 病院でのIC情報の授受・管理の手法
- ・ データ・解析センターとの連携手段

## オンラインプラットフォームによる展開イメージ



(令和2年10月27日 がん全ゲノム解析等連絡調整会議 参考資料5より一部改変;  
<https://www.mhlw.go.jp/content/10901000/000688485.pdf>)

電子化ICの手法・範囲、情報連携の手法など、医療施設での実働性を考えた上での具体的な検討が必要  
→実働上の問題の抽出と開始に向けた具体案の策定を起点として、検討を開始すべき

2. 対象疾患および対象ステージ詳細
3. EP、全ゲノム解析体制の要件など
4. 採取、処理、移送、保管、管理SOP
24. 遺伝カウンセリングなど

## 2. 対象疾患、患者数：対象疾患および対象ステージ詳細

対象患者は再発頻度や学会意見等に従い、ゲノム情報の有用性を考えて対象症例を選出する。具体的には以下の項目を考慮する。

- ・（前向き解析の場合）新たに、患者に対して十分な説明のうえ、同意が得られること。
- ・手術、生検、採血（血液腫瘍）等により全ゲノム解析に必要な質・量の検体が得られること。
- ・手術等において根治の可能性が低い難治がんであること。
- ・解析開始時に生存しており、何らかの治療の提供が期待できる状態であること。
- ・ただし、希少がん・難治がん等で十分なゲノム情報が集積されていないがんについては、この限りではない。

## 4. 保管、管理ルール（採取、保管、移送等のSOP）：

日本病理学会「ゲノム研究用病理組織検体取扱い規程」に従うことを推奨する。

<https://genome.pathology.or.jp/houhou/index.html>

### 第2部 凍結組織検体の適切な採取・保管・移送方法（抜粋）

#### 【採取時間】

可及的に速やかにゲノム研究等に用いる組織検体を採取する（摘出後30分以内）。行えない場合には、手術検体を冷蔵庫（4℃）等に保管するものとする。4℃保管3時間以内を目安にして組織検体を採取する。

#### 【採取量】

病理診断に特段の支障がなく適切な採取部位が確保できる場合は、半小指頭大（1×0.5×0.3cm程度、50mg-100mg程度）の組織を採取する。

#### 【組織の処理】

採取した組織を1辺2-3mm角程度にまで細切する。望ましくは、2-3mm角組織片1個を1チューブに収納する。万一の温度変化に備える（液体窒素の膨張による事故を防ぐ）ため、スクリューキャップ付きのチューブを用いる。

#### 【急速凍結】

作成したチューブを液体窒素に浸漬し、急速凍結を行う。最も望ましくは、手術標本摘出後30分以内に急速凍結を実施する。

#### 【長期保管（一般的に5ないし10年以上）】

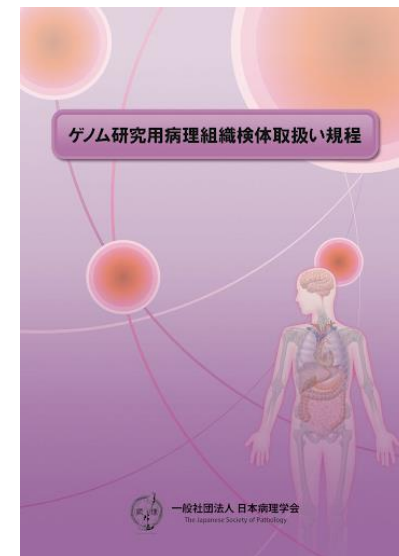
急速凍結したチューブは、最も望ましくは、研究使用に供すまで液体窒素保存容器（-180℃程度）に保管する。液体窒素保存容器（-180℃程度）にかえて、超低温槽（-80℃）を用いることも可能である。

#### 【移送】

温度保持が期待される非密閉移送容器（簡易的には発泡スチロール容器）にドライアイス等を充填し、冷凍便により運送業者に輸送を委託する。

#### 【試料の廃棄】

試料提供者から研究利用への同意の撤回があるときなど、病理凍結組織を廃棄する必要がある場合には、手術等に伴って発生する病理廃棄物（摘出臓器、組織、郭清に伴う皮膚等）と同様に廃棄する。

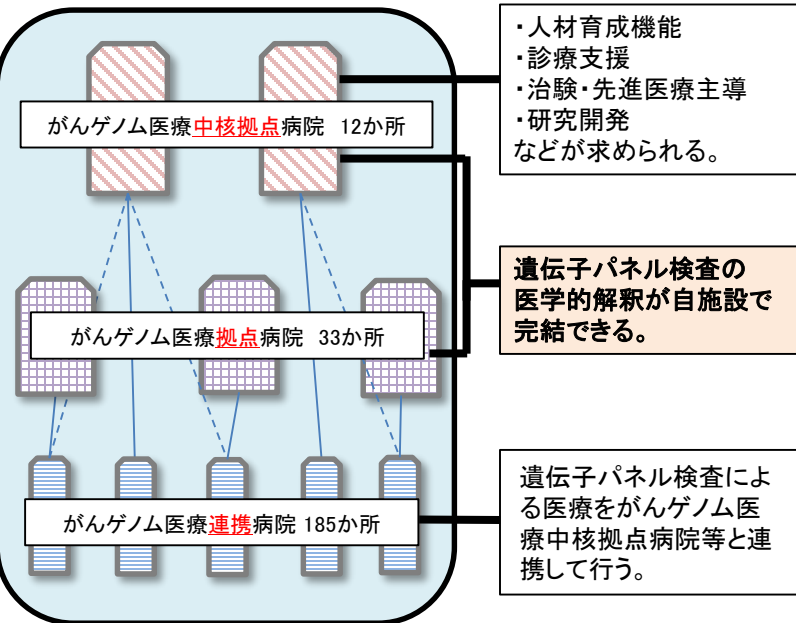


### 3. 医療機関要件：EP、全ゲノム解析体制の要件など

がんゲノム医療中核拠点を主体とし、中核拠点の施設要件をすべて満たしている等、準備\*の整ったがんゲノム医療拠点病院を含める。R3年度中に要件を整備し、R4年度からはこの要件を満たす医療機関で全ゲノム解析等を行う。

#### がんゲノム医療の提供体制 令和4年1月現在

第2回がんゲノム医療推進コンソーシアム  
運営会議(平成31年3月6日)資料1より改訂



#### \*以下が準備されていること

- ・がんゲノム医療に従事する医療者に対して、全ゲノム解析を含んだ必修の研修が行われ、ゲノムリテラシーの向上が図られていること。
- ・全ゲノム解析結果の科学的妥当性を判断できる体制が整っていること(以下の項目に習熟したゲノム研究者が複数名従事していること)
  - ▶ 解析データ(元データ含む)の確認と各種コールエラーの検出
  - ▶ ゲノム変化の解釈と意義づけ
  - ▶ 適切な他の手法によるゲノム変化の検証
- ・患者の同意の下で、臓器横断的に検体が、適切に保管・管理される体制が整備されていること。
- ・臨床研究中核病院、もしくは、それに準じた診療機能体制が整備されていること
- ・臨床試験・治験等の実施について適切な体制を備え、一定の実績を有していること

ゲノム医療を必要とするがん患者が、全国どこにいても、がんゲノム医療を受けられるよう段階的に、全ての都道府県でがんゲノム医療の提供が可能となることを目指す

全ゲノム解析で検出された遺伝子変異に基づいた薬剤投与等の**出口戦略の体制**を急ぎ構築すべき  
→ 前向き臨床試験、産業フォーラム参加企業との連携

### 24. 人材育成・遺伝カウンセリングなど

遺伝子カウンセリング人材の人数等については、患者還元A班の実経験及び次年度以降の全ゲノム解析実施医療機関数と併せて、考えていくべき。

また、本邦の全ゲノム解析について、患者や国民の理解を促すため、**患者や国民の参画の元**、以下の整備を行うべきである。

1. 補助資料としてのパンフレット、動画(概要と詳しいもの)など
2. ホームページ



5. NGS台数、WGS実績等
6. WGSデプス、RNA seq範囲
7. QC方法、タイミング等

## 5. シークエンス受託要件：NGS台数、WGS実績等

ゲノム解析受託施設は、品質保証、かつ、ゲノム解析もしくは検査の実績を重視して選出する。

- ・均一で高品質なシーケンスデータ確保のため実績を重視すべき。
- ・ゲノム解析の先進諸国との国際共同研究でも活用可能なシーケンスデータの取得が可能であること。
- ・遺伝子検査にかかる精度管理を実施している衛生検査所等でシーケンスを行うこと。  
ゲノム検査でISO15189 認定（公益財団法人日本適合性認定協会）、CAP-LAP 認定（米国臨床病理医協会、臨床検査プログラム）、CLIA 認定（CLIA認証検査室改善法）のいずれかを取得していることを重視する。
- ・前向き解析に向けて、検体回収や検体識別番号付けなど、運用面の対応も考慮する。

## 6. 技術的要件：WGSデプス、RNA seq範囲

### 7. QC方法、タイミング：標準手法によるシーケンスの場合

受託企業：ヒトゲノムマッピング前のデータを用いて質・量の評価を行い、基準値を満たすデータを取得する。

解析・データセンター：ヒトゲノムマッピング後のデータを用い、質・量の多面的評価を行う。

機関	受託企業		解析・データセンター	
QCタイミング	ヒトゲノム配列へのマッピング前に行う		ヒトゲノム配列へのマッピング後に行う	
	項目	基準値*	項目	方針
WGS	QV30/20以上の塩基割合	75%/90%以上	-	<ul style="list-style-type: none"> <li>・中央モニタリングに用いるとともに、各サンプルごとの値を研究者及び受託企業に返却</li> <li>・がん種や試料の種類、ライブラリー作成法、受託企業等の条件別に集計***</li> </ul>
	重複リード除去後の塩基数	N: 90G塩基以上 T: 360G塩基以上	-	
	-	-	マッピング率	
	-	-	重複率	
	-	-	インサート長	
	-	-	読み取り深度	
RNAseq	リード数	2,000万リード以上**	-	→次年度以降のデータ追加取得等の方針検討に利用
			アライメント率	
	RIN値	参考情報として収集	-	

\*試料の制限により、標準手法での委託でない際には、それに準じたQC基準を定める。

なお、当該基準値を超えたデータ取得を各研究班の予算内で行うことは可能である。

\*\*ポリA精製ライブラリー調整を標準手法とし、その下限を示す。

また、rRNA枯渇処理ライブラリー調整の際は、上記に見合うmRNA由来リードデータ量の取得を目標値とする。

\*\*\*厚労科研「がん全ゲノム解析等の推進に向けた患者還元、解析・データセンター、ELSI等に係る技術評価、体制構築についての研究」班において、各受託企業のシーケンス精度や、当該集計値及びマッピング前のQC値を用いた精度把握を行う。

8. アカデミア、産業界の役割について
9. データ共有ルール（データシェアリングポリシー）
10. データ利活用審査会

8. アカデミア、産業界の役割について
9. データ共有ルール（データシェアリングポリシー）
10. データ利活用審査会

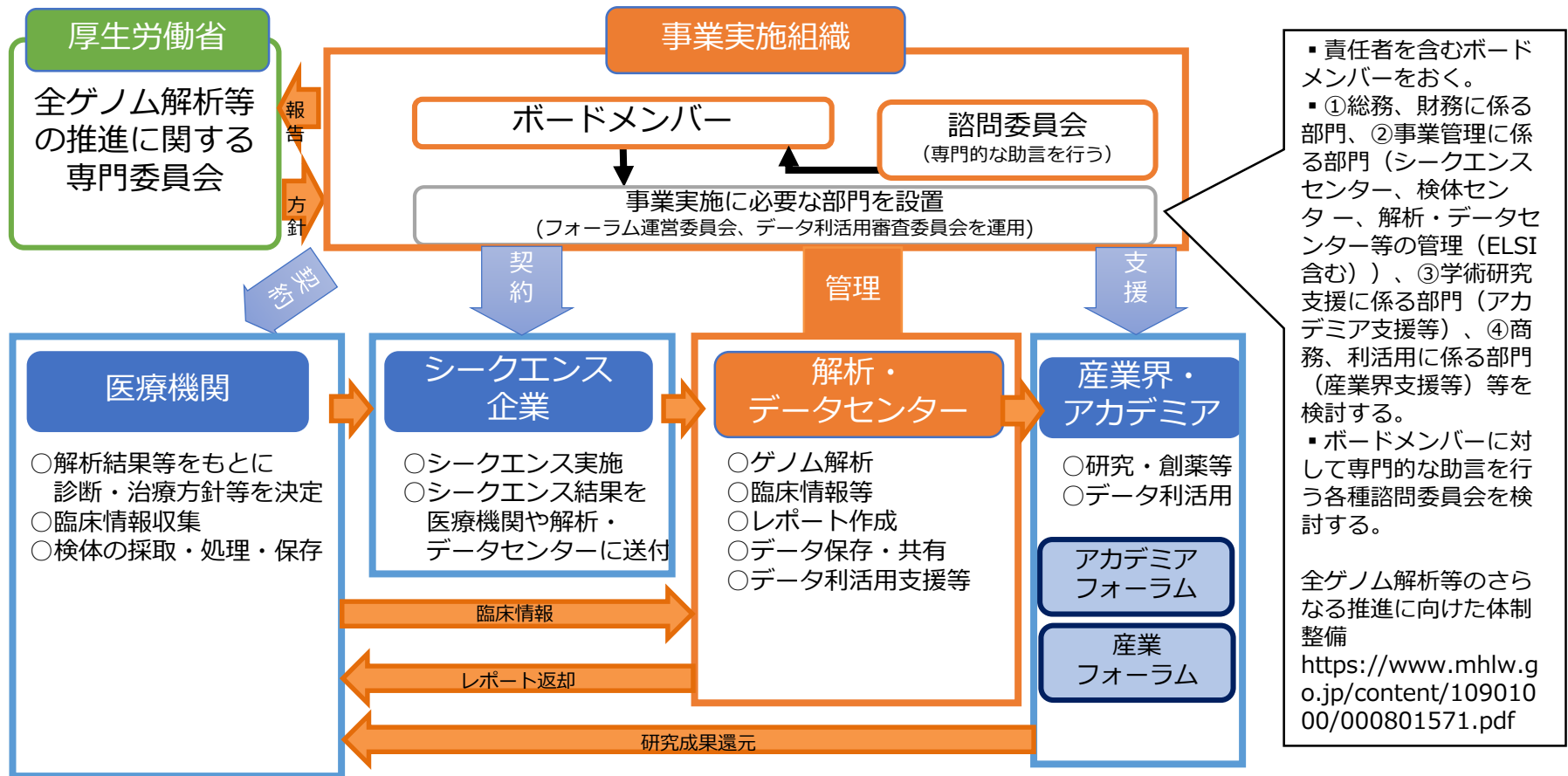
## 本邦のアカデミアフォーラム、産業フォーラム 1/2（案）

項目	アカデミアフォーラム	産業フォーラム
開始	2023年	2023年
概要	創薬や診断技術の研究開発を促進し、患者にいち早く成果を届けるために形成され、産学連携しながら全ゲノム解析等のデータ利活用推進を図る。 <b>事業実施組織は、フォーラム間・企業間の連携等を促す役割を持つ。</b>	
	主としてアカデミアが参画するフォーラム	主として産業界が参画するフォーラム
目標	全ゲノム解析等に係る学術的協議を行う 全ゲノム解析等に基づき、がんの臨床的理解を深め、予防、患者の予後を改善する	全ゲノム解析等に係るデータ利活用による創薬や診断技術、予防の研究および開発等を推進する
参加機関 (利用者)	全ゲノム解析等のデータの収集・解析を実施する研究機関・医療機関 上記の機関と共同研究を行う機関（企業含む）	左以外の医療機関・研究機関・企業（医療産業、非医療産業に関わらず、またベンチャー企業を含めてより多くの企業が参画できる）
利用料金	データ等の利活用については、データ利用者の負担を原則とする。ただし、データ利活用審査委員会がアカデミアのみの利用と認めた場合においては負担を減免するなどの利用者負担制度を検討する。年会費とデータ量毎の利用料等が考えられるが、価格については、運用コストを精査して決定する。	
登録システム (申請→審査→利活用可能となるまでの流れ)	<ul style="list-style-type: none"> <li>（AMED課題でのデータ取得者及び解析者）</li> <li>・フォーラムメンバーとする</li>   <li>（データ取得者及び解析者が新たに参加を求める者）</li> <li>・フォーラム運営委員会で審議・承認</li>   <li>（新規参加希望者）</li> <li>・研究計画概要をフォーラム運営委員会に提出し、審議・承認</li>   <li>・所属機関がフォーラム参加契約書に署名</li> <li>・アカウントの発行</li> <li>・情報ガバナンス研修の受講 （患者のプライバシーを守り、データの機密性とその安全性を維持する方法を理解）</li> </ul>	<ul style="list-style-type: none"> <li>（事業実施組織設立前の参加企業）</li> <li>・フォーラムメンバーとする</li>   <li>（新規参加希望企業）</li> <li>・研究計画概要をフォーラム運営委員会に提出し、審議・承認</li> <li>・契約と利用料支払い</li> <li>・アカウントの発行</li> <li>・情報ガバナンス研修の受講 （患者のプライバシーを守り、データの機密性とその安全性を維持する方法を理解）</li> <li>・計画立案を目的とした全データの俯瞰・簡易解析</li> <li>・詳細解析の研究計画を立案し、データ利活用審査委員会の承認</li> <li>・追加の契約と利用料支払い</li> <li>・詳細解析開始</li> </ul>

# 本邦のアカデミアフォーラム、産業フォーラム2/2 (案)

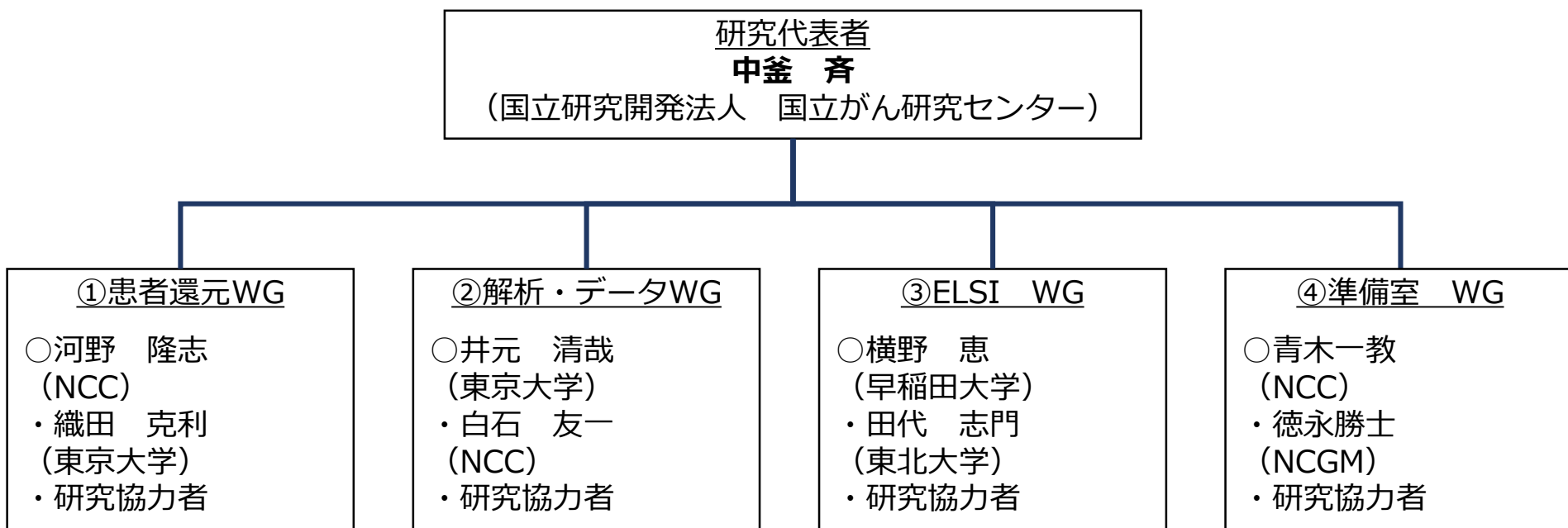
項目	アカデミアフォーラム	産業フォーラム
運営システム (ガバナンス等)	<b>事業実施組織</b> フォーラム間・企業間の連携等を促す役割を持つ。 医薬品開発の専門家を構成員とするなど、創薬を促す体制とする。 フォーラム運営委員会 両フォーラムの運営をつかさどる。両フォーラムへの参加の承認の決定を行う	
	リモートアクセスにより、利用者PCで解析が可能 (テレワークも可) 計画立案を目的とした全データの俯瞰・簡易解析*： 自由に行うことができる *俯瞰：各がん種のデータ蓄積状況を俯瞰すること。 *簡易解析：各がん種における遺伝子変異例陽性例数の把握など、詳細解析の立案に必要なデータ分析を行うこと	
	詳細解析： ・自データの解析：自由に行うことができる ・自データ以外を用いた解析：データ取得者との合意に基づいた共同研究ベースで行う データ利活用審査委員会への報告を必要とする ・利用料については、運用コストを精査して決定する。	詳細解析：年会費で1,000例までは解析可能 ・解析計画に関して、データ利活用審査委員会の承認を必要とする ・データ量毎の利用料 10Kゲノムまで利用可能：3,000-4,000万円/年(税込み) 価格については、運用コストを精査して決定する
共有されるデータ	臨床情報、シーケンス情報を解析・データセンターで詳細に分析し、データベース化したものを共有する ゲノムデータ：FASTQ, BAM/CRAM, VCF, <b>遺伝子変化の基本的な意義付け情報</b> 上記以外のオミクスデータ (予定) 臨床情報：がん種/疾患名、性別、年齢、予後などEDC収集全項目	
共有されるまでの制限期間	起始ポイント*到達後すみやかに提供開始 *100例程度のデータ登録時点 (希少がんは別に定める)	
知財・出願	自由な出願・知財の占有ができる。	
	出願に関して、フォーラムへの届け出を行う	出願に関して、フォーラムへの届け出は不要
成果公表	論文等での成果の公表に関して、フォーラムへの届け出を行う	論文等での成果の公表については、制限期間後とする 但し、出願公開により公表される場合 (出願公開制度による出願後18カ月の公開) を除く 論文等での成果の公表に関して、フォーラムへの届け出を行う

# 全ゲノム解析等の実施体制の将来像（改変案）



事業実施組織の在り方を検討するWGを当厚労研究班内に新たに設置  
 WGはがんと難病の両領域の人員からなるものとする

# 「全ゲノム解析を基盤としたがんゲノム医療の実装に向けた患者還元、解析・データセンター、ELSI体制構築についての研究」班 (厚労科研中釜班) 体制



※①、②、③のWGは、様々な分野の専門家、関係学会、患者団体、データ利活用団体等と意見交換を行い、取りまとめた上、全ゲノム解析等の推進に関する専門委員会に報告する。

※④のWGは事業実施組織発足に向けた**事業実施組織準備室の設置のための検討を行う**。

# 患者還元に関する要点

## 患者還元

- ・ 前向き臨床試験等、結果を患者に還元する仕組み (出口戦略) の構築
- ・ 出口戦略を見据えた患者エンロールと適切な結果開示
- ・ 参加者のエンゲージメントの向上 (動画、ホームページ、eコンセント、患者参画)

## シークエンス解析と検体保存

- ・ がん種等に配慮したSOPに基づく試料採取・出検・解析
- ・ ヒトゲノムマッピング前後のデータを用いたQC
- ・ 解析結果の確認や将来の利活用のための残余検体保存・把握

## データ利活用

- ・ QC済データおよび意義付け変異情報を共有する仕組みの構築
- ・ 公正な利用審査及びデータ利用者への知財の帰属
- ・ 創薬等の研究・開発を促進する事業実施組織の発足



# データ共有ルール（データシェアリングポリシー）

- ・収集したデータ（臨床情報、ゲノムデータ）の迅速な利用を実現するため、必要な事項を定める。  
当該ルールについては、国際的動向や研究の進展にともない、適宜、方針修正等を行う。
- ・収集したデータの共有に際しては、取得された臨床情報、シーケンス情報を解析・データセンターで詳細に分析し、データベース化したものを共有する。  
アカデミアが参画するフォーラム（アカデミアフォーラム）、産業界が参画するフォーラム（産業フォーラム）を形成してデータ共有を行う。
- ・**アカデミアフォーラム**  
データ取得者とデータ解析者で構成される。  
計画立案を目的とした全データの俯瞰・簡易解析が可能である。  
自データ以外を用いた詳細解析には、データ取得者との合意とデータ利活用審査委員会への報告を必要とする。  
自由な出願・知財の占有ができる。  
論文等での成果公表、出願においては、データの公的DB移行の時期によらず、フォーラムへの届け出を行う。

- ・**産業フォーラム**  
計画立案を目的とした全データの俯瞰・簡易解析が可能である(有償)。  
詳細解析には、データ利活用審査委員会の承認を必要とする(有償)。  
自由な出願・知財の占有ができる。出願に関して、フォーラムへの届け出は不要。  
成果を公表できるが、公表まで一定の制限期間（24-30ヶ月）※1,※2,※3を設ける。

※1 各がん種において100例程度のデータ登録時点（希少がんは別途定める）を、起始ポイントとする。

※2 出願公開により公表される場合（出願公開制度による出願後18カ月の公開）を除く。

※3 論文公表等でデータ取得者・解析者が非公表の必要性が乏しいと判断した場合は、制限期間を24か月未満に短縮する。

成果公表においては、データの公的DB移行の時期によらず、フォーラムへの届け出を行う。

- ・一定期間(24-30ヶ月)以降、ゲノムデータおよび基本的な診療情報は公的DBに登録する。  
詳細な診療情報は個人同定の配慮から登録しない。  
詳細な臨床情報の利用には、アカデミアフォーラムまたは産業フォーラムへの参加を必要とする。  
利用者は知財を占有できる。  
論文公表までの制限期間なし。

# データ利活用ポリシー骨子 (1/2)

## 目的

- ・本ポリシーは、収集したデータ（臨床情報、ゲノムデータ）を幅広く、公正かつ円滑に利活用するために必要な事項を定める。

## データ利活用について

- ・データ利活用審査委員会の承認のもと、利用許諾等を行う。
- ・学術研究や医薬品等の開発、科学的なエビデンスに基づく予防を目的とする申請に対し許諾される。
- ・利用者は国内外※の学術研究機関や企業を想定する。  
※日本と同等の水準にあると認められている個人情報の保護に関する制度を有している国または地域
- ・データの利用許諾等を受ける者がポリシーの定めに違反した場合、その者の氏名の公表やデータの利用許諾等の停止、新たな利用申請の拒絶、差止請求、損害賠償請求等の措置を講ずることができる。
- ・データの利用許諾等を受ける者が、自己以外の第三者に当該データの提供や転売等を行うことを認めない。ただし、自己が本ポリシー及びデータの利用許諾に関する契約で負うのと同等の守秘義務を課した上での自己の委託先や関係会社への提供についてはこの限りではない。

## データ利活用審査委員会

- ・第三者への利用許諾等について公平性を担保することを目的に設置する。
- ・データ利活用審査委員会のメンバーは男女両性を含むものとする。そのほかの要件は別途定める。
- ・データ利活用審査委員会のメンバーは、適切なCOI管理を受ける。また利活用審査会メンバーとして知り得た情報については守秘義務を負うものとする。

# データ利活用ポリシー骨子 (2/2)

## データ利活用審査委員会 (続き)

- ・利用許諾等の可否に関して以下の審査項目について審査する。
  - 申請者による利用目的が、学術研究や医薬品等の開発、科学的なエビデンスに基づく予防を目的とするものであること
  - 個人\*・血縁者の同定、血縁関係の存在・不存在の確認、又は個人、小集団そして地域に対して不利益をもたらす可能性のある利用にあたらぬこと
  - \*臨床試験参加など、患者還元を目的とする際はこの限りではない。
  - 研究計画の科学的妥当性及び利用するデータ範囲の妥当性があること
  - 申請者が研究計画を遂行するにあたり、十分な実績又は能力を有していること
  - 申請者が所属する機関が十分な研究設備や情報の保管・廃棄などの管理体制を有していること
  - その他、データ利活用審査委員会メンバーが必要と判断する事項
- ・データの利用者は、データ利活用審査委員会より許諾を受けた利用目的以外の目的にデータを使用してはならない。

## 知的財産権

- ・データの利用から生じる知的財産及び知的財産権は、創出した利用者に帰属するものとする。

## 公表

- ・利用者は、当該データを用いた研究成果を公表することができる。
- ・データのうち患者臨床情報について、成果発表に必要な範囲で論文等に掲載できる。
- ・(特定) 個人の識別につながる可能性等について十分に検討し、その可能性がある場合には発表前に再度利活用審査会に報告し協議することとする。

## 秘密保持

- ・データの利用許諾等を受ける者は、当該データに係る情報を秘密として管理し、第三者に開示又は提供をしてはならない。ただし、データ利活用審査委員会において許可が下りた利用目的の範囲内においては、当該利用者自らが負うのと同等の守秘義務を課すことを条件に解析委託先等にも開示することができる。
- ・データの利用許諾等を受ける者は、データを取り扱うにあたり、不正アクセス、紛失、破壊、漏洩などのリスクに対し、合理的な安全対策を講じなければならない。

## 情報公開

- ・データの利用許諾等を行った相手先の名称等について公開する。

# 解析・データセンターWG

○班長：井元清哉（東京大学医科学研究所）

大項目	項目		第3回	第4回	第5回	第6回	第7回	第8回	
患者還元	ICT/AI技術	1	eコンセント等による参加者の利便性向上			○	●	☆	☆
	対象疾患、患者数	2	対象疾患および対象ステージ詳細	○	●			☆	☆
	医療機関要件	3	EP、全ゲノム解析体制の要件など	○	○	●		☆	☆
検体の保存・利活用	保管、管理ルール	4	採取、処理、移送、保管、管理SOP	○	○	●		☆	☆
シーケンス	受託要件	5	NGS台数、WGS実績等	○	●			☆	☆
	技術的要件	6	WGSデプス、RNA seq範囲	○	○	●		☆	☆
	QC	7	QC方法、タイミング等	○	○	●		☆	☆
データ利活用		8	アカデミア、産業界の役割について				●	☆	☆
		9	データ共有ルール（データシェアリングP）				●	☆	☆
		10	データ利活用審査委員会				●	☆	☆
解析・データセンター	ゲノム解析	11	ゲノムデータベース構築	○	○	●		☆	☆
		12	統一パイプライン	○	●			☆	☆
		13	高度な横断的解析（AI活用含む）			○	○	☆	☆
	臨床情報等の活用	14	臨床情報DB構築（API自動収集）	○	○	○	○	☆	☆
		15	レポート作成システム			○	○	☆	☆
	データ共有システム	16	研究支援システム			○	○	☆	☆
	集中管理システム	17	集中管理システム			○	○	☆	☆
	情報管理・システム構築に関して	18	情報管理			○	○	☆	☆
		19	システム開発や環境構築			○	○	☆	☆
人材育成	20	バイオインフォマティシャン等について			○	○	☆	☆	
ELSI	ICF	21	統一ICF挿入文作成	○	○	●		☆	☆
	ガイダンス	22	IC手法、2次所見・結果還元在り方等	○	●			☆	☆
	患者・市民視点	23	PPIの推進			○	○	☆	☆
	人材育成	24	遺伝カウンセリングなど			○	●	☆	☆

※ ○ 協議継続、●方針決定、☆AMED研究班の経験等を踏まえ必要な検討・修正

## 11. ゲノムデータベース構築

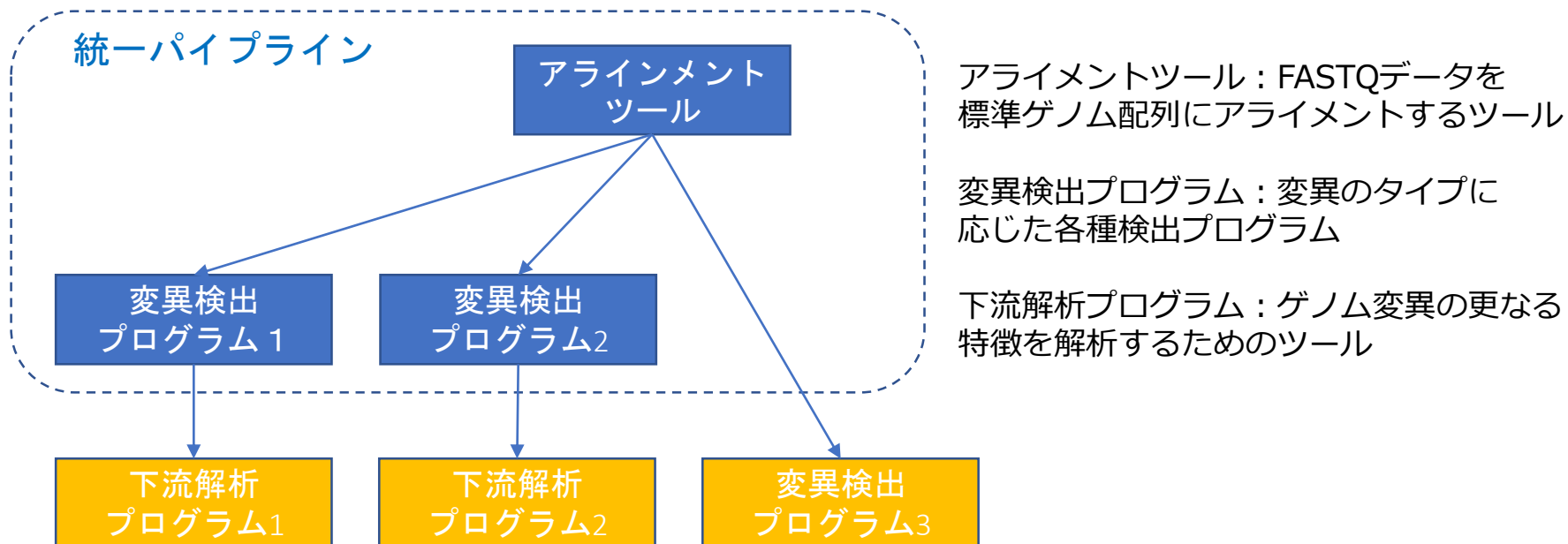
シーケンスセンターからのシーケンスデータ（FASTQデータ）、統一解析パイプラインによって生成されたbamファイル、変異情報（VCFファイル）は、解析・データセンターでデータベース化する。

- データのアップロードなどを管理できるテクノロジー窓口の設置が必要。
- 複数のシーケンスセンター、医療機関とのデータのやり取りを可能とするために、データ転送・フォーマットの方式を規格化を行う。
- 年間1万症例程度の全ゲノムシーケンスデータ（腫瘍120x、正常30x）の一次解析（マッピング、バリアントコール）の結果を保存できる体制・基盤を構築すること。
- 一次解析した結果は、データ提供機関と速やかに共有すること。
- 解析の進捗については、ゲノムデータベースに付随するデータポータルを通してデータ提供機関と共有できる仕組みを検討する。
- 使用頻度に応じたデータ保管形態の工夫による保管コストの削減と分散保管によるセキュリティ確保が必要である。
- **計算リソース**、データストレージについては、既存のリソースを活用しつつ、クラウドへの移行を検討すること。
- 研究・創薬等における利活用についてのシステム、ルールを確立すること。
- 情報管理におけるセキュリティ要件を満たすこと。

## 12. 統一パイプライン

### ツールについて

- 統一パイプラインを制定する目的は、多くの研究者が行う最大公約数的な解析処理を中央で一括して行うことで解析・データ共有の効率化を図ることである。
- 解析のメニュー・パラメータについては、国内外の動向を鑑み、将来のデータ共有、各種データベースとの適合性を図るために、一般的なものを選定すること。
- FASTQデータのアライメント率等の品質に関する項目を**計算**する機能を有すること。
- 統一パイプラインでの処理内容はオープンソースで管理される。
- 統一パイプラインの品質を保つために、中長期的に管理・運営をする枠組みを考案する必要がある。
- 統一パイプラインにおいて使用されているツール群においては、定期的に見直しを検討する必要がある。また、ツールがバージョンアップした際には、統一パイプラインにおいてもバージョンアップする必要性について検討する機会を設ける必要がある。
- 下流解析プログラムが多くの研究者の需要を満たすと考えられ、多数検体で安定的な計算が可能である場合は、プログラムを統一解析パイプラインに組み込むことを検討する。



## 12. 統一パイプライン（続き）

### ゲノム解析について

- 各々の研究班、がん種での個別の解析目的のために、独自の後処理・偽陽性フィルタ等を加えることが想定される。変異検出ツールの結果の後処理・偽陽性を除去するフィルタリング方法などは解析センターと各研究班で連携して協力して構築しつつ、最終的な判断はユーザーが行うことができる体制を構築すること。
- 生殖細胞系列変異においては、データの偏り等に十分注意を払った解析を検討すること。劣性遺伝性疾患に関する報告など、その範囲については、ELSIと十分に検討すること。
- 現在のインフラを有効活用しつつ本事業を進めるため、スタート時はスパコンによる解析を先行で進めるが、順次クラウドに移行する。
- オンプレミスのスパコン上でもGrid Engine等で計算ジョブを並列差配し、処理を分割し複数の計算ノードで解析が行える環境を構築し、スムーズにクラウドに移行できるよう検討すること。イルミナのシステムなどを参考にしながら、サプライチェーンリスクへの対応としてひとつのベンダーに依存しない高度な並列化処理が可能な形式でのシステム構築を進めていくことが必要。
- 解析処理は、同時並行で複数の処理が並列に実施できる実装とする。さらに、ワークフロー言語、仮想コンテナ技術、コンテナオーケストレーション技術等を活用し、効率的な実行形式を検討すること。

# クラウドにおけるゲノム解析（ゲノムデータベース構築・統一パイプライン）システム基盤のプロトタイピングのための技術仕様書案

統一解析パイプラインにおける性能や可用性、セキュリティ及び拡張性の要件を満たしうるクラウドサービスより、複数のサービスを選定し、性能面、費用面を比較検討し、その他課題を抽出すること。

## 性能要件の案

- 解析パイプラインを構成するツールは、コンテナイメージで管理すること。
- 全ゲノムシーケンスデータ解析（T:360G塩基以上、N:90G塩基以上のペア検体）における一次解析、およびデータの品質チェックの作業を、オンプレミス環境の環境と同等程度の時間で実行できること。

## 可用性要件の案

- シーケンスセンターとデータセンター、診療機関とデータセンター間を接続するネットワークは、24時間365日接続可能な状態とする。
- ネットワーク稼働率（ネットワーク全体の稼働）と回線稼働率（接続回線毎の稼働）は共に99.9%以上とする。
- ストレージ、データベースサービスともにスケールアウト（自動での拡張）が可能なこと。

（次ページへ）



(続き)

システム基盤のプロトタイピングにおいては、以下の観点でクラウドサービス毎に比較検討を行うこと。

### 性能検証

- 最低限の性能要件の条件のもと、解析速度を単純比較
- シークエンス量などの条件変更による、計算効率・費用の変化
- 性能向上を阻害する要因の特定

### 費用検証

- データのバックアップ、変異のアノテーションなど種々の用途を考慮したFASTQ、BAM (CRAM) 、VCFなどのゲノムデータ保管費用の比較
- 解析後データをオンプレミスの環境にダウンロードし保管をする場合と、クラウド上にデータ保管を続ける場合の費用を比較
- 月次に解析作業を同時並行で1000 検体程度行うことを条件とし、月次のクラウド費用を比較

### 有用性検証 (定性的)

- システム構築上の課題
- システム運用 (システム監視やデータバックアップ、ユーザビリティ・自動化、リスクなど) における課題整理
- システム障害やセキュリティインシデントへの対応の状況
- システムの証跡管理の状況
- システムの構成管理の状況 (アップデーター更新などの保守)
- 費用管理

### 13. ゲノム解析: 高度な横断的解析 (AI活用含む)

がん種横断的解析など高度な2次解析を研究者が実施するためのシステムの検討・構築を行う。AIの活用を通して患者還元につなげ、更に解析・データセンターの高度化に寄与するシステムにすること。

- 計算機のキャパシティががん種横断的解析のボトルネックにならないよう、クラウドとオンプレミスのハイブリッド解析基盤を構築すること。
- 分散処理やコンテナオーケストレーションなどの先進的な技術を活用し、より解析性能を向上するための調査研究を行うこと。
- サービス事業者が管理して提供しているマネージドサービスや、マネージドサービスとは独立したフレームワークを活用する方式など、種々の方法についてのメリット・デメリットを整理すること。また、解析効率のチューニングを費用・計算時間などの観点から検討すること。
- 上記の項目について、プロトタイプシステムの構築を通じた検証すること。検証においては、解析症例数やシーケンスデータ量などの増加など現実的に予測されるいくつかの設定での検証を行うこと。
- データ転送の方式なども検討し最適な方法を検証すること。
- 全ゲノムシーケンスデータ、RNAシーケンスデータに加え、ロングリードや一細胞シーケンスデータなど最新の技術を用い取得されたデータの解析に耐える頑健な解析基盤を構築すること。
- AI活用は、患者還元につながるテーマを優先して進めること。一方、解析・データセンターにおける人材確保の面からも、魅力ある研究として多くの研究者がAI開発に取り組める基盤を構築すること。そのために、AIが利用可能なようにデータのフォーマットを統一化しておくこと。
- 人工知能の最適化に必要な人工知能維持開発基盤の構築も進めること。
- 各研究班と連携し、がん種の特徴を捉えたAIによる解析精度向上に向けたAIモデルを開発すること。
- AIモデルの質や性能の評価を実施すること。
- 予後や死亡情報などのアウトカムの情報が重要である。アウトカム情報の収集、整備も合わせて考慮すること。
- 患者還元につながるよう変異に該当する薬剤・治療法を自動抽出できる機能等についても検討すること。

## 14. 臨床情報DB構築（API自動収集）

臨床情報収集をfeasibleにするためのAPI自動収集方式、エキスパートパネルにおいて活用される臨床情報データベースのプロトタイピングと課題、改善点の抽出を実施する。

### API自動収集方式

- 臨床情報のデータ形式の標準化をXMLやJSONベースで実施すること。
- 臨床情報の収集においては、臨床現場で feasible であることが必要。API自動収集については、可能な施設より実行を開始する。その際に、REST-APIを活用することを前提とし、本事業に必要な臨床データ項目やデータの送受信を机上ではなく、電子的な構文や送受信の仕組みを開発する。また、特定のアプリケーション技術やライセンスを活用することなく、オープンなAPIとして実現する。そのための、制度的な課題なども研究する。
- API自動収集に取り組む他の研究プロジェクトの動向を踏まえ、開発・構築を進めること。
- 電子カルテから臨床情報を転送するためのオープンなデータソースAPIとデータアクセスを行うAPIを開発すること。
- 特定の電子カルテベンダーに限定することなく、複数の医療機関が多様に保有する電子カルテのデータを標準形式に変換するETLなどのプログラムのプロトタイピングを行うこと。
- 臨床情報の送受信について、プッシュ型とプル型の比較検討を行い、プロトタイプ構築を通じて、課題や改善事項を検証すること。
- 情報管理におけるセキュリティ要件を満たすこと。

### 臨床情報データベース

- 解析・データセンターで収集する臨床情報データベースは、クラウドサービスを利用すること。
- 臨床データベース構造の記述方式について、複数の方式の比較検討を行い、性能要件に合致するものを選択すること。
- リレーショナルデータベースとNoSQLという言葉に代表される非構造化データベースについて、先進的な方式の利活用方式を鑑みつつ、比較検証すること。
- エキスパートパネルが利用しやすいユーザビリティや検索等の機能性能を確保できるものであること。
- プロトタイプシステムを構築し、課題や改善事項を検証すること。
- 情報管理におけるセキュリティ要件を満たすこと。
- 収集する臨床情報項目案については別紙を参照すること。

## 15. レポート作成システム

患者還元班と連携したレポート作成システムのプロトタイピングと課題、改善点の抽出を実施する。

- 各種ゲノム解析結果、臨床情報を抽出するためのAPIの整備を行い、単一の企業によるログインを防ぎつつ開かれた形でレポート生成システムの開発が可能な体制について検討すること。
- 検証にあたり、データセンターと医療機関の役割分担、責任分界点、レポートの目視確認・問い合わせ対応などを含めた想定される運用手順、運用にかかる人員やその役割、体制などをはじめ、課題を抽出すること。
- 変異情報を有効活用して患者還元するためのknowledge baseの構築、またそのアップデートのために必要となる体制、開発担当者の専門性などについての課題を抽出すること。
- VUSについては、機械学習やAI等を含め様々な形式で評価方法を含め検討すること。
- レポート自体に求められる要件としては以下が考えられる。
  - レポート作成システムの構築に当たっては、（1）既存の遺伝子パネル検査等レベル、（2）全エクソン解析等レベル、（3）構造異常や既知の範囲を超えた研究的解析レベル、等にレベル分けを行って段階的に患者還元を行うこと。
  - 当該遺伝子変異に臨床的意味づけを行い、治験情報等を含めたレポートであること。
  - 臨床的意義づけについては、体細胞変異についてはASCO、CAP、AMPによる4段階の分類など、生殖細胞系列変異についてはACMG-AMPガイドラインによる病原性分類なども参考としレポートに反映させることを検討すること。
  - ClinVarやCOSMICなどのデータベース情報についてもエキスパートパネルでの使いやすさを考慮して付与すること。
  - 治験情報については、日本の治験データベース（UMIN臨床試験登録システム、JapicCTI、治験促進センター臨床試験登録システム、臨床研究実施計画・研究概要公開システム）に加え米国ClinicalTrial.govの情報も合わせて検索できるシステムを検討すること。

## 16. 研究支援システム

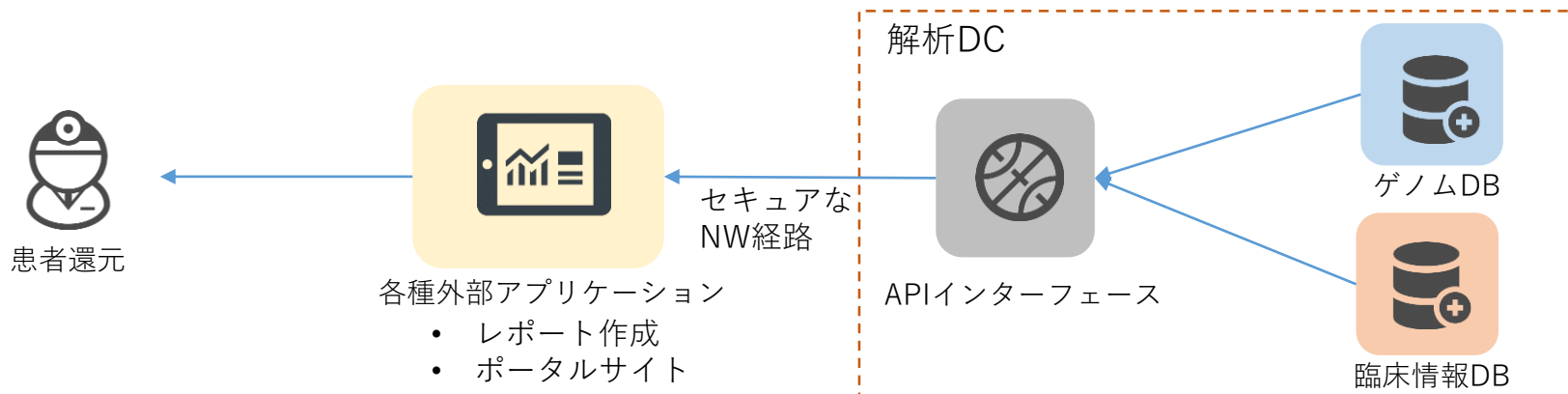
全ゲノム情報を患者還元するためのデータ解析・提供・保管を実施する研究・創薬等の拠点として研究開発が円滑に進むためのシステムを構築すること。

### 研究支援システム

- 研究班やフォーラムが利用可能なデータ共有のシステムをオンプレミスやクラウド上に構築すること。
- 共有データを解析するための環境を整備すること。
- その際、解析プログラムのインストール支援や大規模解析の実行支援、データの利活用促進に努めること。

### オープンAPIによるゲノムデータベース、臨床情報データベースの利活用促進

- データ利活用の最大化を目指すべく、ゲノム・臨床情報DBから必要な情報を抽出するオープンAPIの整備を行い、外部のアプリケーションとの連携を可能とすること。
- 解析・データセンターにおいて、簡便な解析の結果が可視化されたポータルサイト、レポート作成アプリケーションなどのモデル開発を行うこと。
- ポータルサイト、レポート作成アプリケーションなどが満たすべきセキュリティ要件の作成を行うこと。
- 企業やアカデミアが独自の技術やアイデア、データソースによりレポート作成を行うことができるよう、APIの活用を通して参入障壁を下げるように努めること。
- 解析・データセンターは種々の解析リソース・環境の整備により、各々の医療機関が自身の責任をもってレポートを作成・返却できる環境を整えること。



## 17. 集中管理システム

解析・データセンターにおいて、全ゲノムデータ、臨床情報、検体情報等について集中管理できるシステムを構築すること。

- 検体情報（採取日、種類、保存場所、残量等）、全ゲノムデータ（解析・データセンターにおけるデータ受領、QC結果、解析状況、レポート提出状況を含む）、臨床情報、同意の種別（共同研究での試料分譲の可否、産業界単独利用での試料の分譲の可否、意思変更の状況等）を一元的に管理・運営するシステムとすること。
- 現在、全ゲノム等実行計画に沿ってがんと難病のプロジェクトが進められている。特にデータ公開時においては、このような複数のプロジェクトに共通のIDフォーマットを検討すること。
- フォーラムにおけるデータ利用についても、各データセットの起点ポイントを管理し、起始ポイント\*到達後、速やかにフォーラムメンバーからのデータ俯瞰、利用申請やアクセス権付与、利用状況等を集中管理することのできるシステムであること。
- ゲノムデータおよび基本的な診療情報を公共データベースに登録する制限期間、すなわち、起点ポイントから少なくとも24か月、30か月は超えない期間、を管理し、制限期間到達後は公共データベースにデータ登録すること。
- 検体の取り違い防止などの観点から、検体情報が常にゲノム情報、臨床情報のデータベースと紐付けが可能となるように管理する仕組みが必要。
- 研究同意時の本人確認、同意撤回があった際のデータの削除等、トレーサビリティを確保する技術の具体化を検討すること。
- 新規の患者からの検体については、既存の施設を用いて一括管理することが可能な仕組みを構築すること。
- **患者さんが治験に参加する際など必要が生じた際には、検体情報を共有できる仕組みを検討すること。**
- 保存検体から新たなデータを取得する必要がある研究テーマが申請された際には、残検体・余剰検体の状況から研究の実施を支援できる体制の構築が必要。
- **検体の保管や取り出しは、自動化されたシステムを検討すること。**

\*100例程度のデータ登録時点（希少がんは別に定める）

## 18. 情報管理

治療に直結する重要インフラとして臨床情報、ゲノムデータへのアクセスコントロールやID管理、ログ管理、秘密分散技術を用いたデータ保管、セキュリティ検知・対策意思決定プロセスを検討する必要がある。

### 情報管理

- 情報管理に関しては、政府統一基準群や医療安全性ガイドライン、本事業における情報セキュリティに関する取り決め事項に従った情報管理を行う必要があり、定期的な監査を検討する。  
認証機能の強化（多要素認証の導入）、データの無害化の方法、不正アクセスの検知や自動遮断等についての方法論を検討すること。また、基盤運用の負担軽減に必要な運用管理の自動化も考慮すること。
- データのアップロードなどを管理できるテクノロジー窓口の設置が必要。
- 臨床情報やゲノムデータへのアクセスは、データ利用者の管理を徹底すること。そのために、最適な構成も複数案より比較検証し、課題や改善事項を導出すること。

### データ保管

- セキュリティに十分に留意した形式での臨床情報やゲノムデータを保管することを検討すること。そのために秘密分散技術をなど、複数の技術を調査、活用して相互運用性や拡張性などの課題を把握すること。
- ゲノムデータ保管にあたっては、費用面に留意しクラウドとオンプレミスのハイブリッドなど様々な方式を検討すること。
- ゲノムデータの保管に関しては、R3年度においてはオンプレミス環境においてデータは保管する。R4年度には一部クラウドでの保管を検証する。臨床情報に関しては、R3年度よりクラウドにて保管する。

## セキュリティ要件

- 情報管理に関しては、政府統一基準群や医療安全性ガイドライン、本事業における情報セキュリティに関する取り決め事項に従った情報管理を行う必要があり、定期的な監査を検討する。
- 海外リージョンとIPファイアーウォールによりネットワークアクセスを遮断できること。
- ISMAP（政府のクラウド安全性評価）への登録を行っている、もしくは登録申請中であること。
- セキュリティ監視や検知・遮断のために必要なファイアーウォールをクラウド上に配置できること。
- データアクセスに関するログ保管、ログを活用した相関分析ルールによる不審な挙動の監視や発報の仕組みを検討すること。その際、誤検知の最適化のための相関分析の定期的な最適化を自動的に行うための最適な方法を検証すること。
- セキュリティ検知にあたっては、クラウドサービスは、ペイロード（ネットワークを通過するパケット通信の内容）は確認も検知もできないことと、ログベースでの監視であるため、完全なリアルタイムな把握ではないことを回避するための方法も検証すること。
- インシデントが発生してから切り分け、対策、報告を行うためのセキュリティ対策意思決定プロセスをより早く行うための自動化と、誤検知や過検知を回避し、最適な方法で検知を行うための方法も構築し、検証すること。
- 利用者がコンピュータウイルスに感染した場合などのペナルティ、監査について具体的なプロセスを設定すること。
- 臨床に直結する業務フローを扱う領域においては、クラウド（外部保存、外部アクセス）を利用することを考慮し、医療情報システムの安全管理に関するガイドラインで求められる事柄（データアクセスへの多要素認証、VPN、Cookie取得等Webアプリケーションにおけるセキュリティデザインや脆弱性の修正、無害化等）を満たすべく、複数の構成案を比較検証し、事業におけるセキュリティ確保や性能確保、データ品質の確保を徹底できるように、課題や改善事項を導出すること。



## 19. システム開発・環境構築

- 解析・データセンターがコンプライアンス・ガイドラインへ適合しているかを判断する仕組みを設けること。
- 1次解析から2次解析まで解析ワークフローと多様な解析資源の提供をオンプレ・クラウドのハイブリッドも活用して実現すること。
- 医療機関、シーケンスセンター、解析・データセンターとの間には、セキュリティを確保したシステムやネットワーク接続環境の整備が必要となる。また、データセンターは物理的に国内にあるセンターの利用を基本とする。
- **計算リソース**、データストレージについては、既存のリソースを活用しつつ、クラウドへの移行を検討すること。
- 常に新たな技術を取り入れられるように、情報システムは特定の技術に依存しないように、拡張性が高い方式で整備すること。
- 臨床情報の収集や臨床に必要な知識のデータ保管、ゲノム解析基盤（シーケンスセンターとのネットワーク等の接続を含む）、人工知能を活用した変異等の検出、解析データの管理と利活用、システム全体の保守・運用や高度なセキュリティ防御など、各々の分野は技術的専門性が異なる。これらの要素毎に最適な開発技術者の確保と開発体制の整備が必要であり、そのための調達構造やシステム開発・運用のガバナンス維持のためのガイドライン等の整備も行う。

## 20. 人材育成 バイオインフォマティクス等について

- 解析・データセンターの運用、稼働維持、改善のためには、ゲノム解析基盤構築、臨床情報の収集体制の構築、患者レポートの作成業務、ネットワーク・セキュリティの整備、データ利活用のための情報基盤整備など、必要な業務は多岐に渡る。そのために必要となる専門性も生命情報学、医療情報、情報セキュリティ、臨床遺伝学、ハイパフォーマンスコンピューティング、クラウドコンピューティングなど多様である。
- 本事業の継続には、解析・データセンターにおける業務を遂行できる人材の育成が必要不可欠であるが、本邦においては著しく不足していると考えられている。
- 上記の解析データセンターの業務を分類し、今後の解析・データセンターにおける解析症例数の複数の設定において、持続可能な人員を見積もり、必要な人材育成計画を定量的に検討する。
- 解析・データセンターの運営にあたっては、技術開発だけではなく、情報データのロジスティックにかかる多方面への説明・交渉、書類作成、必要な技術・機材の調達など、開発ではないが専門的知識が必要な事務的作業も膨大に発生する。開発人材が事務的作業に忙殺され、技術のアップデートができなくなり損なわれてしまうことを回避するためにも、このような作業を担当する人材について、必要な人員とその確保のための育成計画を検討する。
- 全ゲノムシークエンスデータの解析・解釈はいまだに発展途上である。さらに、ロングリードシークエンス、一細胞シークエンスをはじめとする計測技術、クラウド、各種ウェブ技術、セキュリティ、人工知能など関連する諸分野における革新が続いている。持続的に競争力のあるデータ解析・基盤整備体制を保つためには、解析・データセンターにおいて運用業務のみならず研究を行う機能を持たせるべきである。
- 人材育成の知識体系の整理が必要。
- 人事戦略、ヒューマンリソース戦略の明確化が必要。
- 解析・データセンターの業務に忙殺されることで、研究者が新たな技術の取得し研鑽を積む時間がとれずにキャリアが台無しになることはあってはならない。解析・データセンターの使命に即した現実的なロジスティクス、事務部門の拡充など人材の配置について十分に議論されるべきである。
- 解析・データセンターを持続可能とするためには、その任務を担う能力を有する若い研究者の育成が必要不可欠である。大学院と連携し、解析・データセンターにおいて、データ解析や情報基盤の構築、データ管理などを研究することで学位が取得できるような仕組みを検討すること。
- **セキュリティやアクセスログを管理しリアルワールドデータを活用した人材育成を検討すること。**
- 解析・データセンターにおいて勤務・研究することが企業にとっても有益となる仕組みを検討し、企業から業務委託ではない人材を集めること。**その際には企業秘密の維持や中立性に疑問を生じないことが必要。**
- 長期的な目線でアカデミア、産業界相互の人事交流も通して人材を確保するとともに、人材の流出を防ぐべく、モチベーションの維持、キャリアパスの確保を検討する。

## まとめ（1／2）

### ゲノム解析（11,12,13）

- 年間1万症例程度の全ゲノムシーケンズデータ（腫瘍120x、正常30x）の一次解析（マッピング、バリアントコール）の結果を保存できる体制・基盤を構築し、データ提供機関と速やかに共有すること。
- 処理内容がオープンソースで管理できる統一パイプラインを運用し、適時アップデートを検討すると共にクラウドも活用し解析の効率化、高度化を図ること。
- 最新の技術を用い取得されたデータの解析に耐える頑健な解析基盤とすること。
- 各研究班と連携し、がん種の特徴を捉え患者還元につながるAIモデルを開発すること。そのために必要な人工知能維持開発基盤の構築も進めること。

### 臨床情報等の活用（14,15）

- 電子カルテから臨床情報を転送するためのオープンなデータソースAPIとデータアクセスを行うAPIを開発すること。
- 臨床情報の送受信について、プッシュ型とプル型の比較検討を行い、プロトタイプ構築を通じて、課題や改善事項を検証すること。
- ゲノム変異、臨床情報を抽出するためのAPIの整備を行い、単一の企業によるロックインを防ぎつつ開かれた形でレポート生成システムの開発が可能な体制について検討すること。
- レポート作成システムの構築に当たっては、（1）既存の遺伝子パネル検査等レベル、（2）全エクソン解析等レベル、（3）構造異常や既知の範囲を超えた研究的解析レベル、等にレベル分けを行って段階的に患者還元を行うこと。
- 当該遺伝子変異に臨床的意味づけを行い、治験情報等を含めたレポートであること。

### データ共有システム（16）

- 研究班やフォーラムが利用可能なデータ共有のシステムをオンプレミスやクラウド上に構築すること。
- 全ゲノム情報を患者還元するためのデータ解析・提供・保管を実施する研究・創薬等の拠点として研究開発が円滑に進むためのシステムを構築すること
- オープンAPIを整備し、ゲノムデータベース、臨床情報データベースの利活用を促進すること

## まとめ（2／2）

### 集中管理システム（17）

- 検体情報（採取日、種類、保存場所、残量等）、全ゲノムデータ（解析・データセンターにおけるデータ受領、QC結果、解析状況、レポート提出状況を含む）、臨床情報、同意の種別（共同研究での試料分譲の可否、産業界単独利用での試料の分譲の可否、意思変更の状況等）を一元的に管理・運営するシステムとすること。
- 検体の取り違い防止などの観点から、検体情報が常にゲノム情報、臨床情報のデータベースと紐付けが可能となるように管理する仕組みが必要。
- 研究同意時の本人確認、同意撤回があった際のデータの削除等、トレーサビリティを確保する技術の具体化を検討すること。
- 新規の患者からの検体については、既存の施設を用いて一括管理することが可能な仕組みを構築すること。

### 情報管理・システム構築に関して（18,19）

- 治療に直結する重要インフラとして臨床情報、ゲノムデータへのアクセスコントロールやID管理、ログ管理、秘密分散技術を用いたデータ保管、セキュリティ検知・対策意思決定プロセスを検討すること。
- 臨床情報やゲノムデータへのアクセスは、データ利用者の管理を徹底すること。

### 人材育成（20）

- 解析・データセンターの運用、稼働維持、改善のための多様な業務を担う人材は、本邦においては著しく不足している。人事戦略、ヒューマンリソース戦略の明確化と共に人材育成の知識体系の整理、育成計画を構築すること。
- 任務を担う能力を有する若い研究者の育成については、大学院と連携し、解析・データセンターにおいて、データ解析や情報基盤の構築、データ管理など魅力あるを研究することで学位が取得できるような仕組みを検討すること。
- 長期的な目線でアカデミア、産業界相互の人事交流も通して人材を確保するとともに、人材の流出を防ぐべく、モチベーションの維持、キャリアパスの確保を検討すること。

# ELSI WG

○班長：横野恵（早稲田大学）

大項目	項目		第3回	第4回	第5回	第6回	第7回	第8回	
患者還元	ICT/AI技術	1	eコンセント等による参加者の利便性向上			○	●	☆	☆
	対象疾患、患者数	2	対象疾患および対象ステージ詳細	○	●			☆	☆
	医療機関要件	3	EP、全ゲノム解析体制の要件など	○	○	●		☆	☆
検体の保存・利活用	保管、管理ルール	4	採取、処理、移送、保管、管理SOP	○	○	●		☆	☆
シーケンス	受託要件	5	NGS台数、WGS実績等	○	●			☆	☆
	技術的要件	6	WGSデプス、RNA seq範囲	○	○	●		☆	☆
	QC	7	QC方法、タイミング等	○	○	●		☆	☆
データ利活用		8	アカデミア、産業界の役割について				●	☆	☆
		9	データ共有ルール（データシェアリングP）				●	☆	☆
		10	データ利活用審査委員会				●	☆	☆
解析・データセンター	ゲノム解析	11	ゲノムデータベース構築	○	○	●		☆	☆
		12	統一パイプライン	○	●			☆	☆
		13	高度な横断的解析（AI活用含む）			○	○	☆	☆
	臨床情報等の活用	14	臨床情報DB構築（API自動収集）	○	○	○	○	☆	☆
		15	レポート作成システム			○	○	☆	☆
	データ共有システム	16	研究支援システム			○	○	☆	☆
	集中管理システム	17	集中管理システム			○	○	☆	☆
	情報管理・システム構築に関して	18	情報管理			○	○	☆	☆
		19	システム開発や環境構築			○	○	☆	☆
	人材育成	20	バイオインフォマティシャン等について			○	○	☆	☆
ELSI	ICF	21	統一ICF挿入文作成	○	○	●		☆	☆
	ガイダンス	22	IC手法、2次所見・結果還元在り方等	○	●			☆	☆
	患者・市民視点	23	PPIの推進			○	○	☆	☆
	人材育成	24	遺伝カウンセリングなど			○	●	☆	☆

※ ○ 協議継続、●方針決定、☆AMED研究班の経験等を踏まえ必要な検討・修正

21. ICF（統一ICF挿入文作成）, データ共有
22. ガイダンス（IC手法、2次所見・結果還元在り方等）

## 21. ICFの作成方針

- 統合指針の遵守が前提となる
- 必須とする共通事項の範囲と内容を確認してモデル文案を作成し、各研究班のプロトコールおよびICFへの反映を依頼する

### 必須とする共通事項

- 全ゲノム解析等実行計画の事業概要等
- 本事業の一部として実施する上で必要な研究目的についての記載<sup>1</sup>
- 臨床情報収集の共通項目（右記）
- データの流れ<sup>2</sup>
- 検体の流れ
- 知財の取扱い
- 患者還元

1. 一次利用として想定されるデータの利用目的がカバーできる内容とする
2. 解析機関，データ・センター，公的データ・ベースおよび二次利活用の方針

- 各研究班からのフィードバックを得て中長期的にモデル文案を拡充する
- ICF以外に本事業についての周知・広報（ウェブサイトや説明資材（パンフレット，院内掲示用ポスター等））が必要。ICFでは各研究班のプロトコールの説明が主体になる

### 臨床情報収集の共通項目

- ① 症例基本情報・同意情報
- ② 検体情報
- ③ 患者背景・既往歴・併存疾患・家族歴
- ④ 診断情報
- ⑤ 治療情報（薬物療法等）
- ⑥ 治療経過（最良治療効果，転帰，G3以上の有害事象等）

各研究班から今年度の解析実施におけるIC取得について情報収集を行い（ICFの記載内容や説明・同意取得の方法，周知・広報に関わる課題等），モデル文案の改訂や周知・広報活動に反映させる（令和4年度）

## 21. データ共有方針の整理

- より幅広い機関が柔軟にデータを利用できるようにする必要がある旨のご指摘（第5回専門委員会）を受け、一次利用におけるデータ共有の方針を整理し、新たな整理に基づいてICFモデル文案を修正した
- フォーラムを通じた多様なデータ利用を可能とし、フォーラムの活性化を図る

	利用者	費用
① アカデミア フォーラム	<ul style="list-style-type: none"> <li>● 実行計画に基づくデータ収集を行う機関</li> <li>● データ収集機関と共同研究を行う機関</li> </ul>	医療機関・研究機関 企業 無償
② 産業フォー ラム	<ul style="list-style-type: none"> <li>● 上記以外の機関</li> </ul>	医療機関・研究機関・ 企業 有償

起始ポイント\*  
到達後すみやかに提供開始

\*100例程度のデータ登録時点  
(希少がんは別に定める)

整理のポイント	*詳細は次ページ表参照
<ul style="list-style-type: none"> <li>● ①, ②いずれの枠組みでも医療機関・研究機関・企業による利用が可能</li> <li>● ①の枠組みでは、データ収集機関以外の機関（企業含む）がデータ収集機関との共同研究の形でデータ利用可能であることを明確化</li> <li>● ②の枠組みは、主として企業による利用を想定しているが、研究機関・医療機関も利用可能（有償）</li> <li>● 企業は①の枠組みでアカデミアとの共同研究による利用（無償）、②の枠組みで単独利用（有償）が可能</li> </ul>	



## 21. データ共有方針の整理

### ■ データ共有ルール（データシェアリングポリシー）案を前提としたデータ利用目的の整理

		利用者の範囲	利用可能なデータ	利用方法	データ利用審査	費用	目的	検討事項		
一次利用	1	実行計画に基づくデータの取得・解析を実施する研究機関・医療機関 アカデミアフォーラム	自データ	詳細解析	不要	無償	日本人の全ゲノム配列データベースを構築し、その活用により ① 全ゲノム解析等の成果を患者さんに還元すること ② 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こと ③ 全ゲノム解析等の結果を研究・創薬などに活用すること ④ 上記①～③に関わる人材の育成や保健医療政策の検討を行うこと			
			全データ	俯瞰・簡易解析	不要					
			自データ以外	詳細解析	不要 (データ取得者の合意を得た上で審査会に報告)					
一次利用	2	企業 産業フォーラム	全データ	俯瞰・簡易解析	不要	有償				
			全データ	詳細解析	必要 (事業内で実施)					
二次利用	3	申請による 公的DB	詳細な診療情報を除くデータ (一定期間後移行)	ゲノムデータ and/or 基本的な診療情報を用いた解析	必要 (公的DBで実施)	公的DBのポリシーによる				

## 23. PPIの推進

## 23. 患者・市民参画（PPI）の推進

### これまでの議論

“ 全ゲノム解析等を推進するにあたり、患者・市民参画の仕組みを設けるなどELSI（倫理的・法的・社会的な課題）への対応ができる体制の在り方等について検討する ”

「全ゲノム解析等実行計画（第1版）」

“ 対象患者への周知，説明だけでなく，広く国民や社会に対して継続的 な情報発信を行うとともに，患者や市民参画の仕組みを確保することが重要である ”

“ 患者・市民の視点を取り入れて，データのトレーサビリティや利活用に関する説明や対応方針を策定するなど，実行計画の実施状況の透明性を高める方策を実施する ”

「ロードマップ2021」

### すでに行っている取り組み

- ELSIワーキングに複数のメンバーが患者の立場から参加
- 「全ゲノム解析・患者還元説明文書検討会」の開催

### 今後重要となる取り組み

- 本事業において継続的にPPIを実装していくための仕組み作り
- 患者向け・市民向けの本事業の周知・広報活動

- 参加者向け説明（補助）資料については今年度パイロット版を作成
- 来年度以降は予算確保の上，事業実施組織（準備室）が主体となって拡充

## 23. 本事業におけるPPI推進の必要性・意義

### ■ 事業の規模と高い公共性

数万人の患者さんの参加を想定

単なるシーケンスのプロジェクトではなく、国内における今後のゲノム医療の基盤を構築する役割を担う

### ■ 大規模ゲノム研究の特性

データの複雑な用途

- 長期保管
- シェアリング
- データセットの連結（リンケージ）
- 特定されない多様な用途での利活用（同意が包括的にならざるを得ない）

伝統的な被験者保護システム（IC, 事前の倫理審査）では  
かならずしも十分にデータ利用をコントロールでない

適切なデータ利活用とベネフィット共有のためにはPPIを取り入れたガバナンスが重要

## PPIの推進に関する基本的な考え方

### ■ PPIを事業実施における柱の一つと位置付ける

- ELSIの構成要素の1つとしてではなく、単独で重要な役割を担う
- 参加者パネルおよびパブリックエンゲージメントを主要かつ必須の取り組みとして実施する（次頁）
- 恒常的なものとして組織上明確に位置づける（予算・人員を含む）  
事業実施組織に独立の部門を設け、専門性を持って関連業務に従事できる人材の配置・育成が必要
- PPIを推進するための環境整備として、積極的な情報提供・周知活動を行う
- 患者・市民の参画のために必要な支援をあわせて実施する

# Genomics England の取り組みを参考にした 場合

PPI体制

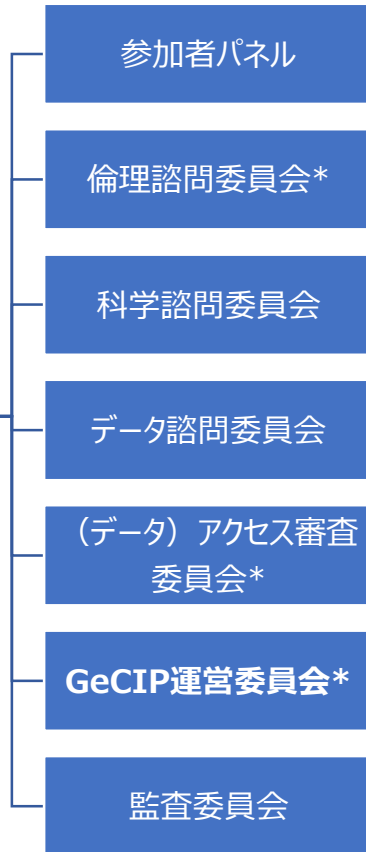
- GELでは4つのチャネル（中央+ローカル）で取り組みを実施
- 事業の性質や国際動向に照らし、**本事業では1と2は必須**
- 3と4はローカルレベルの取り組みであり、将来的な課題として検討

中央ガバナンス体制

諮問組織



運営委員会  
(Board)



\*=参加者パネルから委員選出

1. 参加者パネル（Participant Panel）2016年発足
  - GELの保有データが参加者を尊重した形で管理され、参加者の利益のために使用されるようにすることを目的として活動
  - 運営委員会に対する諮問組織としてガバナンス体制の一部を担う
  - 10万人ゲノムプロジェクトの参加者と家族らで構成（公募実施）
  - 年4回の会合を開催
  - メンバーには謝金・必要に応じたトレーニングを提供
  - 倫理諮問委員会、（データ）アクセス審査委員会、GeCIP運営委員会には、参加者パネル委員枠を設け意思決定に参加

2. パブリック・エンゲージメント
  - セミナーや市民参加イベントの開催を中心に社会全体に情報や議論の機会を提供し、認知の向上を図る

3. 患者市民参画（PPI）ネットワーク
  - 13のNHSゲノム医療センターそれぞれがPPIグループを設置
  - 主な活動は患者向けの情報提供や同意プロセスについてのフィードバック、プロジェクトの同意プロセスに対する評価等

4. GeCIPにおける当事者団体との連携
  - GeCIPはゲノム情報と疾患の関連付けを行うための枠組みで疾患別に42領域ある（主に研究者・臨床医らが参加）
  - 各領域で関連する当事者団体と連携し、当事者代表が参加

## 差別等の社会的な不利益への対応・法律等の整備

- 歴史的に、遺伝に関わる理由に基づいて個人や集団に対する差別的取扱いが、ときには公権力によって政策的に行われ、またそれらの政策の影響によるものも含め、遺伝に関わる差別や偏見が社会に存在してきた
- ゲノム研究・ゲノム医療が社会における差別や分断を助長したり、新たに生み出すことにつながらないような社会環境の整備が必要

基本的な理念は法律等の形で明確化されることが望ましい

個別分野では生命保険等におけるゲノム情報の取り扱いがとくに課題とされている

- 現状では、生命保険等におけるゲノム情報の取り扱いが明確化されておらず、不利益のリスクの具体的評価や研究参加時の情報提供が困難となっている
- 英国・シンガポールでは全ゲノム解析の国家プロジェクト実施に伴い、業界団体協力の下、生命保険等におけるゲノム情報の取り扱いが明確化されている
- 本事業で同様の取り組みが可能か検討し、必要な働きかけを行う

### ゲノム情報に関わるリテラシーの醸成および法律等の制度整備の必要性

- 上記2点は専門委員会で繰り返し指摘されてきた
- リテラシーの醸成はPPI推進の観点からも重要であり、本事業の内外で継続的に取り組む必要がある
- 日本医学会の遺伝学的検査ガイドライン改定案では、遺伝学的検査の結果を他の診療情報と同様に診療録に記載する方針が示されており、遺伝情報を保護し、差別等を防止するための法律等の制度整備が急務となっている
- 厚労科研小杉班の検討課題と連携し、これまでの研究成果を活用して取り組みを実施する

## 100,000 Genomes Project

### Insurance

Find out how taking part in the Project may affect insurance.



Source:

<https://www.genomicsengland.co.uk/understanding-genomics/faqs/insurance/>

## ELSIおよびPPIのまとめ（令和3年度）

本事業が社会の理解と信頼に基づき適切に実施されるためにはELSIに関する適切な検討・対応およびPPIの推進が必要不可欠

### ELSIへの対応

- 本年度は新規検体の解析実施に伴い、データ共有方針の整理を行い、整理の結果に基づきICFモデル文書の作成を実施した
- 今後の事業の本格化に伴い、新たなELSIが生じると予想されることから、事業実施組織においてELSIについて予測・検討を行い、迅速かつ実効的に対応できる体制を整備することが必要。必要な体制について本研究班で検討・提案する
- ゲノム情報による不当な差別・不利益や情報漏洩への対応等は安心してゲノム医療を受けるための社会環境整備として必要不可欠。本事業としての取り組みだけでは十分な対応が困難であるため厚労科研小杉班とも連携しながら法律等の制度整備の提案・働きかけを実施する

### PPI

- ELSIへの対応を含め本事業のあり方を検討する上で、患者・市民の視点の導入は必須であり、そのための仕組みが必要。必要な体制について本研究班で検討・提案する。
- 幅広い周知・情報発信を行うとともに、事業実施組織を主体としてPPIに取り組むための体制を構築する
- ゲノムリテラシーの醸成については本事業外での取り組みも働きかける