

第9回ゲノム医療協議会 議事概要

■日時：令和4年8月3日（水）15時00分～16時40分

■場所：中央合同庁舎第4号館4階共用第2特別会議室

■出席者：

議長代行：西辻 浩 内閣府健康・医療戦略推進事務局長

構成員：奥 篤史 文部科学省研究振興局ライフサイエンス課長[代理]
(Web参加)

伯野春彦 厚生労働省大臣官房厚生科学課長[代理] (Web参加)

中谷祐貴子 厚生労働省健康局がん・疾病対策課長 [代理] (Web参加)

城 克文 厚生労働省大臣官房医薬産業振興・医療情報審議官[代理]

下田裕和 経済産業省商務・サービスグループ生物化学産業課長
[代理]

上野裕明 日本製薬工業協会 副会長

春日雅人 国立研究開発法人日本医療研究開発機構 (Web参加)
ゲノム・データ基盤プロジェクトプログラムディレクター

菅野純夫 国立大学法人千葉大学未来医療教育研究機構特任教授
日本学術会議基礎生物学委員会・統合生物学委員会・基礎
医学委員会合同ゲノム科学分科会委員長 (Web参加)

高木利久 富山国際大学 学長 (Web参加)

門田守人 一般社団法人日本医学会連合会長 (Web参加)

山口 建 静岡県立静岡がんセンター総長

厚生労働省がん対策推進協議会会長 (Web参加)

米村滋人 東京大学大学院法学政治学研究科 教授 (Web参加)

参考人：小崎健次郎 慶応義塾大学医学部 臨床遺伝学センター長 (Web参加)

徳永勝士 国立研究開発法人国立国際医療研究センター ゲノム医科学
プロジェクト戸山プロジェクト長
ナショナルセンターバイオバンクネットワーク・中央バイ
オバンク長 (Web参加)

山本雅之 東北大学東北メディカル・メガバンク機構

吉田輝彦 国立研究開発法人国立がん研究センター中央病院
遺伝子診療部門長

オブザーバー：笠貫 宏 内閣府健康・医療戦略参与 (Web参加)

中釜 斉 内閣府健康・医療戦略参与

三島良直 国立研究開発法人日本医療研究開発機構理事長 (Web参加)

■議 事：

- 1) ゲノムデータのユースケースについて
- 2) 全ゲノム解析等実行計画の進捗について
- 3) その他

■概要：

○長野次長 定刻となりましたので、ただいまより第9回「ゲノム医療協議会」を開会いたします。

本日は御多忙の中、御参集いただきまして、ありがとうございます。私、本日の議事進行を務めさせていただきます健康・医療戦略推進事務局次長の長野でございます。よろしくをお願いいたします。

今回は新型コロナウイルス感染症の感染拡大状況を考慮し、オンラインと対面の併用での開催としております。

本日、構成員は全員御出席予定ですが、今、米村構成員がこの後入られるとお聞きしています。

議事1に関しまして、小崎先生、徳永先生、山本先生、吉田先生、4名の有識者の先生方に御発表いただくこととなっております。なお、小崎先生、徳永先生はオンラインでの参加でございます。

また、オブザーバーとして、笠貫参与、中釜参与、三島AMED理事長にも出席いただいております。

なお、議長である松尾科学技術・イノベーション推進事務局長は別公務のため欠席であり、西辻健康・医療戦略推進事務局長が議長代行を務めます。

最初に、西辻事務局長より挨拶をよろしくお願いいたします。

○西辻事務局長 皆様、どうも御苦労さまでございます。

今日は、構成員の皆様には、大変お忙しいところ御出席いただきまして、ありがとうございます。また、今ほど雷が鳴り始めて、大変足元の悪い中、会場に御参集の先生方には本当にありがとうございます。

今年度から、全ゲノム解析等実行計画が医療実装に向けた段階に移行していくという中で、今日はその詳細を御報告いただけると伺っております。併せて、今日の協議会では、ゲノムデータのユースケースにつきまして、ゲノム研究の分野で特に先駆的な研究をやってこられた4名の先生方から御発表いただけると伺っております。貴重なゲノムデータを幅広い用途で御活用いただき、どのように健康・医療分野におけるイノベーションにつなげていくのかということ、我々も一緒に勉強させていただけるのではないかと考えております。また、ユースケースを受けて、データ利用のニーズを踏まえたCANNDsの構築・運用に生かせるような議論も併せてお願いできればと期待しております。

構成員の皆様には、個別化医療・個別化予防の実現に向けたゲノム医療のあるべき姿を想定した上で、ゲノムデータの利活用をどのように進めるべきかといった観点で、多くの御示唆を賜ればと願っております。

また、各省におかれましても、本日の議論を踏まえて、実効性のある取組を検討いただければと考えております。

本日はよろしくお願いいたします。

○長野次長 それでは、議事に入る前に、本日の資料について事務局より確認させていただきます。

○笠松参事官 担当参事官の笠松でございます。どうぞよろしくお願いいたします。

本日の資料は、議事次第に記載しておりますとおり、資料1-1から資料1-4及び資料2、並びに参考資料1でございます。事前にメールで送付しておりますので、御確認くださいようお願い申し上げます。

○長野次長 それでは、本協議会につきましては公開としてございます。希望する記者の方々に御入室いただいております。

なお、本日配付の資料及び逐語ベースの議事概要を、後日、公開させていただきますので、よろしくお願いいたします。

それでは、議事1に移らせていただきます。4名の先生方から、それぞれの御専門の見地よりゲノムデータのユースケースについて御発表いただきます。最初に、それぞれの4人の先生方から御発表いただいた後に、最後に総合討論あるいは質疑応答という形で時間を取らせていただきたいと思います。

最初に、資料1-1になりますけれども、東北メディカル・メガバンク機構長 山本先生よりよろしくお願いいたします。

○山本参考人 このような機会をいただき、どうもありがとうございます。

それでは、早速始めさせていただきます。「東北メディカル・メガバンクによる日本人全ゲノム解析の意義と活用法」ということで、お手元の資料を御覧いただければと思います。

2ページになりますけれども、私どもは追跡を実施する健康調査、このことを「縦断的な健康調査」と言っているのですけれども、これは国民の健康を守るための重要な基盤であると考えております。体質や生活習慣がどのようにして健康上の問題を引き起こすかを理解するためには、追跡を実施する健康調査、縦断的な健康調査は不可欠です。これにコホートとバイオバンクという事業で挑んでいるのですけれども、コホート事業にゲノム解析を加えたゲノムコホート調査は、個人に合わせた予防医療確立の鍵になると考えています。私どもメガバンク機構は、大規模前向きゲノムコホート調査と複合バイオバンクで個別化予防、個別化医療に挑戦しようと考えている次第です。

3ページを見ていただきますと、このような一般住民のコホートにおいて全ゲノム解析を実施することの重要性というのはどんな点なのでしょう。私は、3点にわたると思います。1点目は、一般住民の全ゲノム解析情報は、患者のゲノム解析を大きく支援します。このことについては後ほど申し上げますけれども、一般住民の全ゲノムの中に全く存在しない、もしくはまれにしか存在しない変異こそが希少難病等の原因になります。ですから、疾患の病因解明のためには大規模な比較対照用データを整備する必要があります。

また、一般住民のゲノム解析情報は、革新的創薬の発展に大きく資するということがご

ざいます。

3点目に、一般住民のゲノム解析情報は、個別化予防や医療の実現に必須であると考えます。遺伝子の多型については、後ほど小崎先生から御説明があると思うのですが、多型は個人の体質や病気になりやすさを規定しています。それで、この多型から一般人の疾患リスクを測る方法の開発につながることができるのですけれども、そのためのしっかりしたデバイスをつくるには、一般住民の全ゲノム解析が必要です。さらに、ゲノム構造は民族ごとに大きく異なりますので、日本人の個別化予防・ゲノム医療を実現するためには、日本人のゲノム構造を正確に決める必要があります。

4ページをお願いします。さらに、こういうことは広く世界中で共有されています。UK Biobankがフロントランナーなのですけれども、20年も前から50万人のコホートをつくり、そして50万人の全ゲノム構造を決めるというプロジェクトが進行しています。一方、米国のAll of USは2016年からなのですけれども、こちらは100万人の全ゲノム配列を決めると宣言しています。UK バイオのほうは、去年の11月に20万人の全ゲノム解析を終了したというアナウンスがありました。一方、All of Usは今年の3月に10万人の全ゲノム解析を終了したと報告されています。

一方、私ども東北メディカル・メガバンクは10万人の全ゲノム解析を目指しています。5万人の全ゲノム解析について6月30日にその結果を報告いたしました。このことについては、後ほど少し説明させていただきます。

それで、All of USは100万人なのですけれども、Fitbitのデータとかスマホで取れるようなデータが多くて、1つの全ゲノム解析結果に乗っている健康情報は非常に薄いものになっています。一方、UK バイオは詳細なベースラインの調査をしていますので、もっと厚いデータが出ているのですけれども、私ども東北メディカル・メガバンクは、それに比べてもさらに追跡調査や家系情報が非常に手厚いので、メガバンクのデータが1つの全ゲノム解析に乗っているものとしては、一番手厚いデータということになります。

さて、5ページを見ていただきますと、私どもは、こういうことで全ゲノムシーケンスを実施してきておりますけれども、これは何のためにやっているかということ、1つは日本人の「全ゲノムリファレンスパネル」というものをつくるのだ。これが世の中の役に立つのだということを信じてやっています。全ゲノムリファレンスパネルというのは何なのかということですが、一般住民における遺伝子多型の位置と頻度を示すものです。一般住民集団の遺伝子多型の正確な頻度情報はゲノム医療の実現に極めて重要です。さらに、解析規模のサイズが大きくなると、頻度情報はさらに正確になります。

先ほど申し上げましたように、私どもは、本年6月30日に5万人のシーケンスを完了して、その中で3万8000人、これは後ほど申し上げますけれども、3世代コホートには近親者もいますので、その分を排除した全ゲノムリファレンスパネルを公開いたしました。さらに、次世代シーケンス解析の参照配列として、日本人基準ゲノムも公開しています。

6ページを御覧いただきますと、先ほど申し上げたとおり、全ゲノムリファレンスパネ

ルというのは、がんや難病に対する大規模な対照情報として、また革新的創薬支援にも活用できます。さらに、個別化予防・医療の促進にも貢献することになります。

幾つかの活用例を御覧に入りたいと思います。

7ページは、後ほど小崎先生の方からも御説明があると思うのですが、私どもは全ゲノムリファレンスパネルを未診断疾患イニシアチブ（IRUD）という、AMEDが実施している事業に優先的に提供してまいりました。私どものリファレンスパネルにある多型は病気の原因にはならないということで、40倍、50倍の疾患の原因となる変異の絞り込みができることになり、疾患のゲノム解析に大いに役立つことが解りました。AMEDのホームページを少し見せていただくと、IRUDのこれまでの6年間で、実に5136家系の解析実績があり、診断確定は43.7%。これはすばらしい成果だと思います。この成果の一部に私どもが貢献できたということで、私たちも喜んでいきます。

さて、8ページを御覧いただくと、クリニカルシーケンスに関連して、さらに、がんのゲノム医療に対してもこのリファレンスパネルが貢献しています。特に、国立がん研究センターのC-CATでがんゲノムパネル解析をやっておられますけれども、ここで見つかった新しい変異については、メガバンクのリファレンスパネルで、日本人集団ではこの変異がどれぐらいの頻度で存在するのかということ、全ての開示文書に記載しています。私どもメガバンクは一般住民集団ですので、比較的頻度高く見つかる変異は、がんの新規ドライバー変異である可能性は低い、というので使っています。

さらに、9ページを見ていただくと、クリニカルシーケンスを支えている解析企業に関して、日本でもベンチャー企業さんや大手の企業さんなど、次々と成立しています。これらの企業に私どもの一般住民の全ゲノムリファレンスパネルをライセンスしています。A社からは、これは使い勝手がいいですよ、使っていますよ、という連絡を四半期ごとにもらうのですが、御覧いただいていますように、2022年度の第1四半期に非常に多くの数が使われています。利用が右肩上がりが増えていくことが御覧いただけだと思います。

一方、10ページを見ていただくと、今後の展開として、創薬研究におけるゲノム情報活用の重要性ということが挙げられます。ヒト遺伝学が薬剤開発の成功確率を上げるのです。このことについては、2015年頃から、大手製薬企業の研究所からの報告が相次いでいるのですが、つい最近出た論文に、2021年度にFDA承認された新規薬剤の3分の2は、遺伝学的な証拠によりサポートされていたという衝撃的なことが出ています。特に、一般住民で分かってくる機能喪失変異というものがあるのですが、これが非常に創薬の鍵になる。この創薬標的によって新薬ができたということもあります。

さらに、11ページを御覧いただくと、最近、よく言われるのがPatient Journeyで、患者さんの記録をしっかり録る。そのことで、創薬標的を探すレバレッジポイントを可視化するという事業があるのですが、今は大抵、左下にありますように、Treatment Journeyで、どんな治療をやっている、どんな帰結になるということが多いです。しかし、私ど

もは一般住民の全ゲノムを使うことによって、これがWhole Life Patient journeyにすることができる。病気になる前からPatient Journeyを追いかけることができるので、これに一般住民の全ゲノム情報が活用できるだろうと考えております。

12ページには、個別化予防の実現のために、国民全部の全ゲノム解析を実施するのはなかなか大変ですけれども、簡易な全ゲノム解析ツールをつくれれば可能になるかもしれない。その解析ツールをつくる場所に一般住民の解析情報が使えると考えています。今後のゲノム医療の重要なターゲットは多因子疾患のリスク予測です。多因子疾患のリスク予測はPolygenic Risk Scoreを使ってできるのですけれども、これはマイクロアレイ法で決めることができます。このマイクロアレイをつくる場所に、私どもの全ゲノム解析の成果を活用することができます。実際に、私どもはジャポニカアレイをつくりました。これをさらに改良して、健診用のジャポニカアレイをつくり、右下にありますような実用化に向かっていきたいと考えます。

さて、13ページを御覧ください。これが最後になりますけれども、実に5万792検体の私どもの参加者の全ゲノム解析が終わりました。それで、申し上げましたように、3世代コホートで親子が入っていますので、近親排除をして、1万1000人ほど除いて、3万8942。また、3万8722というリファレンスパネルを作りました。右側の主成分分析を見ていただきますと、これは国際1000人ゲノム、これには実は1000人以上入っているのですけれども、それと今回の私どもの3万8000人を合わせて、主成分分析したものです。そうすると、国際1000人ゲノムに入っている日本人、東京の人なのですけれども、百人ほど入っているのですけれども、これが見事に東北メガバンクで決めたピンクの集団中におさまっています。

この青い線の内側が私どものリファレンスパネルです。その外側には、韓国人との間の点があったり、中国人との間の点があります。これは、私どもがリクルートした住民票を持っている方でもルーツが異なる方がおられるということだと思いますが、そのことがこういうふういきちんと見られます。しかし、日本人の代表性はしっかりと担保されているということがこれで見ただけだと思います。

まとめです。14ページです。10万人の全ゲノムシーケンス解析を早期に完遂して、英国・米国と肩を並べて、大規模な全ゲノムシーケンス・データ基盤を持つ国になるというのが、私たちにとって非常に重要であると思います。これが個別化予防や革新的創薬の支援体制につながると考えております。

4枚ほど参考資料をつけております。御覧になっていただければと思います。

私の発表は以上です。

○長野次長 山本機構長、どうもありがとうございました。質疑応答については、全ての先生方からの御発表が終わってからにさせていただければと思いますので、後ほどよろしくお願いたします。

続きまして、難病疾患の関係ということで、小崎先生より御発表いただきたいと思います。資料1-2になります。どうぞよろしくお願いたします。

○小崎参考人 慶應大学臨床遺伝学センターの小崎と申します。

私ども、希少疾患あるいは未診断疾患のプロジェクトを進めておりまして、山本先生から先ほど御紹介があったとおりです。これまで約7000家系集まり、2万ゲノムほど集めて、診断率は40%を超えるということで、医療・研究に貢献していると思います。私どもの研究室自身で、10以上の新しい疾患を発見することができています。その際に、コントロールのデータが非常に重要です。そのことについて、少し基本的な話になるのですが、お話しさせていただきたいと思います。

2枚目のスライドをお願いいたします。世界的に「変異」という言葉を使わなくなったのは御存じのとおりで、何が正常かということを決めることはできないので、標準配列ないし参照配列という言葉で言われるようになっていきます。つまり、変異という言葉は機能異常を内包してしまうからでありまして、バリエーションという用語は機能異常を内包せず、中立的であるわけです。

それで、標準的な配列は正常な配列ではなくて、仮に使っている物差しです。基準からの相違の許容範囲を指しているのが、まさにコントロール群のデータでありまして、これが非常に重要で、特に日本人の研究をするためには日本人のデータをできるだけ集める必要があるということです。

次、3枚目に移ります。どのぐらいのコントロールの人数があれば、どのぐらいのバリエーションが得られるかということについて、これは東北メディカル・メガバンクのデータをお借りしていますが、なかなかサチュレートするものではなく、大人数のシーケンスをすればするほど、これまで知られていないバリエーションが得られるということがございます。

次の4枚目に移ります。それで、少しテクニカルな話になって申し訳ないのですが、バリエーションの性質は、対象とする遺伝性疾患がいわゆる優性遺伝病、これは今、顕性遺伝病と呼ぶことになっておりますが、であるか、劣性遺伝病、潜性遺伝病であるかによって少し異なります。全く正常なコントロール集団があったとして、そこに遺伝性疾患がないとするならば、特に優性遺伝病ではコントロール群に認められるバリエーションは病気の原因にはなり得ないと、すぐさま判断することができるわけです。

一方、いわゆる劣性遺伝病では、患者さんがヘテロであって発症していない場合もあるので注意が必要ですが、その場合にも非常にそのようなアレル、バリエーションの頻度が低いということが条件で、これで多くのバリエーションを病気の原因ではないと判断することができるわけです。

一方、いわゆる多因子遺伝病の研究においては、コントロール群の中にも、それから患者群の中にも当該バリエーションが存在する割合が違うということになります。

5枚目に移ります。これは単一遺伝子病の研究のイメージですが、特定の遺伝子のバリエーションがある人は患者群のみに認められて、特定の遺伝子のバリエーションのない人はコントロール群のみに認められる。このように分けることができるわけですが、6枚目、多因子遺伝病ではそうではないということになります。

その次の7枚目が重要になるわけですが、コントロール群の定義として、何ら疾患がないような人たちのグループ、例えば慶應大学で持っている百寿者のように全く病気がない人たちということも考えられますし、それから多因子研究のコントロール群もあるし、多因子研究の患者群もある。

ところが、我々が未診断疾患イニシアチブで調べているような全く新しい病気、しかも小児期に発症するような病気を中心とする場合には、そのような研究には、例えば糖尿病であるとか高血圧の患者群というの、またコントロール群として使うことができるわけです。つまり、コントロール群という言葉の意味合いは、何を患者群にするかによって変わってくるわけでありまして、特に難病研究の患者においては、多因子研究の患者群もまたコントロールとして使い得るということになります。なので、幾つかのコホートを組み合わせてコントロール群として利用可能であるという点が非常に重要であります。

8枚目に移ります。ただ、どのようなコホートを難病研究のコントロールとして利用できるかということについては、個別のサンプルについて、シンプルでもよいのでデジタル化された臨床情報があれば、サンプルデータ、適切なサンプルをコントロールとして抽出することが可能になるだろうということになります。CANNDsなどのプロジェクトが推進されているところですが、ぜひ簡単な表現型でよいので、症状の情報があるとコントロール群の抽出に非常に役立つだろうということになります。

そして、9枚目になりますが、今、申し上げたことのおりなのですが、単一遺伝子病の診断・研究におけるコントロール群を幅広く使って、とにかくnを確保するということが有用だということになります。

スライド10枚目は、先ほど山本機構長がお示しになられたので、省略させていただきます。

スライド11枚目です。では、どうして日本人のコントロールが必要かということになるのですが、ベン図のように描いてあるもののうちの赤いほうは、欧米人を中心とした世界を代表するバリエーションデータベースのgnomADのもので、14万人由来で596ミリオン、バリエーションがあります。

一方、東北メディカル・メガバンク、日本人を中心とするバリエーションデータベースで、3.8万人由来で191ミリオンあるわけですが、そのオーバーラップは非常に少ないんですね。東北メディカル・メガバンクの中に見られたもので、欧米人に見られないバリエーションが3分の2もあるということですので、外国のデータがあるならいいのではないかという考えが間違いであることがよく分かります。

それで、以上のことをまとめたものが12枚目ですが、効果的にフィルター、難病患者の原因となっているバリエーションを探すためには、コントロール集団における頻度が極めて重要で、しかも日本人における頻度が重要で、現在の数万人規模のコントロール群から得られるデータを見る限りでは、まだまだ未知のバリエーションが多数あるということになり、既存の各種コホートのバリエーションを統合・集積して頻度を計算する必要があります。

特に、ここまでは主にSNV、1か所の塩基の変化について話をしましたが、今後重要な課題は構造異常であります。SVの取扱いが重要です。SNVと異なり、SVの検出のためには標準的な手法が未確立であるので、SVの抽出のためにFastqファイルないしBAMファイルが必要です。これらを各種コホートから得られた生データに近いデータを集めてきて再計算する必要があるということになります。

14枚目のスライド、最後のスライドになりますが、結語です。単一遺伝子病の研究・診断では、発症に無関係なバリエントを除外する必要があります。効率的に除外するためには、コントロール集団におけるバリエントの頻度、専門的にはアレル頻度と呼びますが、情報が有用かつ不可欠。

日本人におけるバリエントの研究が必要で、多因子研究のコホートのバリエントデータは、希少疾患の研究におけるコントロール群となり得るということ。そして、その利用のためには臨床情報が必要であるということ。簡単なものでもいいので臨床情報が必要であるということ。

最後に、特にSVなどのデータを計算するためには、各種コホートのバリエントを統合・集積して、改めて頻度を算出する必要があると考えています。

以上でございます。

○長野次長 小崎先生、どうもありがとうございました。

それでは、続けて御発表をお願いしたいと思います。がん領域におけるユースケースということで、国立がん研究センター中央病院の吉田先生より御発表いただきます。資料1-3になります。

○吉田参考人 それでは、よろしくお願ひいたします。

私は、ゲノムのコントロール集団について、どのような利用価値が考えられるかということをご一般論として考えてみましたので、個別の現行の事業などへの意見ではなく、一般的な話となります。ただ、専門が遺伝性腫瘍ですので、その目から見たということになります。また、本日は参考人ということで来ておりまして、がん研究センターの公式見解ではなく、個人の意見として述べさせていただきます。

まず、資料1-3の1ページ目にありますように、コントロールと言ったときに我々が考えるのは、まず基本的な利用として、断面的な、クロスセクショナルなレファレンスのコントロール、頻度データです。

それから、もう一つのレベルの、より高度な利用としては、山本先生の発表でもありました縦断的な一般集団コホート、この2つの意味があると思っております。

2枚目、めくっていただきますと、これは基本的な多因子疾患と単一遺伝子病の違いの模式図で、AさんからEさんのように、持って生まれた緑の遺伝素因と、それから、その後の様々な生活習慣の組合せで、ある疾患になる確率が決まってくる。こういう単純な足し算ではないのですが、両方のファクターが関与しています。

しかし、右側のFさんは例外で、この方は持って生まれた遺伝素因が非常に強いために、

その後、例えば幾らサプリメントなどで頑張っても、例えばがんになるのは避けられないということで、Fさんのような方がいわゆるメンデル型の単一遺伝子病となります。一方、AさんからEさんに関しては多因子疾患として、こちらに関しては、その遺伝素因をまだ探したり、遺伝素因と環境要因との交互作用、あるいはそのエビデンスをいかに持続する予防のプラクティスに結びつけていくかといったことが課題になっていると理解しております。

3枚目は、当時、3年ぐらい前にコントロール群としてどのような使い方があるのかについて、徳永先生にもご指導いただきながら、研究を大まかな類型に向けて考えてみたものです。まず類型Aというのは、先ほど山本先生もおっしゃっていた、がんの体細胞変異を同定するために使うコントロール集団となります。右側の上のほうに患者さんがいらして、ぎざぎざのがん組織の中に赤い変異が見つかった。ただ、それはヘテロ接合で存在するので、青の正常のアレルもある。ただ、この方の左手にクエスチョンマークがついているように、この方御自身の正常細胞をコントロールとして調べない場合は、赤い変異が本当に後天的に起きたがんの体細胞変異なのか、単なる多型なのか分からない。

このようなことは2019年6月に我が国で保険収載されましたがんゲノム医療の、9割以上で採用されている遺伝子パネルでは、腫瘍組織のみを調べるパネルとなっています。そのような腫瘍組織しか調べないパネルでは、コントロール集団において、その変異が一般の人にもよく見られる多型であったら、その人個人の正常組織、すなわちgermlineのゲノムを調べていなくても、見つかった変異は多型だろうと、判断しないとイケないのですね。ということで、コントロール集団のデータが非常に必要になる。

そのため、先ほど山本先生がおっしゃったように、C-CATでは、日本人の信頼できるデータとして東北メディカル・メガバンク機構のデータ等をつけて、各医療機関のエキスパートパネルに報告しており、検出された変異の評価に活用していただいています。現在、年に1万8000人ぐらいが登録され、C-CATには既に4万人近くの患者さんが登録されていますから、ある意味数としては、コントロール集団のデータを一番使わせていただいているところかもしれない。これに関しては、左の水色の矢印に書いてありますように、バリエーションの頻度情報が主体ですので、あまり詳細な臨床情報は必要としない使い方になります。

その次の類型Bは、今度は体細胞変異ではなく、いわゆる生殖細胞系列、germlineの変異、バリエーションで、これは小崎先生が御説明になられたとおりで、ある患者さんのgermline、血液で赤いバリエーションが見つかった場合に、それがコントロール集団でもよく見られるものであれば、まれな遺伝病の原因としては説明しにくいということで、棄却されることになります。

この2つが2019年のゲノム医療実現推進協議会中間とりまとめで、医療の応用に近いと言われた第1グループなのですけれども、その下が第2グループ、着実な推進が必要と言われている、いわゆるCommon diseaseの研究となります。Common diseaseに関しては、この後、徳永先生が御説明されると思いますので、ちょっと省略しますけれども、私の理解

では、Cのような多くのケースコントロールスタディを統合するメタアナリシスのような解析に使ったり、あるいはDのように多段階的にほかの研究で絞り込まれた候補を、さらに自分の症例群と、それから、自分は症例群は持っているけれども、コントロール群を十分持っていない場合に、コントロール群と比較して絞り込んでいく。そのような使い方がされると思います。

メタアナリシスの場合には、臨床情報は比較的制限されていても大丈夫と理解しているのですが、Dのほうについては、より詳細な臨床情報があると、より高い精度での検出が可能となります。

次、4ページを御覧ください。上記のように、単一遺伝子病と多因子Common diseaseを2つに分けるのが基本的な考え方なのですが、ゲノムの情報が蓄積されてくると、必ずしも二峰性のグループじゃなくて、もっと連続的になっているのではないかと考えられます。これは模式図で、グラフの横軸は、その遺伝素因を持っていると疾患を起こす確率、Effect sizeで、縦軸がその遺伝素因の集団中での頻度になりますが、右下は典型的な遺伝性疾患で、遺伝子の効果は強いけれども、集団中には少ない。左上のほうは、遺伝子の効果の一個一個は少ないけれども、多くの人を持っている。ただ、中間のところがあるということも分かってきて、ここで単一遺伝子疾患と多因子疾患の両方の研究がマージしつつあると考えています。

5ページ目を御覧ください。これは遺伝性乳がん、卵巣がん症候群についての外国のデータではあるのですが、我々が外来で必須としている資料になります。これは、BRCA1とかBRCA2に病的なバリエーション、病的変異を持っている方が何歳までに、左側は乳がん、右側は卵巣がんになるかという累積罹患率を示しています。何歳までに、なので右上がりになってくるのですけれども、BRCA1だったら何歳までにこのぐらいの頻度ということで確率として説明するのですけれども、実際には、6ページ目を見ていただくと、このように1本のグラフではなく、幅のあるグラフになってきます。

例えば、左上では、同じBRCA1遺伝子であっても、乳がんの発症率は、このようにグラフの高さが違う。このばらつきは何かということになりますが、その一因が右のほうに書いてありますけれども、Polygenic Risk Score、先ほど山本先生から御説明があったこととなります。強い単一遺伝子病の原因に加えて、多くのSNPが影響している。こういうことが続々と分かってきているということになります。

このPolygenic Risk Scoreに関して、私が注目している最近の論文2つをちょっと簡単に御紹介します。

7ページ目を御覧ください。これは、いわゆる一般集団のゲノムと臨床情報から単一遺伝子病の個人差のばらつきを説明するPolygenic Risk Scoreを抽出できる場合があるというフィージビリティを示した論文です。

右上のほうに箇条書きで書いてあるものを簡単に見ていただきますと、我々の遺伝性腫瘍の臨床の問題点は、同じ遺伝子、同じバリエーション、同じ家系でさえ表現型が随分違う。

がんになる、ならないも違うし、がんの種類も違う。これが分からないのですね。そのことを明らかにするために、稀少疾患なので、しかもバリエーションごとに症例を集めようとすると、これはとても無理だということになります。

そこで、この研究者たちは仮説を立てました。それは、メンデル型の遺伝病というのは、一般集団の同じような病気、同じような表現型の病気の極端な、エクストリームな偏り値であると考えたのです。左上のグラフにあるグレーが一般集団の表現型の分布で、その中に右のほうに紫のグラフとして遺伝性の病気の方の場合の分布が書いてありますけれども、このように分布は基本的に同じだけれども、偏っている場合だと。

こういう仮説が成り立つ場合は、構造化された大規模な臨床情報を活用することで、かつ右下のフロー図にありますように、50万人のUK Biobankと120万人のUCSFの臨床情報を使ってお互いに検証することで、左下にありますように、 $\alpha 1$ -アンチトリプシン欠乏症のような遺伝子疾患でも、このように強い既知の原因遺伝子にPolygenic Risk Scoreを加えることで層別化が可能でした。

こういうふうに、単一遺伝子病の従来型の遺伝子の研究に加えて、大規模コントロール集団の、臨床情報・ゲノム情報を使うことで、多因子疾患と単一遺伝子病の研究がマージしていくといった重要な使い方が、コントロールの有用なところかと思います。

もう一つが8ページです。このようにPolygenic Risk Scoreは非常に注目されており、多くはレトロスペクティブに検証されることが分かっているのですが、このスライドに示すように、前向き臨床試験でPolygenic Risk Scoreを患者さんに返しながらか検証していくということがClinicalTrials.govに登録されるような臨床試験も始まっています。

ここで注目されるのは、真ん中の円の内側に赤い点線でつけたのですが、Polygenic Risk Scoreに加えてメンデル型の遺伝病のバリエーションの評価についても、見つかった場合には報告されることになっています。既に、このように単一遺伝子病と多因子疾患と両方併せてゲノム情報を患者さんに返しながらか、前向きに検証していくという時代に入っています。以上が2つ、注目される動きだと思います。

9ページ目、最後、まとめです。このように2つのレベル、正常対照群のデータとして、比較的限られた臨床情報で使うことに関して、ユーザーとして一般的に期待する方向性としましては、小崎先生がおっしゃったように、構造バリエーションも含めた情報が欲しい。

それから、新たな、今まで注目されていなかった領域を注目する研究がどんどん出てきています。例えばテロメアの情報とか。こういうためには、元データを必要に応じて提供、アクセスできるようにしていただきたいです。

それから、ゲノムデータ解析visiting環境を使って、元データをダウンロードせずに解析する方式が多く使われるようになっておりますけれども、こういったことに関しては、一定のセキュリティーが担保されるという条件のもと、研究者が自施設の計算機からアクセス可能にしていきたいというような希望がありました。

同じく海外サービスについても、海外の様々なゲノム研究のプラットフォームが出てお

りますので、そういうものを活用するのも、国際レベルのデータシェアリングという意味では効率的だと思います。

それから、2つ目のレベルの一般集団コホートのほうは、ゲノムデータについては上記正常対照群と同様ですが、より高度な、構造化された臨床情報、がん登録等の予後情報へのリンクが重要であると考えます。

ということで、最後ですけれども、下のピンクのところですが、Caseについては、1つのやり方としては、C-CATのように国民皆保険の中でリアルワールドデータをビッグデータ化していくということは可能である。しかし、コントロール集団については、国の基盤として積極的な維持が必要であろうと思います。

今後、鍵となるのは臨床情報とその構造化、それから法的整備が必要と了解はしているのですが、各種登録事業などのデータへのアクセスということになります。これは、国全体が取り組むべき、我が国の生命線になっている最後の資源といたしますか、非常に重要な資源と考えていて、例えば、この前、閣議決定されました骨太方針2022に書かれておりますような、全国医療情報プラットフォーム等における電子カルテ情報の標準化といった枠組みで進めていく必要があるだろうと思います。そしてそのためにも、国民が安心・信頼して情報共有・利活用できる枠組みの整備などが必要であると考えます。

以上です。

○長野次長 吉田先生、ありがとうございました。

次に、国立国際医療研究センター研究所の徳永先生より御発表をお願いいたします。資料は1-4になります。徳永先生、よろしくをお願いいたします。

○徳永参考人 ありがとうございます。よろしく申し上げます。

私、多因子疾患の研究において大規模なゲノムデータを共有することの意義を御紹介させていただきたいと思います。

2枚目のスライドは、ゲノム全域解析の主な方法として、DNAマイクロアレイあるいはSNPアレイと呼ばれるもの、それから次世代シーケンサーを使うものという2つの技術があるということ、本当にイントロダクションとして御紹介したいと思います。多因子疾患の場合は、これまでは比較的SNPアレイを用いてきた。

3枚目を見ていただきますと分かりますように、多因子疾患の場合は、一つ一つの遺伝要因のリスクあるいは効果というものが比較的小さなものですから、非常に多数の検体について分析しなければいけないということで、勢いコストをある程度下げるといいますか、抑えるためにSNPを解析する場合が大部分でございました。国際共同研究も含めて、共同研究の規模をどんどん大きくするということが行われてきたわけです。

4枚目以降は、3つのトピックに分けて、まず最初に意義をまとめて、そして事例を紹介するという形で進めていきたいと思います。

4枚目、トピックの1は、国内外のデータ共有による大規模解析の意義ということで、多くの遺伝要因が特定され、より正確なリスク予測が可能になる。発症機序の理解が進み、

創薬標的の候補も見つかる場合があるということでございます。

5枚目に移っていただきますと、1つの例として私どもが参加してきた原発性胆汁性胆管炎のGWASを取り上げます。これは日本全国多施設共同研究によって、アジア初のGWASをしばらく前に行いました。患者の数が487名、健常者がほぼ同数ということで、3つのゲノム領域が検出されたということが分かると思います。これが大体2倍の規模になったときに8領域になりました。さらに、イギリスのグループと私ども日本のグループが中心になって、ほかの幾つかの国に声をかけて、患者の数1万人、対照者が2万人という規模のメタ解析を行いますと、一挙に56領域見つかるという結果になったわけです。

6枚目を見ていただきますと、この56領域の遺伝子がどんなたんぱくをコードしているかというものを図示化しているわけです。たんぱくが、この場合は免疫系のいろいろな細胞の中で、どんな分子パスウェイに所属しているかというのを模式化、図示しているわけです。そうすると、各種のT細胞活性化において重要な役割をするものが、すみません、言い忘れかもしれませんが、赤で塗り潰してあるものがGWASで見つかった遺伝子の産物です。

そして、右下の図で言いますと、パターン認識と呼ばれるTLR(トールライクレセプター)という遺伝子がつくるたんぱくの、細胞内のシグナル伝達、分子パスウェイの重要なたんぱくが幾つか赤で塗られていることも分かると思いますし、右側にTNFレセプターのたんぱくが幾つか赤く塗られて、またその下に細胞内のシグナル伝達に関わる分子群が赤く塗られているのが分かっております。つまり、T細胞活性化とかパターン認識受容体、TNFシグナルが、この疾患の発症に重要な役割を果たしているということが見えるわけです。

一方、薬につきましては、こういった分子パスウェイに影響するといいますか、そこに効果を持つ薬として知られている薬が、ひょっとするとこれまでは原発性胆汁性胆管炎の治療には使われていなかったけれども、効果を持つかもしれない。そういった可能性も提案したという例です。

さらに、7枚目のスライドは、近年話題になっておりますCOVID-19感染症の重症化に関わる遺伝要因の探索研究です。下に書いてありますように、多くの国の研究グループの共同で、5万人弱の患者さんのゲノム解析を行った結果、実際には13遺伝子領域が検出されたのですが、この規模にしては余り大きな効果を持った遺伝要因は見つからなくて、当初予想したほど遺伝要因は強くなかったという結果になっております。

次の話題に移ります。8枚目です。全ゲノム解析データの価値として、今日、小崎先生、吉田先生から遺伝病／希少疾患あるいはがんにおける重要性をお話いただきました。多因子疾患においても、これから大いに役立つだろうということをお話しさせていただきたいと思っております。

9枚目に移りますが、多因子疾患のゲノム研究の課題が明らかにあります。上のほうは、既にお話がありました。コントロール群は、意外と個々のプロジェクトごとに収集することは簡単ではありません。ということと、下のほうが重要ですが、レアバリエント・構造変異の検出はSNPの解析では大変困難であるということ、これを解決するのが次の10枚目。

完全に解決することは、まだ技術的に難しいのですが、全ゲノム解析であれば、ある程度その課題が解決できる。

ただし、コストがかかりますので、大規模なバイオバンクが全ゲノム解析を実施し、そのデータを提供する例が増えている。UK Biobankがその典型ですが、我が国でも東北メディカル・メガバンク、バイオバンクジャパン、先ほど山本先生から御紹介がありました。そして、私どもNCBNの全ゲノム解析を最後に少し御紹介しますが、そういうデータ共有が始まっているということでございます。

右の図は、海外のグループによる総説の図をそのまま借りてきたのですが、全エクソーム解析あるいは全ゲノム解析を使ってGWASを行っている論文の出版数も着実に増えている。ただ、トータルの数はまだまだ少ないという状況がお分かりかと思えます。

11枚目に移っていただきますと、先ほど小崎先生から御指摘のあった点に類似しますが、欧米のデータベースは日本人には適切ではないという点です。アジアの諸集団の大規模な全ゲノム解析に私どもも参加して、その結果が「Nature」誌の表紙になったわけですが、大変興味深い結果というのは欧米のデータベースとの比較です。世界的なデータベース、HGMDとかClinVarに病気の原因変異であると登録されているものの中の数百という数のバリエーションについては、アジアの集団ではごく当たり前のものであった。つまり、欧米の人たちを解析していると極めて珍しいために、これこそ病気の原因だろうと誤って考えられたというものになります。日本人のゲノム医療を実現するためには、日本人の標準バリエーションデータが必要であるという1つの根拠になります。

それから、12枚目も基本的には同じことでありますけれども、私どもが担当しているMGeNDというデータベース、日本人で見出された疾患関連バリエーションを登録する公開データベースです。上の表は時間の関係で省略させていただいて、下の図を見ていただきたいのですが、MGeNDの1年ぐらい前のデータで恐縮ですが、右下の数字です。6万3000ぐらいのバリエーションを登録しておりますが、それに比べて、世界一のアメリカのClinVarは100万を超えたバリエーションが登録されております。しかしながら、たかだか6万のMGeNDの中のバリエーションの中の74%、赤いエリア、つまり4分の3はClinVarには見られないものです。このぐらい集団差があるということでありまして、日本人のバリエーションデータが必要であるということのもう一つの根拠と言えるかと思えます。

13枚目に移っていただきますと、私どもも全ゲノム解析を行った多因子疾患のGWASを行いまして、これは重症の薬剤過敏症でありますスティーヴンス・ジョンソン症候群の解析結果でございます。従来、私ども自身がSNPを用いてGWASで行ってきた既知の関連SNPがもちろん見出されますけれども、それに加えて、新たに構造変異、先ほどのSV、構造バリエーションやマイクロサテライトの関連解析も行うことができますし、またレアバリエーションが集積する遺伝子も検出できたということでございます。

3番目の最後の話題に移りますが、大規模なデータサーバ・計算環境の提供が必要になるということでございます。

15枚目です。多因子疾患の全ゲノム解析に必要なデータ解析環境。コントロール群もケース群も、最初に申し上げたように非常に検体数を多く解析する必要がありますので、ゲノムデータの量が膨大になり、もはや輸送とか転送が困難という状況です。さらに加えて、こういうGWASあるいはPolygenic Risk Scoreに必要な計算リソースが膨大になります。ということで、大規模なデータベース、UK Biobankを筆頭として、こういったデータベースがvisiting環境に移行しているということです。

国際共同研究に必要な環境も、読んでいただくとお分かりになるかと思います。こういった状況に変化しているということが言えます。

16枚目は、私たちのナショナルセンターバイオバンクネットワーク（NCBN）の紹介をさせていただきます。これは、国立がんセンターをはじめ、私が所属する国立国際医療研究センターなど6つのナショナルセンターが、それぞれ特色ある疾患患者さんの治療をする病院を持っておりまして、その病院に来られる患者さんの中で同意をいただいた方々の検体をバイオバンクに登録しているわけです。高度な専門性、詳細な臨床情報を伴い、追跡可能性もあるバイオバンクになっておりますので、提供実績も非常に豊富であるということがお分かりかと思います。

この12万余りの登録者数の中から、比較的健康な方、あるいはありふれた多因子疾患の患者さんの試料を選んで、最後の17枚目ですけれども、コントロール群全ゲノム解析プロジェクトに参加いたしました。NCBNの目標は、その黄色にありますように8000人ですが、実際には、下の文字を見ていただきますと、9850検体の全ゲノムシーケンズ解析を完了することができました。これは、将来、CANNDsに格納されて幾つかの情報を加えて、研究者コミュニティで制限共有されることになっております。がん、難病等のゲノム医学研究のためにデータを共有していただくということになります。

ただ、既にデータが出そろっておりますので、この9850検体のデータを使いまして11の共同研究をもう開始しております。その中で利用される先生方は、まずはそこにありますCANNDsに予定されている基本臨床情報で十分であるということでスタートしておりますが、将来必要が生じれば、より詳細な臨床情報を各ナショナルセンターから提供することが可能であるということで、既に利活用していただいているという状況でございます。

私の発表は以上です。ありがとうございました。

○長野次長 徳永先生、どうもありがとうございました。

それでは、4名の有識者の先生から発表いただいたところですが、構成員のメンバーの方、ほかの方も含めて、御質問、御意見などございましたら、よろしくお願いたします。私のほうからちょっとコメント申し上げますと、今日、それぞれの先生方の様々な専門の関係からお話いただいたところですが、割と共通的に出てきたかなと思われる点、そこからの示唆については3つぐらいあるかと思います。

1つ目が、まさに日本人のゲノム研究・診断のためには、日本人のゲノムデータの基盤が非常に重要であろうという点。

2つ目は、基本の臨床・属性情報のみが見える場合であっても、規模がきちんとあれば、ゲノムデータを大規模に扱って、それによって可能となる研究開発というものがあるだろうということで、ゲノムデータを大規模に集めて、基本的な臨床情報・属性情報のみであっても、それが非常に有効であるということ。それから、その際、コントロールデータという場合には、それぞれの対象に使う場合には、その疾患に応じて使われるコントロール群が違ってくるでしょうし、あるいは一般人そのものの研究のためにコホート群と言われるデータを使うことができるだろうということ。

3番目が、ゲノムデータに付随する臨床情報については、研究によってはより多くの臨床情報が利用できる環境を整えば、より多様な研究ができるのではないかといったことがあるかなと、私としてはコメントさせていただきたいと思います。

構成員のほうから。高木構成員、よろしくお願ひいたします。

○高木構成員 高木でございます。

徳永先生中心にお聞きしたいのですが、徳永先生の最後のスライドで制限共有の話が出てまいりましたけれども、制限共有と制限公開あるいはvisitingあるいは頻度情報を完全にオープンにする。この辺りの使い分けというのでしょうか、今後どうなっていけばいいとお考えでしょうか。もしお考えがあればお聞かせください。

○徳永参考人 ありがとうございます。

大規模なデータに関しては、visiting環境を用意するというところで利活用するのが、現実的にもうそれしかできないという状況になりつつあると思います。そういう意味では、これまでの制限共有という言い方にはなるかと思うのですが、実際にデータベースにアクセスしていただいて、そのデータベースが提供する計算環境を使っていただいて計算して、その結果だけ取り出すという形になるので、従来イメージしていた制限共有というものと少し感じが違ってきているのかなと思っていますが、ある意味、かなりセキュアな環境を提供側がしっかり持っていれば、利活用しやすい状況になってきているかなと思います。

○高木構成員 もう少しいろいろ詳しく、また今度お聞かせいただこうと思います。ありがとうございます。

○徳永参考人 はい。

○長野次長 ありがとうございます。

ほか、いかがでしょうか。よろしいでしょうか。

それでは、せっかくこういった形で有識者の先生方から御発表いただきましたし、コントロール群の利用も含めて、一般人の方のゲノム利用の環境ということもありますので、AMEDのほうから何かこれに関係して、CANNDsを今まさに構築中だと思いますけれども、三島理事長、ございますでしょうか。

○三島理事長 ありがとうございます。三島でございます。

先生方、今日、いろいろと教授していただきまして、本当にありがとうございました。非常に勉強になりました。

私どもといたしましても、ユースケースがあつてのデータ基盤事業ということでございますし、あとは属性情報、健康・臨床情報を取り扱うに当たっては、種々解決しなければならない課題があるかと思いますが、意義ある研究開発を可能とするように、CANNDsの開発・運営にしっかりと反映させてまいりたいと思う次第でございます。

ありがとうございました。

○長野次長 ありがとうございます。

それでは、また改めて全体を通じて御議論もあるかと思ひますけれども、まずは議事2に移らせていただきたいと思います。議事2、厚生労働省のほうから「全ゲノム解析等実行計画の進捗について」ということで、報告のほうをよろしくお願ひいたします。

○城審議官 厚生労働省の医薬産業振興・医療情報審議官の城でございます。先日まではAMEDの立場で本協議会にお世話になっておりましたが、このたび厚生労働省に移りました。よろしくお願ひいたします。

本日は、全ゲノム解析等実行計画の進捗について、私から御報告いたします。資料2を御覧ください。

おめぐりいただきまして、2ページからになります。こちらには全ゲノム解析等実行計画の推進につきまして、本年度決定された政府方針などをまとめております。上に記載の経済財政運営と改革の基本方針、いわゆる骨太の方針2022、それから新しい資本主義のグランドデザイン及び実行計画関係でございますが、「臨床情報と全ゲノム解析等の情報を連携させ搭載する情報基盤を構築し、その利活用に係る環境を早急に整備する。新たな治療法を患者に届ける取組を推進する。」との文言が掲載されております。引き続き、厚生労働省としても、こちらの文言も踏まえ、実行計画を着実に推進していきたいと考えております。

3ページ目を御覧ください。令和元年12月に策定されました全ゲノム解析等実行計画におきましては、下の図にありますように、最大3年程度を目途に先行解析を行い、そして、その後、速やかに本格解析を行うとされておりました。令和3年度までの先行解析におきましては、主にバイオバンク、解析拠点で保存されております検体を活用して、がん、難病の全ゲノム解析を実施いたしました。令和4年度からは、全例、新規の症例を対象とした本格解析を開始しております。臨床情報を含め、データを蓄積して研究・創薬への利活用を推進していきたいと考えております。

4ページ目を御覧ください。これまでの解析実績と令和4年度の解析の予定を記載しております。

表の中央、解析実績の列を御覧ください。ここに記載がありますように、令和2年度におきましては、がん領域については、難治性のがん、希少がん、遺伝性のがんを対象に、合計して約3,800症例、また、難病領域においては、単一遺伝子性疾患、多因子疾患、診断困難な疾患を対象に、約2,500症例の解析を実施しております。令和3年度は、がん領域については約9,900症例、難病領域については約3,000症例の解析を実施しました。一番右側

にお示ししている令和4年度については、解析予定数としては、がん領域は新規の患者さんで約2,000症例、難病領域については約2,500症例の解析を予定しております。来年度以降の症例数につきましては、実際の解析の実績数を見ながら設定していきたいと考えております。

5ページ目を御覧ください。これまでの解析実績例でございます。全ゲノム解析による病的変異の同定の例をお示ししております。これはがんの全ゲノム解析により病的変異を同定した例でございます。上段は膵がんにおいて全ゲノム解析を実施したところ、長期の生存例において、塩基挿入・欠失に特徴的な偏りを示す高度な変異集積を有する患者集団が存在することが明らかとなりました。全ゲノム解析に基づいた患者層別化の例でございます。下段は、オミックス解析により、スプライシング異常が検出された事例であります。具体的には、乳がん、卵巣がんの発症に関与するとされる遺伝子におきまして、ゲノムの変異により、通常であればスプライシングの際に切り離されるイントロン部分が切り離されない異常が起こって、これらが発がんに関与している可能性が示唆されたという例でございます。これらの知見は、今後の新たな個別化医療の開発につながる可能性が示唆されます。

6ページ目を御覧ください。こちらは、難病の未診断患者に対する、疾病特定のための全ゲノム解析の可能性をお示ししています。難病領域におきましては、臨床所見や既存の遺伝学的検査では診断に至らない患者さんが存在し、その一部については、研究における全エクソーム解析や全ゲノム解析により疾患の特定に至ることもございます。そのための対象範囲や解析手法を確定させることが重要と考えられます。未診断患者さんの疾病特定のための全ゲノム解析等の可能性として、今後、医療実装を検討していく上で、これまで解析した症例をレビューし直すとともに、前向きな解析症例を組み合わせることで、例えば全ゲノム解析でないと、診断のつかない症例とはどういうものかということや、それは全体でどれぐらい存在するのかを明らかにしていきたいと考えております。

7ページ目を御覧ください。全ゲノム解析の実施体制でございます。全ゲノム解析の実施体制につきましては、令和4年度におきましても引き続き、真ん中にお示ししている「全ゲノム解析等の推進に関する専門委員会」において方針の決定を行い、専門的事項の検討を行う厚生労働科学研究班、それから全ゲノム解析の実施を担当するAMED研究班と連携を取りながら進めていきたいと考えております。

8ページ目を御覧ください。「全ゲノム解析等実行計画2022（仮称）」の案の目次をお示ししています。現在、この「全ゲノム解析等実行計画2022（仮称）」（案）につきましては、全ゲノム解析等の推進に関する専門委員会において検討を進めておりまして、来月に策定する見込みです。その目的でございますが、次のページ以降に、抜粋してお示ししております。9ページ目を御覧ください。「全ゲノム解析等実行計画2022（仮称）」（案）における、患者還元が可能な情報及びその還元方法についてでございます。まずは、骨太等に記載がございましたが、情報基盤を構築して、その利活用に係る環境を早急に整備する

ことにより、一番上の記載の通り、研究・創薬等における成果に関する情報として、医薬品開発による新規治療法等の提供を目指したいと考えております。2点目として、引き続き、日常診療に導入できる情報として、医学的意義が明確な情報の提供、適応がある場合の診断・治療等の提供をしていきたいと考えております。3点目は、新たな個別化医療等に関する情報として、臨床研究、治験等への参加機会の提供など継続していきたいと考えております。

10ページ目を御覧ください。「全ゲノム解析等実行計画2022（仮称）」（案）の内容の詳細について、本文からの抜粋をお示ししております。まず1つ目の「対象患者」ですが、がん領域におきましては、中ほどにポツが幾つか並んでいる中の2ポツ目になりますが、手術等において根治の可能性が低い難治性のがんの患者様、また、3ポツ目に記載のある、解析開始時に生存しており、今後何らかの治療の提供が期待できる状態であることとしております。

難病領域においては、引き続きですが、単一遺伝子疾患、多因子疾患、診断困難な疾患と整理しております。

それから、2つ目の「患者還元を行う医療機関」については、がん領域においては、医療機関に求められる要件を定めまして、毎年1回、新たに要件を満たした医療機関の審査・承認を専門委員会において実施し、また、医療機関の体制、実績等に係る評価を、毎年1回、専門委員会において実施したいという形の記載をしております。

続いて、11ページ目になりますが、「解析・データセンター」、「検体集中管理センター」、「アカデミアの役割について」、「産業界の役割について」を記載しております。特にアカデミアの役割については、アカデミアが主体的に全ゲノム解析等に係る学術的協議を行うためのオールジャパンの学術組織として、アカデミアが主催する「アカデミアフォーラム」を構築すると記載しております。アカデミアフォーラムには、全ゲノム解析等に係るデータを共有し、幅広いデータ利活用の権限を付与し、そのかわり、領域別に専門家によるグループを設置し、高度な横断的解析等によって新たに指摘された変異等の知見についての臨床的意義、病理学的意義を協議し、必要なデータを取りそろえた後、患者還元に値するかどうか判断する等といった役割が求められています。また、一番下の「産業界の役割について」でございますが、全ゲノム解析等により得られたデータをもとに、新たな診断技術や治療薬の開発等を目指すということで、そのために、産業界が主体的にデータ収集の過程を含めて、本事業の計画検討の当初から参画できるよう、産業界が主体の「産業フォーラム」を構築するとしております。具体的には、医療産業、非医療産業にかかわらず、またベンチャー企業等も含め、多くの企業が参画できるような組織として、令和4年度中の発足を目指しています。データの利活用の具体的な運用ルールについても定めるということでございます。

12ページ目を御覧ください。ここは「倫理的・法的・社会的課題（ELSI）」、それから「患者・市民参画（PPI）」について記載しております。

ELSIにつきましては、特に留意が必要な点として、統一的な同意説明文書、患者さんへの説明に当たっての必要な体制の整備、遺伝カウンセリングの機会の確保・拡充、全ゲノム解析に係る相談実施体制の整備等について記載しております。「患者・市民参画」としましては、本実行計画について、対象患者への周知・説明だけではなくて、広く国民や社会に対して継続的に情報発信を行い、患者・市民参画の仕組みを構築して透明性の確保と患者・市民の視点の導入に努めることが必要だということを明記しております。

13ページ目を御覧ください。事業実施組織の発足に向けた中長期的スケジュールでございます。上段の「患者還元」につきましては、今後も検査としての有効性の確立に関する検討を継続し、また、出口を見据えた臨床試験・治験等を推進することにしております。次に、2段目の「データ基盤としてのインフラ整備」としましては、システム構築に着手するとともに、解析データ・臨床情報の蓄積を並行して継続してまいります。そして、事業実施組織の発足に合わせたデータ提供による利活用開始を目指すとしております。それから、一番下の「法人」については、現在、準備室WGとなっておりますが、本年度中には、事業実施準備室を発足させ、そして、民間とも協力しながら、事業実施組織そのものの発足に向けての検討を進めるとしております。そして、当面は最適な法人の在り方についての検討を行って、必要があれば法改正を行うなどの対応を行っていきたいと考えております。

以上が厚生労働省における、全ゲノム解析等実行計画の進捗についての御報告になります。以降の資料は参考でございます。私からは以上でございます。

○長野次長 ありがとうございます。

それでは、今の厚生労働省からの説明に対しまして、御質問、御意見などありましたら、よろしく願いいたします。

菅野構成員、よろしく願いいたします。

○菅野構成員 御説明ありがとうございます。

もう専門委員会のほうで検討されていることかもしれないのですがけれども、がんの本格的なゲノム解析のほうでちょっとサゼスションがあるのであれば、がんというのは長い時間をかけて体の中で発達してきて、最終的に臨床がんとして発見されるというプロセスを取っていますので、比較的最後にでき上がったものというのは、それまでのプロセスで何が起きているかということは類推できるけれども、分かりにくいところがある。

お話を伺っていると、本格研究では新規に様々なサンプルを収集されてシーケンスされるということでしたら非常に難しいとは思いますが、いわゆる前がん病変とか非常に高率にがんが発生するだろうと思われるような病変が幾つも知られていると思いますので、そういうものの中での変異がどういうふうになっているかということも含めて解析されてもいいのではないかと。そういう選択肢については、今、検討されているのかどうかも含めて、ちょっとお伺いしたいのですが、いかがでしょうか。

○長野次長 厚生労働省、お願いします。

○城審議官 現在、実行計画では、がん、難病と診断された患者さんを対象としていますが、そういった前がん病変等における解析の意義については、個別の各研究班で検討いただくのかと思います。

○菅野構成員 そういうサンプルもシークエンスしていく可能性があるかと理解してよろしいのでしょうか。

○長野次長 厚生労働省のがん・疾病対策課が入られていますから、何かコメントがあれば。

○中谷課長 厚生労働省健康局がん・疾病対策課長の中谷でございます。今日は健康局長の代理で出席させていただいております。がんの全ゲノム解析の研究にも関わるところですので、当課から御回答させていただきます。

全ゲノム実行計画の方針に基づきまして、複数の研究班が走っているところです。それで、今、令和4年度からは出口戦略のための臨床研究を臓器別に複数立ち上げておりまして、それらの中でいかに意義のある情報を見つけていけるかというところを、まさに研究班が研究しておりますので、今、先生のいただいたコメントも適切にフィードバックさせていただければと思います。御指摘ありがとうございました。

○菅野構成員 どうぞよろしくお願いいたします。

○長野次長 それでは、米村構成員、よろしくお願いいたします。

○米村構成員 米村でございます。

御報告いただきまして、ありがとうございました。全体像としては、大変意欲的な計画を立てていただいているのではないかと考えて拝見したのですが、私に多少関わりのあるELSIの課題のところについて1点だけ質問させていただきたいと思います。12枚目のスライドに関わるところです。

ここで挙げられている検討課題というのは、基本的に統一的な説明文書を用いるとか、インフォームドコンセントの取得過程をきちんと整備するという内容になっていまして、以前もこのゲノム医療協議会の場で私、発言したような気がするのですが、これまでの既存のゲノムバイオバンクの活用というのがかなり制限されてきた一番の理由が、それぞれのバイオバンクごとに説明同意文書のばらつきがかなりあって、それによって実際に幅広い利活用ができなくなっている。特に、民間企業の研究とか、研究に当たらない民間企業の利用といったものにおいて、なかなか使いにくい形になっているというのがかなり大きな問題として認識されてきたのではないかとという気がします。

私のほうでELSIの課題ということで、先般、このゲノム医療協議会、2回にわたってお話しをさせていただいた際には、こういった同意の内容に依存する利活用の在り方というものにそもそも問題があって、ゲノムというのは、全国民あるいは全世界にとっての重要な問題なのだから、一人一人の同意に依存して大幅に利活用の在り方が変わるというのは困るので、客観ルールをきちんと明確にして、国全体で統一したルールの下に使っていくという方向性を考えていくべきじゃないかという話をしたと思います。

今、仮に例えば民間企業については、一定の説明同意文書の下で一応使えることになっているとしても、将来的にもずっと同じ形で使えるかどうかというのは分かりませんし、同意時点で将来、先々を見通した同意が取れないというのは、このゲノムの研究利用を含めたバイオバンクのガバナンスにおいて本質的な問題でありまして、何十年も使っていくべき資料や情報を集める以上は、何十年先の実際のユースケースとか科学水準というものを把握して同意を取ることは不可能ですので、同意の内容に依存して、その先、何十年もの利活用の方向性が決まってしまうということをする、という形であれ、今後常にバイオバンクの利活用が制限されることになると思います。

その問題について、きちんと検討を進めていただきたいというのが私の希望であります。私自身は、先ほど申しましたとおり、このゲノム医療協議会の場で、そういった客観ルールの整備という方向で利活用の仕組みをつくっていただきたいということをお願いしたわけで、その方向で検討していただくということで皆さんの御了解をいただいたと思っておりますので、厚生労働省においても、ぜひそのことを踏まえて、今後のこういった実行計画の中で、ELSIの課題として、そういったことをきちんと検討課題に挙げていただきたいというのが私の要望でございます。その点について、現状、どのような御検討をされておられるのかをお伺いしたいと思います。

以上です。

○長野次長 ありがとうございます。

厚生労働省からコメントございますか。

○城審議官 厚生労働省でございます。

御指摘の通り、現状この実行計画に定めている対応としては、今後状況が変化した場合などは、統一フォーマットで同意の取り直しをすることと考えておりますが、御指摘のように、そもそも医療情報について、本計画で収集している情報も含めた利用の在り方は、本計画の範囲を超えた制度改革そのものにもつながる話であり、もっと大きな制度の検討になるかと思えます。

○長野次長 米村構成員の問題提起は非常に重い問題提起だと思いますので。

○城審議官 制度改革そのものにつながる話でございます。

○長野次長 ゲノム医療実現との関係でも十分に関係者で議論が進められたらと思いますので、厚生労働省もよろしく願いいたします。

すみません、予定の時間がほぼ来ている中で、挙手されている先生方がたくさんおられるので、まとめてそれぞれお聞きして、その上でコメントをいただきたいと思えます。

まず、上野構成員のほうからよろしく願いいたします。

○上野構成員 ありがとうございます。製薬協副会長の上野と言います。

今回の実行計画2022で患者還元の方法ということで、医薬品開発による新規治療法等と明記していただいたのは、私どもにとっても非常にありがたく、これに対していかに貢献できるかというのを、これから検討を進めていきたいと思えます。その実行においても、

産業界の役割について明確に書いていただきまして、そこに産業フォーラムを設置して議論していく。特に、計画の当初から参画できるということは、従前から申し上げてきた要望がここに盛り込まれたということで、感謝申し上げます。

その点からいったときに、13ページに書かれています中長期的なスケジュールの中でいきますと、真ん中のデータ基盤としてのインフラ整備の中で、赤い矢印のデータ提供による利活用が、令和4年度から点線では書かれていますのですけれども、これは我々としては、なるべく早い段階から試行運用の中から我々も参画させていただいて、なるべく早く本格運用に至って、そこで創薬につないでいきたい。創薬というのは、スタートしてから実際には時間がかかるものなので、そういう中でいろいろ試行錯誤しながら、データ構築の中でもいろいろ意見を申し述べていきたいと思っておりますので、その点、どうぞよろしく願いいたします。

以上です。

○長野次長 ありがとうございます。

続けて、山口構成員、よろしく願いいたします。時間の都合上、それぞれ手短にいただくと大変ありがたく存じます。申し訳ございません。

○山口構成員 よろしく願います。

この協議会の主なテーマで、利活用のところ、データの一本化という問題が和泉前議長のときから語られてきたと思うのですけれども、全ゲノム解析に関しては、CANNDsなのか、あるいは厚労省に新たな組織をつくるのかという形で進んできたと思うのですが、今日、お話を聞いていると、コントロール群はCANNDsで、各がんと難病に関しては新たな厚労省の組織でという整理がなされたように伺ったのですが、そういう理解でいいのかというのが第1点と。

2番目に、R4年度までに、既のがんに関しては1万例近くのデータが集積されていますので、厚労省の事業実施組織の立ち上がりはかなり前倒ししないと、出口戦略に大きな影響を与えるのではないかと思いますので、これはコメントです。その整理はそういう形なのかというところを御回答いただければありがたいです。

以上です。

○長野次長 ありがとうございます。

回答はまとめてさせていただきたいと思っております。

次に、笠貫参与、お願いいたします。

○笠貫参与 ありがとうございます。

1つは、全ゲノム解析、ゲノム・オミックス、クリニカルシーケンス、そして臨床表現型情報を電子カルテに統合した医療情報システムをどう実現するか。それから、環境・生活習慣情報とモバイルヘルス情報も入ってくると、生命・医療情報が全て包括されるという意味では、今回の骨太に入りました総理大臣を長とする医療DX推進本部で、国家としてのグランドデザイン策定の中で、この医療ビッグデータの情報共有とデータ解析につい

て検討すべきではないかと感じました。

もう一つは、ELSIの話が先ほど出ました。全ゲノム解析、オミックス解析情報は、例えば保険加入のときの差別の問題とか個人特定の問題は、臨床診療情報と異なる特殊性があるのではないかと。そういう意味で、この前の会議でイギリスのように、罰則を設けるという話が出ています。ガイドラインは早急に検討すべきだと思います。

先日、ある企業からDTCの全ゲノム解析遺伝子検査キットのプレスリリースがありました。民間ではDTCで進められていることも踏まえた上で、社会のニーズは非常にスピード感があるという意味で、ゲノム解析に関しては、国としてELSIの検討を早急に進めるべきではないかと思えます。以上です。

○長野次長 ありがとうございます。

それでは、門田構成員、お願いいたします。

○門田構成員 門田です。

私、2点だけコメントとして申し上げます。

1点は、今、笠貫先生と米村先生がおっしゃられたことです。これは毎回、私、発言しているような気がして、ちょっと引けるのですけれども、今、内閣府の中において、このことをやっているということは、もっと大局的な視点がなくてはならないということをも思っているのですが、このELSIの問題というのは個別の問題よりもはるかに大きな問題。にもかかわらず、個別のいわゆる光の当たるの部分のことばかりが話題になるということは大いに反省すべきだということを毎回申し上げているような気がする。ぜひこれは、今、皆さんが仰っておられるように、大きな課題として大上段に構えた形のを早急にスタートするというのをやっていただきたいということが1つです。

もう一つは、先ほどアカデミアの役割について。私、医学会の立場から参加しておりますので、あえて申し上げているのですけれども、アカデミアというのは我々が最も協力できるところではないかなと思うのですけれども、このアカデミアフォーラムについても文章にはいろいろ書かれていますけれども、なかなか理解しがたい。もう少し具体的に、何を、どこを、どういう形でやっておられようとしているのか、あるいはそういうことを我々はどのような形で協力できるのかというのをもう少し具体的に示していただき、協力できるような形にしていきたいという2点を申し上げたいと思います。

以上です。

○長野次長 ありがとうございます。

それでは、申し訳ありません。時間が過ぎておりますけれども、回答さしあげるところを回答さしあげたいと思います。

まず、山口構成員からありましたCANNDs、それから全ゲノム解析実行計画の体制の話ですが、議題1のほうでもございましたが、ToMMoとかBBJやNCBNといったそれぞれのバンクでのゲノムの事業がございます。それらについても、それぞれデータの利活用を進めていくということだと思います。

それと同様に、かなり大きな規模で全ゲノム解析計画ということで、がん、難病のゲノムについて解析を進め、利活用を進める。さらには事業化を目指して、医療に向けた活用をしていくということで計画されていると認識しています。

CANNDsというのは、それと同様な、別のデータベースをつくるというものではございませんで、バンク、全ゲノム解析等実行計画などで、それぞれをつないでいって、各バンクそれぞれで使われる方はそれぞれにアクセスして使えばよろしいですし、複数のものを横断的な検索なり、プレリサーチ検索などをされる場合には、CANNDsを入り口にしていただくという、つないで入り口になって、ゲノムデータの利活用をさらに促進されるような取組ということで、CANNDsを私どもとしては考えてございます。また、全体の考え方ですとか体制、それから今後のことについて、次回以降の協議会でも御報告申し上げ、御議論いただければと思います。

それから、ほかの先生方も含めて、ELSIの問題などについては以前よりコメントいただいていることを承知しており、これから、この協議会も含め、各省と連携して、どういった議論ができるかということをお相談させていただければと思います。

アカデミアフォーラムの関係、それから全ゲノム計画の中での、あるいはゲノム医療に向けてのELSIの問題などについてもコメントございましたが、厚生労働省から何か簡単にコメントございますでしょうか。

○城審議官 アカデミアフォーラム、それから産業フォーラムにつきましては、詳細は検討している最中で、決まっておきませんので、今日いただきました御意見も踏まえ、Genomics England等の学術組織も参考にしながら、きちんと御相談しながら進めていければと思っております。

今日の段階では、現段階の構想の御紹介にとどめさせていただきたいと思っております。

引き続きよろしくお願いたします。

○長野次長 ありがとうございます。

それで、議事進行の不手際で大変申し訳ありません。時間が過ぎておりますので、その他、次回の協議会などについて事務局から。

○笠松参事官 次回の協議会につきましては、事務局より改めて日程等について、お伺いの御連絡をさせていただきます。

○長野次長 それでは、これもちまして、第9回「ゲノム医療協議会」を終了させていただきます。

本日は、貴重な御意見等を賜りまして、ありがとうございました。御多忙のところ、どうもありがとうございました。