

令和 5 年度厚生労働科学研究費補助金
政策科学総合研究事業（倫理的・法的・社会的課題研究事業）事業
分担研究報告書

ゲノムデータの持つ個人識別性に関する研究
研究分担者 石川 俊平
国立大学法人東京大学 大学院医学研究科衛生学分野 教授

研究分担者 荻島 創一
国立大学法人東北大学 大学院医学系研究科 医科学専攻 ゲノム医療情報学分野 教授

研究分担者 鎌谷 洋一郎
国立大学法人東京大学 大学院新領域創成科学研究科 メディカル情報生命専攻 複雑形質ゲノム解析分野 教授

（法的観点からの検討） 殿村 桂司
長島・大野・常松法律事務所 弁護士

研究要旨

本研究で扱うゲノムデータの個人識別性に関する調査研究については過去にも同様の趣旨のものが存在するが、本研究において改めて調査研究を行った理由としてはゲノムデータの利活用に関する背景が変化していることが挙げられる。具体的には個人識別性や社会受容性の考え方が変化していることや、全ゲノム解析等実行計画の進展、遺伝子パネル検査の保険収載ともなっており包括的なゲノムデータが国内で蓄積されていることなどがあげられる。またゲノム科学の進展によりこれまでの解釈が馴染まない場面もでてきている。本研究では、国内外の利活用状況、個人識別性の範囲に関する調査を行い、科学的観点や個人識別性等の解釈や利活用における論点に関して総合的に整理を行った。さらにそれらの論点に関して、現行の科学的、社会的解釈によって解決策が見出すことが可能なものについては具体的な提言を行い、また中期的に解決すべき課題と考えられたものについては解決にいたる考え方を提示している。前者については、がん等における後天的に発生する体細胞変異に関する情報は、現在の生物学的もしくは技術的な観点から、特定の個人の身体の一部の特徴を示す情報としては可変性の程度が大きく、個人識別符号には該当しないと考えられること、および、単一遺伝子疾患等における生殖細胞系列の遺伝的バリエーションにおける配列情報は、個人識別符号の定義に当てはまらない場合は、個人識別符号には含まれないと明示する必要があることを確認し、これらをガイドラインや Q&A へ明記することで利活用を促進することが期待されると考えられることを挙げた。また学術機関におけるゲノムデータの利活用に関する整理と周知が挙げられる。後者については公衆衛生目的の例外規定の整備やゲノムデータの利活用を推進するための新たな法整備が考えられた。現行の規制に対して短期的、中期的にこれらの考えを取り入れることで個人情報保護を考慮しつつゲノムデータの利活用が促進されると考えられた。

A. 研究目的：

ゲノムデータの利活用のあり方については、平成 28 年 10 月のゲノム情報を用いた医療等の実用化推進タスクフォース「ゲノム医療等の実現・発展のための具体的方策について（意見とりまとめ）」や平成 29 年 4 月の厚生労働科学研究費補助金「ゲノムデータの持つ個人識別性に関する研究（研究代表者;国立感染症研究所名誉所員 吉倉廣）」によ

り、国際的な議論動向を含め当時の科学技術水準における個人識別性などに考え方が取り纏められているものの、当時から一定期間も経過する中、科学技術の発展や社会受容性も変化していることからゲノムデータの個人識別性などに考え方に幾ばくかの変化があることなども考えられる。

国内においてはバイオバンク・ジャパン、ナショナル・センター・バイオバンク・ネットワークや

東北メディカル・メガバンク計画など国内の大規模バイオバンクにおける全ゲノムレベルのシーケンシングや SNP タイピング、また未診断疾患イニシアチブ (IRUD) の取り組みにおける全エクソームシーケンスによってゲノムデータの蓄積が進んでいる。また全ゲノム解析等実行計画の進展やがん患者の遺伝子パネル検査の保険収載によりがんや難病患者のがん、難病の症例に対して研究目的だけでなく医療目的で取得したゲノムデータが蓄積しつつある。これらのゲノムデータは、疾患の診断、治療のための利活用だけでなく、利活用のための周辺技術の開発など様々な目的のために二次利用することが求められている。そのために現在の科学的、社会的な現状に鑑みてゲノムデータの持つ個人識別性の考え方を整理し、ゲノムデータの提供元である個人を適切に保護しながらゲノムデータの利活用の基準をより明確化することの必要性が高まっていると考えられる。

本調査研究の目的はゲノムデータの持つ個人識別性に関する調査を行い適切な利活用に向けた規制や指針の提言を行うことである。具体的には、ゲノムデータの利活用状況についての調査・分析を行い、国内外の動向を把握し、特に個人情報の保護に関する法律（以下「個人情報保護法」という。）下での学術研究や公衆衛生目的による例外規定に関する倫理的・法的・社会的な観点から検討し具体的な課題や留意事項を抽出する。またゲノムデータの個人識別性や体細胞変異の扱いに関して、科学的な観点や国内外の議論を総合的に解釈し、適切な方策を提案することが含まれている。

B. 研究方法：

調査研究の手法として、以下の様々なアプローチで行なった。まず国内外のゲノムデータの利活用状況を文献調査やデータベースの分析を通じて具体的な事例や動向の把握を行った。次にゲノムデータの個人識別性や体細胞変異の扱いについて、科学的な研究結果や先行研究のレビューを行い、個人識別性の判断基準や規制の必要性を検討した。また学術研究例外や公衆衛生例外について、法的文献の分析や関係者へのインタビューを通じて適用範囲や具体的な規制の解釈の調査を行った。得られた情報や意見を総合し、ゲノムデータの持つ個人識別性や適切な利活用に関する提言の骨子となる論点を整理した。また論点のなかで現行の科学的、社会的解釈によって解決策が見出すことが可能なものについては具体的な提言を行い、また中期的に解決すべき課題と考えられたものについては解決にいたる考え方を提示することとした。

（倫理面への配慮）

本研究は、文献調査と有識者へのヒアリングに基づくものであり、研究の対象としては倫理に関

わる内容ではあるものの、研究の遂行については求められる特段の倫理面への配慮はないと判断した。

C. 研究結果：

研究結果については、1.体細胞変異の個人識別性について、2.単一遺伝子疾患等の遺伝的バリエーションの個人識別性について、3.包括的な DNA の配列データの匿名・仮名加工の難しさ、4. ゲノムデータの匿名・仮名加工の困難さを前提とした利活用のための方策、5.公衆衛生目的による例外規定の利用、6.学術機関におけるゲノムデータの利活用の順番で記載する。

1.体細胞変異の個人識別性について

ゲノムデータの利活用においてその運用が明確に定まっていなかったものの一つはがんなどの疾患における体細胞変異情報である。体細胞変異は受精後に細胞の DNA に発生する変異であり、子供に遺伝して生涯にわたって不変な生殖細胞系列バリエーション情報とは医学・医療上異なる扱いがされている。現状では日本の個人情報保護法には生殖細胞系列バリエーションと体細胞変異の区別についての明示的な記載はないが、現在運用されている「人を対象とする生命科学・医学系研究に関する倫理指針」として統合される前に運用されていた「ヒトゲノム・遺伝子解析研究に関する倫理指針」（参考文献 1）においては『一方、がん等の疾病において、病変部位にのみ後天的に出現し、次世代には受け継がれないゲノム又は遺伝子の変異を対象とする研究（いわゆる体細胞変異（somatic mutation）を解析する研究をいい、変異の確認のために正常組織を解析する場合を含む。）、遺伝子発現に関する研究及びたんぱく質の構造又は機能に関する研究については、原則として本指針の対象としない。』として体細胞変異は指針対象外として別の扱いをしていた経緯がある。学術研究においても、大規模ながんゲノムデータベースにおいてはそれぞれの研究機関、医療機関での倫理委員会の承認の元で、個人ごとの体細胞変異についてはインターネットで誰でもアクセスして利用可能な状態で公開されているなど、生殖細胞系列とは区別した運用がされてきた（参考文献 2）。

併せてがんの体細胞変異については、既存の分子標的治療や免疫治療のための患者層別化や新たな分子標的を探索する際に有用性が高いことから、多数の患者のゲノムデータを共有して解析するという利活用の形態が社会的にも重要と考えられる。

がんの体細胞変異についての個人識別性に関しては、先行研究「平成 28 年度厚生労働科学特別研究事業 ゲノムデータの持つ個人識別性に関する研究」の分担研究報告書「がん研究におけるゲノムデータの個人識別性について」（参考文献 3）におい

ては、がん研究等で用いるゲノムデータの特徴を、①個人と情報との結びつきの程度（一意性等）、②可変性の程度（情報が存在する期間や変更の容易さ等）、③本人到達性の観点から整理している。一意性については、『ホットスポット以外の遺伝子変異は、特定の個人に限定的なデータである可能性があり、他の研究者が有する電算記号化されたそのゲノムデータのコピーと照合することにより』一意性があるとしている。②可変性については、『治療やがん自体の進展の影響により、現存する生体の特徴としての体細胞変異情報は、変化する可能性があり可変性を持つ』としている。③本人到達性については、『ゲノムデータのみから特定の個人に到達するには別途その本人の生体試料を用いてゲノム解析を行うか、あるいは当該個人のゲノムデータの全てあるいは一部が公開あるいは別途取得されている必要があり、現実的には③本人到達性は低い』としている。これらを総合して『ゲノムデータが適確かつ正確に取得されていることを前提とし、かつ過去の病歴としてのゲノムデータではなく、現存する生体個体の特徴としてのゲノムデータの個人識別性を考える場合』、少なくとも体細胞変異のゲノムデータおよびホットスポット型生殖細胞系列バリエーションは個人識別符号とは考えられないという見解を示している。

欧州の一般データ保護規則 (GDPR) では、genetic data を、健康に関するデータ (data concerning health) と同様に、personal data の special category として定義している (GDPR 第 9 条第 1 項)。個人識別符号とは定義しておらず、生殖細胞系列バリエーションと体細胞変異は区別していない。適法に処理するためには、一般の個人データの処理に求められる適法性の根拠 (GDPR 第 6 条第 1 項) とは異なる根拠が求められており、「データ保護の権利の本質的部分を尊重し、また、データ主体の基本的な権利及び利益の安全性を確保するための適切かつ個別の措置を定める EU 法又は加盟国の国内法に基づき、第 89 条第 1 項に従い、公共の利益における保管の目的、科学的研究若しくは歴史的研究の目的又は統計の目的のために取扱が必要となる場合」に genetic data を取り扱うことができる (GDPR 第 9 条第 2 項)。

韓国の保険福祉部及び個人情報保護委員会による「保険医療データの活用に関するガイドライン (2022 年)」(参考文献 4) においては (以下日本語訳) 『2) 生殖細胞変異情報を削除した新生物 (neoplasm) 固有の体細胞変異情報』に関して『生殖細胞系列バリエーションを除去した体細胞変異情報は、癌の原因となる変異情報のみを含むことになるため、個人識別の可能性はない』としてがんにおける体細胞変異については通常のゲノムデータと異なり特別な配慮を必要としない運用が記載されており、2024 年の改正 (参考文献 5) においては

この記載そのものは見られないが、以前の内容を否定する記載は認められていない。

日本においては、上述のように個人情報保護法には生殖細胞系列バリエーションと体細胞変異の区別についての明示的な記載はないが、次世代医療基盤法においては、「次世代医療基盤法 ガイドライン II. 認定作成事業者編」(参考文献 6) において同じく上述の先行研究「平成 28 年度厚生労働科学特別研究事業 ゲノムデータの持つ個人識別性に関する研究」の分担研究報告書「がん研究におけるゲノムデータの個人識別性について」(参考文献 3) と同様の見解を記載しており『「互いに独立な 30 未満の SNP から構成されるシーケンスデータ、がん細胞等の体細胞変異、単一遺伝子疾患の原因遺伝子の (生殖細胞系列の) ホットスポット変異」については「個人識別性がほぼ無いと判断できる」レベルとされており、これらを考慮して匿名加工の方法を検討する必要がある。』と記載されており、体細胞変異についての個人識別性についてはほぼ無いと判断できるレベルであると見解が述べられている。

以上のような科学的観点と社会的な現状に鑑み、体細胞変異については個人識別符号に当たらないという見解が可能と考えられた。これまでも学術研究等においては体細胞変異については、生殖細胞系列バリエーションと異なり制限のないアクセスが可能で例が多く見られたが、個人情報保護法のガイドラインをはじめとする関連の法規に明確な記載がなかったため、運用に関して生殖細胞系列の変異と異なる扱いをして良いか不明であり利活用が十分に進んでいなかった可能性が考えられた。このような状況に鑑み個人情報保護法のガイドラインや Q&A 等に体細胞変異の扱いについて個人識別符号としないという旨を盛り込むことなどが考えられる。現行の個人情報保護法 ガイドライン 通則編 2-2 個人識別符号 (法第 2 条第 2 項関係) にはゲノムデータに関する個人識別符号の定義として『ゲノムデータ (細胞から採取されたデオキシリボ核酸 (別名 DNA) を構成する塩基の配列を文字列で表記したもの) のうち、全核ゲノムシーケンスデータ、全エクソームシーケンスデータ、全ゲノム一塩基多型 (single nucleotide polymorphism : SNP) データ、互いに独立な 40 箇所以上の SNP から構成されるシーケンスデータ、9 座位以上の 4 塩基単位の繰り返し配列 (short tandem repeat : STR) 等の遺伝型情報により本人を認証することができるようにしたもの』と記載されているが、遺伝型情報という言葉も科学的観点ではその使用法が少しずつ異なってきている。遺伝型という用語は英名 genotype の日本語訳である。国際的な用法について確認すると、例えば米国国立遺伝学研究所では、genotype とは『ゲノム上の特定の場所における配列の差異 (バリエーション) の

内容 (A genotype is a scoring of the type of variant present at a given location (i.e., a locus) in the genome.)』を示す用語としている (参考文献 7)。近年、生殖細胞系列ゲノムだけではなく、体細胞変異を多数含むがん組織の全ゲノム、あるいは分化した体細胞変異を含む末梢組織のゲノムデータ、さらには一細胞レベルのゲノムデータを得られるようになってきた。これらのゲノムごとに **genotype** を得ることは可能であって、実際に報告されている。そのため、遺伝型という用語のみでは体細胞変異を含めた解釈がなされる可能性も十分考えられる。そのため当該部の「遺伝型」という言葉が親から子に遺伝するものに限定している旨をガイドライン又は **Q&A** 等に記載することが有効と考えられた。また保険承認されている遺伝子パネル検査のデータについては今後継続的に蓄積し利活用の需要が高いことからその運用基準についても **Q&A** などで明記することにより利活用が促進されることが考えられた。遺伝子パネル検査におけるがん組織やがん細胞のシーケンスデータそのものは体細胞変異情報及び生殖細胞系列のデータの両方を含むため、そのままでは個人識別符号に該当する可能性があると考えられる。非腫瘍部のシーケンスデータを取得して生殖細胞系列の遺伝型情報を除外するか、人口集団における遺伝的バリエーションの頻度データベースや体細胞変異の変異アレル頻度 (**Variant allele fraction** または **frequency: VAF**) からの推定などにより生殖細胞系列バリエーションを除去することが考えられる (参考文献 8)。いずれの方法を用いても理論的には極少数の生殖細胞系列バリエーションが含まれることがあり得るが、体細胞変異の取り出しにあってはその時点における科学的に合理的な方法を用いて行われることが必要と考えられ互いに独立な 40SNP 以上になると個人識別性を持つということを考慮して、そのリスクに沿った適切な運用が望まれる。

体細胞変異はがんにおけるものの他に、体細胞変異を持つ血液細胞がクローン性に増殖しているもの (**Clonal hematopoiesis of indeterminate potential; CHIP** など) の研究が進み、さらには一般に胎内から始まって生涯にわたりあらゆる細胞に発生するという考えがデータからも確認されている (参考文献 9)。本報告書の対象は、これらを含み、がんに限らず体細胞変異一般である。がんにおいて知られているように、生殖細胞系列以外に生じた体細胞変異は親から子に遺伝しないため、生殖細胞系列バリエーションと異なると考えることが可能であり、がんにおける体細胞変異と同様に個人識別符号には該当しないと考えることができる。こうした個人識別符号の基準の明確化にあたっては欧州の **GDPR** に基づく十分性認定への影響の有無という点も視野に含める必要がある。**GDPR** の **special category** の **Personal Data** の定義には体

細胞変異について特別に生殖細胞系列バリエーションと別に扱う記載はないが、現状でも **GDPR** 上の **Personal Data** には含まれるが、日本の個人情報保護法上は「個人情報」に含まれないものがあり、今回の基準の明確化により直ちに十分性認定に影響があるとは考えにくいことが想定された。

2. 単一遺伝子疾患等の遺伝的バリエーションの個人識別性について

体細胞変異とは異なり、生殖細胞系列バリエーションは個人識別符号と判断されることがある。すなわち生殖細胞系列ゲノムデータは一定の情報量があるときに個人識別性があるとされる。わが国においては、前述の通り個人情報保護法ガイドラインにおける通則編 2-2 個人識別符号 (法第 2 条第 2 項関係) において個人識別符号の定義は『ゲノムデータ (細胞から採取されたデオキシリボ核酸 (別名 **DNA**) を構成する塩基の配列を文字列で表記したもの) のうち、全核ゲノムシーケンスデータ、全エクソームシーケンスデータ、全ゲノム一塩基多型 (**single nucleotide polymorphism: SNP**) データ、互いに独立な 40 箇所以上の **SNP** から構成されるシーケンスデータ、9 座位以上の 4 塩基単位の繰り返し配列 (**short tandem repeat: STR**) 等の遺伝型情報により本人を認証することができるようにしたもの』とされている。

このように生殖細胞系列ゲノムデータは一定の情報量があるときに個人識別性があるとされるが、そのみで本人に到達するわけではないと述べている研究報告を参考として念頭に置くべきである。先行研究「平成 28 年度厚生労働科学特別研究事業ゲノムデータの持つ個人識別性に関する研究」の分担研究報告書「個人特定性とゲノムデータ・遺伝的識別性の関係について」では、ゲノムデータから本人に到達するには、『個人名などの情報と共に登録されていて利用できるという条件がないと個人の特定につながらない』とされている (参考文献 10)。

平成 28 年度報告の時点と比べると、遺伝学的検査を実施可能な遺伝子疾患の数が大きく増えている。体細胞変異と同様に、このような遺伝子検査結果についてもその運用が明確に定まっていない。ゲノムデータの医療応用・研究応用を単一遺伝子疾患領域で進める際には、既知の単一遺伝子疾患の原因バリエーションと病名を蓄積し、使用制限のないデータベースとして国際的に活用することが極めて重要かつ必須と認識されるようになった。例えば、米国のデータベース **ClinVar**、英国のデータベース **HGMD** は、国境を越えて公衆衛生学的な見地から活用されている。遺伝性疾患の発症には国際的な見地からは地域性があり、わが国においても同様のデータの蓄積が必須であるが、その際に手元のデータを登録するため研究者や医師がどの

ような手順を踏めばよいか我が国においては明確な記載はなされていない。

わが国においても公的研究費による難病の大型研究によりこのような疾患原因のバリエーションと病名の情報が蓄積し、さらに、健康保険による遺伝学的検査の拡充にしたいが、遺伝子検査会社の報告書等により相当量のデータが生成されつつあり、このデータを国内外単一遺伝性疾患の診療や研究に供し、世界の公衆衛生に貢献することが潜在的に可能な状況となっている。ここで、単一遺伝子疾患の原因バリエーションの個数は、単一遺伝子疾患という言葉の指す通り単一の遺伝子による疾患であることから、個人について1箇所、もしくは複合ヘテロ接合体による発症様式である場合に2箇所である。また、バルデー・ビードル症候群のように3箇所と報告される遺伝子疾患もある(参考文献11)。これらの単一遺伝子疾患等の遺伝子検査の結果として得られる遺伝的バリエーションの配列情報について、互いに独立な40箇所未満のSNPから構成されるシーケンスデータや、9座位未満の4塩基単位の繰り返し配列(short tandem repeat : STR)である等、上述の現行個人情報保護法ガイドラインの個人識別符号の定義のいずれにも当たらない場合は、個人識別符号ではないと考えられることをここに再確認するために、「生殖細胞系列の遺伝的バリエーションにおける配列情報について、個人情報保護法ガイドラインの個人識別符号の定義に当たらない場合は、個人識別符号には含まれないと判断できる」、と改めて明示すべきである。この点についてガイドラインやQ&Aへ明記することで利活用を促進することが期待される。

なお、同様の考え方は、先行研究「平成28年度厚生労働科学特別研究事業 ゲノムデータの持つ個人識別性に関する研究」の分担研究報告書「がん研究におけるゲノムデータの個人識別性について」(参考文献3)においても記載があり、「次世代医療基盤法 ガイドライン II. 認定作成事業者編」(参考文献6)においても、この先行研究の記載と同様の見解を用いて『「互いに独立な30未満のSNPから構成されるシーケンスデータ、がん細胞等の体細胞変異、単一遺伝子疾患の原因遺伝子の(生殖細胞系列の)ホットスポット変異」については「個人識別性がほぼ無いと判断できる」レベル』と記載されている。本報告書ではさらに、「ホットスポット変異」であるかどうかにかかわらず、個人について単一遺伝子疾患の遺伝子検査結果として得られる遺伝的バリエーションについては、現行個人情報保護法ガイドラインの個人識別符号の定義のいずれにも当たらない場合は、そのみでは個人識別性がないと考えられることを改めて確認する。

ただし、現行の個人情報保護法によれば、これら単一遺伝子疾患等の遺伝的バリエーションは、氏名や生年月日など、特定の個人を識別する他の情

報と一体となっている場合には、個人情報となる。また、単一遺伝子疾患等について診断的価値を持っているバリエーション(病的バリエーション)の配列情報が特定の個人を識別する他の情報と一体となっている場合は要配慮個人情報に該当する可能性があることに留意が必要である。

3. 包括的な DNA の配列データの匿名・仮名加工の難しさ

これまで体細胞変異や生殖細胞系列の単一遺伝子疾患等の原因バリエーションにおける配列情報についてその個人識別性について述べてきたが、全ゲノムデータ、全エクソームデータ等の包括的なDNAの配列データについての個人識別性について述べる。次世代医療基盤法においては個人情報の匿名加工や仮名加工についてのガイドラインが存在し、ゲノムデータについても「13-5-2 ゲノムデータ」として『ゲノムデータ(細胞から採取されたデオキシリボ核酸(別名 DNA)を構成する塩基の配列を文字列で表記したもの)のうち、全核ゲノムシーケンスデータ、全エクソームシーケンスデータ、全ゲノム塩基多型(single nucleotide polymorphism : SNP)データ、互いに独立な40箇所以上のSNPから構成されるシーケンスデータ、9座位以上の4塩基単位の繰り返し配列(short tandem repeat : STR)等の遺伝型情報により本人を認証することができるようにしたもの』をいう。このような要素が一連のシーケンスデータの中に含まれる場合には、「個人識別符号」(個人情報保護法第2条第2項)に該当するものとして削除する必要がある。』との記載がされている(参考文献6)。しかしながら、ヒト疾患の原因遺伝子やバリエーションの探索に必要とされる情報を失わずに、匿名加工、仮名加工する具体的な方法に関しては記載されていない。人為的に特定のゲノム領域をマスクすることは可能であるが、疾患の原因遺伝子や領域の探索などの場合は、ゲノム上のどの部分に疾患関連の変異が存在するかが事前に予測がつかず全ゲノム領域の情報をマスクせずに解析することが本質的に必要となる。また解析時点において変異や多型が報告されていない領域であっても、将来的に新しい人口集団や疾患患者のゲノムデータが蓄積した時点で、個人識別や疾患のリスクに重要なゲノム領域が新たに出現するため、現在マスクの必要がなくても将来的に必要性が生じる可能性がある。ゲノムデータの匿名化の手続きや技術については多くの議論があるものの、GDPRでは匿名化プロセスがどのように実施されるべきか、または可能であるかに関して手続き的な指針を提供していない(参考文献12)。以上のことから、現時点での科学的な観点から、全ゲノムデータ、全エクソームデータ等の包括的なDNAの配列データを疾患や

その他の形質の探索的研究に必要な情報量を保ったままで、現状の個人識別符号の定義に該当しない匿名加工、仮名加工を行うことは極めて難しいと判断される。

4. ゲノムデータの加工の困難さを前提とした利活用のための方策

前項においてはゲノムデータの中でも包括的な DNA の配列データの十分な情報を保持したままでの匿名加工・仮名加工の難しさを論じた。ゲノムデータの利活用についてはその公衆衛生学上、保健衛生学上のメリットは大きく、そのため包括的な DNA の配列データ自体の匿名加工・仮名加工が技術的には難しいという前提でも被験者の保護を前提としつつゲノムデータの利活用の促進を促す手段について考慮するのが良いと考えられる。考え方の一つとして、ゲノムデータの特殊性に鑑み、一般的な個人情報とは異なりデータそのものの加工を行わず、その代替措置として被検者保護のための追加の規制を設けることにより、仮名加工情報相当とすることが考えられる。追加の規制については、ゲノムデータに特異的な個人識別行為やセキュリティに関わる具体的な禁止項目等の追加、その他データの存在場所当に関する管理項目の追加、罰則の強化などが考えられる。より具体的には様々な配列検索ツールをインターネット上で個人の DNA の配列データを入力して検索を行うこと、個人以外にも個人の血縁者のゲノムデータが他に存在するかどうかを検索する行為、個人の詳細な人種的背景を特定する行為、就職・昇進における差別、生命保険での不適切な使用などがあげられる。追加の規制の中核は、ゲノムデータの転々流通を防止する機密性の担保による被検者や患者の保護であり、参考となる英国や EHDS (European Health Data Space) では、ゲノムデータを含む医療データを、visiting 環境である Trusted Research Environment で解析することで機密性を担保し、そこでは一定程度の個人特定性があることは許容されている (参考文献 13)。

さらに包括的な DNA 配列データの匿名加工・仮名加工が原則難しいという特殊性に鑑み、ゲノムデータ・情報を個人識別符号としたまま利活用を促進するための新しい法規 (特別法) を作るという考えもあり得る。現在のゲノムデータを含めた医療情報の規制は情報の加工や同意取得など最初の提供段階で入口規制をかける考えが基本となっているが、実際のエンドユーザーがどういう目的でどれだけの情報を使うのかによって規制する出口規制の考えも取り入れることも考慮に値する。またゲノムデータは通常の個人情報と異なり、他者と共通のデータを持っているという共有性があり、通常の個人情報・プライバシーのルー

ルでは適切な問題解決にならないという考え方も存在する。これに関わる基本的な考え方として UNESCO 「ヒトゲノムと人権に関する世界宣言」 1997 年 (参考文献 14) がある。ここでは「第 1 条「ヒトゲノムは、人類社会のすべての構成員の根元的な単一性並びにこれら構成員の固有の尊厳及び多様性の認識の基礎となる。象徴的な意味において、ヒトゲノムは、人類の遺産である。」第 15 条「国は、人権、基本的自由及び人間の尊厳の尊重を保障し、公衆の健康を保護するため、この宣言に述べられている諸原則を十分に尊重しつつ、ヒトゲノムに関する研究の自由な実施のための枠組みを提供するための適切な措置を講ずべきである。」第 18 条「国は、この宣言に述べられている諸原則を十分かつ適切に尊重しつつ、ヒトゲノム、ヒトの多様性及び遺伝学的研究に関する科学的知識の国際的普及を引き続き促進し、そのことに関し、科学的、文化的協力、特に先進国と開発途上国の間のそのような協力を促進するため、あらゆる努力を払うべきである。」と記載され、ヒトゲノムに関する共有の財産という考え方にに基づき、公衆の健康の向上のメリット、国際的貢献という視点で活用する趣旨が述べられている。以上のような背景から、国レベルの十分な利活用が可能で国民が十分にメリットを享受できるゲノムデータに特化した新たな法規制の枠組みを策定することが考えられる。並行して国民の理解を得るために日本においては、セキュリティに加えて国レベルでデータを使うことの公衆衛生上、保健衛生上のメリットについて、社会的な認知を促進することも必要となる。このような概念は令和 5 年 6 月 16 日に施行された「良質かつ適切なゲノム医療を国民が安心して受けられるようにするための施策の総合的かつ計画的な推進に関する法律」(いわゆるゲノム医療法) に対する日本医師会・医学会の提言にも同様の趣旨のことが盛り込まれている (参考文献 15)。具体的には「2. 遺伝医療・ゲノム医療実現に向けた研究を格段に推進するためには、患者や被験者のプライバシーを尊重しつつも、医療・医学に関わる個人情報、例えば受診履歴、診療情報、レセプト情報、個人遺伝情報等を研究開発に柔軟に利活用できる枠組みの整備が必要である。そのために、総務省と厚生労働省は、改正個人情報保護法を一部阻却し当該分野を所管する特別法 (個別法) を制定する意義について検討を進めるべきである。」とし、より具体的に特別法の在り方に際し「改正個人情報保護法を、同意取得のみに頼らず、利活用の用途や様態で規制していく抜本的な変革を行うことにより、ゲノム研究やゲノム医療の推進のために多様な個人情報をゲノムデータと結びつけて利活用することを可能とすることが必要である。その際、出口規制を強化し、不適切な利用に対する厳罰化

を検討すべきである。」と提言している。

ここまで述べてきたように、我が国では、現行の個人情報保護法の下でゲノムデータの利活用を十分に推進することが困難な状況がある。一方、GDPRに基づく英国やEU各国では、活発なゲノムデータ利活用が実現している。実際に英国のUK BiobankやGenomics Englandでは、十分な被験者保護措置を講じることで、世界中の研究者または企業に対し、visiting環境でゲノムデータ提供をしており、参加者から個別の提供先への明確な同意を得ることなく包括同意の元を実施している。我が国においても、個人情報保護法におけるゲノムデータ単体の取り扱いの見直しなどの方策を取ることによって、英国や欧州と同等の個人情報保護とゲノムデータ利活用の両立を図る必要がある。これにより、ゲノムデータの利活用を一層推進し、医療や研究の発展に貢献することが期待される。

5. 公衆衛生目的による例外規定の利用

個人情報保護法においてはいわゆる公衆衛生目的における例外規定があり、「公衆衛生の向上又は児童の健全な育成の推進のために特に必要がある場合であって、本人の同意を得ることが困難であるとき。」においては、あらかじめ本人の同意を得ないで、個人情報を当初の利用目的の達成に必要な範囲を超えて取り扱うこと（第18条第3項第3号）、あらかじめ本人の同意を得ないで、要配慮個人情報を取得すること（第20条第2項第3号）、あらかじめ本人の同意を得ないで、個人データを第三者に提供すること（第27条第1項第3号）等が許容されている。6.に述べる学術研究に係る例外規定と併せて現行の例外規定のような考え方によってゲノムデータの利活用の促進の余地があるかについても検討した。

「個人情報の保護に関する法律についてのガイドライン」に関するQ&A（参考文献16）においては、製薬企業の行う探索的研究開発等について、「製薬企業が行う有効な治療方法や薬剤が十分でない疾病等に関する疾病メカニズムの解明、創薬標的探索、バイオマーカー同定、新たな診断・治療方法の探求等の研究は公衆衛生の向上に特に資するもの」等が公衆衛生の向上に特に資すると解釈されている。また「本人の同意を得ることが困難な場合」についても、例えば、「医療機関等が保有する患者の臨床症例に係る個人データを、観察研究のために他の医療機関等へ提供する場合であって、同意を取得するための時間的余裕や費用等に照らし、本人の同意を得ることにより当該研究の遂行に支障を及ぼすおそれがあるときには、同号の規定により本人の同意なしに製薬企業に提供可能」と記載され、費用や時間に照らして本人同意を得

ることが現実的でない場合にも「本人の同意を得ることが困難であるとき」と解釈されている。これらガイドラインやQ&Aの解釈により、ゲノムデータを適切な扱いによって公衆衛生例外の考え方の下で学術研究機関や医療機関から民間企業に提供して利活用することも可能と考えられる。

一方で、Q&Aに応用例は少数の事例があるのみで、具体的な適用の可否を判断する基準が明瞭であるとはいいがたい。具体的には既存薬を使った適用拡大などはこれまでのunmet needsを満たすという点で公衆衛生の向上に資すると考えられるがこれが公衆衛生例外の適用になるのか、また「結果が広く共有・活用されていくことで」という文言が、学会、論文等の結果公表など以外に、具体的な製品の普及という意味での共有・活用が含まれるかなどは明らかにすることが必要と考えられる。Q&A以上の判断根拠が明記されているものではなく、聞き取り調査においてもその運用基準が明確でないでこれを根拠にして個人情報を扱うことの難しさが明らかになった。現在は医療機関や企業など使用者が独自に判断するため使用の範囲を規制するのが困難であり、真に必要な範囲への線引きが難しくなり、患者や被験者の保護が課題となると考えられる。公衆衛生目的における例外規定については学術研究における三省倫理指針のような運用指針や学術機関における倫理委員会のような審査や助言を専門に行う組織が整備されている状態ではないと考えられる。個別の医療機関や民間企業が対応できるような現実的な運用指針と共に頻繁に問題点となりやすい点についてはより明確な運用指針の作成、具体的なユースケースのような形でガイドラインやQ&Aに盛り込むと共に、審査や助言を行う組織の設置など、運用の促進をはかるための施策が必要と考えられた。

6. 学術機関におけるゲノムデータの利活用

個人情報保護法においては、利用目的による制限、要配慮個人情報の取得制限、第三者提供の制限等について学術研究は例外とされて、代わりに「人を対象とする生命科学・医学系研究に関する倫理指針」による運用が行われている。学術機関同士の個人情報のやり取りについては、第27条1項に記載の学術研究における第三者提供が行われていることが多いと考えられる。第27条第1項第6号の考え方をを用いると「学術研究機関等が個人データを提供する場合であり、かつ、当該学術研究機関等と共同して学術研究を行う第三者（学術研究機関等であるか否かを問わない。）に当該個人データを学術研究目的で提供する必要がある場合（当該個人データを提供する目的の一部が学術研究目的である場合を含み、個人の権利利益を不当に侵害するおそれがある場合を除く）」については、第三者への個人データの提供にあたって本人の同意は不

要である（参考文献 17）。これについては特に学術機関の現場において運用上問題になることは少ないと考えられる。むしろ学術機関と民間企業との共同研究に伴うゲノムデータを含む個人情報のやり取りにおいて混乱が起こりやすい。上記の「当該個人データを提供する目的の一部が学術研究目的である場合を含み、」という記載については、企業と学術機関が共同で行う研究においてはその全体もしくはその一部を学術研究と考えられるケースも存在すると思われるが、これらの運用基準をより明確化した上で、個人の権利保護に留意しつつゲノムデータを含む個人情報の範囲を明確にした上で共有する考え方も可能である。

学術機関から民間企業へのゲノムデータを含む個人情報の提供については、適切な本人同意の有無が議論になりやすいが、上述の学術研究における例外規定以外に個別の本人同意を必要とせずに第三者提供が可能な代表的なケースとして委託によるものと共同利用によるものがある（個人情報保護法第 27 条第 5 項）。「人を対象とする生命科学・医学系研究に関する倫理指針」ガイダンス（参考文献 18）においても、個人情報保護法のこの該当部の趣旨に沿った記載が存在しガイダンスの「第 81 (3)他の研究機関に既存試料・情報を提供しようとする場合」該当部には①委託に伴って提供する場合（個人情報保護法第 27 条第 5 項第 1 号参照）、②共同利用に伴って提供する場合（個人情報保護法第 27 条第 5 項第 3 号参照）には、第 8 の 1 の (3) の手続を履行する必要はないとされている。ただし、①委託に伴う提供の場合、提供先の機関は、委託された業務の範囲内で試料・情報を取り扱う必要があり、また、②共同利用に伴う提供の場合、提供先の研究機関は、その特定された利用目的の範囲内でのみ試料・情報を取り扱う必要がある。このなかの「特定された利用目的の範囲」について、個人情報保護法では、第三者提供先を必ずしも個別の機関として特定する必要はなく「製薬企業」「医薬品の開発」のような提供者に伝わるように利活用範囲を限定した表現が可能であり、このような記載を同意取得の際の説明文書に盛り込むことにより同意取得ときに具体的になっていなかった提供先についても個別の再同意取得を必要とせずに共同利用という考え方にに基づき、個人の保護措置を講じた上でゲノムデータを含む個人情報の提供が可能と考えられる。

近年ゲノムデータを用いた学術論文等では学術的検証や幅広い利活用のために公共データベースにゲノムデータを入れることが求められるようになって来ている。個人識別符号となるゲノムデータを入れた公共データベースは原則コントロールアクセスと呼ばれる審査を受けて認められるユーザーのみがアクセス可能であるが、研究開始もしくは提供者の同意取得時には想定していなかった

多くのユーザーの 2 次的な利活用が行われこのなかには民間企業も含まれる場合もある。このようなデータベースにおける 2 次利用についても上述の考え方をうけて提供先（2 次的な利用者）を必ずしも個別に特定する必要はなく「製薬企業」「医薬品の開発」のような提供者に伝わるように利活用範囲を限定した表現を用いて包括同意として取得することで、提供したデータベースから試料やゲノムデータを学術機関や民間企業に提供して利活用することが可能となる。なお、この場合に利用するデータベース又はその運営主体をどの程度特定すべきかについても、明確化が望まれる。この点において説明文書、同意書に関して、倫理的視点だけでなく企業法務から見ても利用が可能であることが明確に読み取れる内容になっていることが利活用の促進に必要と考えられる。

この項でのべた学術機関におけるゲノムデータの利活用法、第三者への提供方法については、各医療機関や学術機関にそのような使用法が知識として普及しているとはいいい難い状況であり、個別の研究者や倫理担当部署の知識の普及の程度によって利活用が進んでいない状況が考えられる。そのため学術機関におけるゲノムデータの利用についてユースケースとしてわかりやすく提示することが有効であることが考えられた。これまで学術研究機関のみで当初の目的に沿って使われていたゲノム研究が多く、企業での利活用やバイオバンクにおける幅広い第三者利用に対応した同意取得等の手続きがなされていないケースが存在する。今後の民間企業を含めた利活用の推進のためには、学術研究でない企業利用を想定し企業法務の視点でも問題のないインフォームドコンセントのモデル文書の整備を行うことで普及を促進することも必要となると考えられる。

D. 考察:

ゲノムデータの利活用は公衆衛生学上、保健衛生上のメリットが大きいと考えられるが、調査の結果現行の法令や指針において、その条件や範囲が明確でないことにより企業や学術研究機関における利活用の阻害要因になっていることが考えられた。そのためにゲノムデータが利活用できる条件、範囲を現行の法令や指針により明確に記載するとともに、各機関の研究者や倫理担当部署を含めてその知識普及に努めることが必要と考えられた。

体細胞変異は受精後に細胞の DNA に発生する変異であり、生殖細胞系列バリエントとは異なる扱いがなされてきており、個人識別符号に該当しないという解釈が可能と考えられた。日本の個人情報保護法には生殖細胞系列バリエントと体細胞変異の区別についての明確な記載はないが、先行する「ヒトゲノム・遺伝子解析研究に関する倫理指

針」では体細胞変異は対象外として扱われていた経緯がある。また、韓国では体細胞変異情報は個人識別性がない運用をおこなっており、次世代医療基盤法のガイドラインにも同様の見解が記載されている。さらに、単一遺伝子疾患等の遺伝子検査で得られるような遺伝的バリエーションにおける配列情報については、現行の個人情報保護法ガイドラインの定める定義に当てはまらない場合は、個人識別符号に該当しない旨を明示すべきと考えられた。実際に米国や英国など国際的には単一遺伝子疾患の原因バリエーションと病名を蓄積したデータベースが活用されており、このような取り扱いは国際的なプライバシー保護の現状とも一致し、かつ診療や研究、公衆衛生に貢献することが明らかである。一方で全ゲノムデータ、全エクソームデータ等の包括的な DNA の配列データの仮名加工は技術的に困難と考えられた。全ゲノムデータや全エクソームデータなどの包括的な DNA の配列データは、重要な領域を解析前に予見することが困難であり、一部をマスクすることが難しい。また、個人識別性や疾患リスクと関わる領域が経時的に変化する可能性があるため、現時点での科学的観点からは、包括的な DNA の配列データについて匿名加工や仮名加工を行うことは難しいと判断される。このような考え方のもと、ゲノムデータの利活用のためには、新しい規制を作ることと有効であると考えられた。ゲノムデータ自体の加工を行わず、被験者保護のための追加の規制を設けることで仮名加工情報相当とすることや、ゲノムデータ・情報に特化した特別法を策定し、出口規制の考え方を取り入れることが検討に値すると考えられた。このような新しい法規制の推進においては国レベルでゲノムデータを使うことの公衆衛生上や保健衛生上のメリットについて社会的な認知を促進することが必要と考えられる。本調査研究においては、公衆衛生目的による例外規定や学術研究に係る例外規定といった現行の規制の解釈によってゲノムデータの利活用の促進の余地があるかについても検討した。公衆衛生目的による例外規定の利用も、ゲノムデータの利活用を促進できると考えられる。個人情報保護法には公衆衛生目的による例外規定があり、本人の同意を得ることが困難な場合には、特定の条件下で個人情報を当初の目的を超えて利用することや第三者に同意なく提供することが許容されている。ただし、この例外規定の具体的な適用基準や運用指針が明確でないため、運用の促進を図るためには具体的なユースケースや運用指針を作成し、審査や助言を行う組織の設置が必要と考えられた。学術研究においては、共同利用の考え方や説明文書の記載法の周知で利活用を促進できる状況が存在すると考えられる。学術研究機関同士の個人情報のやり取りについては第三者提供の際の本人同意が不要な場合も多く、また共同利用の仕

組みを利用することで一定の条件のもとで再同意を必要とせずに企業利用が可能となる。また、企業法務の視点でも問題のないインフォームドコンセントのモデル文書の整備が必要であり、このような取り組みが利活用の促進につながると考えられる。

これらの利活用のために必要な知識については、各医療機関や学術機関の研究者や倫理担当部署に普及しているとはいい難いく利活用が進んでいない要因の一つと考えられたため Q&A やユースケースとして普及を図ることが有効である。

さらにはここで述べた利活用については個人情報の提供者、例えば患者、個人の権利利益を不当に侵害しないということが前提条件であり、特に様々な例外規定や共同利用等で再同意取得なしで第三者に渡ったゲノムデータが規定された範囲内で適切に共有されているかを保証する仕組みを合わせて整備し、提供者側のゲノムデータの散逸に対する不安のないような運用を行う必要がある。

【まとめ】

- ・がん等における後天的に発生する体細胞変異に関する情報は、現在の生物学的もしくは技術的な観点から、特定の個人の身体の一部の特徴を示す情報としては可変性の程度が大きく、「遺伝型情報により本人を認証することができるようにしたもの」に該当しないため、個人識別符号には該当しないと考えられる。
- ・単一遺伝子疾患等における生殖細胞系列の遺伝的バリエーションにおける配列情報のうち、「ゲノムデータ（細胞から採取されたデオキシリボ核酸（別名 DNA）を構成する塩基の配列を文字列で表記したもの）のうち、全核ゲノムシーケンスデータ、全エクソームシーケンスデータ、全ゲノム一塩基多型（single nucleotide polymorphism : SNP）データ、互いに独立な 40 箇所以上の SNP から構成されるシーケンスデータ、9 座位以上の 4 塩基単位の繰り返し配列（short tandem repeat : STR）等の遺伝型情報により本人を認証することができるようにしたもの」に該当しないものについては、個人識別符号には含まれない旨を改めて明示する必要がある。

E. 結論:

現行の法令や指針では、ゲノムデータの利活用条件や範囲が明確でなく、これが企業や学術研究機関の活用を妨げていると考えられた。体細胞変異や、単一遺伝子疾患等の原因バリエーションのうち現行ガイドラインで示される個人識別符号の定義

に当てはまらない配列情報については、現在の科学的・社会的観点から見ても、さらには国際的な取り組みの状況を参照しても、個人識別符号とはみなされないとする考えが可能である。これらに関し法規制の内容を明確化することは、学術機関におけるゲノムデータの利活用に関する整理と具体的な Q&A やユースケースの提示と併せて、短期的にゲノムデータの利活用を推進できると考えられた。公衆衛生目的の例外規定の運用に係る整備やゲノムデータの利活用を推進するための特別法のような新たな法整備についても中期的に解決を図ることで利活用が大きく推進されると考えられる。第三者に渡ったゲノムデータの規定された範囲内の中での適切な管理を保証する仕組みも併せて検討が必要となる。

F. 健康危険情報 なし

G. 研究発表 なし

1. 論文発表 なし

2. 学会発表 なし

H. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得 なし

2. 実用新案登録 なし

3. その他 なし

I. 謝辞

本報告書の作成にあたり、次の先生方から貴重なご意見を伺いました。深く御礼申し上げます(五十音順): 安中良輔先生(日本製薬工業協会)、石浦浩之先生(岡山大学)、板倉陽一郎(ひかり総合法律事務所)、井上悠輔先生(京都大学)、小崎健二郎先生(慶應大学)、加藤尚徳先生(KDDI 総研)、鈴木正朝先生(新潟大学)、鬼頭正博先生(日本製薬工業協会)、白神昇平先生(日本製薬工業協会)、高木浩光先生(産業技術総合研究所)、徳永勝士先生(国立国際医療研究センター)、中釜斉先生(国立がん研究センター)、松本直通先生(横浜市立大学)、水澤英洋先生(国立精神・神経医療研究センター)、武藤香織先生(東京大学)、森田朗先生(次世代基盤政策研究所)、山本雅之先生(東北大学)、吉野孝之先生(国立がん研究センター)、米村滋人先生(東京大学)

文献

1. ヒトゲノム・遺伝子解析研究に関する倫理指針(平成20年12月1日一部改正)

<https://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/genome/0504sisin.html>

2. がん体細胞変異等のデータベースcbiportal

<https://www.cbiportal.org/>

3. 平成28年度厚生労働科学特別研究事業 ゲノムデータの持つ個人識別性に関する研究 分担研究報告書 がん研究におけるゲノムデータの個人識別性について https://mhlw-grants.niph.go.jp/system/files/2016/161031/201605001A_upload/201605001A0007.pdf
4. 韓国保健福祉部 保険医療データの活用に関するガイドライン(2022年) https://www.khis.or.kr/board.es?mid=a10301000000&bid=0001&list_no=808&act=view
5. 韓国保健福祉部 保険医療データの活用に関するガイドライン(2024年) https://www.mohw.go.kr/board.es?mid=a10501010200&bid=0003&act=view&list_no=1480106
6. 次世代医療基盤法 ガイドラインII. 認定作成事業者編 <https://www8.cao.go.jp/iryuu/hourei/pdf/guideline.pdf>
7. 米国国立遺伝学研究所Webサイト”genetics-glossary”ページ、2024年5月17日改訂 <https://www.genome.gov/genetics-glossary>
8. A computational approach to distinguish somatic vs. germline origin of genomic alterations from deep sequencing of cancer specimens without a matched normal. PLoS Comput Biol. 2018 Feb 7;14(2):e1005965.
9. The origins and functional effects of postzygotic mutations throughout the human life span. Science 2023; 380(6641):eabn7113.
10. 平成28年度厚生労働科学特別研究事業 ゲノムデータの持つ個人識別性に関する研究 分担研究報告書 個人特定性とゲノムデータ・遺伝的識別性の関係について https://mhlw-grants.niph.go.jp/system/files/2016/161031/201605001A_upload/201605001A0010.pdf
11. Triallelic Inheritance in Bardet-Biedl Syndrome, a Mendelian Recessive Disorder. Science 2001; 293(5538): 2256-2259.
12. Re-identifiability of genomic data and the

GDPR。 EMBO Reports (2019) 20: e48316

13. 第2回健康・医療・介護情報利活用検討会医療等情報の二次利用に関するワーキンググループ2024（令和6）年1月11日資料
<https://www.mhlw.go.jp/content/10808000/001188758.pdf>
14. UNESCO「ヒトゲノムと人権に関する世界宣言」1997年
<https://www.mext.go.jp/unesco/009/1386506.htm>
15. 日本医学・日本医師会：「良質かつ適切なゲノム医療を国民が安心して受けられるようにするための施策の総合的かつ計画的な推進に関する法律」に関する提言（令和6年3月13日）
https://www.med.or.jp/dl-med/teireikaiken/20240313_3.pdf
16. 「個人情報の保護に関する法律についてのガイドライン」に関するQ&A
https://www.ppc.go.jp/all_faq_index/
17. 個人情報保護委員会 令和3年改正個人情報保護法について
https://www.soumu.go.jp/main_content/000790352.pdf
18. 人を対象とする生命科学・医学系研究に関する倫理指針ガイダンス 令和3年4月16日（令和5年4月17日一部改正）
<https://www.mhlw.go.jp/content/001087864.pdf>