



ライフサイエンス分野の統合データベース整備の  
第二段階のあり方について(報告)

平成 24 年 7 月

独立行政法人 科学技術振興機構

バイオサイエンスデータベースセンター運営委員会

NBDC

## 目次

1. はじめに	1
2. NBDC 発足の経緯	2
3. 第一段階における活動と成果	3
4. 第二段階のあるべき姿	6
5. バイオインフォマティクス分野における人材育成について	9
6. おわりに	12

### 【参考資料】

ライフサイエンス分野の統合データベース整備の第二段階のあり方について（報告）（概要）	13
NBDC 運営委員会 委員名簿	16
NBDC 運営委員会 開催実績	16
NBDC における平成 23 年度の主な取り組み	17

## 1. はじめに

我が国におけるライフサイエンス分野のデータベースに関する取り組みとしては、平成 13 年度に独立行政法人 科学技術振興機構（JST）に設置されたバイオインフォマティクス推進センター（BIRD）によって、データベースの構築、及び情報解析技術の開発が進められた。さらに、データベースの統合に向けた取り組みについては、大学共同利用機関法人 情報・システム研究機構（ROIS）に設置されたライフサイエンス統合データベースセンター（DBCLS）を中核機関として、文部科学省においてライフサイエンス分野の統合データベース整備事業（統合データベースプロジェクト）が平成 18 年度から推進された。

データベースを統合的に提供するナショナルセンターを整備することにより、データを提供する側にとっては研究プロジェクトが終了した後も恒久的にデータベースを提供することが可能になり、また必要に応じてデータを二重化して持たせることにより、事故や自然災害等によるデータの消失を防ぐことが可能になる。一方、データを利用する側から見ると、様々な機関で長い期間に亘って蓄積されてきた研究の成果であるデータベースを、簡単に探し出して利用することができ、さらには統一されたフォーマットで提供された複数のデータベースを同時に解析したり、自分の実験データと組み合わせで解析したりすることにより、単独のデータベースでは得られなかった知見を得ることが可能になる。その結果としてライフサイエンス分野の研究が活性化し、関連する産業の国際競争力が高まることが期待される。

我が国における恒久的かつ一元的な統合データベースの必要性については、平成 21 年 5 月に内閣府総合科学技術会議 ライフサイエンス PT によって取りまとめられた「統合データベースタスクフォース報告書」に謳われている。その実現に向けて、先行する取り組みの成果を引き継ぎ、ライフサイエンス分野のデータベース統合を進めることによってデータの価値を最大化する中核機関を目指し、平成 23 年 4 月に JST にバイオサイエンスデータベースセンター（NBDC）が設置された。

上記報告書では、平成 23 年度から平成 25 年度を第一段階と定め、統合データベースのためのセンターを JST の新たな組織として整備することが提言されている。その一方で、平成 26 年度以降の第二段階については、第一段階で整備した成果を考慮しつつ、バイオサイエンスデータベースセンター運営委員会等の議論も踏まえ、引き続き総合科学技術会議で検討していくことが必要であるとされている。

平成 26 年度からの第二段階を円滑にスタートするための予算要求、体制の整備等を考慮すると、第二段階に向けた統合データベースの基本的な方向性については平成 24 年度中には結論を得る必要があると考えられる。そこでバイオサイエンスデータベースセンター運営委員会（NBDC 運営委員会）では、NBDC が発足して 1 年余りの時期ではあるが、NBDC 発足の経緯、これまでの活動状況を踏まえて、第二段階に向けた検討を行った。

## 2. NBDC 発足の経緯

### (1) BIRD (平成 13 年度～平成 23 年度)

平成 12 年にヒトゲノムの概要配列が決定され、いわゆるポストゲノムシーケンス研究が激しい国際競争の中で本格化してきた。こうした国際的な動きの中で、平成 12 年 11 月に、科学技術会議 ライフサイエンス部会 ゲノム科学委員会報告書「ゲノム情報科学における我が国の戦略について」が取りまとめられた。本報告書を受けて、JST (当時は科学技術振興事業団) は平成 13 年 4 月にバイオインフォマティクス推進センター事業を開始し、「データベース構築」及び「情報解析技術開発」を推進した。なお、「人材養成」については科学技術振興調整費により大学や研究所での人材養成に対する支援が実施された。

### (2) 統合データベースプロジェクト (平成 18 年度～平成 22 年度)

その後、ライフサイエンス分野の著しい発展により膨大な量のデータが産出されるようになると、我が国のライフサイエンス分野のデータベース整備に関する取り組みの遅れが認識されるようになった。こうした背景の下、平成 18 年 5 月に文部科学省ライフサイエンス委員会 データベース整備戦略作業部会による報告書「我が国におけるライフサイエンス分野のデータベース整備戦略のあり方について」がまとめられ、文部科学省は平成 18 年度から 5 年間のプロジェクトとして、統合データベースプロジェクトを開始した。ROIS に設置された DBCLS は中核的な機関としてこのプロジェクトを推進した。

## 統合データベース整備のロードマップ

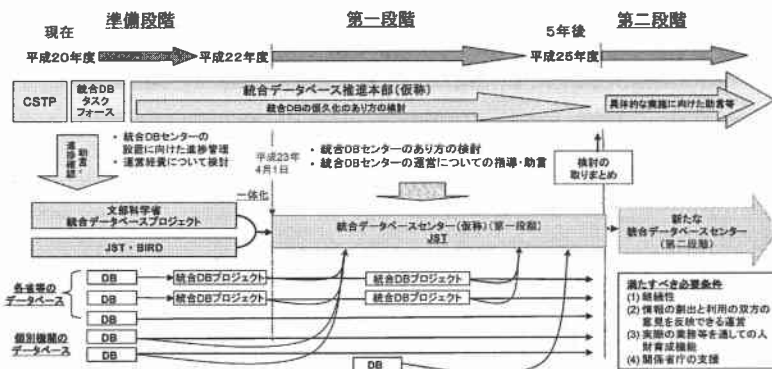


図1 統合データベース整備のロードマップ  
(統合データベース タスクフォース 報告書 (H21.5.27) より)

### (3) NBDC (平成 23 年度～)

平成 20 年 10 月に科学技術・学術審議会 研究計画・評価分科会 ライフサイエンス委員会にライフサイエンス情報基盤整備作業部会が設置された。作業部会の報告書である「ライフサイエンスデータベースの統合・維持・運用の在り方」(平成 21 年 1 月 13 日) では、ライフサイエンス分野のデータベースの整備を進めるに当たって、DBCLS と BIRD の一体的な運用を、JST に設置する新たな組織で行うことが提言としてまとめられた。

さらに、総合科学技術会議 ライフサイエンス PT 統合データベース タスクフォースの場で、データベースの統合に向けた具体的な制度設計や行動計画が議論され、「統合データベース タスクフォース報告書」(平成 21 年 5 月 27 日) では統合データベース構築のための体制整備、ロードマップ等が取りまとめられた(図1)。これらの検討に基づき、平成 23 年 4 月に NBDC が発足した。

## 3. 第一段階における活動と成果

図1にも示されているとおり、統合データベースの整備については、「統合データベース タスクフォース 報告書」の中で、平成 23 年度から平成 25 年度までを第一段階として JST の新たな組織を立ち上げ、平成 26 年度以降の第二段階については、NBDC 運営委員会等の議論も踏まえ、引き続き総合科学技術会議において検討していくことが必要であるとされている。

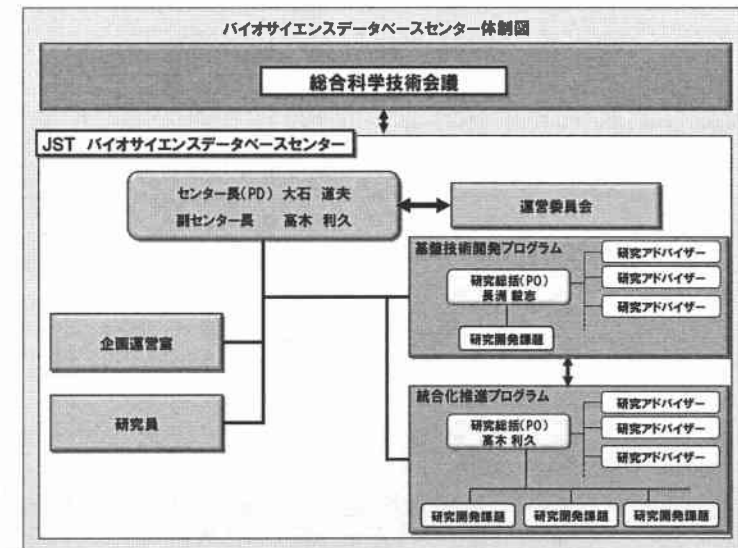


図2 NBDC 体制図

第一段階における NBDC の体制は図 2 のとおりである。NBDC では、以下の 4 つの柱に基づき事業を推進している。

- 1) 戦略の立案
- 2) ポータルサイトの構築、運用
- 3) データベース統合化基盤技術の研究開発
- 4) バイオ関連データベース統合化の推進

平成 23 年度における成果、ならびに第一段階中に予定している活動としては主に以下が挙げられる。

- 1) 戦略の立案
- a) ヒト由来データの取り扱いに関する検討

ヒト由来のデータは、ライフサイエンス分野のデータベースとして重要なコンテンツであるばかりでなく、倫理面から他の生物種のデータとは異なる配慮が求められる。我が国においては、データベースにおけるヒト由来データの取り扱いに関する統一的な指針が未だ整備されていない状況であり、また、平成 23 年度に統合化推進プログラムで採択した研究開発課題の内、3 課題がヒト由来データを取り扱うデータベースであるため、ヒト由来データの取扱いは NBDC の問題として積極的に取り組んでいく必要がある。

そこで、NBDC 運営委員会に倫理分科会を組織し、ヒト由来データの公開・共有のあり方について我が国における統一的指針の先鞭をつけ、標準を作ることを目指して議論し、ヒト由来データに関するデータ保管・管理、データ共有・公開、ガイドラインの作成等に関する基本的な方向性について合意した。具体的内容については、データ共有分科会を設置し、引き続き審議していく予定である。

- b) オールジャパンでのデータベース統合への取り組み

我が国のライフサイエンス分野のデータベースの統合を進めるためには、データベースを所有している府省が連携することが重要である。平成 23 年 5 月には、総合科学技術会議 ライフサイエンス PT に設けられた統合データベース推進タスクフォース会合において、NBDC の事業の推進状況について報告すると共に、その進め方について専門家委員ならびに関係府省からの意見を聞いた。また、同年 12 月にはライフサイエンス分野のデータベースを有する 4 省（文部科学省、厚生労働省、農林水産省、経済産業省）のデータベース統合の取り組みについて、総合科学技術会議 政務三役と総合科学技術会議有識者議員との会合で発表し、有識者議員からも日本にとって非常に重要な取り組みであり、今後も着実に推進してもらいたいとの発言を得ている。

- 2) ポータルサイトの構築、運用
- a) NBDC ポータルサイトの構築・運用

各府省や研究機関、大学に散在しているデータベースについて、約 900 のデータベースの所在や概要を紹介するカタログ情報を提供し、約 300 のデータベースを対象として横断検索を可能とし、約 50 のデータベースについてデータがダウンロードできるように

NBDC のポータルサイトを構築し、さらに対象となるデータベースを拡充している。

今後は、文部科学省の各種プロジェクトや厚生労働科学研究費、JST の CREST、さきがけ等のファンディングプログラムで作り出されるデータの受け入れを進め、ゲノムコホート研究のデータの受け入れについても倫理面に配慮しながら可能な限り受け入れる方向で検討を進める。

- b) 4 省合同ポータルサイト (integbio.jp) の構築

これまではライフサイエンス分野のデータベースを有する 4 省（文部科学省、厚生労働省、農林水産省、経済産業省）によって個別にデータベース統合が進められてきたが、統合の次の段階として NBDC が中心となって 4 省合同ポータルサイトを構築し、昨年 12 月に公開した。今後も各省の協力を得てサイトのコンテンツの拡充を進める予定である。

- 3) データベース統合化基盤技術の研究開発
- 4) バイオ関連データベース統合化の推進

事業の柱の 3)、4) については、ファンディングプログラムとして公募を行い、3) については基盤技術開発プログラムで 1 課題、4) については統合化推進プログラムで 11 課題を採択している（平成 23 年度と平成 24 年度の合計）。研究開発課題の推進に際しては、研究総括が進捗報告会等で各課題の状況を把握・確認し、課題の推進について助言を行っており、各課題とも順調に進捗している。

今後は、分野別に統合されたデータベースをより上位で統合化することにより、一層有効にデータベースを活用することが可能になるように、分野を超えたデータベースの連携についても検討を進める。

- 5) その他

10 月 5 日を「トーゴーの日」としてシンポジウムを開催した。これは、統合データベースプロジェクトにおいて開始されたものであるが、平成 23 年度は 2 つの基調講演の他、農林水産省におけるデータベース統合の取り組みなど 9 件の講演、43 件のポスター発表が行われ、約 230 名の参加者があり活発な議論が行われた。

また、8 月 21 日～26 日に、先端技術を用いてシステムやプログラム開発を行なっている現場の研究開発者が参集し、合宿形式で分野横断的に問題解決にあたる、ユニークな形式の国際開発者会議 BioHackathon (バイオハッカソン) を開催した。平成 23 年度は、「LinkedData によるデータリソース公開とその周辺技術開発」をテーマとし、95 名（海外 27 名、国内 68 名）の参加者を得て、データベース統合に関する技術的な問題点などについて議論を重ねた。

上記シンポジウムや国際会議については今年度以降も引き続き実施する予定である。

#### 4. 第二段階のあるべき姿

##### (1) 第一段階開始以降の状況の変化について

現時点(平成24年7月)においては、第一段階を検討した時(平成22年度)とは状況が激変し、東北メディカル・メガバンク事業などのゲノムコホート研究において、新規に対応を必要とする諸事業が展開されることとなったため、現在の予算・体制では対応不可能な要請が生じている。第二段階を検討するに際しては、この状況は第一に考慮すべきと考えている。

ゲノムコホート研究に関しては、これまでに基礎研究の成果として蓄積されたゲノム、トランスクリプトーム、プロテオーム、メタボロームなどのデータベースと組み合わせることで、疾病の再分類や、再定義、さらには疾患のメカニズムを解明し有効な治療法・治療薬の発見につながることを期待されている。この種の研究ではデータの共有が非常に重要なため、NBDCからのデータ公開が求められている。NBDCを取り巻く状況は具体的に以下のような動きとなっている。

##### 1) 統合化推進プログラム

京都大学が実施している、ながはまコホート研究に関しては、平成23年度に統合化推進プログラムの1課題として採択され、データ共有に向けた取り組みが開始されている。また、統合化推進プログラムではヒト由来のデータとして、ヒト脳疾患画像データベースやヒトゲノム多様性関連データベースの統合化を進めている。

##### 2) 個別プロジェクトとの連携

東北メディカル・メガバンク事業や、科学技術システム改革事業において進められている国立がん研究センターのゲノムコホート研究と、データベース構築・公開での連携が予定されている。

##### 3) その他のゲノムコホート研究

それ以外のゲノムコホート研究についても、今後、日本国内のゲノムコホート研究で蓄積されるデータを研究者が統合的に利用できるような仕組みを作る必要がある。日本学術会議ゲノムコホート研究体制検討分科会においても、全国のデータをNBDCで受け入れてもらいたいとの意見が出ている。

研究環境に関する変化としては、次世代シーケンサによるデータ爆発の時代が到来し、研究者が容易にデータを産出することができるようになったことが挙げられる。夥しい勢いで産出されるシーケンサデータを蓄積し、必要なデータを選び出し、解析するためにハードウェア、ソフトウェア両面での最新技術を活用し、データベースを整備することが、研究の成果を左右する。また、今後大量に出てくるデータとして、イメージングデータがある。今後イメージデータを上手に取り扱うために、「イメージ・バイオインフォマティクス」の発展が期待される。さらに、ヒト由来データに関しては、データ共有のガイドライン、情報セキュリティのガイドラインの作成や、コントロールドアクセスデータの審査な

どの役割もNBDCに求められている。

これらの変化に対応し、日本のライフサイエンス分野の情報基盤整備を推進するために、第二段階では予算・人員・体制を拡充し、統合データベースセンターとしての充実化を進める必要がある。

##### (2) 担うべき機能について

第二段階のあるべき姿については、総合科学技術会議の場において、①引き続きJST内の組織としてさらに充実強化を図り組織を編成していくこと、②分野的に近い機関に移行すること、③新たな機関を設置すること、などの実現可能性について今後検討されることとなっている。いずれの形で運営されることになったとしても、前述の4つの柱、

- 1) 戦略の立案
- 2) ポータルサイトの構築・運用
- 3) データベース統合化基盤技術の研究開発
- 4) バイオ関連データベース統合化の推進

に基づいてデータベースの統合を引き続き推進し、(1)で述べた状況の変化に対応していくことは必須である。

オールジャパンでデータベースの統合化を推進するには、本来的には省庁から独立した新たな機関を設立することが望ましいものの、現在の国を取り巻く状況を踏まえると、第二段階が始まる平成26年度までに実現可能な方策はなく、また、NBDCがスタートしてから3年で分野的に近い別の機関に移行することも困難であることから、引き続きJSTが引き受け、中核機関として恒久化を図ることが、現時点では現実的な選択であると考えられる。

なお、①となった場合には、第二段階においては複数年に亘ってJSTとROISの下で一体的な形で運営されていくことが最も効率的であり、JSTとROISとの間で共同研究開発契約などを締結して進めることも検討されるべきである。

データベース統合を、これまでの成果を継続・発展させて今後も円滑に進めるためには、第一段階について、JSTの新たな組織として整備すべしという旨の提言を総合科学技術会議が出したのと同様に、第二段階はJSTとROISが一致協力して進めるべしという総合科学技術会議からの提言、あるいは関係省庁のコンセンサスに基づいて体制を整備することが望ましい。また、総合科学技術会議の指導の下、省庁連携をさらに推進するべきである。

##### (3) 予算、人員について

海外でNBDCと類似の役割を果たしている機関の状況は以下のとおりである。

##### a) NCBI (National Center for Biotechnology Information : アメリカ)

組織形態: NIH (National Institute of Health) 傘下のNLM (National Library of Medicine) の1部門

予算: 8,350万ドル=約64億円 (2009年) ※1ドル=77円

人員: 約600名、うち正規職員230名 (サービス: 研究: その他=55:30:15)

b) EBI (European Bioinformatics Institute : ヨーロッパ)

組織形態: EMBL (European Molecular Biology Laboratory) の傘下の非営利学術機関

予算: 4,000万ユーロ=約43億円 (2010年) (約半分はEMBLから) ※1ユーロ=108円

人員: 約500名

c) NBDCの現在の状況

組織形態: JSTの1部門

予算: 約16億円 (2012年度)

人員: 約20名

d) 第二段階におけるNBDCの予算、人員

第二段階におけるNBDCの体制については、JSTとROISによる一体的な運営を前提とし、また、NCBI、EBIの例も参考にして必要な予算と人員について試算を行った。なお、東北メディカル・メガバンク事業等の個別プロジェクトのデータ共有・公開のためのサポートについては、連携においてNBDCが担う具体的な役割や範囲は決定していない。従って、現時点では業務量を見積もることが困難なため今回の試算には含めておらず、別途検討が必要となる。NBDCはこのような個別プロジェクトのサポートに積極的に関与すべきであり、他のプロジェクトとの共同のための予算・運用について別途方策を作る必要がある。

予算: 約30億円

人員: 約100名 (JST+ROIS)

<予算の内訳> (単位: 億円)

	1) 戦略立案	2) ポータルサイトの構築	3) 基盤技術の研究開発	4) 統合化の推進	合計
予算	1	6	8	15	30
内容	<ul style="list-style-type: none"> <li>各種ガイドラインの策定</li> <li>国内外の動向調査</li> <li>省庁連携のための調整</li> <li>国際連携</li> <li>広報活動</li> </ul>	<ul style="list-style-type: none"> <li>ポータルサイトの構築・更新</li> <li>システム管理</li> <li>サービス用ハードウェア</li> </ul>	<ul style="list-style-type: none"> <li>ROISによる研究開発</li> <li>研究開発用ハードウェア</li> <li>研究開発成果の実装</li> </ul>	<ul style="list-style-type: none"> <li>ファンディングによる推進</li> <li>個別プロジェクトのサポート</li> <li>イメージデータの受け入れ</li> <li>分野を超えた統合</li> </ul>	

<人員の内訳> (単位: 人)

	研究・開発	アノテーション	企画・管理	合計
JST	25	10	15	50
ROIS	35	10	5	50
合計	60	20	20	100

5. バイオインフォマティクス分野における人材育成について

NBDC だけでは解決が困難であり我が国全体として取り組むべき問題として、バイオインフォマティクス分野における人材育成についてここで取り上げる。

(1) 人材の育成、キャリアパスの確立

次世代シーケンサの開発やバイオイメージング技術の発展などにより、研究の成果として得られるデータが大量になり、その種類も多様化しており、これからのライフサイエンス分野の発展には、上記のような、いわゆるビッグデータを活用することが欠かせない。そのためにはデータをデータベース化し、解析し、そこから新しい知識を見つけ出せるバイオインフォマティクス人材が多数必要となる。我が国においては、バイオインフォマティクス人材の育成は一部大学において実施されているが、今後のデータ量の増加に対応できる育成システムとはなっていない。

バイオインフォマティクス人材の不足は研究や臨床の現場だけではなく、ライフサイエンス関連の産業界における情報技術の応用・普及も欧米と比べて遅れており、ソフトウェア開発、情報技術開発の人材が不足している。

産業界における情報分野の人材育成への取り組みとしては、(社)日本経済団体連合会(経団連)高度情報通信人材育成部会における事例がある。平成17年に経団連として高度情報通信人材育成に関する提言を行い、その実現に向けて平成17年度に大学に対して高度ICT人材養成拠点の公募を行った。平成18年度には重点支援拠点を決定し、大学院修士課程の新コースの立ち上げを支援した。平成19年には産学官から構成されるナショナルセンターの設立に向けた提言を行い、さらには、平成21年に、それまでの経団連の取り組みを引き継ぎ、発展させるために、日本経団連の有志企業11社によって、特定非営利活動法人 高度情報通信人材育成支援センター (CeFIL) を立ち上げている。

大学、研究機関、企業、病院などで優秀な人材の確保が困難な状況がある一方で、パーマナントなポストが少ないことは、キャリアパスの確立という課題の解決を困難にしている。JSTにおいても任期制職員に対して他の研究機関との連携を進め、転入・転出が容易にできるようにしたり、大学や企業からの出向を受け入れたり、アカデミックキャリアの形成を支援したりするなどの取り組みを進めるべきであるが、バイオインフォマティクス分野の人材の育成、キャリアパスの確立については、研究活動が可能な機関との人材交流が重要であり、1つの機関のみで解決できる問題ではない。そこで、複数機関によるバイオインフォマティクス人材育成のためのネットワークを構築し、今後のライフサイエンス分野の研究開発の発展に必須であるバイオインフォマティクス人材の育成を安定的に推進する仕組み「バイオインフォマティクス育成推進ネットワーク」(図3)を立ち上げることを提案したい。

### バイオインフォマティクス育成推進ネットワーク

- ・バイオインフォマティクス人材の育成
- ・バイオインフォマティクス人材の循環の促進

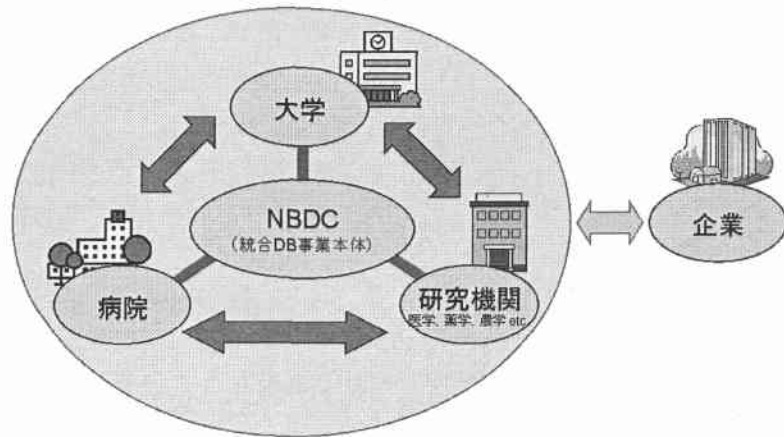


図3 バイオインフォマティクス育成推進ネットワーク

#### (2) バイオインフォマティクス育成推進ネットワークの狙い

ライフサイエンス分野のデータベースを効果的に活用するためには、データベースの作成、ならびに利用の場においてバイオインフォマティクスが不可欠である。バイオインフォマティクス育成推進ネットワークではNBDCがコアとなり、バイオインフォマティクス研究を行う大学や研究機関と共にバイオインフォマティクス育成推進ネットワークを形成する。このネットワークが「人材の育成」と「人材の循環の促進」の2つの機能を持つことにより、人材育成とキャリアパスの確立の2つを実現することを目指す。

人材の育成では、情報科学の知識を持った人材が、医学、薬学、農学などの様々な知識を身につける、あるいは、逆に医学等の知識を持った人材が情報科学の知識を獲得することによって「π型人才」となり、それぞれの分野で活躍できるような教育システムを構築する。

人材の循環の促進では、ネットワーク内で人材が一定年限毎に循環することにより、スキルを身につけ、経験を積み、大学、研究機関や企業などで管理職クラスのポジションにつけるような仕組みを立ち上げる。

NBDCは、ライフサイエンス分野のデータベースの省庁を超えた統合を推進しており、データの統合や解析のためのツール開発の実績がある。また、ライフサイエンス分野のデータベースやツールの使い方を動画で紹介するサイトを提供しており、さらにはバイオイ

ンフォマティクスの基礎的な知識やデータベースの使い方に関する講習会を開催した経験もある。これまでに蓄積したこれらのノウハウを活用してバイオインフォマティクス人材をOJT的に育成するシステムを構築することが可能であり、バイオインフォマティクス育成推進ネットワークの中核機関の役割を担うにふさわしいと考える。

バイオインフォマティクス育成推進ネットワークについても、JSTとROISの一体的な連携協力の下に運営されることが望ましく、関連する諸機関との連携を深めていくことが重要である。循環型の人材育成プログラム構築については経団連高度情報通信人材育成部会等との連携についても検討を進める必要がある。

#### (3) ネットワークの活動内容

- ① これからバイオインフォマティクスになることを目指す人で生物科学の素養を持つ人には、バイオインフォマティクスの教育カリキュラム(実習、講義)を提供し、また、Web上で学べる教材を配信する。また、情報科学の素養を持つ人には、生物科学の講義を提供し、また、実験研究者との出会いの場を提供する。
- ② 研究機関や民間企業への就職を目指す人には認定試験による資格認定を行う。また、人材マッチング情報を提供する。
- ③ 研究職としてのキャリアを目指す人には、データ解析のみならず、バイオインフォマティクスの基礎となる理論や方法論を学ぶ機会を提供し、ポジション獲得の支援をする。
- ④ 既に各機関でバイオインフォマティクスとして働いている人には、問題解決、スキルアップのためにバイオインフォマティクス同士が情報交換できる場を提供する。
- ⑤ アノテーターには、統合データベース作成の実務に参画する機会を与え、OJTを通じたスキルアップを支援する。

#### (4) ネットワーク参加機関に求められること(例)

- ① ネットワークに参加する大学または研究機関は、ネットワーク事業で育成される人材の5年任期のポジション(以下、「人材育成ポジション」と呼ぶ)を提供する。
- ② 年間のネットワーク事業費の一部を機関の経常費あるいは民間企業等からの基金、寄付金などで最低10年間程度拠出する。(マッチングファンド)
- ③ バイオインフォマティクス関連の教育ができる非常勤教員をNBDCに派遣する。
- ④ 人材育成ポジションの採用選考はNBDCと各ネットワーク機関に属する責任者が構成する委員会で行う。
- ⑤ 人材育成ポジションの給与体系はネットワーク機関間で原則的に均一になるような制度的配慮を行う。

#### (5) 人材育成ポジションを得た人のメリット

- ① 様々なネットワーク参加機関・大学からのスタッフによる教育を受けることができる。→情報科学、生物科学両方の知識を習得することができる。
- ② 期間中に一定の評価を受けると、ネットワークの委員会から修了証明書が授与される。

→新しいポジションの獲得にプラスとなる。

- ③ 10年間ネットワーク内で循環就労することができる（ただし、10年を待たずにネットワーク以外の職に付くことも可）。

以上は、バイオインフォマティクス育成推進ネットワークについて現時点での検討内容をまとめたものであるが、詳細については引き続き検討を行いたい。このような仕組みを構築することによって、バイオインフォマティクス育成推進ネットワークを構成する拠点が積極的に企業連携を実施し、人材の育成に資するネットワークを社会に広げていく必要がある。

なお、我が国の厳しい経済状況を考えた時に、「バイオインフォマティクス育成推進ネットワーク」の運営に際しては、人材を必要としている民間企業等からの資金提供や基金の導入などにより、産学官共同で人材を育成していく枠組みが必要であると考えられる。

## 6. おわりに

世界はビッグデータの時代を迎えており、ライフサイエンス分野においても研究の成果として得られるデータの洪水の中で、日本が活力のある国として存在し続けるためには、大量に産出されるデータを統合し、共有することが重要である。これにより医学、薬学、農学分野等での研究開発を促進し、国内産業を活性化することにより、海外の製品だけに頼らない国作りが可能となり、さらには、医療費の抑制にもつながることが期待される。

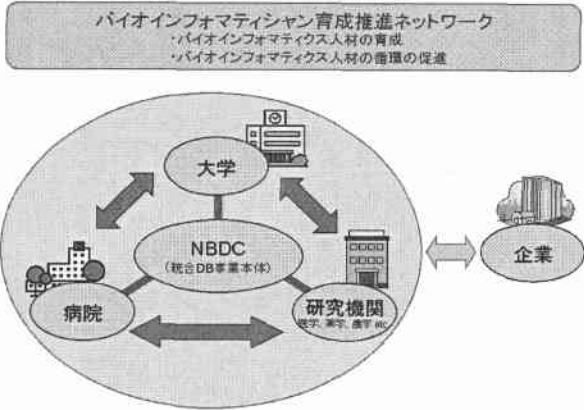
我が国におけるライフサイエンス分野のデータベース統合は、本来であれば省庁から独立した機関を設置して推進すべきところであるが、NBDC 運営委員会では、これまでのNBDC の運営状況を踏まえた上で、現実的な方策として引き続き JST 内の組織としてさらに充実強化を図り組織を編成していくことを仮定して議論を行った。本報告書は、その検討結果をまとめたものである。

統合データベースの第二段階のあり方については、文部科学省、更には、総合科学技術会議の場で検討がなされることと認識しており、省庁連携を含めたオールジャパンでの推進方策が早い時期に決定されることを望むものである。



ライフサイエンス分野の統合データベース整備の第二段階のあり方について(報告)(概要) その2

3. バイオインフォマティクス分野の人材育成=我が国全体で取り組むべき課題



産学官共同で人材を育成していく枠組み

- ・ ライフサイエンス分野のデータベースを活用するには、データベースの作成、ならびに利用の場においてバイオインフォマティクスが不可欠。
- ・ バイオインフォマティクス育成推進ネットワークではNBDCがコアとなり、バイオインフォマティクス研究を行う大学や研究機関と共にバイオインフォマティクス育成推進ネットワークを形成する。
- ・ 「人材の育成」と「人材の循環の促進」の2つの機能を持つことにより、人材育成とキャリアパスの確立の2つを実現することを目指す。
- ・ 人材の育成では、情報科学の知識を持った人材が、医学、薬学、農学などの様々な知識を身につける、あるいは、逆に医学等の知識を持った人材が情報科学の知識を獲得することによって「π型人材」となり、それぞれの分野で活躍できるような教育システムを構築する。
- ・ 人材の循環の促進では、ネットワーク内で人材が一定年限毎に循環することにより、スキルを身につけ、経験を積み、大学、研究機関や企業などで管理職クラスのポジションにつけるような仕組みを立ち上げる。

ライフサイエンス分野の統合データベース整備の第二段階のあり方について(報告)(概要) その1

1. 第一段階(平成23年度~25年度)  
 ライフサイエンス分野のデータベース統合を進めることによってデータの価値を最大化することを目指し、平成23年4月にJSTにバイオサイエンスデータベースセンター(NBDC)が設置された。

2. 第二段階(平成26年度以降)  
 ①引き続きJST内の組織として更に充実強化を図り組織を編成していくこと  
 ②分野的に近い機関に移行すること  
 ③新たな機関を設置すること  
 などの実現可能性について、総合科学技術会議の場において今後検討される予定  
 しかし、現在の国を取り巻く状況を踏まえると、  
 ③については第二段階が始まる平成26年度までに実現可能な方策はなく、  
 ②については、NBDCがスタートしてから3年で分野的に近い別の機関に移行することも困難である  
 ①の形で引き続きJSTが引き受け、恒久化を図ることが、現時点では現実的な選択であると考えられる。

第一段階の体制  
 予算 約16億円、人数 約20名

総合科学技術会議の助言の下、オールジャパンでのデータベース統合を進める

- 1) 戦略の立案
- 2) ポータルサイトの構築・運用
- 3) データベース統合化基盤技術の研究開発
- 4) バイオ関連データベース統合化の推進

第二段階の体制  
 予算 約30億円 人数 約100名  
 JSTとROISによる一体的運営

- 【新たな要因】
- ・ゲノムコホート研究への対応
  - ・ビッグデータへの対応 (次世代シーケンサ、イメージデータ...)

NBDC 運営委員会 委員名簿

委員長

末松 誠 慶應義塾大学 医学部長

委員(五十音順)

春日 雅人 独立行政法人 国立国際医療研究センター 総長

金岡 昌治 大日本住友製薬株式会社  
常務執行役員 研究本部長 兼 知的財産・IT企画  
推進担当

久原 哲 国立大学法人 九州大学大学院  
農学研究院 生命機能科学部門 教授

小原 雄治 大学共同利用機関法人 情報・システム研究機構 理事  
国立遺伝学研究所 所長

篠崎 和子 国立大学法人 東京大学大学院  
農学生命科学研究科 応用生命化学専攻 教授

城石 俊彦 大学共同利用機関法人 情報・システム研究機構  
国立遺伝学研究所 系統生物研究センター センター長  
哺乳動物遺伝研究室 教授

大力 修 新日鉄ソリューションズ株式会社  
フェロー (常務取締役待遇)

武田 英明 大学共同利用機関法人 情報・システム研究機構  
国立情報学研究所 情報学プリンシプル研究系 教授

松田 秀雄 国立大学法人 大阪大学大学院  
情報科学研究科 バイオ情報工学専攻 教授

NBDC 運営委員会 開催実績

- 第1回 平成23年10月11日
- 第2回 平成23年12月5日
- 第3回 平成24年2月29日
- 第4回 平成24年5月11日
- 第5回 平成24年6月5日

NBDC における平成23年度の主な取り組み

平成	5月23日	4省関連機関による連携のための打合せ
23年	5月25日	内閣府統合データベース推進タスクフォース
	6月10日	「基盤技術開発プログラム」「統合化推進プログラム」合同キックオフミーティング
	6月～	CREST、さきがけとの連携 ・平成23年度発足領域の募集要項中にNBDCへのデータ提供の協力に関する文言を記載 ・研究総括との打合せを実施
	7月15日	統合データベース講習会
	8月21～26日	バイオハッカソン
	10月5日	トーゴの日のシンポジウム
	10月11日	第1回運営委員会
	11月～	厚生労働科学研究費補助金との連携 ・平成24年度公募要項中にNBDCへのデータ提供の協力に関する文言が記載された
	12月5日	第2回運営委員会
	12月12日	4省合同ポータルサイト「integbio.jp」を開設
	12月13日～16日	分子生物学会年会への出展
平成 24年	1月11日	平成24年度統合化推進プログラム公募開始
	1月16日	第1回倫理分科会
	2月10日	統合データベース講習会
	2月22日	第2回倫理分科会
	2月28日	内閣府統合データベース意見交換会
	2月29日	第3回運営委員会
	3月2日	統合データベース講習会
	3月26日	内閣府統合データベース意見交換会

提言

ヒト生命情報統合研究の拠点構築  
—国民の健康の礎となる大規模コホート研究—



平成24年（2012年）8月8日

日本学術会議

第二部

ゲノムコホート研究体制検討分科会

この提言は、日本学術会議第二部ゲノムコホート研究体制検討分科会の審議結果を取りまとめ公表するものである。

### 日本学術会議第二部ゲノムコホート研究体制検討分科会

委員長	浅島 誠	(連携会員)	日本学術振興会理事
副委員長	本庶 佑	(連携会員)	京都大学大学院医学研究科特任教授 静岡県公立大学法人理事長
幹事	小原 雄治	(第二部会員)	情報・システム研究機構理事 国立遺伝学研究所所長
幹事	菅野 純夫	(連携会員)	東京大学大学院新領域創成科学研究科教授
	斎藤 成也	(第二部会員)	情報・システム研究機構 国立遺伝学研究所 集団遺伝研究部門教授
	山本 雅之	(第二部会員)	東北大学大学院医学系研究科教授 東北メディカル・メガバンク機構長
	山本 正幸	(第二部会員)	公益財団法人かずさDNA研究所長
	巖佐 庸	(連携会員)	九州大学大学院理学研究院教授
	春日 雅人	(連携会員)	独立行政法人国立国際医療研究センター総 長
	高木 利久	(連携会員)	東京大学大学院新領域創成科学研究科教授
	辻 省次	(連携会員)	東京大学大学院医学系研究科脳神経医学専 攻長, 神経内科学教授, 医学部附属病院ゲノ ム医学センター長
	大江 和彦	(特任連携会員)	東京大学大学院医学研究科教授

報告書の作成にあたり、以下の方々にご協力いただきました。

清原 裕	九州大学大学院医学研究院基礎医学部門社会環境医学講座環境医学 分野教授
津金 昌一郎	国立がん研究センター予防研究部部长
深尾 彰	山形大学理事・医学部医学科社会環境予防医学部門公衆衛生講座教授
松田 文彦	京都大学大学院医学研究科附属ゲノム医学センター教授
森 英郎	独立行政法人科学技術振興機構研究開発戦略センターフェロー

本件の作成に当たっては、以下の職員が事務を担当した。

事務局 相原 佳奈 参事官(審議第一担当) 審議専門職付

## 要 旨

### 1 作成の背景

先進諸国においては、疫学研究の重要性を認識し、大規模コホート研究が実施されている国が多く、自国民の健康・医療状況を判断するための情報蓄積とその活用が進められている。このような健康に関するデータは、生活習慣や集団による差異が存在するものであり、我が国の研究として、自国民のデータを取得する必要がある。それゆえ、我が国でも、数多くのコホート研究が実施されてきたが、その規模、内容において、時代の要請に応じたものであるかと言うと、不十分であると言わざるを得ない。

2010年6月に閣議決定された新成長戦略で示された「ライフ・イノベーションによる健康大国の実現」のため、総合科学技術会議では「平成23年度科学・技術重要施策アクション・プラン」のための施策として「ゲノムコホート研究と医療情報の統合による予防法の開発」を推進することとされ、近年著しい進歩を見せるゲノム解析を組み込んだ新しいコホート研究の推進が提言された。大規模ゲノムコホート研究は、将来的な健康と高度医療を確保するために、欠くべからざる基礎データを提供するものである。本分科会では、我が国の国民の健康・医療に大きな恩恵をもたらす、あるべき大規模ゲノムコホート研究の姿について具体的に議論した。

### 2 現状及び問題点

2011年9月には科学技術戦略推進費「ゲノム情報と電子化医療情報等の統合によるゲノムコホート研究の推進」によるパイロットプロジェクト「大規模分子疫学コホート研究の推進と統合」が、国立がん研究センターにより開始された。また、「東日本大震災からの復興の基本方針」（2011年7月29日東日本大震災復興対策本部）ならびに「日本再生の基本戦略」（2011年12月24日閣議決定）に基づき、大規模ゲノムコホート研究を含む「東北メディカル・メガバンク計画」[1]が立案された。これらを中心に、我が国の大規模ゲノムコホート研究が進められることになるが、どのように全国的な規模まで拡大してゆくか、ゲノムなどの最新情報をどのように研究に組み込むかについての具体的な方策は未だ明確になっていない。

大規模ゲノムコホート研究において、血液や尿を分析・解析して得られるゲノム、代謝物などのオミックス情報と、環境・生活習慣情報等を統合解析することは、とりもなおさず新しいヒトを対象とした総合的な研究を目指すことになる。また、近年の新しい物理化学的計測技術の進歩は、侵襲なくヒトを直接解析する様々な手法を提供している。このような新しい基盤技術から得られる解析データを、ゲノムコホートに統合することにより、ヒトの健康や医学を総

合的に研究する「ヒト生命情報統合研究」が可能になろうとしている。しかしながら、このような研究を実施するためには医学、生物学の研究者のみならず、理学、薬学、工学、情報科学の研究者を動員し、新しい計測機器による微量分析の技術開発、得られた膨大な情報を効率的に扱う情報科学の構築、多様な情報や異種のデータを統合して疾患との関連を統計学的に解析する新たな生命情報解析理論の構築など、克服すべき課題が多く、我が国の科学技術全般に亘る大きな挑戦的課題である。また、個人情報保護への配慮、倫理対応、組織構築、制度設計および解析技術開発など事業に関わる課題を明らかにした上で、国のとるべき姿勢を明確にし、事業を適切に実施する必要がある。

### 3 提言等の内容

本分科会における議論の結果、ゲノムコホート研究内容をさらに先進のものとした新しい研究内容とその体制、すなわち統一基準による全国体制の「ヒト生命情報統合研究」の実施が必要と結論し、その実施に向けた提言をまとめた。

#### (1) ヒト生命情報統合研究の創出

人間集団を対象とするコホート研究では、単に病気だけでなく、人間の持つ健全形質の多様性（表現型情報）にも目をむけるべきである。そのような大規模コホート研究で収集すべき表現型情報としては、新しい物理化学的計測技術による解析データ、すなわち血液のメタボロームといった生化学的データ、細胞生物学的データ、生理学的データ、MRIやX線CTなどによる画像診断データなどを含み、さらには心理学や行動学的データもあげられる。これら膨大なデータを個々人から長期間にわたって収集する長期観察と、最新の情報科学を用いた膨大なデータの解析研究から得られる結果は、健康・医学の進展に大きく寄与するだろう。また、生物が共有する基本的生命現象を発見する大きな可能性も秘めている。本分科会では、大規模ゲノムコホート研究をさらに先進的なものとするため、最新の物理化学機器から得られる多様な表現型情報を融合し、得られる巨大な情報を統合的に解析するまでを一貫して実施する新しい研究分野すなわち、「ヒト生命情報統合研究」を創出すべきであると提言する。

#### (2) 医療情報基盤のさらなる整備

大規模コホート研究あるいは、「ヒト生命情報統合研究」への研究参加者<sup>1</sup>が、期間中の任意の時期に自分の意思で受診する医療機関や健診機関での健康医療情報を電子的にかつ標準化された形式で網羅的に収集でき、同一個人ごとに統合できる情報基盤の整備を行う必要がある。そのためには国は次の3点を目指すことが必須である。

<sup>1</sup> 本提言においては、「研究の趣旨を理解し、研究対象者として参加する人」という意味で使用

- ① 標準化対応の医療情報システムの開発と導入
- ② 新たな「国民保健番号（仮称）」の制度と法令の整備
- ③ 同意研究参加者の医療情報追跡基盤の構築

### (3) ヒト生命情報統合研究の拠点整備とその機能

100万人規模のバイオバンクの構築と生体試料、臨床情報、疾患罹患情報を統合した疾患解析による多因子疾患の原因解明と予防・治療法の開発を目標に、中核拠点を軸とした組織構築、制度設計を行い、既存のコホートをを用いたパイロット研究を経て、コホート事業の開始に向けた解析基盤、情報基盤の整備を行う。そして、事業の三年目をめどに、全国で数カ所の実施拠点を構築し、統一基準による全国体制下の事業組織によるヒト生命情報統合解析の実施基盤を構築する。その基盤は、①中核拠点、②地域研究拠点、③データ解析センター、④生体試料バンク、からなる。また、事業の成果を産業へ結びつけるための仕組みとして、⑤産学連携コンソーシアムを形成する必要がある。

本研究の発展には膨大な臨床情報とゲノム情報や中間形質情報の統合解析技術を駆使できる人材育成が急務となっている。この人材育成は、実データ解析を通して疾患研究者と日々議論できる研究フィールドにおいてなされる必要がある。さらに人材育成はキャリアパスを形成するために恒久的な組織で行われる必要もある。そこで、臨床現場に近い医育機関に臨床ゲノム情報技術人材の育成に専念する専門教育組織を全国数カ所に恒久的に整備すべきである。

### (4) 包括同意にもとづく研究を可能とすることに対する国民の理解醸成と研究倫理指針の改訂

ヒト生命情報統合研究は国民全体の健康増進といった、より大きな成果を求め、大規模化、長期間にわたる前向き研究を志向している。このような研究では、網羅的な分析と生活情報・臨床情報の統合が成果を上げる上で重要である。特に、分析技術は日進月歩で進歩しており、新しい解析方法を積極的に取り込んでいく必要がある。このような研究においては、研究参加者の包括同意を得ることが必須である。ヒト生命情報統合研究から得られる成果とその公共性を鑑みて、国は、一定条件を満たす研究については包括同意を可能とすることに対する国民の理解醸成を積極的に行い、関連する研究倫理指針を改訂して包括同意にもとづく研究遂行を可能にするとともに、様々な制度面の整備を行う必要がある。

### (5) 研究拠点整備と提言実行のための推進協議会（仮称）の設置

ここまで述べた研究拠点の整備等の実現を図るため、専門家による推進協議会を設置することを提言する。

## 目 次

1	はじめに .....	1
2	海外におけるゲノムコホート研究 .....	4
3	我が国におけるコホート研究の歴史 .....	6
4	ゲノムコホート研究からヒト生命情報統合研究へ .....	9
5	提言 .....	12
(1)	ヒト生命情報統合研究の創出 .....	12
(2)	医療情報基盤のさらなる整備 .....	12
(3)	ヒト生命情報統合研究の拠点整備とその機能 .....	14
(4)	包括同意にもとづく研究を可能とすることに対する国民の理解醸成と研究倫理指針の改訂 .....	19
(5)	研究拠点整備と提言実行のための推進協議会（仮称）の設置 .....	20
	<参考文献> .....	21
	<参考資料 1> 対象集団を 100 万人規模とする根拠について .....	23
	<参考資料 2> ゲノムコホート研究体制検討分科会審議経過 .....	25



## 1 はじめに

我が国ではこれまでも、集団を対象として対象者の生活習慣、生活環境等に関する追跡調査を行う数多くのコホート研究が実施されており、その中には、発症者を対象とし、疾患とゲノム多型との関連解析によって疾患感受性遺伝子を同定しようとするゲノム研究を含むものも存在する。しかしながら、糖尿病、高血圧などの生活習慣病や認知症といった頻度が高く、国民生活とその医療にとって病因の解明と予防・治療法の確立が極めて重要な多因子疾患は、環境・生活習慣と複数の遺伝的変異の影響を受ける複雑なもので、従来の規模と仕組みによるコホート研究では疾患関連因子の同定が困難である。複雑な環境要因と遺伝子要因と疾患の関係を解き明かすためには、大規模な健常者集団の精緻な長期観察データを基盤とする前向きの大規模ゲノムコホートの構築が必須である。

ゲノム情報が比較的容易に得られるようになった今日、ゲノム情報を軸として、バイオマーカー情報、環境・生活習慣情報や疾患に関する高精度の臨床情報を重層するゲノムコホート研究は、発症前介入による疾患の予防や発症の遅延を目標とした最も効果的な戦略であると考えられるようになった。世界的に見ると大規模ゲノムコホート研究はすでに開始されており、例えば英国の UK バイオバンクではすでに 50 万人の登録が完了し、また UK バイオバンクと同じプロトコールで開始した中国の China Kadoorie Biobank は、UK バイオバンクに先んじて 2008 年に 50 万人の目標を達成した。

我が国でも、以下に述べる社会的要請、学術的要請から、健常人を対象とした大規模なゲノムコホート研究を、国策として推進する必要性がこれまでになく高まっている。2010 年 6 月に閣議決定された新成長戦略で示された「ライフ・イノベーションによる健康大国の実現」のため、総合科学技術会議では「平成 23 年度科学・技術重要施策アクション・プラン」を策定した。同プランでは、予防医学の推進による罹患率の低下を重点課題とし、そのための施策として「ゲノムコホート研究と医療情報の統合による予防法の開発」を推進することとした。そこでは、ゲノム情報に基づくコホート研究から得られる詳細な疫学情報とデータベース化された医療情報との統合により、様々な疾患に関する原因因子の解明を行い、予防医学のさらなる発展を目指すことが謳われた。より具体的には、健常人よりなる 10 万人規模のコホート研究においてゲノム解析・バイオマーカー測定を組み込んだ研究を実施するとともに、データベース化された研究参加者の医療情報との統合的情報解析を行い、疾患罹患リスクにおけるゲノムと環境因子との関連や疾患メカニズムの解明、薬剤の効果や副作用に関わる因子の探索や評価が期待されている。

このような背景のもとに、2011 年 9 月には科学技術戦略推進費「ゲノム情報と電子化医療情報等の統合によるゲノムコホート研究の推進」によるパイロツ

トプロジェクト「大規模分子疫学コホート研究の推進と統合」が、国立がん研究センターにより開始された。また、東日本大震災に伴う被災地域住民の健康を守るという立場からも、「東日本大震災からの復興の基本方針」（2011年7月29日東日本大震災復興対策本部）ならびに「日本再生の基本戦略」（2011年12月24日閣議決定）に基づき、大規模ゲノムコホート研究を含む「東北メディカル・メガバンク計画」が立案された。これは、東日本大震災被災地の住民の健康診断、医療機関の受診、ゲノムの塩基配列等の情報と生体試料を統合したバイオバンクを構築し、長期の健康モニタリングから得られる研究成果を被災地の住民の健康増進に還元することを目指す事業である。本事業では、東北大学、岩手医科大学が実施する被災地域を対象とした15万人規模の住民ゲノムコホート、ゲノム情報等の解析が計画されている。2012年6月に文部科学省「東北メディカル・メガバンク計画検討会」より公開された「東北メディカル・メガバンク計画検討会提言」には、我が国の叡智を結集して実施計画を具体化するべきとしており、実際に関係者により具体化に向けた方策が進められつつある。

ゲノムコホートでは、大規模な健常人集団を対象に、個々人の血液や尿を分析・解析して得られるゲノム、代謝物などのオミックス情報、環境・生活習慣情報等を統合解析するものであり、とりもなおさず新しいヒトを対象とした総合的な研究を目指すことになる。従来の疾患研究の多くが、細胞や動物モデルを用いた解析によって行われてきた。その主たる理由はヒトへの侵襲を最小限にとどめ、その情報を分析することの技術的な限界があったためである。ヒトから採取できる試料は、ほとんどが血液や皮膚などの一部の組織に限られ、そこから得られる情報は限定的であった。ところが、近年の新しい物理化学的計測技術の進歩、例えば侵襲のほとんどない画像解析、血液成分からの極めて微量な物質の定量的な測定、「1000ドルゲノムの時代」と称されるようにコストが著しく低下した全ゲノム塩基配列の決定方法などにより、この状況は一変しつつある。このような解析基盤の登場で、技術的にはすでにヒトの疾患を細胞やモデル生物でなく、ヒトの体内に存在する多様な分子の網羅的解析を通してみる「ヒト生命情報総合研究」が可能になろうとしている。しかしながら、このような研究を実施するためには医学、生物学の研究者のみならず、理学、薬学、工学、情報科学の研究者を動員し、新しい計測機器による微量分析の技術開発、得られた膨大な情報を効率的に扱う情報科学の構築、多様な情報や異種のデータを統合して疾患との関連を統計学的に解析する新たな生命情報解析理論の構築など、克服すべき課題が多く、我が国の科学技術全般に亘る大きな挑戦的課題である。本分科会では、大規模ゲノムコホート研究をさらに先進的なものとするため、新しい物理化学的計測技術がもたらす膨大なヒトの解析データを積極的に取り込み、得られる巨大な情報群を統合的に解析する新しい研究分野すなわち、「ヒト生命情報統合研究」を創出すべきであると結論した。ま

た、この「ヒト生命情報統合研究」では、研究参加者が自分の意思で任意の医療機関や健診機関を受診した際に得られる検査結果やそれにもとづく健康診断情報などを、受診機関を横断的にかつ長期間にわたり必要に応じてゲノム情報と個人ごとに統合できる必要がある。このためには、任意の医療機関や健診機関での医療情報が電子的に記録され、統合可能な標準的な形式で収集できる医療情報基盤の整備とそれを活用した医療システムが不可欠である。特に、任意の複数医療機関からの医療情報を個人ごとに統合できるようにするためには、同一個人を確実に同一であると同定できる個人識別番号が不可欠であり、これが医療システムに導入される社会基盤整備も必須である。さらに重要なことは、ゲノム情報や、疾患罹患情報などは個人のプライバシーに関わる情報である。そのような情報を網羅的に扱う本研究においては、いかに研究参加者の人権を守りつつ包括的な研究を推進するのかという、倫理的に極めて重要な問題に対する最善の解決法を見出す必要がある。すなわち、研究参加者が将来かかる病気は予見できないため、研究参加者の個人情報の厳格な保護に加え、研究活動の透明性、説明責任等の倫理的妥当性を担保した「個人情報を最大限に保護しつつ幅広く医学的研究に用いることを認める」という包括同意が必須である。比較的頻度の低い疾患まで含めて解析を実施するためには、100万人規模の集団を設定し、100万人の一人ひとりから包括同意を得る必要があり、この事業の最も困難な課題の一つと考えられる。

前述の科学技術戦略推進費によるパイロットプロジェクト、東北メディカル・メガバンクにおける住民ゲノムコホートでの事業成果や、既存のゲノムコホートの経験を最大限に活用することで、統一基準による全国体制の「ヒト生命情報統合研究」の効率的な実施が期待されている。しかしながら、上に述べたように、100万人の研究参加者の長期にわたる追跡と生命情報の網羅的分析が必須の「ヒト生命情報統合研究」の推進における課題は、十分に整理されているとはいいがたい。このような我が国の現状を鑑みるに、「ヒト生命情報統合研究」の実施において国のとるべき姿勢を明確にし、組織構築、制度設計および解析技術開発など事業に関わる課題を明らかにした上で、事業を適切に実施可能とする提言を政府、社会に行うことは極めて重要なことである。そこで、日本学術会議において、「ヒト生命情報統合研究」を進めるための課題について集中的に検討を行い、その結果を提言書として、ここにまとめた。

## 2 海外におけるゲノムコホート研究

ヒトゲノム計画の終了前後からヒトゲノムの配列情報、特に一塩基多型を利用して、疾患の遺伝的要因を網羅的に解析しようという機運が高まった。その中心は症例対照研究を手段とする Genome-Wide Association Study (GWAS) であった。GWAS では、多くの成果が上がっているものの、症例対照研究としての固有の限界を有しており、特に、疾患の発症に寄与する環境因子の同定や時間的経過を明らかにして疾患の予防につなげる点で弱点を持っている。これに対し、コホート研究（前向きコホート研究）は補完的な役割を果たし、それ以外の方法では得にくい予防医学上の貴重な知見を得られる方法であり、ゲノム解析と組み合わせることでゲノムに基づく個別予防医学といった新しい分野を切り開くものとして、今、世界で注目され、それへの取り組みが始まっている。

コホート研究（前向きコホート研究）において、環境要因に加え、疾患の発症等に関わる遺伝学的要因を網羅的に解析しようという研究計画（すなわち、ゲノムコホート研究）が最初に具体化されたのは英国であり、プランニングはヒトゲノム計画の終了前の 1999 年にさかのぼる。UK Biobank[2]とよばれるこの研究は 2001 年に立ち上げ宣言が行われ、2003 年には計画の概要が定まり、2006 年まで 4000 人規模のパイロット研究が行われた。2007 年に研究参加呼びかけと生体試料の収集が本格化し、2011 年中に予定の 50 万人の研究参加者からの生体試料の収集を終了した。3 年という短期間に 50 万人という規模の研究参加者を得られたのはコホート研究の滑り出しとしては大成功と考えられている。研究参加者は 40 歳から 65 歳までの成人であり、観察期間は 20~30 年を予定している。血液と尿のサンプルを収集しているが、ゲノム解析は直ちに行わず、いつどの規模で解析を行うかは研究の進展にしたがって判断していく予定となっている。UK Biobank を手本として、類似のゲノムコホート研究がいくつかの国で立ち上がっている。スウェーデンでは LifeGene[3]プロジェクトが、カナダでは Partnership for Tomorrow[4]プロジェクトが 50 万人規模を目標に研究参加者を募集している最中である。中国では、UK Biobank に関係した研究者が中国 CDC と共同して China Kadoorie Biobank[5]を 50 万人規模で立ち上げ、本家よりも早く 2008 年に研究参加者が 50 万人に達した。

また、既存のコホート研究にゲノム解析要素を組み入れる動きも多い。EU では、50 万人規模の EPIC (European Investigation into Cancer and Nutrition[6]プロジェクトが、ゲノム解析をあとから組み入れたものでは規模の大きいものである。米国も、現状では既存のコホート研究を利用してゲノムコホート研究を進める方向である。そもそも米国は、NIH だけで年間 4 億ドルを超える予算で各種コホート研究をおこなっており、NIH-AARP Diet and Health Study[7]や National Children's Study[8]、American Cancer Society Cancer Prevention Study[9]といった大規模のコホート研究のみならず、Veterans Affairs Genomic

Medicine Program[10]やVanderbilt BioVU[11]のように大量の生体試料を保存する事業を含むプログラムが実施されており、ゲノムコホート研究に組み替えやすい。ただ、同意書や生体試料収集の関係で、既存のコホートをゲノム解析に使用するには制約が多いのも事実である。2010年に開かれた”New Models For Large Scale Prospective Studies”というシンポジウム[12]で、NIH 所長のFrancis Collins は、既存のコホート研究の拡張だけでなく、新規の大規模ゲノムコホート研究が必要であると発言している。

このような、欧米の動きに対し、アジアでもゲノムコホート研究への動きが見え始めた。コホート研究は症例対照研究に比べても安定な医療体制が必要であり、我が国を除くアジア諸国ではコホート研究が少ない。またあっても、欧米の研究の一部として進められる場合が多かった。現在、ゲノムコホート研究を実施する基盤が整っているのは、我が国以外では、まずシンガポールと韓国、台湾があげられ、次に中国とインドが続く。韓国では25万人規模を目指すKorea Genome and Epidemiology Study (KoGES) [13]が、2004年にスタートし、12の地域コホートと12の医療機関コホートが参加し、2007年12月時点で13万人の研究参加者をとっている。またイギリスとの共同研究として実施されている中国のChina Kadoorie Biobank[5]では、ゲノムコホート研究を目指して約50万人が参加しているが、ゲノム解析の詳細については未だアナウンスがない。これらのゲノムコホート研究が、今後安定的に継続できるか未知の部分が多い。

### 3 我が国におけるコホート研究の歴史

コホート研究とは、固定した集団を一定期間追跡し、要因と疾病発生の関連などを調べる観察的研究である。具体的には、ある要因への暴露を起点として、前向きに追跡している集団の中の個人の計測パラメーターにどのような変化が起こるかを見ている。一般的なコホートの研究参加者は健常人である。それに対して、疾患の罹患者、あるいは患者予備群と推定されるヒトを集めてその人たちの追跡調査を行い、その研究データを集める前向きコホートを疾患コホートと称する。2000年以降は、コホート研究にゲノム解析を取り入れるものも生まれ、それらはゲノムコホートと称されている。コホート研究に対し、症例対照研究は、後ろ向きである。疾病に罹患した集団を対象に、過去にさかのぼりどのような危険因子にさらされたことがあるのかを調べる。

我が国で行われた初めての本格的コホート研究としては、放射線影響研究所による広島・長崎の被爆者を対象とした生涯にわたる健康影響調査研究[14]があげられる。前身は1947年に米国資金で設立された原爆傷害調査委員会(ABCC)であるが、その研究計画は1955年に大幅に見直され、生涯調査の基礎が築かれた。約9万4千人の被爆者と、約2万7千人の非被爆者から成る約12万人を、明確に定義された固定集団とし、主に原爆放射線が死因やがん発生に与える長期的影響の調査を行っている。最近では、被爆者のがん病理組織の分子生物学的解析も実施している。

1961年には、福岡県久山町の地域住民を対象とした脳卒中の疫学実態調査研究[15]が始まった。当時の我が国の死因の第1位は脳卒中であったが、その内訳が欧米と大きく異なっていたため、死因調査の正確性を期すための死亡者全例剖検が行われ、現在に至っている。1961年からの第1集団(1,618名、剖検率80%)のデータでは、脳出血による死亡率は脳梗塞のわずか1.1倍であり、コホート研究開始以前の脳出血が脳梗塞の12.4倍という死亡診断書群は、病型診断を誤ったものを多く含んでいたであろうことが示された。久山町コホート研究では5年ごとに40歳以上の住民を対象にした新しい集団を設定して、生活習慣の移り変わりの影響も追跡している。2002年の集団(3,772名)から、遺伝子解析(SNPs)も加えて、ゲノムコホートを我が国で初めて開始した。遺伝子検査結果を将来的研究に使用することに関する包括同意を96%の対象者から得て進めており、ゲノムコホート研究としての成果も出ている。

2000年以前に開始された我が国の大型コホート研究としては、大気汚染と肺がんとの関連を検討のためにスタートしている三府県コホート研究(宮城県・東北大学衛生学・公衆衛生学教室、愛知県・愛知県がんセンター、大阪府・大阪府立成人病センターが実施)[16]、日本人の生活習慣ががんとどのように関連しているかを明らかにすることを目的としている文部科学省科学研究費がん特定領域大規模コホート研究[17]、日本人における循環器疾患の危険因子・発

症・死亡の関係を明らかにすることを目的とした Japan Arteriosclerosis Longitudinal Study (公益信託・日本動脈硬化予防研究基金の助成で、国内各地で行われている循環器コホート研究の個票データを統合し、メタアナリシスを実施) [18]などがあるが、DNA の解析は行われていない。

国立がんセンターにより 1965 年から 17 年間実施された「計画調査」[19]は、約 26 万人を対象としたがんに関する大規模コホート研究である。緑黄色野菜のがん予防効果、間接喫煙によるがんリスクの上昇を明らかにした。さらに、国立がんセンターでは 1990 年から、多目的コホートに基づくがん予防など健康の維持・増進に役立つエビデンスの構築に関する研究 (Japan Public Health Center-based prospective, JPHC Study) [20]を全国 11 ヶ所の保健所とともに、約 14 万人を対象に進めている。どのような生活習慣をもつ人が、がん・脳卒中・心筋梗塞・糖尿病などになりやすいのか、あるいはなりにくいのかを明らかにすることを目的としている。生活習慣や健康に関するアンケート調査情報、健康診断データを集めるとともに、血液を採取し、10 年以上にわたる長期追跡を行っている。ゲノム情報収集に関しては、付加可能となっているが、現時点では生活習慣病に関連した疾患に限定した研究が一部で行われているのみである。同様に生活習慣病に限定して DNA 情報利用を行っているコホート研究として山形大学を中心とした地域連携コホート[21]がある。また我が国のコホート研究の特徴として、小規模な疾患コホート研究が数多く実施されてきている。高血圧治療中の患者を対象とした高齢者高血圧コホート研究 (J-CHEARS Study) [22]が例として挙げられる。これらの研究は、大なり小なり特定の疾患に注目したコホート研究に DNA 解析を加えたものであり、疾患ゲノムコホート研究と分類される。

疾患の遺伝的要因をゲノム情報により解析する研究の歴史という視点では、これまでは GWAS に代表される症例対照研究が主流であり、我が国は GWAS 研究に巨額の研究資金を投入して来た。しかしながら、症例対照研究から得られる情報では、疾患の発生率、5 年生存率などを予測することができない。また、限定された対象者 (症例と対照群) の比較に依存するので、結果にバイアスの影響が入り込みやすいという欠点もある。また、影響を受ける遺伝因子が多い疾患については明確な統計的有意な差を捉えることは困難である。症例を調整選抜することが非常に難しいことから症例対照研究には限界があると考えられる。

2000 年以降では、コホート研究にゲノム情報を本格的に取り入れる試みが始まっているが、ゲノム情報から個人を特定できるという問題があり、個人情報保護という倫理面からの議論が続いている。「個人情報保護法」を順守しつつ、「ヒトゲノム・遺伝子解析研究に関する倫理指針 (ゲノム倫理指針、文部科学省・厚生労働省・経済産業省の三省合同指針)」に基づき、データを公的利益に

つながる形で公開・活用するという対応が必要とされる。ゲノム倫理指針は、現在改定作業が進められており、これまでの個人情報保護偏重の姿勢から、個人情報保護とデータの公共性とのバランスを取る形が模索されている。現在の改定前ゲノム倫理指針では、データの公開などに関するインフォームド・コンセントを取るための説明を1対1で行うことになっており、10万人を超える大規模ゲノムコホートでの実施は、現実的には困難である。誰が何のためにデータを利用し、どのような公的メリットが導き出されるかを明確にしたうえで、個人情報を保護しつつデータの活用を考える現実的な制度が必要である。

このような現状の中で、ゲノムデータを統合的に活用する有用性を認め、コホート研究参加者との包括同意のもとに研究を進めている我が国のゲノムコホートの代表例は、前述の久山町コホート研究[15]と次に説明するながはま0次予防コホート事業[23]があげられる。ながはま0次予防コホート事業は、京都大学医学研究科と滋賀県長浜市が連携して2005年から計画され、京都大学が受ける公的研究費を主な財源として2007年から研究参加者のリクルートを開始した。30歳から74歳のおおむね健康な長浜市民を対象としており、追跡研究の導入となる0次健診には、目標の約1万人の参加を達成している。多因子疾患の発症における、遺伝的感受性と、種々の環境要因および生物学的要因の相互作用を明らかにすることを目的としている。そのため、提供された血液と尿を保存したサンプルバンクを形成することにより、将来新たなゲノム解析手法が開発された場合に、そのサンプルバンクが利用できるように「ながはまルール」を策定し、研究参加者の包括同意を得ている。原則として遺伝子名と解析方法を明記した上で同意を得るとする現在のゲノム倫理指針では、将来的な遺伝子利用までを含み同意を取る「包括同意」の扱いが明文化されていないので、「ながはまルール」を策定した上で、研究参加者から包括同意を得るという手順が取られている。

我が国の本格的コホート研究は、ここまで振り返ったように、被爆者を対象とした生涯健康影響調査に始まり、現在はゲノムコホート研究へと発展してきている。また、ゲノム決定コストが安価になった現在、ゲノムコホート研究は国民全体の健康増進といったより大きな成果を求め、大規模化を志向している。その場合に問題となるのは、包括同意である。ゲノム科学は日進月歩で進化しており、1年後の研究手法を予測することさえ困難である。そのような状況下での大規模コホート研究遂行は、包括同意のもとに実施せざるを得ない。大規模ゲノムコホート研究から得られる成果とその公共性を鑑みて、ゲノム倫理指針の改定など、様々な制度面の整備を行う必要性がある。



#### 4 ゲノムコホート研究からヒト生命情報統合研究へ

国民の高齢化、生活習慣の変化により患者数が急速に増えている生活習慣病や認知症などは、家族を含めた社会の負担増や医療費の激増などの極めて深刻な社会問題を引き起こしている。高齢者の疾病の多くは、罹患率が高く有病期間が長いことに加えて根治法がない多因子疾患であるため、疾病の早期発見と発症前の予防的介入が危機的な少子高齢化時代において健康で活力ある長寿社会を構築する唯一の手立てとなる。したがって、21世紀の医学の目標は、「病気にかからない」あるいは「病気との平和共存」の医療開発である。しかしながら、こういった疾患に対する患者を対象とした臨床研究は、明確な疾患概念の確立と診断基準の策定に多大な貢献があったが、革新的な予防・治療法の発見には至っていない。また、細胞や動物モデルを用いた病因の探索が世界中で精力的に行われてきたにもかかわらず、その成果をヒトに応用し効果的な薬が開発された例は少ない。このことは、ヒトの多様性を十分に考慮した戦略なしでは、多因子疾患の克服は非常に困難であることを如実に表している。

一方、最近のゲノム、代謝物、タンパク質などの生体分子の研究（オミックス研究）の進歩は、分子レベルで疾患の遺伝素因を明らかにし、また臨床症状が現れる前の発症前診断も一部の疾患で可能となった。また、画像診断における分析技術の向上は、より精度の高い検診の実現と多くの疾患の早期発見につながった。

そういった画期的な技術開発により、世界の疾患研究の流れは、動物モデルや試験管内での疾患解析からヒトの疾患そのものを直接解析する研究へと比重を移しつつある。多因子疾患の発症機構の解明は、伝統的なモデル動物研究では限界がある。欧米では、健康長寿社会の実現に向けて、強力な解析技術の開発を武器に、疾患を統合的に解析する大規模バイオバンクの構築と、それを利用して多数の健常者を長期観察し、疾患の原因を探る集団対象（コホート）研究を国家プロジェクトとして展開している。この戦略は、ヒトの多様性を意識しつつ、病気を「生体分子を通して身体全体で見る」ことが次世代の医学と医療に道を拓く、という考えに根ざしている。

これに対し我が国は、欧米追従の医学研究、時代の変貌に対処できない旧態依然の保健医療システム、世界最先端から立ち後れた生体分子解析の技術開発に加え、予防医学の浸透率の低さのために、疾患解析研究が出遅れた。このままでは、激増する長期療養者や国民医療費の問題に対し、診断・予防・治療方法で解決するための有効な手が打てないばかりか、先行する欧米に予防法や治療薬の開発を軒並みさらわれ、国民のライフラインである保健医療分野の市場が海外企業から席卷され国益を著しく損ねる危機に瀕している。この観点から、先端分析技術を活かした疾患解析基盤の構築は、国家戦略として緊要の課題である。

しかし、この逆境の中でも、我が国は、先端技術の融合と技術改良を通して最高精度のシステムを構築し、研究成果を実用化・汎用化する応用技術では世界でずば抜けた力を持つ。世界に誇る应用能力をもって国際的競争力のある大規模バイオバンクを構築し、長期の追跡によって継続的に得られる生体分子の網羅的かつ詳細な分析・解析情報と質の高い疾病罹患情報を統合した、究極のヒューマンバイオサイエンスとしての「ヒト生命情報統合研究」を実施することで、世界一の長寿国で、世界に一步先じた高齢化社会の健康長寿モデルの構築が可能である。例えば、認知症、脳卒中による寝たきりが2割減れば、介護保険だけで年間3兆円近くの経費節減と試算されており（厚生労働省「社会保障の給付と負担の見通し」より）、「病気にかからない、医療費のかからない医療」がいかに社会に裨益するかは明白である。また、予防に関する情報を用いた新たなヘルスケア産業の創出や保健医療情報のIT化による新時代の保健医療システムの構築も超高齢化社会の我が国にとって極めて重要である。また、日本人の疾患情報は欧米人からは得られないが、アジア人は遺伝的に似通っているため、21世紀の大市場、中国に対しても知的財産上のアドバンテージを得ることになる。

この全く新しい予防医学のアプローチは、医学研究の一分野である疫学で行うコホート研究をはるかに凌駕するものである。まず、このヒト生命情報統合研究においては、今日の最先端の生体試料分析技術による血中成分の定量的測定はもとより、画像解析、細胞動態の分析技術、生体代謝産物から食物摂取状態の推定も可能になると思われる。そのためには、血液、尿、便など非侵襲的に得られる検体から、可能な限りの情報を得られるような新しい分析技術の開発が不可欠であり、理学、薬学、工学の研究者を総動員してその技術開発にあたるべきである。もちろん、医学的通常検査情報も不可欠であるが、これについては登録された個人が自分の意思で病院にかかる際の医療情報等を網羅的に収集できる医療情報ネットワークシステムの構築も不可欠である。そして、施設間で比較可能とするための検査情報の標準化やデータ形式、ひいては医療情報システムの標準化と導入普及策の推進も必須である。大規模研究では、高精度な臨床情報をいかに蓄積するかが重要であり、疾患毎に最適化された標準化プロトコールに基づいた高精度臨床情報を効率よく集積する情報システムや、それを有効に機能させる制度整備が必要である。

集団をUK Biobankと同様に青壮年の世代に設定し20年間の追跡を実施すると仮定して、現在の我が国の疾患発症率をもとに冠動脈疾患や糖尿病などの生活習慣病より頻度の低いアルツハイマー病などの病因に迫ることを考慮すると、対象集団の規模として100万人程度が理想と考えられる。地域がん登録[24]と秋田県立脳血管研究センターのデータ[25]を活用し、年齢階級の罹患率を見ると、40～69歳男女コホート100万人では、胃がん、肺がん、大腸がん、乳がん、

子宮がん、前立腺がん、肝がん、膵臓がん、膀胱がん、腎がん、食道がん、悪性リンパ腫、脳卒中が解析範囲に入り、白血病、脳腫瘍などが外れる。胃がん、肺がん、大腸がんなどは、病理組織型からさらに細かく分かれるが、100万人であればその細分類にも対応可能である。ただし、将来的にはデータ取得対象の年齢層を青壮年の世代に限る必要はなく、ヒト生命情報統合解析という観点ならびに、小児で特に問題となる疾患（自閉症、アトピー性疾患など）の発症率を検討する必要もあり、全年齢をカバーすることも今後期待される。このような大規模な集団の解析はひとつの拠点において行うことは不可能であり、また地域間の環境や生活習慣の差異も考慮すると、全国10カ所程度で行われることが適切と考えられる。しかしながら、このような分散型集団拠点を構築する際に最も重要なことの1つに、様々な測定規格の統一がある。それと同時に倫理的な基準の統一が不可欠である。このような研究では、網羅的な分析、生活情報、臨床情報の統合が成果を上げる上で重要である。また、分析技術の発展により新しい解析方法も積極的に取り込んでいく必要がある。さらには、研究参加者の包括同意を得ることが必須である。研究参加者への説明、同意の取得を経た、包括同意に基づく研究を、個人のプライバシーを適切に保護しながら推進していくためのシステムを設計する必要がある。そのために、国として統一的な基準を構築し、国が一体となってこのプロジェクトを推進することにより、予防医学の発展とともにヒト生命科学の新しい地平を開拓し、我が国の生命科学を飛躍的に強化するものとなることが期待される。

長期にわたるヒト生命情報統合研究では、事業期間中に学術研究成果のみならず、創薬シーズの探索と新薬の開発、新たな分析・解析機器の開発と評価、ITの基盤技術の獲得など、産業界にも大きなイノベーションをもたらす。しかしながら、国家予算だけではこれだけ大規模な研究事業の長期永続的な運営は困難である。また、我が国の国力の回復には、産業育成が不可欠である。そのため、公的研究資金を基本としつつも、産業界からの積極的な研究協力を促し、成果を知的財産として共有する制度を確立することで、多様な分野からの運営資金の調達を目指すことが望まれる。

## 5 提言

### (1) ヒト生命情報統合研究の創出

ゲノム配列の決定にかかる費用と時間は、サンガー法を用いた第一世代に比べて、現在使われている次世代型機器では格段にどちらも縮小した。さらに、きわめて長い塩基配列をさらに高速かつ低コストで決定できる時代が近づいている。この意味で、ヒトゲノム配列の決定自体は、きわめて簡単になりつつある。加えて、近年の新しい物理化学的計測技術の進歩、例えば侵襲のほとんどない画像解析、血液成分からの極めて微量な物質の定量的な測定、なども著しい。

したがって、これからの人間集団を対象とするコホート研究では、単に病気だけでなく、人間の持つ健常形質の多様性にも目を向けるべきである。具体的な表現型の収集は今後の検討を待つ必要があるが、そのような大規模コホート研究で収集すべき表現型情報としては、新しい物理化学的計測技術による解析データ、すなわち血液のメタボロームといった生化学的データ、細胞生物学的データ、生理学的データ、MRI や X 線 CT などによる画像診断データなどを含み、さらには心理学や行動学的データもあげられる。また、通常の間人ドックをはじめとして脳ドックなど様々な検査体制が整いつつある。これらの情報を総合できるシステムが開発されれば、100 万人規模のコホート研究を数十年単位で今後行うことが可能になるだろう。

これら膨大なデータを単一個体から長期間にわたって収集できることは、ヒトが研究対象として極めて優れている点である。ここに、ヒューマンバイオサイエンスの観点から哺乳類あるいは脊椎動物が共有する基本的生命現象を発見できる大きな可能性がある。

ゲノムコホート研究に加え、個々人の多様な表現型情報を長期収集し、収集した膨大なデータを解析する総合的研究から得られる結果は、健康・医学・生命科学の進展に大きく寄与し、我が国の国民の健康を増進するための知的基盤となる。また、生物が共有する基本的生命原理を発見する可能性も秘めている。そこで本分科会では、ゲノムコホート研究をさらに先進的・大規模なものとし、最新の物理化学機器から得られる多様な表現型情報を融合し、得られる巨大な情報を最新の情報科学を用いて統合的に解析するまでを一貫して実施する新しい研究分野すなわち、「ヒト生命情報統合研究」を創出すべきであると提言する。

### (2) 医療情報基盤のさらなる整備

研究参加者が、期間中の任意の時期に自分の意思で受診する医療機関や健診機関での健康医療情報を電子的にかつ標準化された形式で網羅的に収集でき、同一個人ごとに統合できる情報基盤の整備を行う。

そのためには国は次の3点を目指すことが必須である。以下、それぞれについて説明する。

- ①標準化対応の医療情報システムの開発と導入
- ②新たな「国民保健番号（仮称）」の制度と法令の整備
- ③同意研究参加者の医療情報追跡基盤の構築

#### ① 標準化対応の医療情報システムの開発と導入整備

現在多くの企業により販売されている電子カルテシステムや検査情報管理システムは、データ出力規格がバラバラで、容易に多施設データの収集も統合もできない状況にある。国は、標準化されたデータ形式（SS-MIX2形式とよばれる標準規格[26、27]が存在する）で情報を出力できる機能を装備することが、医療情報システムにとって必須の要件とするガイドライン等を整備するとともに、ガイドライン準拠の医療情報システムを導入することが医療機関や検査機関にとって必須となるような制度や共同開発導入体制支援などの環境整備を行う。また既存のすべての医療機関の情報システムにもこの標準化機能を装備するよう経費面での支援を含めた導入推進策を実施する。

#### ② 新たな「国民保健番号（仮称）」の制度と法令の整備

任意の複数医療機関からの医療情報を個人ごとに統合できなければヒト生命情報統合研究における医療情報の統合は実現不可能である。そのためには、同一個人を確実に同一であると同定できる個人識別番号が収集される電子医療情報とリンクできることが不可欠であり、これが医療システムに導入される社会基盤整備も必須である。2012年7月時点で国会に提案されている「行政手続における特定の個人を識別するための番号の利用等に関する法律案」（いわゆるマイナンバー法案）で導入が準備されているマイナンバーは、医療やコホート研究での個人識別番号としての利用はできない。また、マイナンバーは医療のようなセンシティブ（取扱いに機微を要する）な情報との連結（ひも付け）には使われるべきではない。そこでマイナンバー制度が持つ同一人同定の基盤技術を活用しつつも、マイナンバーとは別体系の、医療やコホート研究で使用可能な個人識別番号である「国民保健番号（仮称）」を国は新たに創設すべきである。

「国民保健番号（仮称）」を医療や研究でのみ限定的かつ安全に使用できるようにするには、目的外使用の禁止やその罰則規定を適切に整備した上で、国等が実施するヒト生命情報統合研究や大規模疫学研究において、同意した人に対してはこうした番号を使用できることを明確にするよう、個人情報保護関連法令と関連倫理指針の整備と改訂が必要不可欠である。

### ③ 同意研究参加者の医療情報追跡基盤の構築

長期間にわたる100万人規模の前向き研究では、個人の住居地の移動、連絡手段の変更、受診医療機関の分散化などにより、長期間追跡率を高水準に維持することが非常に困難であり、追跡できない脱落者が年を追って増える。前向き研究において最も重要な結果（アウトカム）指標は、将来の疾患の発生と死亡である。しかし、脱落者が増えていくことによって、これらの重要な結果指標を把握できない患者数が、観察年数を経るに従って増えていく。この問題を解決するには、同意研究参加者の医療情報追跡基盤の構築が必須である。具体的には、研究参加者の事前の包括同意のもとで、住民票記載事項の変更を伴う住所変更、死亡届け提出、加入する医療保険者に集積する医療機関受診など、研究参加者の追跡に手がかりとなる重要な出来事が発生したという情報だけは、確実かつ効率的に研究実施組織が取得できるよう、制度整備とシステム構築をすべきである。また、高齢者の医療の確保に関する法律に基づいて国が収集するレセプト情報は極めて網羅性の高い医療機関受診情報を含んでいるにもかかわらず、現状では保険者の段階で匿名化されているため利用できない。国等が実施するヒト生命情報統合研究や大規模疫学研究では、医療機関受診に関する情報（いつどの医療機関を受診したかに関する情報）の把握だけでも同意研究参加者については可能となるよう基盤構築をすべきである。以上のように、個人のプライバシーを十分に保護しながら、医療情報追跡基盤を適切に利用する仕組みを整備するとともに、研究参加者に対して医療情報追跡について十分な説明と理解のもとに同意を得ることが必要である。

### (3) ヒト生命情報統合研究の拠点整備とその機能

100万人規模のバイオバンクの構築と生体試料、臨床情報、疾患罹患情報を統合したヒト疾患解析による多因子疾患の原因解明と予防・治療法の開発を目標に、中核拠点を軸とした組織構築、制度設計をおこない、既存のコホートをを用いたパイロット研究を経て、コホート事業の開始に向けた解析基盤、情報基盤の整備を行う。そして、事業の三年目をめどに、全国で数カ所の実施拠点を構築し、下図の事業組織による統一基準による全国体制でのヒト生命情報統合解析の実施基盤を構築する。以下に、①中核拠点、②地域研究拠点、③データ解析センター、④生体試料バンク、の順に提案する事業組織の詳細を説明する。また、事業の成果を産業へ結びつけるために、産業界との連携を積極的に行う必要がある。そのための仕組みを、⑤産学連携コンソーシアム、として説明する。

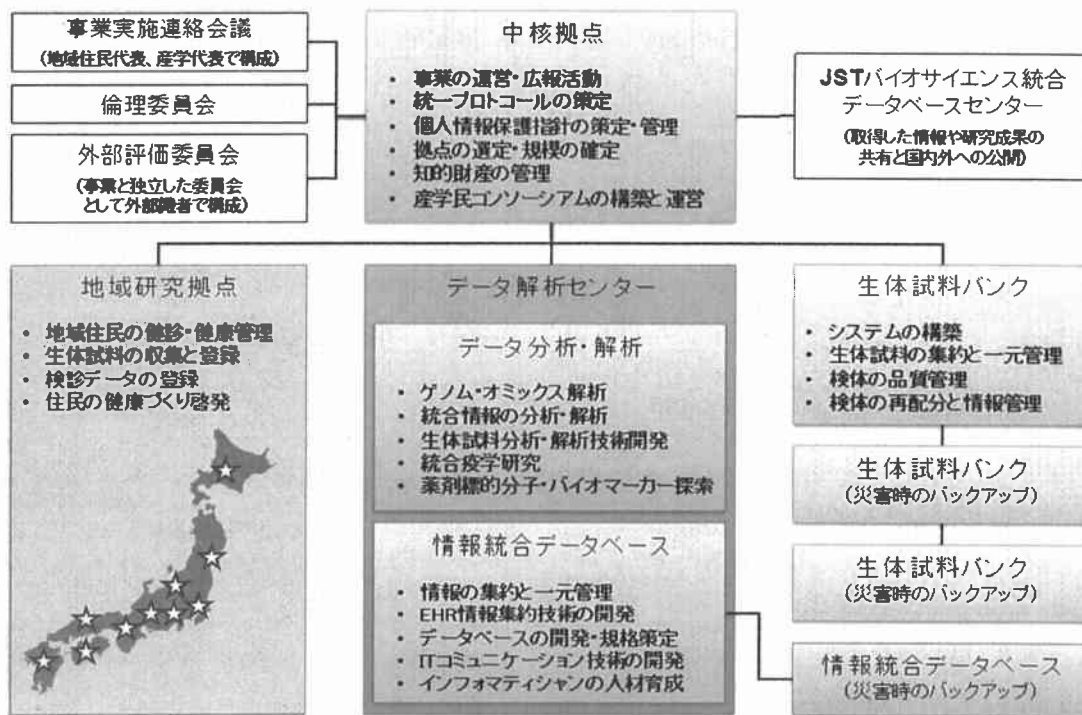


図1 ヒト生命情報統合研究事業 組織図

① 中核拠点ーヒト生命情報統合研究中核拠点の役割

事業全体の司令塔として、事業の統括運営を行う。計測項目や基準の統一と倫理規約の統一は最も重要な役割である。主な業務として、事業組織の構築、制度設計と事業計画の立案を行う。

ア 事業組織の構築

事業の全体計画に基づき、本事業の意思決定機関である事業運営委員会を組織し、事業計画の立案、標準プロトコル策定および地域拠点の選定と事業実施規模の確定を行う。運営委員会の定期的開催で各地域拠点での事業の進捗の把握につとめ、情報の集約と方針決定を行う。事業の倫理委員会を立ち上げ、事業及び付随する研究の倫理審査を実施する。また、実施組織外に外部評価委員会を設置し、事業の第三者による評価を行う。

イ 事業に関連する制度設計

事業の円滑な実施には、研究参加者の個人情報の厳格な保護に加え、研究活動の透明性、説明責任等の倫理的妥当性を担保した包括同意が必須である。そこで新ゲノム指針をふまえつつ、本事業により適合するかたちの倫理規定を策定し、事業実施主体と研究参加者個人の間の契約に基づく登録制度を設計する。また、データ取得に関して必須となるプロトコルの標準化を行う。データや研究成果は、公開条件・安全対策等を検討した上で、JSTバイオサイエンスデータベースセンターを通して公開する。また、

事業の長期間の安定的運営のための産学民コンソーシアムを構築し、産業界からの積極的な研究協力を促し、成果を知的財産として共有する制度を確立する。

### ウ 事業の運営

生体試料バンクの運営と試料の管理、データ解析センターとの連携による、データ分析・解析と情報統合データベースの構築、国内外の関連研究機関（海外のバイオバンク事業、JST バイオサイエンスデータベースセンターなど）との間の連携を統括する。研究者コミュニティが、データを広く利用できるように情報統合データベースを構築し、情報の公開・共有を進める。また、事業で生じた知財の管理や事業の広報活動を行う。

	2012	2013	2014	以後
事業の運営・広報活動	事業実施連絡会議の定期開催			
	生体試料バンクの設計・構築		運営・管理と生体試料の集積	
	データ解析センターの設置		試験的運用	本格稼働
	外部評価	外部評価	外部評価	外部評価
統一プロトコル策定	プロトコル案の策定	試験研究	改訂	策定終了
個人情報保護	倫理委員会の設置			
	倫理委員会の定期開催		個人情報管理システムの構築・運用	
地域研究拠点	個人情報ガイドラインの策定		個人情報管理システムの構築・運用	
	候補地域・機関の調査	拠点選定と実施規模の確定	拠点間の連携体制の確立	
知的財産	知財の取り扱いに関する検討			
	ガイドライン策定		検討・改訂	
産学民コンソーシアム	企業の意識調査		コンソーシアムの設立と運営	

図2 ヒト生命情報統合研究における中核拠点の役割

### ② 地域研究拠点

各地域における事業の円滑な実施・推進のために、まずは地域住民、自治体などとの密接な関係の構築につとめ、健康づくりの土壌を育てる。そして、定期的な健診を実施し、地域住民の健康管理を通じた予防医学研究を行う。具体的実施項目として、統一プロトコルに則った生体試料（血液、尿ほか）の収集と登録、バイオマーカー測定、生理学・身体機能検査などの健診データの登録があげられる。また、大規模研究への参画に加えて、各拠点での独自の研究を実施し、成果を公表するとともに、健診データを利用した住民の健康づくり啓発につとめる。10万人規模のコホートを、10か所程度の地域研究拠点で分担して進めることにより、100万人規模のコホートを実施する。



### ③ データ解析センター

得られた生体試料の分析・解析を実施するとともに、情報統合データベースの構築により、臨床情報や、疾患罹患情報の蓄積を行う。また、それらの多様な情報と疾患との関連を解析する。それらを実行するため、以下の活動を行う。

#### ア データの分析・解析

最新技術を用いたゲノム・オミックス解析を実施し、得られた多様な情報の統合と分析・解析、新たな生体試料分析・解析技術の開発を行う。また得られたデータを用いた情報統合解析研究による疾患関連因子の同定、薬剤標的分子・診断マーカーなどの探索を推進する。

#### イ 情報統合データベース

健診や生体試料の分析・解析で得られる情報の集約と一元管理をおこない、こういった多様な情報、あるいは診療・健康情報とオミックスデータといった異種データを統合するデータ標準化形式の策定とデータベースの構築を行う。また、EHRを用いた疾患罹患情報の集約技術を開発し、情報を集積して分析・解析データと統合する。加えて、ICT技術の開発と、それを用いた研究参加者との双方向コミュニケーションシステムの構築を試みる。解析によって得られた情報や研究成果は、我が国固有のヒト生命情報としてバイオサイエンスデータベースセンターを通して公開し、内外の研究者の利活用を促す。

#### ウ 人材育成

本研究の発展には膨大な臨床情報とゲノム情報の統合解析技術を駆使できる人材育成が急務となっている。この人材育成は、実データ解析を通して疾患研究者と日々議論できる研究フィールドにおいてなされる必要がある。さらに人材育成はキャリアパスを形成するために恒久的な組織で行われる必要もある。これまでは、3年から5年といった時限の教育プログラムが一時的に導入されることが行われてきたが、これでは優秀な人材を時間をかけてフィールドで研鑽を積みせつつ育てていくという教育環境が提供されない。そこで、臨床現場に近い医育機関に臨床ゲノム情報技術人材の育成に専念する専門教育組織を全国数カ所に恒久的に整備することが重要である。本事業は、膨大なライフサイエンスデータを扱う実践教育の理想的な環境であることから、最新のゲノム医学の広範な知識と豊富な経験を持つバイオインフォマティシャン、統計解析研究者の育成に力を注ぎ、我が国のヒト生命情報統合研究における人的資源の確保を目指す。

### ④ 生体試料バンク

得られた生体試料の保管、管理、再分配を行う。適切な管理の下に、連

結可能匿名化を行い、必要な臨床情報とともに生体試料の再分配を行う。各活動は、個人情報保護法などの関連法規制に従って実施されるため、厳格なセキュリティー管理と組織の継続性が必要である。バンクにおいては、バンキングシステムの構築と、LIMS (Laboratory Information Management System) の導入による、各地域研究拠点で収集された生体試料の集約と一元管理および収集された生体試料のデータベース登録と品質管理を主要な業務とする。また、非常時に備えた生体試料の分散管理と事業継続計画の策定や、生体試料の再配分とそれに関わる情報の管理も行う。

#### ⑤ 産学連携コンソーシアム

長期間の追跡をとまなうにわたるヒト生命情報統合研究では、事業期間中に新たなバイオマーカーや先端技術を用いた分析・解析法が創出される可能性が高い。また、疾患発症追跡のための医療情報ネットワーク構築には必ず匿名化のステップが含まれるため、安全性の高い情報集約システムの導入が必要である。こういった研究開発の実施はアカデミアの力のみでは困難であり、製薬・医薬機器、分析・測定・診断機器、情報・通信など様々な産業分野で強い開発能力を有する企業との連携が必須である。加えて、我が国の国際競争力の回復には、新規産業の創成と育成が不可欠である。そこで、以下のような目標設定のもと、産学連携を推進する。

- ア 長期の医療情報の集積とセキュリティーの高い情報管理
- イ 研究者と企業が法的規制を遵守し活用できる制度設計
- ウ 研究参加者に対する研究・開発の透明性を保ったマネジメント
- エ 長期に安定した事業資金を確保し、孫子の世代までプラスを生むシステム
- オ 我が国の先端技術を利用した新たな医療産業の育成
- カ 成果が健康づくりに速やかに応用できる健康産業の創出
- キ 人材の育成と安定した雇用を生む実践的な教育プログラム
- ク 情報統合のノウハウを世界に輸出できるモデルの構築

産学連携の推進にあたっては、企業が参画しやすい体制を備える必要がある。また、研究の体制は、集積された情報へのアクセス、アカデミアとの共同研究、生体試料や情報を利用した企業単独の研究など、多様な形態が存在する。そのために、以下のような条件のもとでの連携体制を構築する。

- ア 企業が研究開発に必要な情報の提供
- イ 生体試料を用いた分析・解析の受託
- ウ 研究開発段階でのマイルストーンは不要
- エ 資金を提供した研究には優先開発権を付与

- オ 知的財産使用料は製品・サービス開発後に回収
- カ 技術供与等も投資として認める

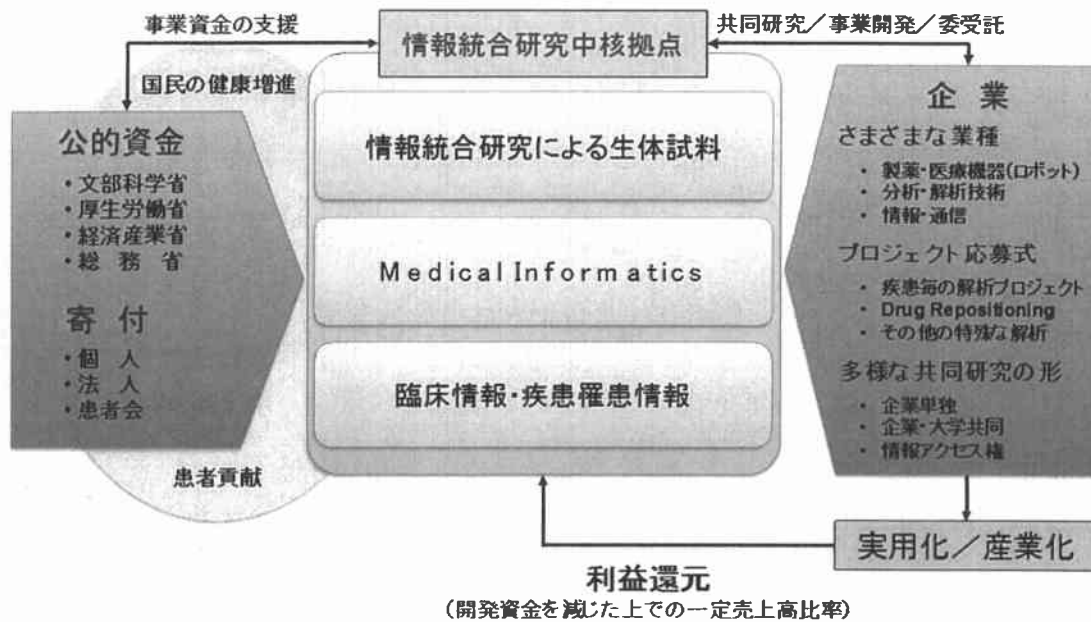


図3 ヒト生命情報統合研究における企業参画のしくみ

#### (4) 包括同意にもとづく研究を可能とすることに対する国民の理解醸成と研究倫理指針の改訂

ヒト生命情報統合研究は国民全体の健康増進といったより大きな成果を求め、大規模化、長期間にわたる前向き研究を志向している。このような研究では、網羅的な分析と生活情報・臨床情報の統合が成果を上げるために重要である。特に、分析技術は日進月歩で進歩しており、新しい解析方法を積極的に取り込んでいく必要があるため、このような研究においては、包括同意を得る必要がある。しかしながら、ゲノム解析においては、個人のプライバシーに関する情報が含まれることから、プライバシーの保護にも十分な配慮を行い、研究参加者の理解を得なければならない。研究内容、意義に対する研究参加者の十分な理解を得た上で、包括同意のもとでの長期間研究への協力が必須となる。ヒト生命情報統合研究から得られる成果とその公共性を鑑みて、国は、一定条件を満たす研究については包括同意を可能とすることに対する国民の理解醸成を積極的に行い、関連する研究倫理指針を改訂して包括同意にもとづく研究遂行を可能するとともに、様々な制度面の整備を行う必要がある。

ここで提案されている研究は、100万人を対象とすることからも、その規模において前例のない大規模研究となり、必然的に予算規模もきわめて大きくなる。研究の実現には、国民的理解を得ることが前提となる。本研究の実現には、

研究に対する国民的な理解の醸成と支援を得ることが何より重要となり、倫理的・法的・社会的課題を検討し、国民的理解の醸成を積極的に推進する体制を充実させることが重要である。

#### **(5) 研究拠点整備と提言実行のための推進協議会（仮称）の設置**

ここまでに述べた研究拠点の整備等の実現を図るため、専門家によるヒト生命情報統合研究推進協議会（仮称）を設置する。この協議会においては、全国規模で行われる大規模ゲノムコホート研究の標準化、研究対象となる集団・地域の選定、中核機関と地域拠点の連携方法、既に進行している他のコホート研究プロジェクト（環境省・エコチル調査など）との連携あるいは棲み分けなどを議論し、事業方針を決定する。事業の開始後あるいは、拠点の設置後は、該協議会は事業の評価を行い、必要な場合には事業の再設計を行う。

協議会のメンバーは、現在実施されているコホート研究の関係者、医療関係者（看護系、保健師等も含む）、基礎生物学研究者、情報学関係者、倫理問題の専門家などから形成される。

## <参考文献>

- [1] 東北メディカル・メガバンク機構 : <http://www.megabank.tohoku.ac.jp/>
- [2] 英国 UK Biobank: <http://www.ukbiobank.ac.uk/>
- [3] スウェーデン LifeGene プロジェクト: <http://lifegene.ki.se/>
- [4] カナダ Partnership for Tomorrow プロジェクト:  
<http://www.partnershipagainstcancer.ca/priorities/research/strategic-initiatives/canadian-partnership-for-tomorrow-project/>
- [5] 中国 China Kadoorie Biobank: <http://www.ckbiobank.org/>
- [6] 欧州連合 European Investigation into Cancer and Nutrition プロジェクト:  
<http://epic.iarc.fr/>
- [7] 米国 NIH-AARP Diet and Health Study: <http://dietandhealth.cancer.gov/>
- [8] 米国 National Children's Study:  
<http://www.nationalchildrensstudy.gov/Pages/default.aspx>
- [9] 米国 American Cancer Society Cancer Prevention Study:  
<http://www.cancer.org/Research/ResearchProgramsFunding/Epidemiology-CancerPreventionStudies/CancerPreventionStudy-3/index>
- [10] 米国 Veterans Affairs Genomic Medicine Program:  
<http://www.cdc.gov/genomics/translation/GAPPNet/meeting/file/print/slides/Przygodzki.pdf>
- [11] 米国 Vanderbilt BioVU: <http://dbmi.mc.vanderbilt.edu/research/dnatabank.html>
- [12] New Models For Large Scale Prospective Studies シンポジウム:  
<http://conferences.thehillgroup.com/largeststudiessymposium/index.html>
- [13] Yoo KY, Shin HR, Chang SH, Choi BY, Hong YC, Kim DH, Kang D, Cho NH, Shin C, Jin YW. Genomic epidemiology cohorts in Korea: present and the future. *Asian Pac J Cancer Prev.* 6:238-243, 2005.
- [14] 広島・長崎の被爆者を対象とした生涯にわたる健康影響調査研究:  
<http://www.rerf.or.jp/intro/establish/index.html>
- [15] 久山町コホート研究:  
<http://www.med.kyushu-u.ac.jp/psychiatry/cn11/pg117.html>
- [16] 三府県コホート研究: <http://www.pbhealth.med.tohoku.ac.jp/node/314>
- [17] 文部科学省科学研究費がん特定領域大規模コホート研究:  
<http://www.aichi-med-u.ac.jp/jacc/index.html>
- [18] Japan Arteriosclerosis Longitudinal Study : <http://jals.gr.jp/>
- [19] Hirayama, T. Non-smoking wives of heavy smokers have a higher risk of lung cancer: a study from Japan. *British Medical Journal.* 1981, 282:183-185.

[20] JPHC Study: <http://epi.ncc.go.jp/jphc/index.html>

[21] 山形大学を中心とした地域連携コホート:

<http://gcoe.id.yamagata-u.ac.jp/jp/cohort/>

[22] 高齢者高血圧コホート研究 (J-CHEARS Study) :

<http://jchears.hosting.p1d.unet.dion.ne.jp/index.html>

[23] ながはま0次予防コホート事業:

<http://www.city.nagahama.shiga.jp/index.cfm/9,3709,19,158.html>

[24] 地域がん登録統計データ:

<http://ganjoho.jp/professional/statistics/statistics.html>

[25] 秋田県立脳血管研究センター統計データ:

<http://www.pref.akita.lg.jp/www/contents/1287488551823/files/data.pdf>

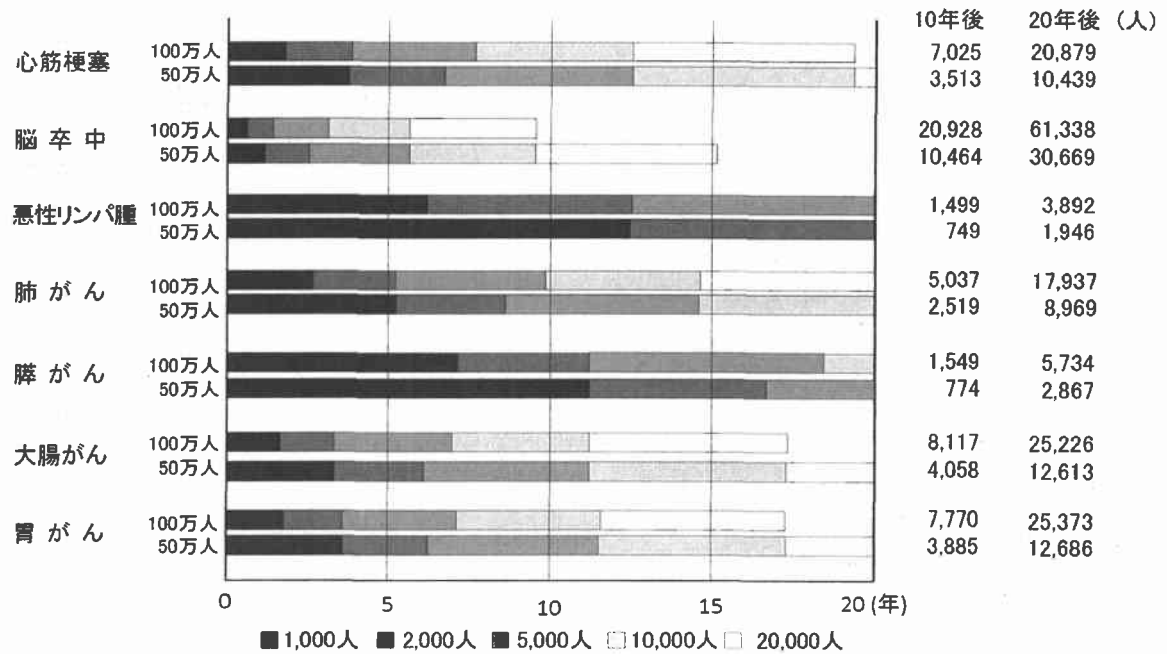
[26] SS-MIX 普及推進コンソーシアム: <http://www.hci-bc.com/ss-mix/>

[27] 日本医療情報学会: <http://www.jami.jp/jamistd/index.html>

<参考資料1>

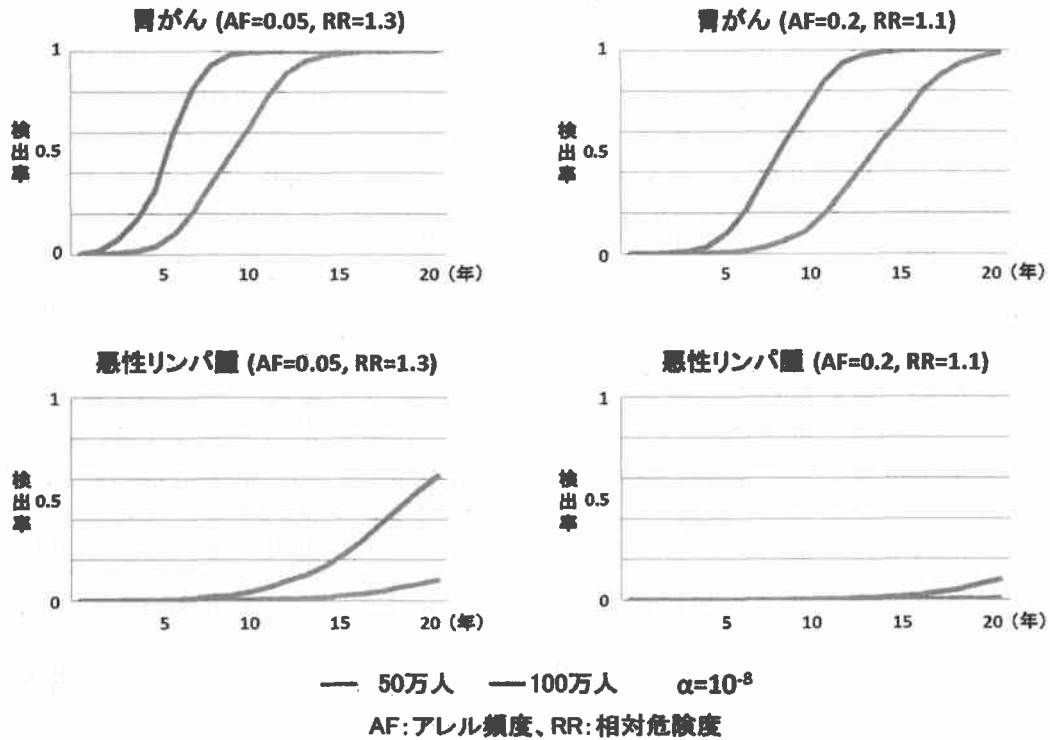
対象集団を100万人規模とする根拠について

図1



50万人と100万人の集団での各疾患の累積発症者数を、追跡の時間軸で表したものの。母集団の規模を2倍にすると、10年で一万人の患者が集まるものが多い（心筋梗塞、肺がん、大腸がん、胃がん）ため、短い追跡期間で疾患解析が可能となる。（[23]のデータをもとに計算）

図 2



頻度の高い胃がんと、低い悪性リンパ腫での検出力の推定で、いずれの疾患においても 100 万人規模の場合大きな検出力に早く到達する。

例えば左上の胃がんの場合、アレル頻度が 0.05 の遺伝的多型で、相対危険度が 1.3 のものを統計的な有意水準が  $10^{-8}$  で得られる可能性は、約 8 年後に計算上は 100%になる。([23] のデータをもとに計算)



## <参考資料2>ゲノムコホート研究体制検討分科会審議経過

平成 24 年

- 1月27日 日本学術会議幹事会（第144回）  
○ゲノムコホート研究体制検討分科会設置
- 2月20日 日本学術会議幹事会（第146回）  
○委員決定
- 4月4日 ゲノムコホート研究体制検討分科会（第1回）  
○役員を選出  
○今後の進め方について  
・本庶佑委員から「ゲノムコホート研究について」紹介  
・菅野純夫委員から「ゲノムコホート研究の国際的な状況」の紹介  
・これらに基づき提言作成に向けた活動予定を策定。
- 5月7日 ゲノムコホート研究体制検討分科会（第2回）  
○各機関における関連活動報告のヒアリング  
・山本雅之委員から「東北メディカル・メガバンク (ToMMo) 構想」の紹介  
・高木利久委員から「統合データベースプロジェクト」の紹介  
・本庶佑委員から「ゲノム・環境情報統合を基盤とした Human Biological Science の世界への発信」の紹介  
○今後の進め方について審議
- 5月28日 ゲノムコホート研究体制検討分科会（第3回）  
○日本におけるゲノムコホート研究の現状と課題についてヒアリング  
・松田文彦参考人から「ながはまゲノムコホートについて」紹介  
・清原裕参考人から「久山町におけるゲノム疫学研究」の紹介  
・辻省次委員から「ヒトゲノム・遺伝子解析研究に関する倫理指針の見直しについて」の紹介  
・本庶佑委員から「ゲノム・環境情報統合研究における中核拠点の形成スケジュールと役割」の紹介  
○提言の構成について審議

- 6月28日 ゲノムコホート研究体制検討分科会（第4回）  
○日本におけるゲノムコホート研究の現状と課題についてヒアリング  
・深尾彰参考人から「地域に根ざした山形ゲノムコホート創成」の紹介  
・津金昌一郎参考人から「我が国において、大規模ゲノムコホート研究を実施するにあたっての諸問題」の紹介  
○提言の構成について審議
- 7月9日 ゲノムコホート研究体制検討分科会（第5回）  
○提言案について審議
- 7月24日 ゲノムコホート研究体制検討分科会（第6回）  
○提言案について確定
- 7月27日 日本学術会議幹事会（第155回）  
○ゲノムコホート研究体制検討分科会提言「ヒト生命情報統合研究の拠点構築—国民の健康の礎となる大規模コホート研究—」について承認

## 医療・健康等に関する情報の利活用についての意見書

2012年8月28日

竹内 誠

疾病構造の変化や少子高齢化等に伴い医療需要は増大し、また多様化してきている。このような背景の下、有効性・安全性・効率性等の観点から質の高いサービスを如何に提供していくかが重要な課題であると考え。そのためには、エビデンスに基づく医療や医療の標準化を進めることが重要な視点である。

一方、ICT (Information and Communication Technology) 技術の進歩により、長期的かつ正確に情報を収集・蓄積することが可能になっている。これらの情報を活用した医学研究を推進することにより新たな医療エビデンスを創出し、それに基づいて、今後の医療政策やその関連施策を提言、企画・立案することが可能になってきている。

欧州等においては、国の政策として医療情報を蓄積し、利活用することが行われている。我が国においても医療情報を蓄積し、利活用することにより医療イノベーションを実現することが喫緊の課題であり、明確なグランドデザインの下に、実際に活用できる医療情報データベース(DB)を構築することが求められる。しかし、現有の医療情報 DB は標準化がなされていないため互換性がなく、使い勝手が悪い。そこで今後構築される医療情報 DB に関しては標準化を進め、互換性のあるものとして確立すべきであると考え。また、既存の医療情報 DB についても必要な情報をやり取りできる仕組みを作り、利活用できる状態にすることが望まれる。

更に、医療情報は高度な個人情報を含み、基本的に保護されるべきであるが、公益的なもの、即ち公衆衛生上必要性の高いものは例外的扱いとし、匿名化等保護のための必要な措置を講じた上で医療情報の利活用を認めるべきである。

また、これらの医療情報を利活用する側の視点も重要である。実際に用いられて期待される成果を挙げるためには、企画・設計段階から利活用者である産業界関係者も参画し、どのような目的でどのような医療情報を収集・蓄積し、どのように利活用するのか、の観点から構築の検討に加わる必要があると考え。

このような認識の下、医療情報の利活用に関して、以下の4点を提案する。

#### 1. 医薬品の安全対策や研究開発に向けた医療情報 DB の利活用の推進

英国では同様の施策が既に開始されているが、現在、我が国では電子化された大規模医療情報 DB は産業界等で利活用できる環境にない。より海外に先行した形で我が国の強みである ICT や医薬品開発を確立するためには、医療情報 DB の産業界での利活用は必須と考える。

産業界が利活用する目的には安全対策と研究開発がある。

医薬品の安全性・有効性の確保は医療の質の向上に不可欠であり、現場の医療関係者のみならず、企業も責任を有している。このため、必要に応じ企業側も利活用できるような仕組み作りが必要である。

また、研究開発に関しては、疫学情報、臨床情報、オミックス情報等を基礎研究とリンクさせることにより、疾患メカニズムの解明や新規創薬ターゲットの発見等に結びつくことが期待され、結果的には製薬企業における医薬品の研究開発が促進される。

したがって、国の施策として医療情報 DB を構築し、利活用を推進するためには企画・設計段階から利活用者としての産業界の参画が望ましいと考える。

## 2. 医療情報 DB リンケージの推進

医療等分野における識別子である「医療等 ID (仮称)」ができ、医療情報 DB 間のリンケージが可能になれば、医薬品の暴露と有害事象の因果関係が解明されやすくなり、結果的には我が国の医薬品の安全監視の飛躍的改善(薬害のない国)をもたらすことが期待される。

例えば、ナショナルレセプト DB とセンチネル DB のリンケージが一例である。また、ナショナルレセプト DB とがん登録のリンケージにより、医薬品と発癌性の因果関係も同様に解明されやすくなる。

上記のごとく医療情報 DB リンケージが、DB 研究の進展および疫学の進展をもたらし、医療政策の立案や産業界における研究開発戦略策定等に活用することにより、医療への多大な貢献が期待される。

## 3. 電子カルテの普及促進

医療等分野における安全かつ効率的な情報連携を可能にするためには、「医療等 ID (仮称)」とそれに基づく認証・認可機能、情報提供ネットワークシステムとの接続機能を中心とする「医療等情報中継 DB (仮称)」の構築と電子カルテの普及が重要である<sup>1)</sup>。海外でも検討されているように電子カルテを普及させるための国家的な施策が必要である。

1)電子カルテの導入率(JAHIS 資料):診療所 12.8%、病院 14.3%、400床以上の病院 46.5% (2010年)。臨床医薬 28 巻 5号 (5月) 2012

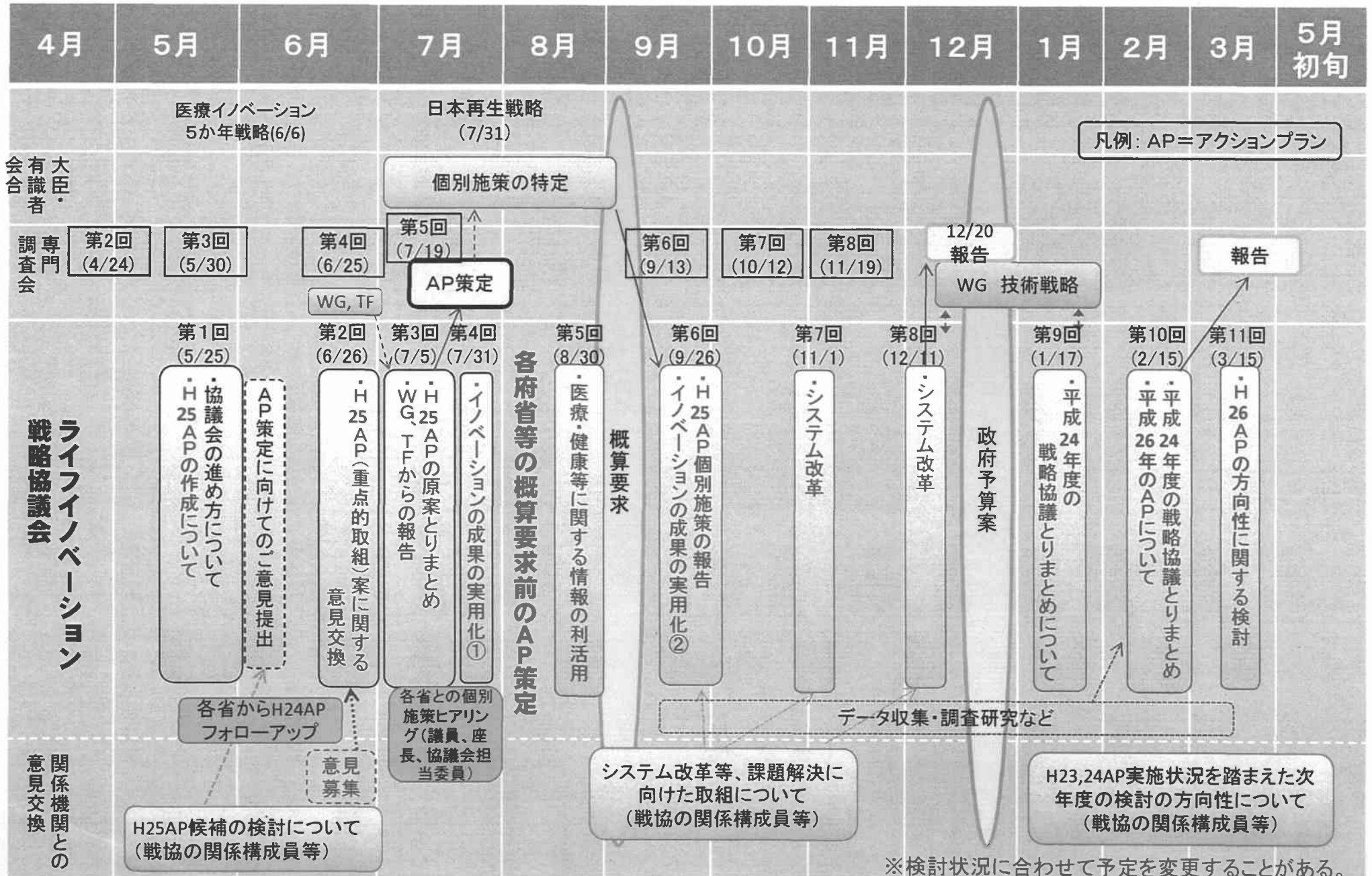
## 4. 人材育成

現在、我が国には医薬品と有害事象の因果関係や有用性等を研究する薬剤疫学専門家が殆ど存在しない。医療情報を利活用するためには、各種医療情報を利活用しやすいデータとして変換する薬剤疫学専門家を早急に育成することが望まれる。

以上

# ライフイノベーション戦略協議会の年間スケジュール予定

資料3



## バイオバンクの現状と展望

医薬産業政策研究所 主任研究員 南雲 明

人類を苦しめる多くの病気（がん、心臓疾患、脳血管疾患など）の大部分は、生活習慣などの環境要因と遺伝要因が複雑に相互作用することによって発症・進展すると考えられている。これらの要因を明らかにすることは、有効な医薬品・治療法の開発や予防医療等を進める上で極めて重要である。主に環境要因に関しては、疫学的手法を取り入れた研究が古くから行われていたが、近年、ゲノム医科学研究の進歩を背景に、従来の疫学研究にゲノム解析情報を取り入れて遺伝要因を探るゲノムコホート研究が注目されている。「バイオバンク<sup>1)</sup>」は、これらの研究を行う上で必要不可欠な研究資源であり、世界各国でその整備が進められている。本稿では、バイオバンクに関する背景と現状を概観し、その課題と展望を考察する。

### コホート研究とケース・コントロール研究

まず、バイオバンクの重要性を理解する上で必要となる疫学研究の基本事項について記載する。疫学研究の代表的手法として、コホート<sup>2)</sup>研究（追跡調査研究）とケース・コントロール研究（症例対象研究）がある。コホート研究は、一定の研究参加者集団を長期にわたり追跡調査する手法で、一般的には現在から未来へと時間軸に沿って追跡する。一方、ケース・コントロール研究は、疾病

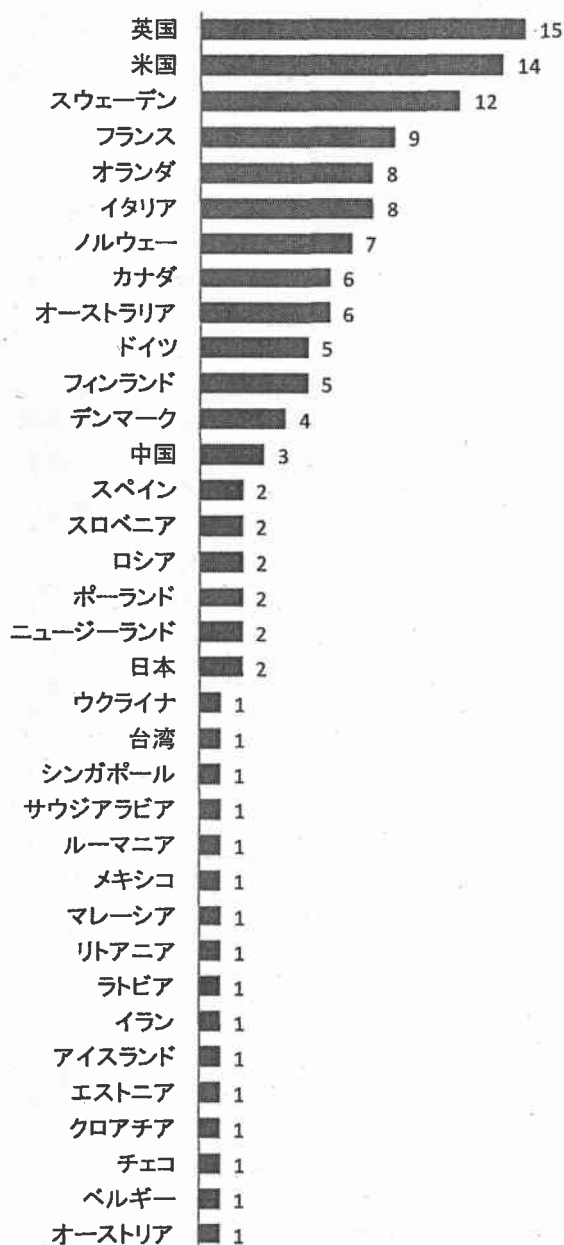
の有無によって患者集団（ケース群）と非患者集団（コントロール群）を設定し、それぞれの過去に遡って要因を解析する。世界的に整備が進められているバイオバンクの多くは、コホート研究のためにデザインされており、対象となる研究参加者は原則として健常人である。研究参加者の将来を長期的に追跡調査し、あらかじめ設定した調査項目（環境・医療・ゲノム情報など）と疾病の発症・進展状況や治療に対する反応性などを前向きに解析する。疫学研究を実施する上で、バイオバンクは非常に重要な役割を担っている。OECD（経済協力開発機構）によれば、バイオバンクとは「集団または集団中の小集団に関する生体物質とそれに関連したデータ・情報が、高度に管理されたシステムで保管されたコレクション」と定義されている<sup>3)</sup>。すなわちバイオバンクには、単に生体試料（血液・細胞・組織など、以下「試料」）が保管されているだけでなく、それに付随する様々な個人情報（環境・医療・家系情報など、以下「情報」）が厳格なガバナンスの下で保管・維持・管理される。

### バイオバンクの現状

次に、世界におけるバイオバンクの現状を概観する。ここでは、世界の主要バイオバンク100件を

- 1) バイオバンク：広義には微生物・細胞などの研究試料バンク、血液・組織・臍帯血・骨髄、幹細胞などの治療目的バンク、病理診断用組織切片などの診断目的バンク、精子・卵子バンクなどが含まれるが、本稿の対象は疫学研究に利用されるヒト生体試料バンクとする。
- 2) コホート（cohort）の語源：元々は古代ローマ軍の歩兵隊一連隊をさす言葉。前進するローマ軍の姿から「前向き調査研究」を指すと言われる。
- 3) OECD Glossary of statistical terms Definition (<http://stats.oecd.org/glossary/detail.asp?ID=7220>)

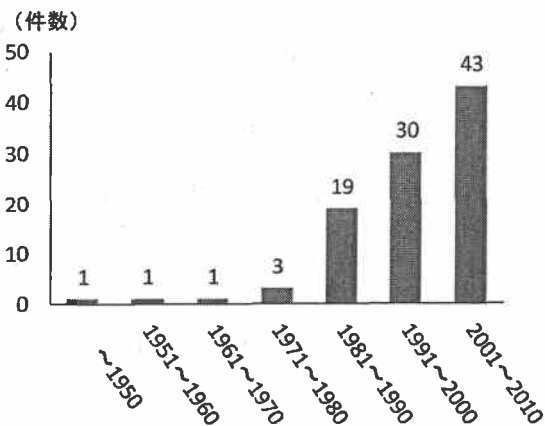
図1 国別バイオバンク数



注1：複数国が関与するバイオバンクについては、重複してカウントした。

出所：Public Population Project in Genomics (P<sup>3</sup>G) のWeb サイト、Biobanks in Europe : Prospects for Harmonisation and Networking, European Commission-Joint Research Centre (2010) 及び関連文献をもとに作成 (2012年5月現在)。

図2 年代別バイオバンク設立数



注1：開始年度不明のバイオバンク (1件) と2011年以降に開始されたバイオバンク (1件) は含まない。

出所：図1に同じ。

リストアップし<sup>4)</sup>、その内容について調査した。図1には、国別のバイオバンク数を示す。

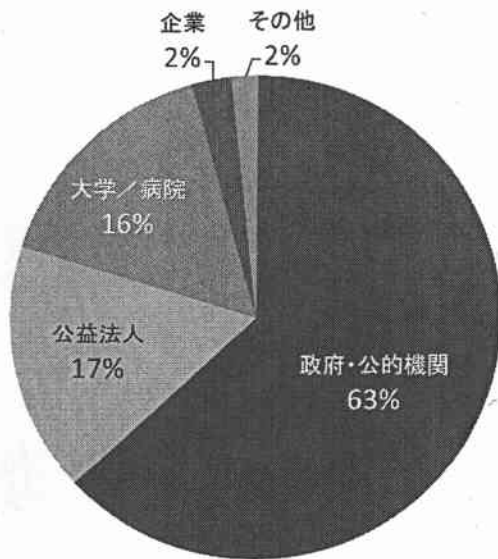
バイオバンクが最も多い国は英国 (15件) で、次いで多いのは米国 (14件) であった。また、全体の7割にあたるバイオバンクが欧州に存在し、特に北欧諸国 (スウェーデン12件、ノルウェー7件、フィンランド5件、デンマーク4件) に多くのバイオバンクが見られた。中でもスウェーデンは、英国や米国に次いで第3位のバイオバンク数を示した。スウェーデンでは、古くからヒト由来試料が研究資源として保存されていた。また、個人識別システム (個人識別番号) や健康情報登録制度など、バイオバンクに必要な社会基盤が整備されており<sup>5)</sup>、これらがバイオバンクの設立を促した一因と考えられる。

図2は、バイオバンクの設立数を年代別に示したものである。1970年以前に設立されたバイオバンクは3件で、最も古いものは1948年に設立された米国の Framingham Heart Study だった。1980年代に入るとバイオバンクの設立数は19件と急速に増加し、ゲノム医学研究が本格化した2000年代には43件に達している。

4) バイオバンクに関する国際コンソーシアム P<sup>3</sup>G (Public Population Project in Genomics) に登録されたバイオバンクのうち必要な情報が揃う94件、及び EC/JRC (European Commission/Joint Research Centre) のレポート “Biobanks in Europe : Prospects for Harmonisation and Networking (2010)” 等より選んだ特徴的な6件を加えて100件とした。

5) 増井徹、「スウェーデンのバイオバンク」、科学研究費補助金 (基盤 (B)) 「生命科学・医学の発展に対応した社会規範形成 - 生命倫理基本法の構築」班長：位田隆一 分担研究報告書 (平成18年度)。

図3 スポンサー別バイオバンク数



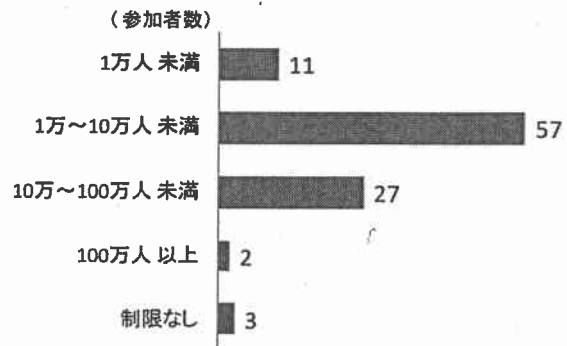
注1：複数のスポンサーが関与するバイオバンクについては、重複してカウントした。  
出所：図1に同じ。

図3にはスポンサー別のバイオバンク数を示す。バイオバンクのスポンサーとしては、政府や国公立研究機関などが60%以上を占め、非営利団体などの公益法人と大学または病院によるバイオバンクがそれぞれ16~17%を占めた。また、民間企業が単独でスポンサーとなっているバイオバンクは、アイスランドのIcelandic Biobank (deCODE)のみであった。

図4に示すバイオバンクの参加者数を見ると、1万~10万人が約6割、10万~100万人が約3割を占め、参加者数100万人を超えるバイオバンクが2件存在した(英国 Million Women Study：130万人、米国 Million Veterans Program：100万人)。

図5には、「試料」及び「情報」について種類別にバイオバンク数を示す。「試料」に関しては、血液のみを収集するバイオバンクが35件、血液に加えて他の体液(唾液・尿・臍帯血)と口腔細胞・毛髪・爪などを収集するバイオバンクが43件であった。さらに、これらの「試料」に加えて脂肪組織・脳組織・がん病変組織などの体組織を収集するバイオバンクが16件であった。「情報」に関しては、環境情報(居住地・教育歴・出生国・家族構成・生活習慣等)および医療情報(身体所見・病

図4 バイオバンクの規模



注1：参加者数は原則として予定数または最終目標数を示す。  
出所：図1に同じ。

図5 バイオバンクの「試料」・「情報」の種類

<試料>

血液のみ	血液+体液+他	血液+組織+他	その他
35	43	16	6

<情報>

環境	医療	環境+医療	環境+医療+家系
14	4	67	15

注1：「血液+体液+他」は血液・唾液・尿・臍帯血・口腔細胞・毛髪・爪などを含む。「血液+組織+他」は上記に加えて脂肪組織・脳組織・がん病変組織などを含む。ただし、各バイオバンクがそれぞれ全ての試料を含むわけではない。

注2：「環境」は居住地・教育歴・出生国・家族構成・生活習慣などの情報、「医療」は身体所見・病歴・診療録などの情報、「家系」は家系図情報を含む。ただし、各バイオバンクがそれぞれ全ての情報を含むわけではない。

出所：図1に同じ。

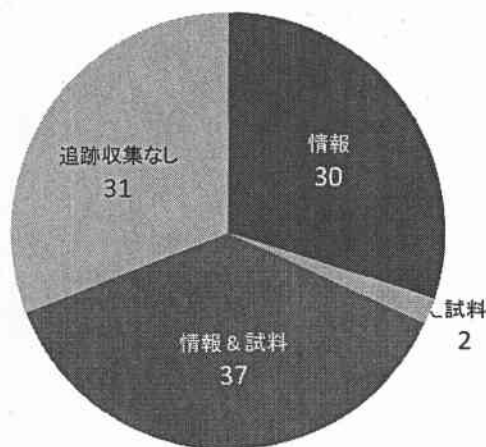
歴・診療録等)をあわせて取得するバイオバンクが約7割(67件)を占め、これらに加えて家系情報を収集するバイオバンクが15件であった。

図6には、「試料」及び「情報」の継続的収集を設定しているバイオバンク数を示す。約7割のバイオバンクで、「試料」または「情報」のいずれかまたは両方について収集のフォローアップが計画されている。

表1は、バイオバンクで保管・管理されている「試料」及び「情報」について、外部研究者への提供の可能性を示したものである。アカデミア研究



図6 「試料」・「情報」収集のフォローアップ



出所：図1に同じ。

表1 外部研究者への試料・情報の提供

	情報 (+ 試料)	
	アカデミア	企業
外部研究者への提供可	77 (71)	61 (55)

注1：( )は情報及び試料を提供可能なバイオバンク数。  
出所：図1に同じ。

者に対しては77件のバイオバンクで「情報」の提供を認めており（うち71件は「試料」提供可）、企業研究者に対しては61件のバイオバンクが提供を認めていた（うち55件は「試料」提供可）。ただし、提供にあたっては、ほとんどの場合、各バイオバンク独自の審査などの条件が設定されている。

#### 特徴的な世界のバイオバンク

世界のバイオバンクのうち、特徴的なものについて、以下にその概要を紹介する。

○Framingham Heart Study（フラミンガム心臓研究）（米国）：

米国における心血管合併症増加への対応を検討するため、1948年に米国心臓研究所<sup>6)</sup>によって開始されたコホート研究。マサチューセッツ州フラ

ミンガム市の住民5,209人によって開始され（オリジナルコホート）、その子供と配偶者を追跡するオフスプリングコホート（5,124人）、さらにその子供たちを対象とした第三世代コホート（4,095人）を加えた追跡調査が現在も進行中。本研究は、当時感染症研究の手法であった「疫学」を初めて循環器疾患分野に応用したもので、疾患の危険因子（リスクファクター）という概念が提唱され、心血管疾患の予防と治療に多大な影響をもたらした。研究開始当時には想定されなかった研究項目に備え、血清の凍結保存が行われるなど、その先見性が高く評価されている。また、データ解析のために新たな統計解析手法（多変量解析など）が開発され、後に続く研究にも大きな影響を与えた。現在では、保管された「試料」・「情報」を世界中の研究者に提供しており、バイオバンクのプロトタイプと考えられている。

○Kaiser Permanente Research Program on Genes, Environment and Health（米国）：

カリフォルニア州を拠点とする米国最大の保健維持機構<sup>7)</sup> Kaiser Permanente によって、2005年に開始された大規模コホート研究。喘息・アルツハイマー病・がん・鬱病・糖尿病・心疾患・生殖障害など、多岐にわたる疾患の発症に関与する遺伝因子・環境因子の同定を目的とする。Kaiser Permanente が提供する健康保険の加入者を対象に参加者を募り、50万人からの「試料」・「情報」収集を目標としている。本研究の特徴は、健康保険加入者の医療記録を直接利用できる点にあり、疾患発症に関する遺伝因子・環境因子同定の他に、新薬の効果を評価する目的でも利用できるものと期待されている。保管された「試料」・「情報」は一定の条件下で外部研究者に提供される。

○UK Biobank（英国）：

英国人口の約1%にあたる50万人を対象とし

6) 米国心臓研究所（National Heart Institute）：米国衛生研究所（National Institute of Health）傘下の研究所で、現在は米国肺血液研究所（National Heart, Lung and Blood Institute）。

7) 保健維持機構（Health Maintenance Organization）：米国の保健医療システムにおいて、企業が提供する健康保険のうちの1つ。

て、2003年に予備的研究が始まり、2007年から本格的にスタートした大規模コホート研究。生活習慣病を含む一般疾患の罹患率が上昇する40歳から69歳の英国国民を対象に、自由意思による参加者を募り「試料」・「情報」が収集される。また、参加者には今後20～30年にわたる追跡調査への協力が要請されている。「情報」のフォローアップは、定期的なアンケートの他、英国国民健康サービス(NHS)<sup>8)</sup>が集中管理している医療情報データベースなどを通じて行われる。長期間にわたり必要となる巨額の資金は、Wellcome財団、英医学研究諮問委員会(MRC)、英国保健省(DH)、NHSなどから出資される。収集された「試料」・「情報」は、厳格なガバナンスの下で管理され、UK Biobankの倫理審査委員会による審査を通れば、外部研究者や企業も利用することができる。

#### ○Icelandic Biobank (deCODE) (アイスランド):

1996年に設立された民間企業 deCODE Geneticsにより運営されているバイオバンク。アイスランドは、北大西洋上に位置する人口約30万人の島国。移民の流入がほとんどなく、少数祖先の遺伝子プールを保ったまま現在の人口まで増加したため、国民の遺伝的同一性が非常に高い。また、医療水準が高く、全国的に均一な医療提供体制が敷かれており、各医療機関には長期にわたる国民の診療記録が保存されていた。さらに先祖代々の家系図を保つことが誇りとされており、1000年以上に及ぶ家系図記録が残されていた。1998年に成立した「保健医療分野データベース法(Health Sector Database Act)」により、国内の全医療情報が1つのデータベースに集約され、個人の遺伝情報や家系図情報との照合も可能となった<sup>9)</sup>。この法律は、これらの「情報」をアイスランドにおける健康政策に還元することを目的としたが、同時に

deCODE Genetics社が独占的に「情報」を企業活動に利用することも許可した。現在までに、アイスランド国民112,500人から血液が収集され、それらを用いた疾患関連遺伝子探索などの事業が展開されている。外部研究者からバイオバンク内の「試料」・「情報」への直接アクセスは認めず、主に共同・委託研究契約などによる研究が行われている。

#### 日本のバイオバンク

本稿の調査でリストアップした100件のうち、日本のバイオバンクは2件(多目的コホート研究、Biobank Japan)であった。ここではこれらに加えて、近年日本において整備が進むバイオバンクの概要を併せて以下に紹介する。

#### ○多目的コホート研究(Japan Public Health Center-based Prospective Study)(日本):

日本におけるがんや心血管疾患の危険因子を明らかにし、生活習慣病対策を推進することを目的とした大規模コホート研究。1990年に開始され、これまでに約14万人から「情報」の収集が行われており、そのうち約6万人からは血液も収集されている。本研究は、2009年度まで厚生労働省がん研究助成金による指定研究班として実施されていたが、2010年度以降は国立がん研究センターによって実施されている。本研究は、全国11保健所、国立循環器病研究センター、大学、研究機関、医療機関などとの共同研究として実施されており、その他の研究者への「試料」・「情報」の提供については不明。2011年12月より、国立がん研究センターにおける新たな試みとして、保存血液を用いた「オミックス研究」<sup>10)</sup>が開始されている。

#### ○Biobank Japan(日本):

ミレニアムプロジェクトの一環として、2003年より実施された「オーダーメイド医療実現化プロ

8) 英国国民保健サービス(National Health Service, NHS): 患者の医療ニーズに対して公平なサービスを提供することを目的に1948年に設立された英国の国営医療サービス事業。利用者の健康リスクや経済力にかかわらず、完全無料で利用が可能。

9) 林かおり、「アイスランドの「保健医療分野データベース法」及び「バイオバンク法」、外国の立法、vol.218、92-135(2003)。

10) オミックス研究: 生体機能分子(遺伝子、タンパク質、代謝産物等)を網羅的に調べる研究。ここでは血液中のタンパク質の量や遺伝子多型などを多層的に組合せる多層オミックス技術により、生活習慣による疾病リスクへの影響を検討する。

ジェクト」において整備されたバイオバンク。基本的にはケース・コントロール研究としてデザインされているが、疾患重症化の過程を分子レベルで研究するために、参加者の血清および臨床情報が毎年継続的に収集されている<sup>11)</sup> ため、コホート研究の側面も持つと思われる。2011年1月現在、47疾患を対象に200,001人(320,219症例)から「試料」・「情報」が収集されている<sup>12)</sup>。「試料」は、東京大学医科学研究所に保管されており、外部研究者も利用が可能<sup>13)</sup>。「情報」に関しては、配布試料の対象疾患選択に必要な試料群分け情報のみが提供され、個別の「試料」ごとに付随する「情報」は提供されない。

#### ○内閣府ゲノムコホート研究(日本):

2010年7月に総合科学技術会議が策定した「平成23年度 科学・技術重要施策アクション・プラン」<sup>14)</sup>において、ライフイノベーション推進のための施策の1つに「ゲノムコホート研究と医療情報の統合による予防法の開発」<sup>15)</sup>が記載され、その中に10万人規模のバイオバンク整備計画が盛り込まれた。本施策の実質的な活動は、国立がん研究センターにより進められている。科学技術戦略推進費<sup>16)</sup>で実施される科学技術システム開発事業<sup>17)</sup>の2011年度プロジェクトとして、「大規模分子疫学コホート研究の推進と統合(国立がん研究センター)」が採択された。本プロジェクトは、国際的評価にも耐えうる質の高い大規模分子疫学コホート研究の実現を達成目標とし、既存コホート研究を統合するための標準化とノウハウ開発を進め、

最終的には数十万人規模(最低30万人)のバイオバンク構築を目指している。

#### ○東北メディカル・メガバンク(日本):

2011年3月11日に発生した東日本大震災からの復興を前提に計画された施策。被災地における医療の再生と医療機関の復興にあわせ、被災地を中心とした大規模ゲノムコホート研究を行うことにより、地域医療復興への貢献と次世代医療体制(予防医療、個別化医療など)の構築を目指している。被災地域である宮城県、岩手県の住民を対象に、約15万人から「試料」・「情報」が収集される予定。事業内容は、大規模ゲノムコホート研究の実施と医療情報ネットワークの構築を柱に、Biobank Japanや既存コホート研究との連携を視野にいれた活動が計画されている。保管された「試料」・「情報」は、必要な審査・手続きを経て希望する外部研究者にも提供される予定。実施期間は当面10年で、2012年度までに総額214億円の予算が投入されている(文部科学省:2011年度第三次補正予算158億円、2012年度予算56億円)。

#### バイオバンクの課題と展望

疫学研究では、解析結果の信頼性を向上させるため、より大きなサンプルサイズが要求される。地域住民を対象としたコホート研究の場合、参加者数の世界標準は10万人以上となっており、更に複数のコホート研究によるメタ解析あるいは総合解析が盛んに行われている<sup>18)</sup>。現在、世界に散在するバイオバンクの間で、「試料」・「情報」の共有

11) 田中敏博、「オーダーメイド医療実現化プロジェクト」、実験医学、vol.27、1836-1841(2009)。

12) 「オーダーメイド医療実現化プロジェクト」シンポジウム、バイオバンク・ジャパンの全貌—その可能性と未来—(2011年10月2日)パンフレットより([http://www.biobankjp.org/info/s111002/2011\\_Panf.pdf](http://www.biobankjp.org/info/s111002/2011_Panf.pdf))。

13) 試料提供の条件は、国内に法人格を持つ機関が国内で使用すること。試料はBiobank Japanでの審査を経た上で配布される。

14) 平成23年度 科学・技術重要施策アクション・プラン:2010年6月に閣議決定された新成長戦略の方針(グリーン/ライフ・イノベーションの推進)に従い策定された。

15) 大目標を「予防医学の推進による難治性疾患の罹患率の低下」とし、短期的に実現すべき目標として「大規模ゲノムコホート推進体制の完成とバンキング事業の実施」「横断的解析からの疾患関連マーカーの同定」「疾患にかかわる遺伝、環境因子の同定と相互作用の解明」としている。

16) 科学技術戦略推進費:総合科学技術会議が自ら策定した科学技術イノベーション政策を戦略的に推進するために2011年度に新設された予算。予算の事務的対応は文部科学省が担当する(「競争的資金以外の研究費」として計上されている)。

17) 科学技術システム開発事業:文部科学省の委託を受けて、科学技術振興機構(JST)が公募により採択する。

18) 吉田輝彦、「がん研究基盤としてのバイオバンクの整備の意義と展望」、血液内科、vol.63、140-150(2011)。

表2 バイオバンクの国際ネットワーク

機関・組織名	スコープ
International Society for Biological and Environmental Repositories	バイオバンキングにおける高度標準化とイノベーションの推進
OECD : Guidelines for Human Biobanks and Genetic Research Databases	指針・勧告の作成に関する国際政府フォーラム
Public Population Project in Genomics (P <sup>3</sup> G) Observatory	集団ゲノム科学における連携推進のための非営利団体
Biomolecular and Biobank Research Infrastructure	欧州及び世界における生物医学・生物学研究のための汎欧州レベル基盤
Marble Arch Working Group	バイオバンキングのハーモナイゼーションに向けた研究者主導の国際シンクタンク
Forum of International Biobanking Organizations	ヒトバイオバンクを扱う組織における世界規模での関係強化に向けたコミュニケーション・フォーラム

出所 : Hewitt, R et al., "Biobanking in a Fast Moving World : An International Perspective", J. Natl. Cancer Inst. Monogr., 42, 50-51 (2011) をもとに作成。

に向けた動きが活発化している。そのための国際ネットワークが既に複数設立されており(表2)、本稿で情報源としたP<sup>3</sup>Gはその1つである。これらの国際ネットワークでは、バイオバンクにおけるガバナンス・デザイン・サンプル処理・品質管理・品質保証等の標準化と、統一された規制・倫理要件などが検討され、連携促進に向けた活動が進められている。わが国においても、「試料」・「情報」のGLPに準拠した品質管理・品質保障など、国際連携を見据えた対応が望まれる。

疫学研究は科学技術の進展を背景に、新たな段階を迎えており、その対応も非常に重要である。次世代型コホート研究と称される「分子疫学コホート研究」は、オミックス技術を駆使して生体の状態を分子レベルで把握し、コホート研究に活

かそうとするものである。しかし、これを実現するためには、分析技術や情報処理・解析技術等の分野で飛躍的進歩が必要とされる。また、倫理的課題への対応、国民理解の醸成、安定資金の長期的確保など多くの課題があり、国家レベルの対応が望まれる。

バイオバンクは、医科学研究領域における基礎研究の成果を臨床応用に結び付けるトランスレーショナル・リサーチにおいて、極めて重要な研究資源である。しかし、バイオバンクの成果が実を結び、臨床応用されるまでには、まだまだ多くの労力と長い時間を必要とする。従って、バイオバンクは次の世代に残す貴重な財産と考えるべきであろう。様々な課題を適切に解決し、次世代にとって最大限有用な贈り物となることを願いたい。