

トーゴーの日シンポジウム2011
2011年10月5日 日本科学未来館

大規模ゲノム疫学研究の統合情報基盤の構築

京都大学医学研究科附属ゲノム医学センター

松 田 文 彦

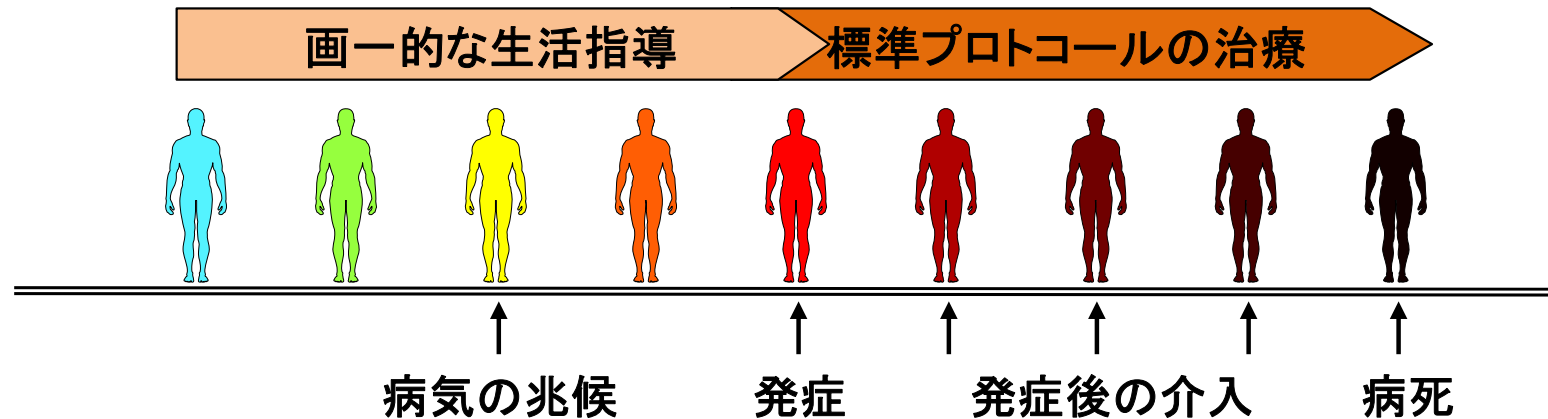
「治す」医療から「予防する」医療へ

～ 病気の予防は、21世紀の医療の中心的課題 ～

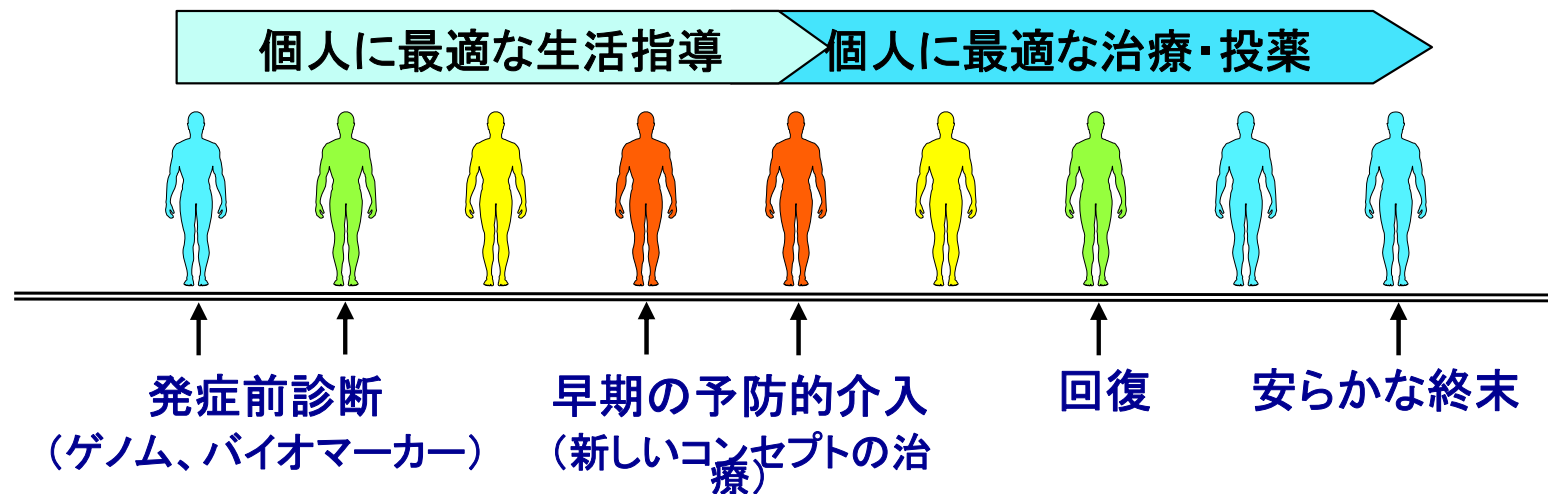
- ・ 病気を治す医療は、現代医学の発展で目覚ましく向上し、近い将来、多数の病気の完治が期待される
- ・ 長寿化により、認知症や脳血管障害などの疾患が増加
 - － 家族を含めた社会の負担増、寝たきり老人の増加による医療費の激増などの極めて深刻な社会問題を引き起こし、健康長寿を阻害する要因
- ・ 生活習慣病に代表される複合遺伝性疾患は、患者が多いうえ有病期間が長く、病態が多様で根治法がない
 - － 疾病の早期発見と予防が危機的な少子高齢化時代において健康で活力ある長寿社会を構築する唯一の手立てとなる

先制医療 (Preemptive Medicine) とは

従来の予防： 集団全体を対象としてリスク因子を解明し、それを避けることにより発症を防ぐことを目的する



先制医療： 遺伝情報を基盤とした「個の医療」であり、発症前に高い確度で診断し、介入して発症の防止・遅延を目指す



複合遺伝性疾患の遺伝要因を明らかにするのは簡単ではない

- ・ 病気の本態が極めて複雑で、患者／健常者の定義が困難
- ・ 病気の発症には、遺伝、環境、生活、地域特性等が関与するが、これらを総合的に扱う遺伝解析はまだ始まったばかり
- ・ 多数の変量を扱う、新たな統計遺伝学的手法が必要

大集団の長期的観察に基づく疫学研究とゲノム・オミックスを融合した学際的研究が世界各国で開始されている

ゲノムコホート研究とは

- ・ 従来の疫学研究では、研究対象が比較的マクロな因子に限られ、遺伝因子や分子レベルのバイオマーカーを加味した研究は技術的制約から困難であった
- ・ 疾病リスク、薬剤応答性などは人種によって大きく異なるため、日本人を対象とした疫学研究が不可欠
- ・ ゲノム・バイオマーカー・疫学・臨床情報を融合し、長期的な集団の観察と膨大な情報の精緻かつ効率的な分析・解析による新世代の予防医学研究
- ・ 分析・解析技術に加え、大規模データ利用の情報基盤やバイオインフォマティクスを用いた解析基盤の整備が必須

疫学という医学研究の一分野のみで担える研究活動ではなく、多分野の協力と連携で実現する学際的研究

研究開発の目標・ねらい

- **ゲノム疫学研究の情報基盤の構築と公開**

「ながはま0次コホート研究」の一万人の生活習慣・環境情報、臨床情報、ゲノム・オミックス情報を標準化し、データベースを構築する。

集積した情報を、個人情報保護のもと、医学・生命科学研究者に提供する。

- **データベースの枠組みの提供と情報の連結**

これをモデルケースとして、同様の研究をおこなう際に即時活用可能なかたちで、分子疫学研究者にデータベースの枠組みを提供する。

他の研究で蓄積された遺伝型・表現型データを連結、共有することで、個別の研究で得られた情報の一元化によるそれらの再利用を促す。

- **ゲノム情報科学の若手研究者の育成**

バイオインフォマティクス、遺伝統計学の若手研究者に教育訓練(OJT)をおこない、これらの分野の将来における中心的研究者の育成をはかる。

**「パーソナルヘルスレコード」の情報提供先として機能できる
汎用性の高い健康情報管理システムを提案**

滋賀県

長浜市



湖北地域の地方中堅都市
人口 約120,000人

- 65歳以上老年人口割合は18~20%
- 住民の移動が少ない
- 健康推進員*というボランティアが市内各地域で活躍(合併後555人)

*食生活推進員、運動推進員、母子保健推進員を兼ねる。県が制度を作り、市が養成。

基幹病院

市立長浜病院 (676床)

長浜赤十字病院 (549床)

長浜市立湖北病院 (153

床)

長浜市を中心に20万人以上の診療圏

なごはま0次予防コホート

～ 地域に根ざした未来型健康づくりの試み ～

コホート事業

- 長浜市民一万人の参加のもと10年以上の計画で事業を推進
- 5年ごとの基本健康調査による、臨床マーカーの変化や発症前の病態の把握
- 医療情報や人口動態情報を用いた、有症候性の疾患情報の取得

分析・研究

- ゲノム・メタボロームの解析と、成果の先制医療への応用
- 新たな分析・解析技術の積極的な導入

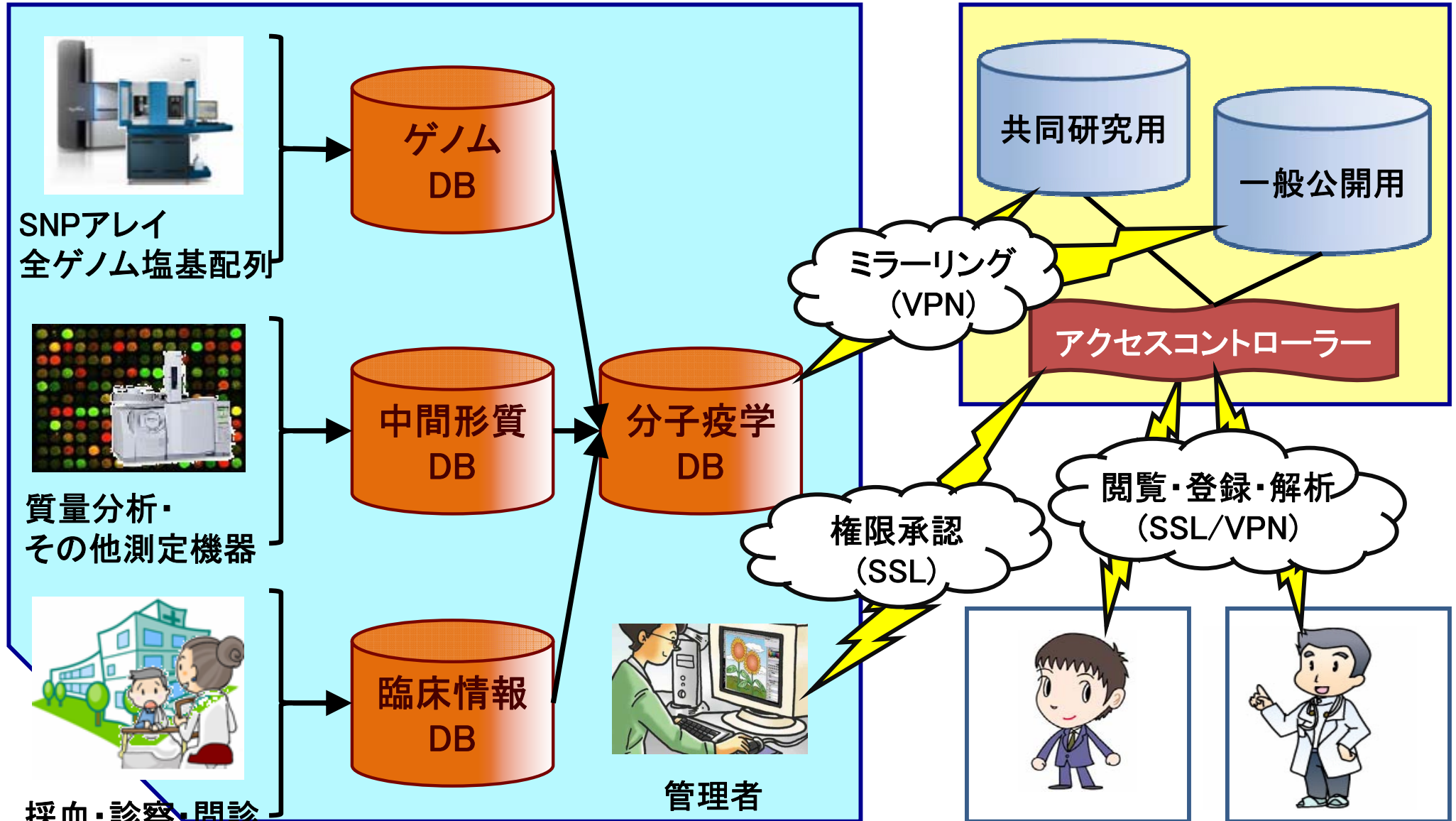
研究成果の公開と利活用

- NBDCを通して一定の条件下で国内外に公開し、研究・開発に利活用

データベースシステム概念図


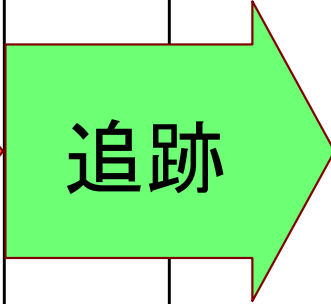

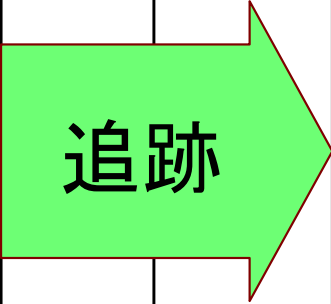
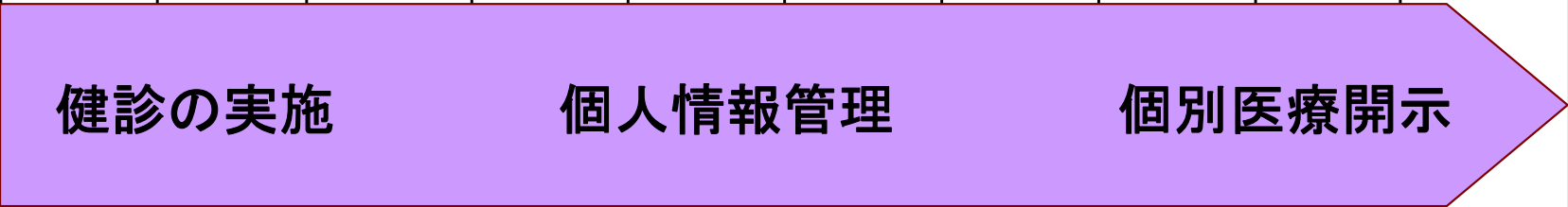

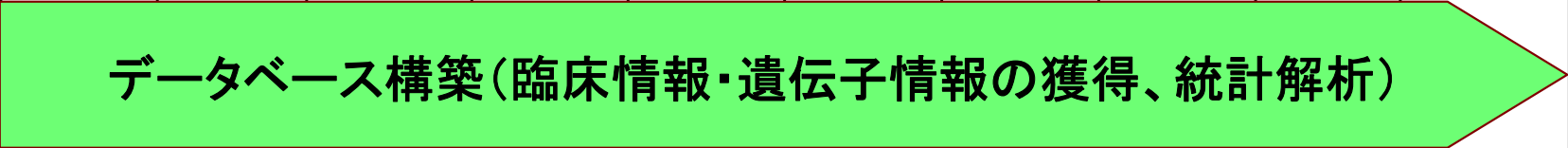

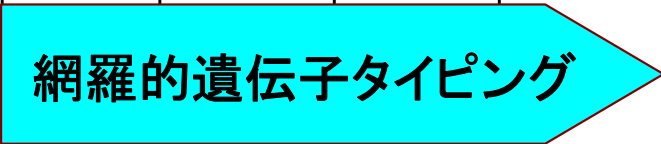
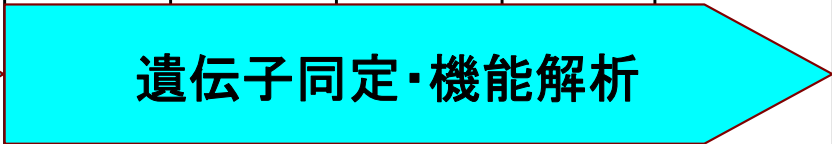

京 都 大 学

NBDC



データベース利用者

ながはま0次コホート事業のロードマップ

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017			
京都大学 <ul style="list-style-type: none"> • 健診、質問表調査による調査 • 特定健診項目の実施 	 <p>初回調査 個人の健康情報の入手</p>			 <p>追跡</p>		 <p>再調査 各疾患の罹患率の把握</p>			 <p>追跡</p>				
長浜市 <ul style="list-style-type: none"> • 医師の採用 • プロジェクト参加者の勧誘、事業の説明 • DNA検体採取 	 <p>健診の実施 個人情報管理 個別医療開示</p>												
ゲノム疫学研究 <ul style="list-style-type: none"> • 検体の収集・保管 自動ストッカーによる一元管理 • データベース 臨床情報、遺伝子情報の統合 • 遺伝解析 全ゲノムスキャンによる網羅的解析 	 <p>検体収集・保管</p>			 <p>データベース構築 (臨床情報・遺伝子情報の獲得、統計解析)</p>							 <p>検体収集・保管</p>		
	 <p>網羅的遺伝子タイピング</p>					 <p>遺伝子同定・機能解析</p>							
	 <p>全ゲノムシーケンシング・オミックス解析</p>												

健診における測定・調査項目

- 1 自記式質問票による健康情報の取得(742項目)
- 2 身体計測と生理学的検査
- 3 血液・尿検査(40mlの採血、血清・血漿・DNAを保存)

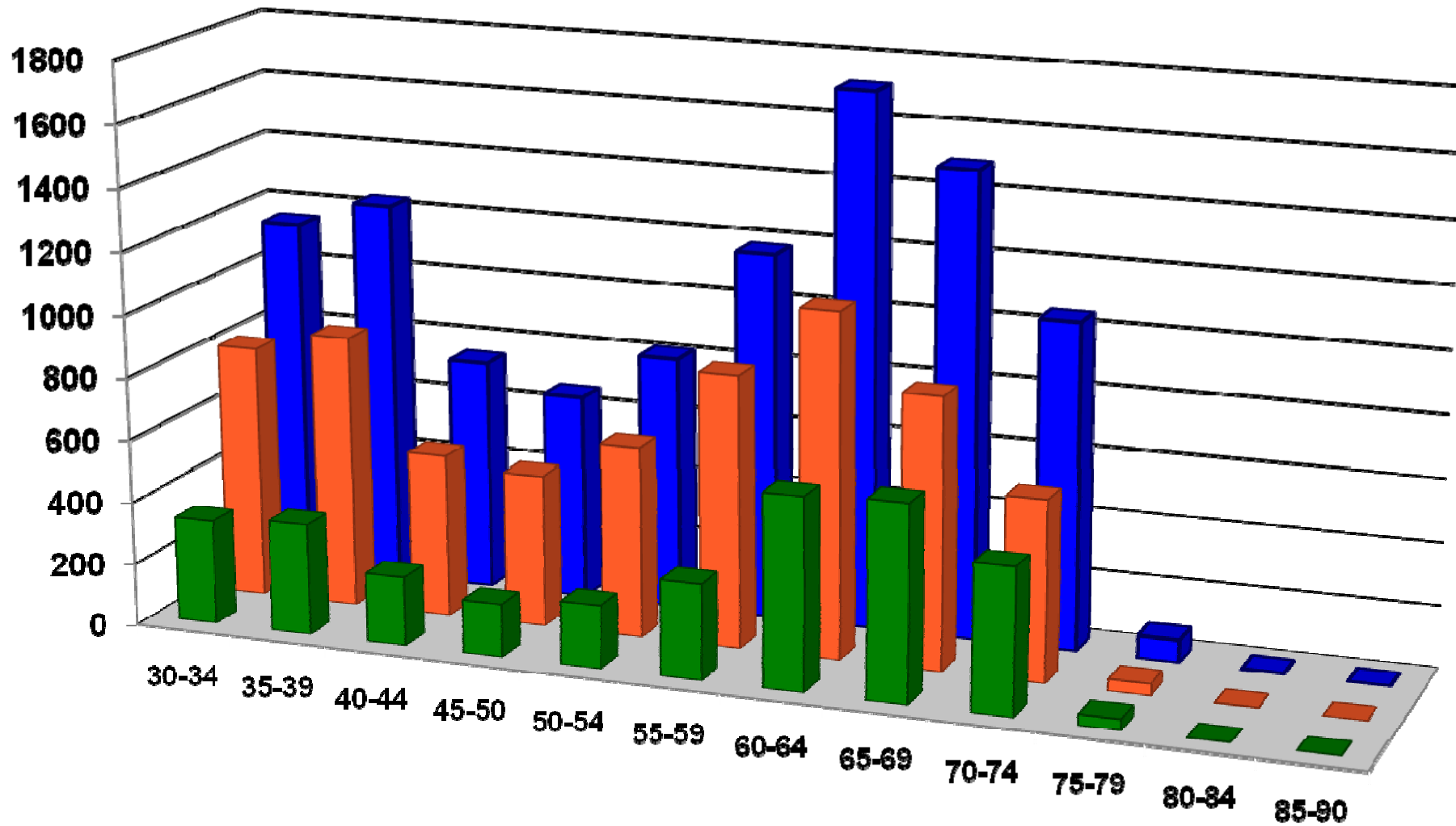
所要時間 約90分

いずれも、国保レベルの集団検診より、はるかに詳細な質問項目、測定、検査を施行

長期保存の血清・血漿では測定困難な項目を優先して測定。

検査項目は今後の研究状況に応じて追加、変更される。

参加者数と年齢分布 (2010年度健診終了時)



■ 男性 3,333人

■ 女性 6,747人

■ 全数 10,080人

ベースライン情報

1. 環境・生活習慣情報（質問票による742項目）
2. 健診による臨床情報（生化学、生理学、血液学など）
3. ゲノム情報
ゲノムスキャン（5,000検体 2011年度）
Exomeシーケンシング（500検体 2011年度）
4. オミックス情報
メタボローム解析（10,000検体 2012年度）
トランスクリプトーム（500検体 2012年度）

身体計測・生理学検査

身長	特定健診
体重	特定健診
腹囲	特定健診
BMI	特定健診
中心血圧	健康情報学
心電図	健康情報学
眼底検査	糖尿病内科
眼軸長測定	眼科
胸部X線(直接撮影)	呼吸器内科
CAVI	健康情報学
呼吸機能(スパイロ)	呼吸器内科
内臓脂肪	内分泌内科

血液検査（抜粋）

非特異IgE	呼吸器内科
IgEイネ	呼吸器内科
IgEカビ	呼吸器内科
IgE雑草	呼吸器内科
IgEハウスダスト	呼吸器内科
IgEダニ	呼吸器内科
IgEネコのフケ	呼吸器内科
IgEイヌのフケ	呼吸器内科
IgEトリコフイトン	呼吸器内科
百日咳抗体	呼吸器内科
クラミジアIgG	呼吸器内科
クラミジアIgA	呼吸器内科

網羅的ゲノム多型解析の目的

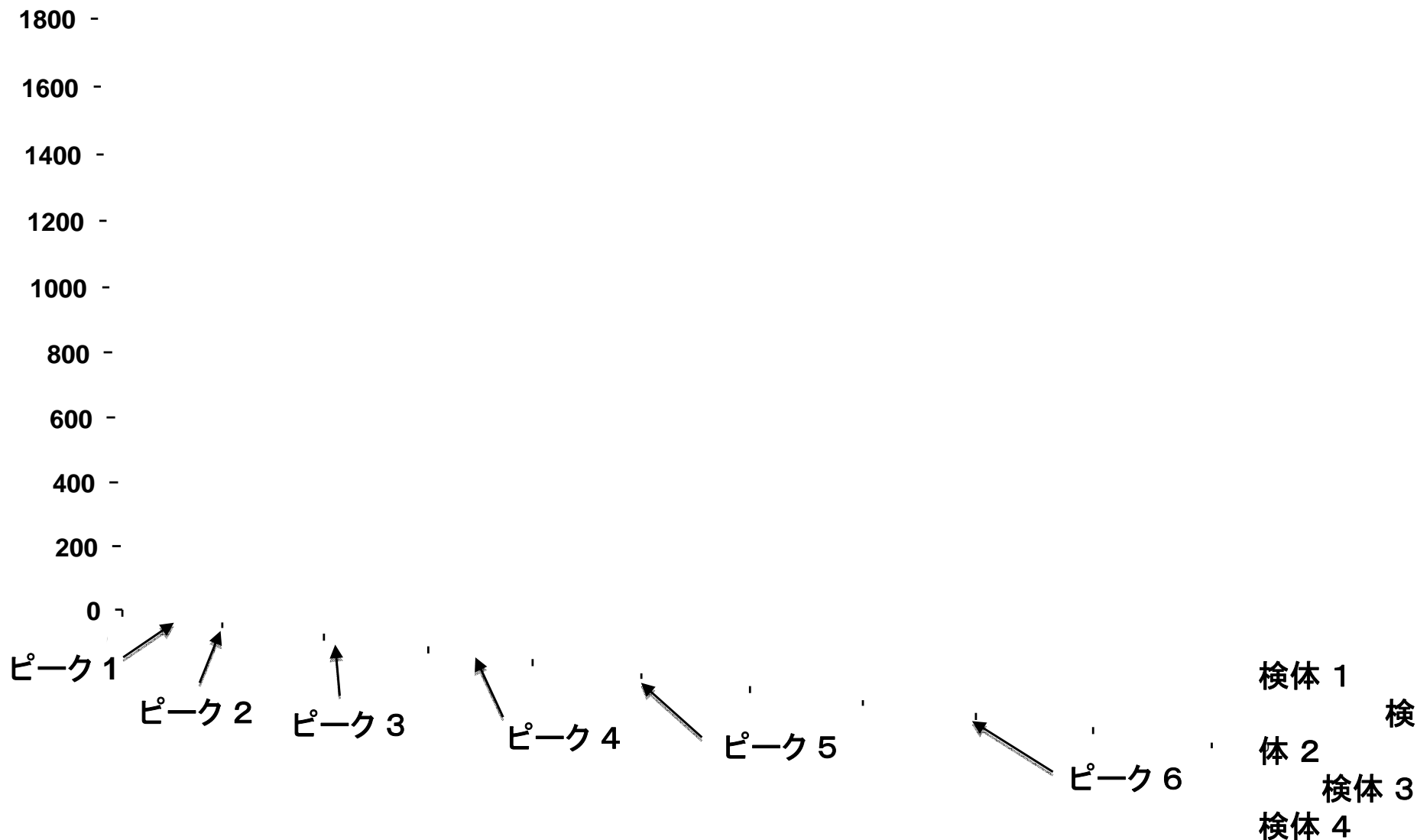
- 我が国の網羅的ゲノム解析に良質の対照群の多型情報を提供
- 関連解析で見いだされた疾患感受性遺伝子の前向き
の証明
- 種々のバイオマーカー、生理学的検査値などとゲノム
多型の量的関連(QTL解析)
- 網羅的プロテオーム、ペプチドーム、メタボロームとゲ
ノム多型の量的関連(QTL解析)

こういった解析を通して、新たな疾患関連遺伝子、
バイオマーカーを探索

網羅的ゲノム多型解析のストラテジー

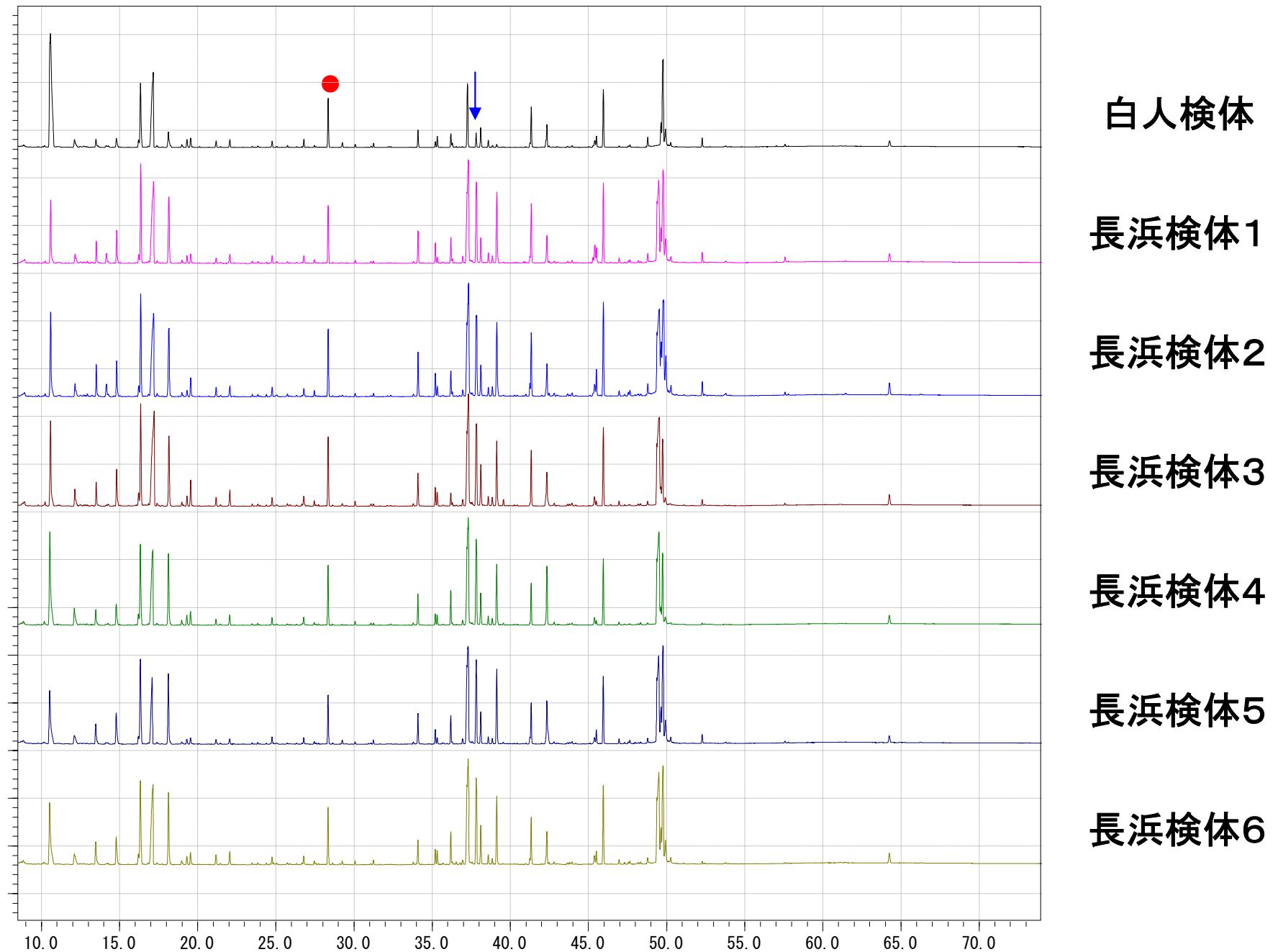
1. 10,000人のDNA検体のうち、半数の5,000検体のゲノムスキャン（Illumina社 HumanHap610K）
2. 種々の量的形質との関連解析
3. 関連が見いだされた遺伝子多型を、残り半数のDNA検体を用いて再現性検証
4. 量的形質の網羅的解析と遺伝子多型の関連解析とカタログ化
5. 候補領域の塩基配列決定による、疾患と真に機能的関連を持つ遺伝子／多型の同定

網羅的メタボローム解析による疾患関連マーカー探索



疾患に関連する特定の代謝物とゲノム多型の関連が
明らかになれば、新たな疾患関連遺伝子／マーカーの同定が可能に

GC-MSを用いた血漿の代謝物解析



● = 内部標準 (2-isopropylmalic acid)

追跡による参加者の疾患罹患情報

1. 質問票による調査

自己申告情報の確度に問題

2. 特定健診の情報

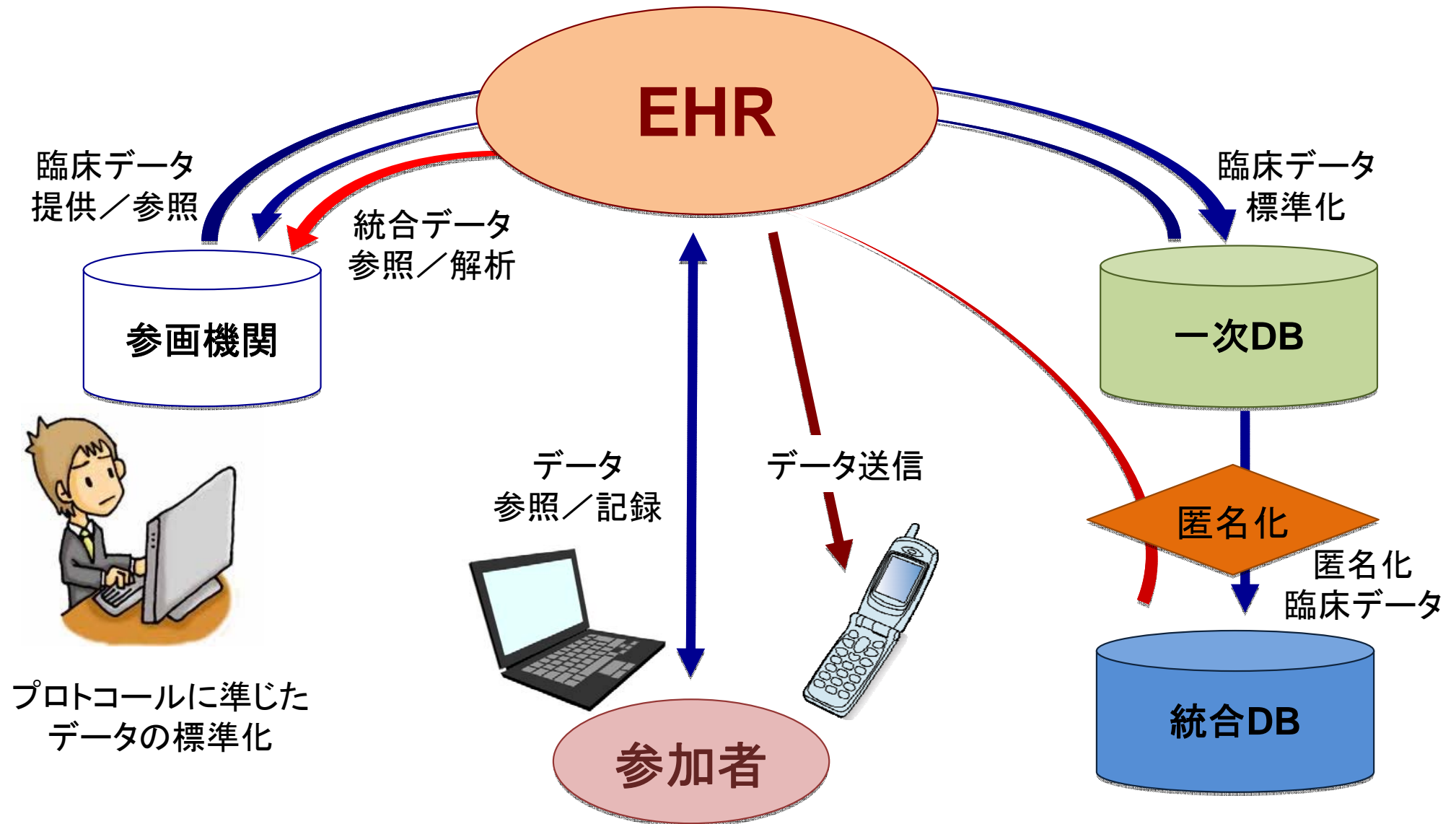
受診率の低さに問題

3. 地域の疾病登録制度の利用（がん登録など）

4. 移動・死亡情報（住民基本台帳、死亡小票の閲覧）

5. EHRを利用した疾患関連情報の取得

大規模コホートとEHRの統合による 統合疫学データベース構築



アクセスコントロールレベルの設定

- 利用者に対し、検体群ごと(疾患1、疾患2…、対照群)に権限を付与
- それに基づく利用者の閲覧・利用の制限(Webインタフェースを変更)

高 ↑ 限権利用の能機 ↓ 低	レベル5.	個人情報に伴うデータの継続的登録
	レベル4.	個人に属する個々の情報の閲覧と利用
	レベル3.	個人に属する情報を用いた解析
	レベル2.	自己所有データの一時アップロード それらと集団の集約情報とによる解析
	レベル1.	集団の集約情報の閲覧と利用



一般ユーザー
(レベル1~2)

集団の集約情報の閲覧と、
それに基づく解析

SSLで一般公開用DBにアクセス



共同研究者
(レベル3~5)

利用権限を持つ検体群の情報による、高度で緻密な解析

VPNによる高セキュリティ通信



管理者

権限申請の承認