

2030年の挑戦

Human-centric デジタルツインが
目指す未来社会

日時 : 2024年5月15日 (水)
14:00~17:00
会場 : オンライン (Zoomウェビナー)
参加費 : 無料
主催 : 国立研究開発法人科学技術振興機構



前田 英作氏
総司会



村上 善則氏
本格研究講演

共同研究者 (敬称略)

鎌谷洋一郎、鈴木亨
津田宏治、中谷明弘 (以上、東京大学)
三宅美博 (東京工業大学)
佐藤健吾 (東京電機大学)
井上真奈美 (国立がんセンター)
瀬山倫子 (NTT ライフサイエンス)
笠原要、塩見なぎさ (NTT物性研)

多層的生体情報の統合による 疾患予防デジタルツインの構築

日本医科大学 先端医学研究所
村上 善則

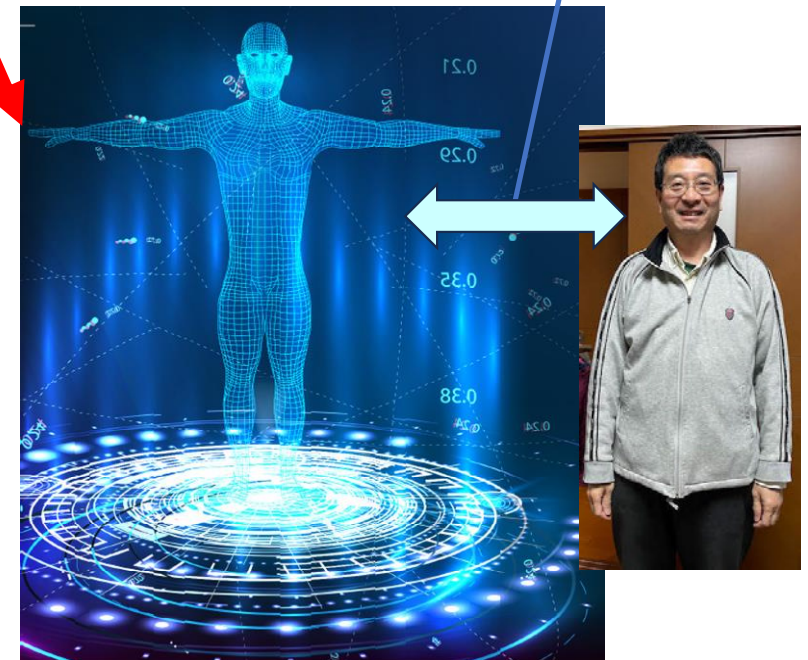
1. 背景

保健・医療領域におけるデジタル・ツイン(DT)とは

1. 組織の構築・運営
→ 病院経営の合理化
2. バーチャル・リアリティー
→ 患者の臓器、組織を画像で再構築し、手術訓練
3. 各臓器、組織レベルでのコピーの複製
→ iPS細胞からの細胞、臓器再生、
薬剤感受性などの検討
4. 個々人の生体、臨床情報の統合によるDT
→ アバターに条件を付与して、
将来の健康をシミュレーション
→ 疾患予防、診断、治療

仮想空間、
仮想時空間で
別人生を体験

生体情報の
統合



(NTT基礎研究所HPより)

1. 背景

疾患は予防することが第一

未来社会：人生 100年時代の到来、だが、健康維持が必須
疾患予測、予防は個々人の未来の最大関心事の一つ

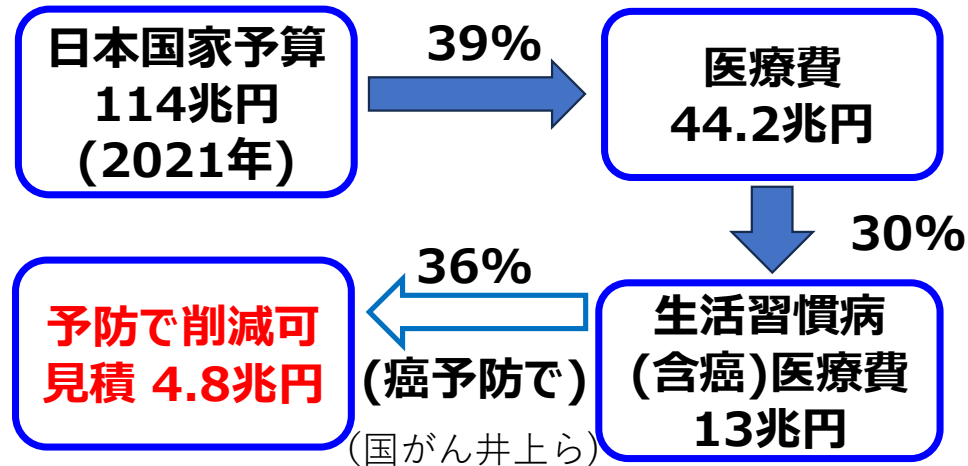
↔ 成人の65% が生活習慣病に罹患、**早期**からの疾患『**予防**』が必須

『**予防**』は最大の医療

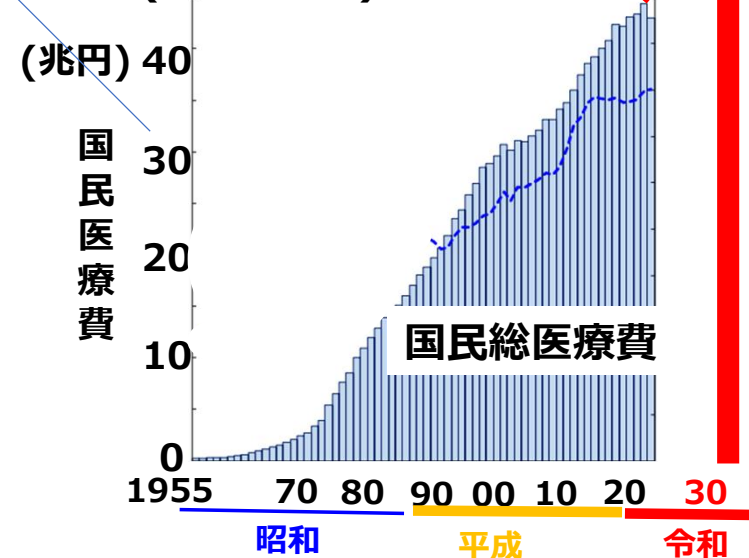
成功例：歯科領域の虫歯予防
「仕上げはお母さん♪」
昭和後期以降、関連医療費はほぼ横ばい

日本独自の**健康診断**のデータとゲノム情報、生体情報を組み合わせ、
最先端の情報解析技術を活用して個別化疾患予防を実現する

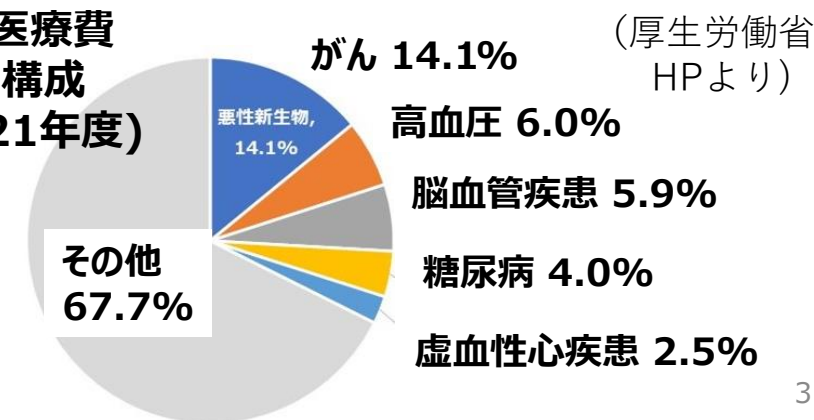
疾患予防のインパクト



国民医療費の年次推移
(2021年度)



一般医療費
の構成
(2021年度)

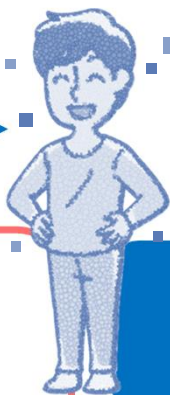


疾患リスク予測の基盤となる健診コホート、データベース

病院受診
or
経過観察



デジタルツイン構築



仮想空間



個人データ

あくまでも
個人情報



ビッグデータ

本格的高度AI解析
(未開拓領域)

疾患リスク予測
シミュレーション

行動変容
精密検査

疾患の予防

健診結果

個人経年データ

データベース

500項目×20万人×22年間

観測

(記録・収集)

収集

(自律学習)

受診者個人

勤続〇年

健診コホート

20万人の集団

匿名化

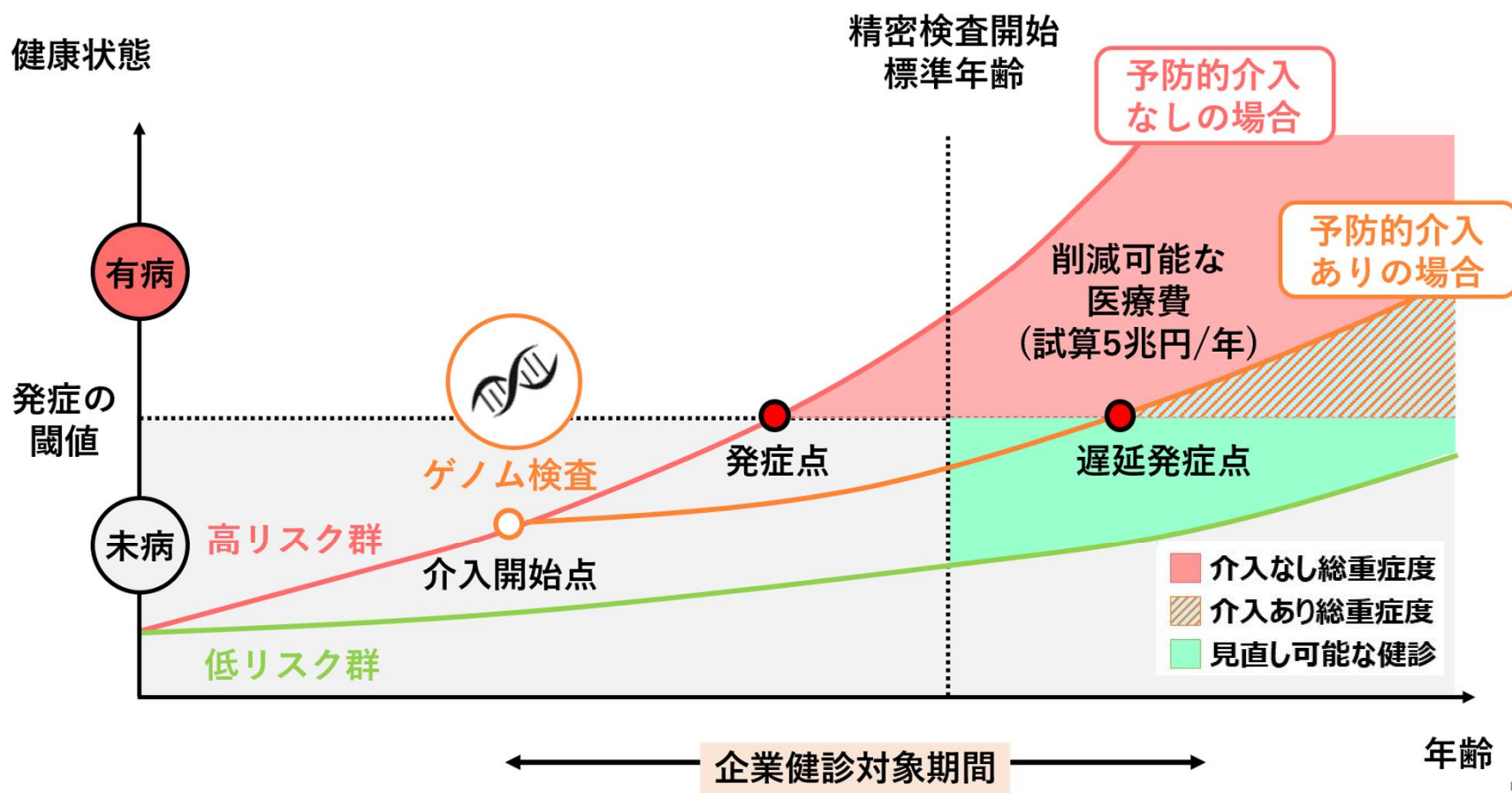
匿名化



健診・ゲノム情報による個別化疾患リスク予測の全体像

- ・情報科学、ゲノム・オミクス科学の爆発的发展
- ・生体情報、特にゲノム情報による疾患リスク予測が可能となりつつある

コホート研究：
ある要因の曝露・非
曝露集団を追跡し、
疾患発症率を比較し
て、要因と疾患発生
の関連を検討する



日本独自の企業健診を基盤とする健常人コホート

次世代健常人企業コホートの特徴

- ・企業健診と連動 – 離脱者が少ない、参加者が自動的に増加する
- ・高い参加同意率 – 90%[<]
- ・正確で精緻な健診情報、付随情報
- ・長期の経時的情報の追跡が可能 – 最長34年、平均16年
- ・参加者との双方向性情報交換が随時可能
- ・受益者負担の運営

2. コホート、データ生成

予防は最大の治療: 網羅的な疾患予防が重要

バイオバンク・ジャパン(BBJ)
第1コホート20万人登録時
BMIの疾患ごとの割合
(平成18年国民健康利用調査
との比較、年齢調整後)

肥満者が多い疾患
だけでなく、
やせた人が多い
疾患もある

■ BMI ≥ 25.0 ■ 18.5 ≤ BMI < 25.0
■ BMI < 18.5 ■ BMI 不明

男性上位5疾患

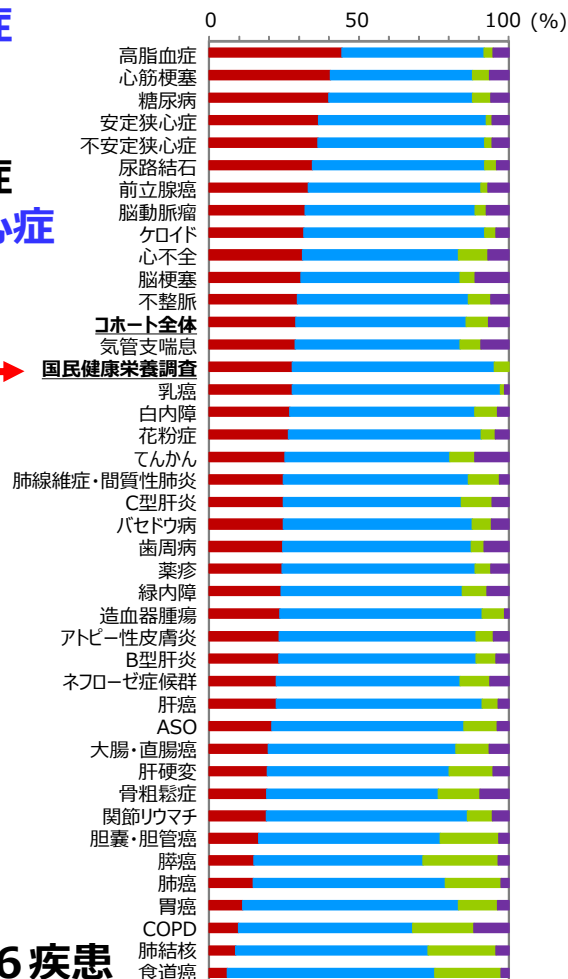
- ・脂質異常症
- ・心筋梗塞
- ・糖尿病
- ・安定狭心症
- ・不安定狭心症

全国平均 →

- ・膀胱がん
- ・肺がん
- ・胃がん
- ・COPD
- ・肺結核
- ・食道がん

男性下位6疾患

男性



女性上位5疾患

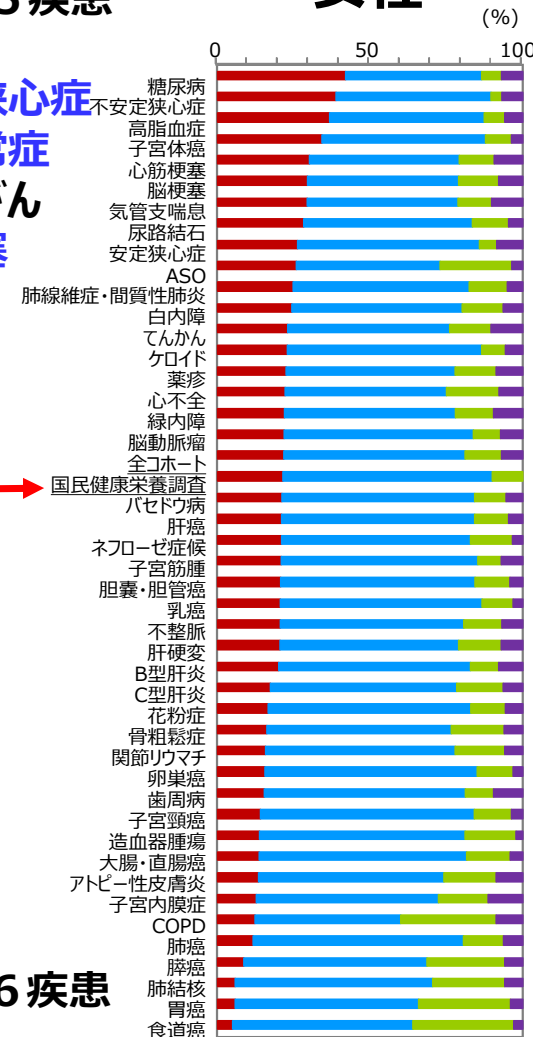
- ・糖尿病
- ・不安定狭心症
- ・脂質異常症
- ・子宮体がん
- ・心筋梗塞

全国平均 →

- ・COPD
- ・肺がん
- ・膀胱がん
- ・肺結核
- ・胃がん
- ・食道がん

女性下位6疾患

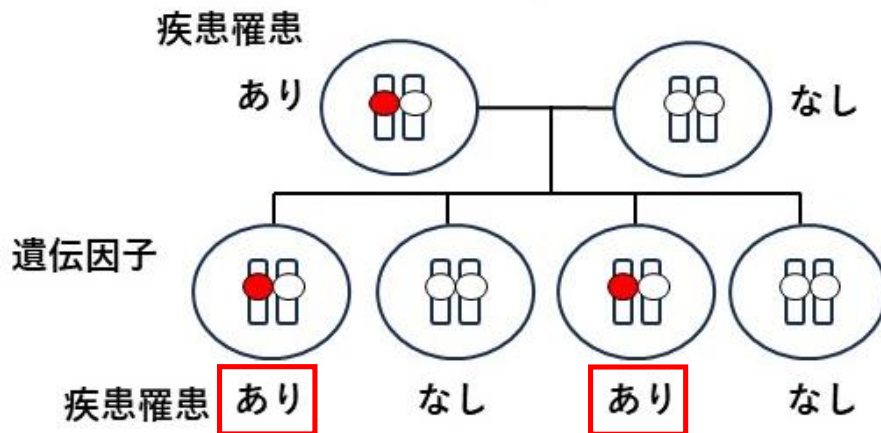
女性



多因子疾患の成因となる遺伝因子の解析

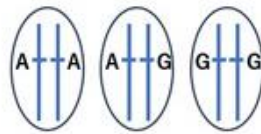
多因子疾患の成因：遺伝因子 x 環境因子 x 加齢
(多数) (多数)

単一遺伝子疾患

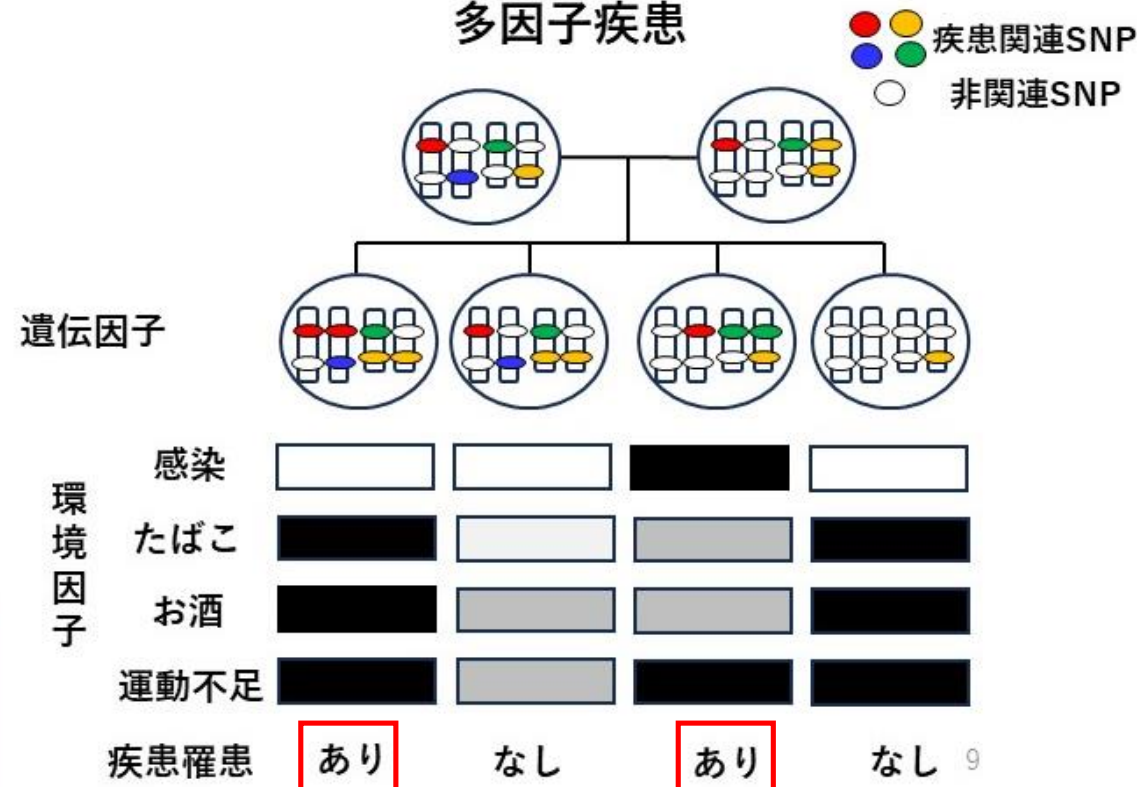


● 変異遺伝子
○ 野生型遺伝子

SNP(1塩基多型)：
ヒトゲノムの約0.1%を占め
個人の体質の原因となる



多因子疾患



ポリジェニック・リスクスコア(PRS)による多因子疾患のリスク予測

PRS (ポリジェニック・リスクスコア) :

PRS の推定

$$\sum_i^{risk\ SNP} \beta_i x_i$$

各多型のリスク $\times$ 各多型のジェノタイプ

オッズ比
5倍<の例
も見られる

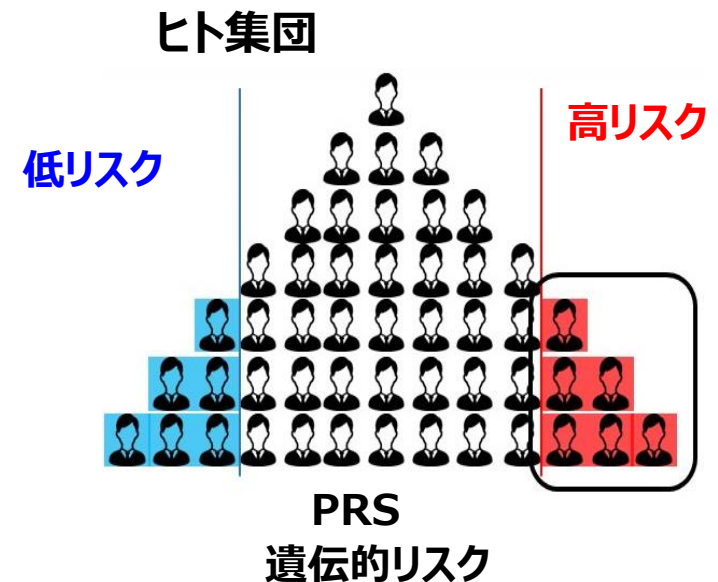
定義 :

数十～数十万の疾患関連遺伝子多型の重みをつけた総和。

特徴 :

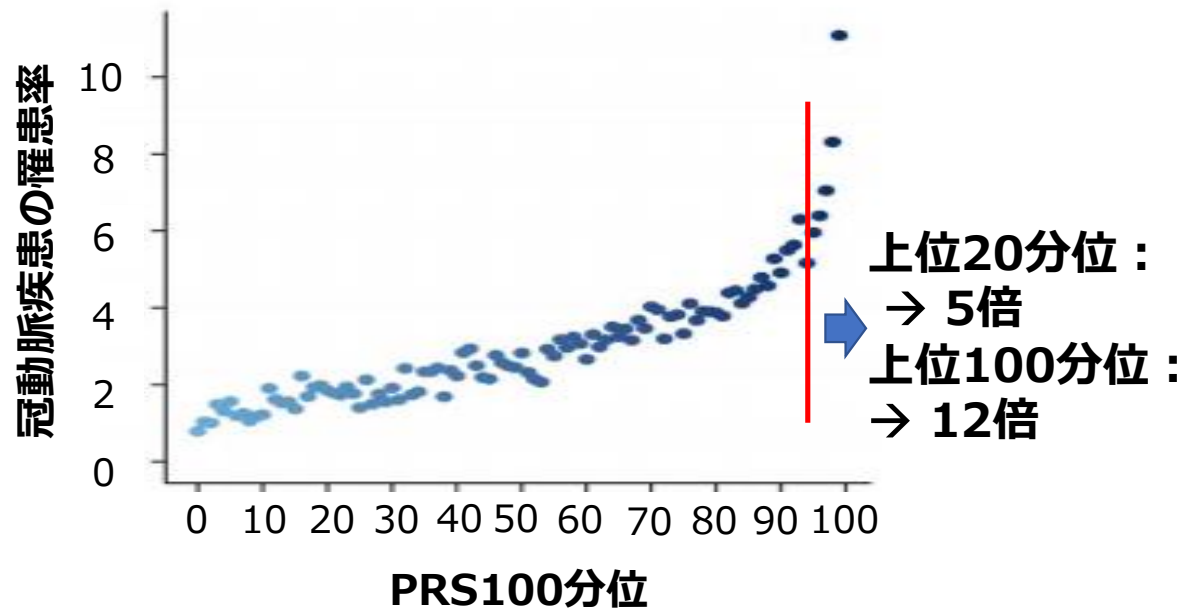
- ・各人が各疾患に対し1つのPRS。
- ・集団内の疾患の高リスク群の抽出に有効。
- ・PRS法により、あらゆる疾患のリスクを評価できる。
- ・計算手法に諸法あり、現在も改良が進められている。

PRSは高リスク群の同定に有効



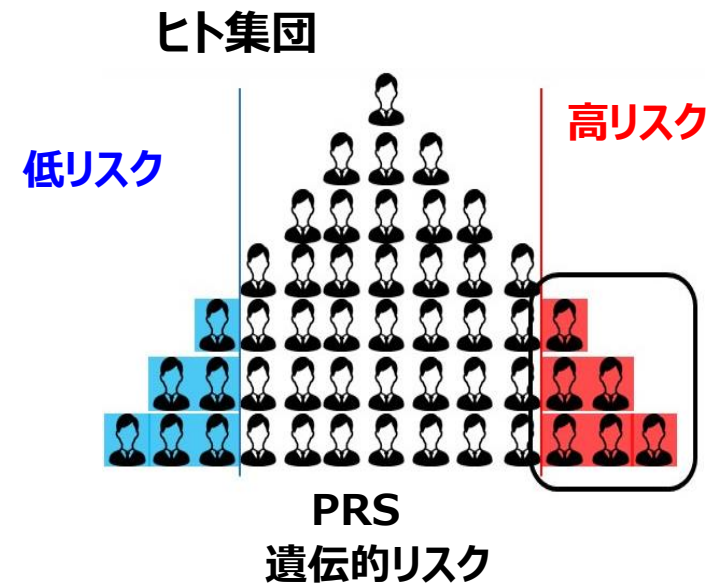
ポリジェニック・リスクスコア(PRS)による多因子疾患のリスク予測

冠動脈疾患のPRS 100分位ごとの罹患率



(Khera A et al, Nat Genet 2018)

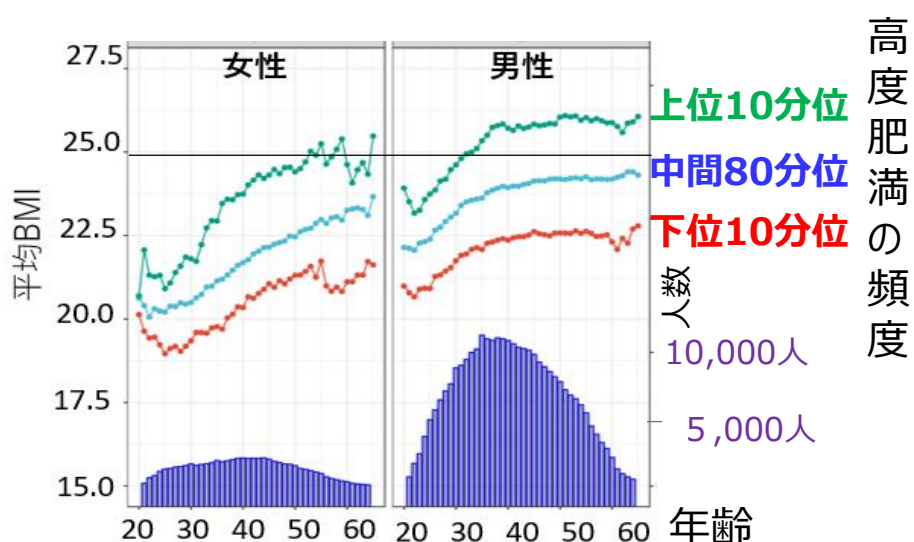
PRSは高リスク群の同定に有効



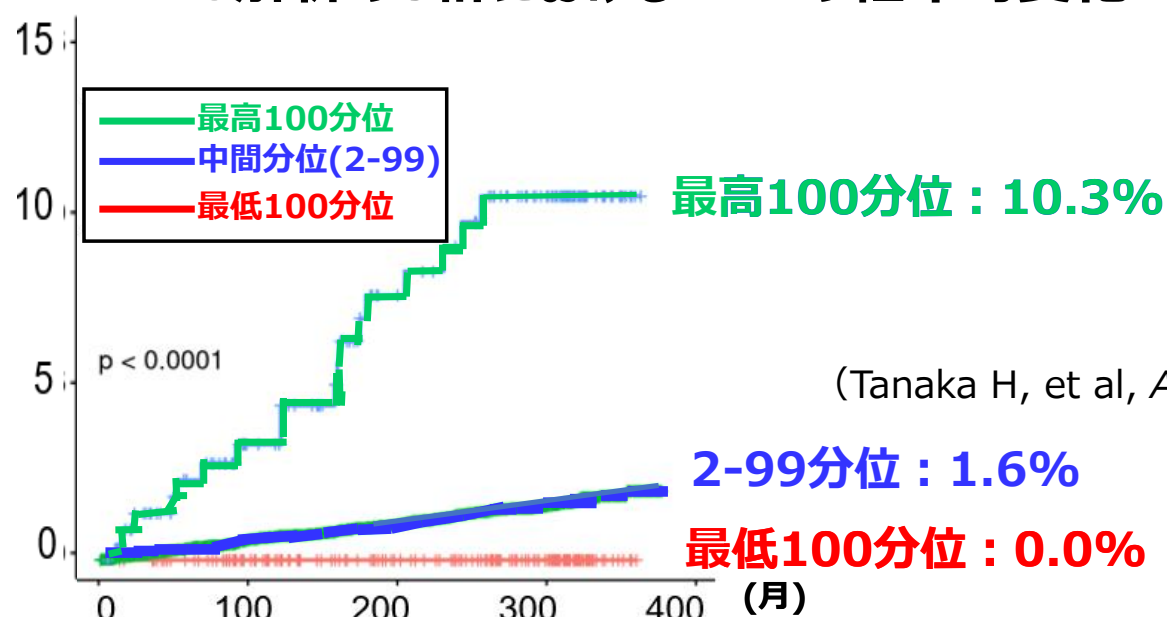
2. PRS, データ解析

PRS による肥満 (BMI) の高リスク群の抽出 (企業健診集団)

PRS解析による3群の BMIと年齢の分布



PRS解析の3群におけるBMIの経年的変化



(Tanaka H, et al, ASHG2023)

(Tanaka H, et al, ASHG2023)

(東大鎌谷ら)

BMI PRSにより、高度肥満者(BMI 35<)を予測できた



- ・行動変容の推奨 (一次予防)
- ・健診内容の層別化 (二次予防)

38疾患のどれかで

- ・上位10%に入る人：約 98%
- ・上位 1%に入る人：約 32%

脂質異常症、2型糖尿病、心房細動、消化性潰瘍などでも、高リスク群を同定

エピゲノム(DNAメチル化)の大規模網羅的解析による疾患リスク予測

エピゲノム (DNAの次に安定な指標)

- ・遺伝子発現を制御
- ・ゲノムより不安定
- ・他の指標より安定
- ・生活習慣を反映

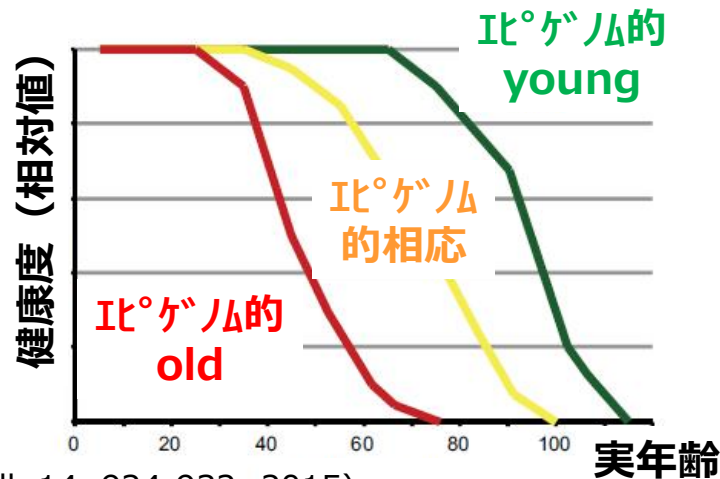


疾患予防の指標に適する
老化時計としての指標

- ・加齢
- ・慢性炎症/慢性感染
(肝炎ウイルス、H.ピロリ
潰瘍性大腸炎など)
- ・喫煙

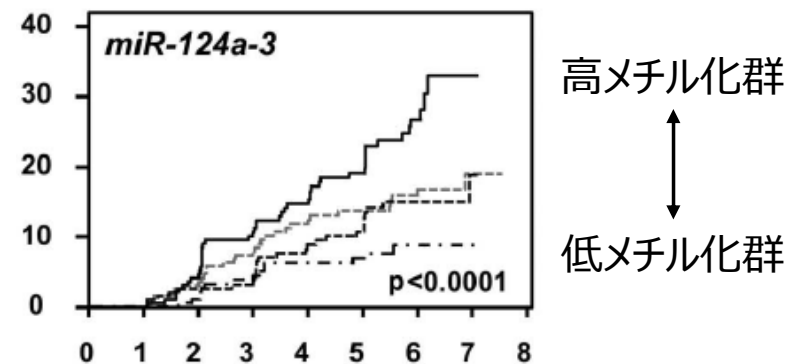
- ・網羅的解析技術 (メチル化アレイ) の進歩
- ・現在、100 例程度の小児、老人疾患解析が主流

エピゲノム状態が加齢を反映する可能性



(Aging Cell, 14, 924-932, 2015)

H.ピロリ感染胃の2次がん発生と 胃粘膜エピゲノム(メチル化)異常



(Gut, 66, 1721-2723, 2015)

3. 情報解析

健診データを用いた疾患リスク予測アルゴリズムの構築

個々人の時系列の健診データから、
・疾患の発症リスク
・N年後の発症確率
を予測するアルゴリズムを構築
(機械学習、多変量解析等)

+ PRS

アルゴリズムの予測精度のさらなる向上

脂質異常症では
予測精度は
0.86-0.88 (NTT 笠原ら)

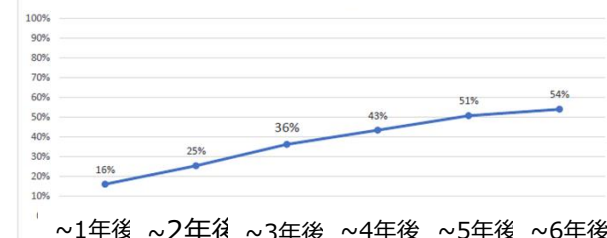
- ・行動変容を促す
- ・健診の層別化、精密検査を促す

本研究では38疾患に展開

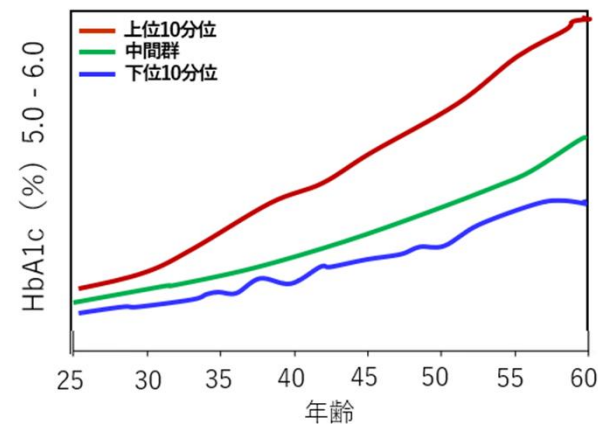
糖尿病リスクの 解析 (例)

受診者ID:	
検査データ取得日	
身体計測	検査時年齢
	性別
	身長
	体重
	BMI
血圧	腹囲
	血圧1回収縮期
	血圧1回拡張期
尿	尿糖定性
	尿蛋白定性
血液	赤血球
	血色素量
	ヘマトクリット
	肝機能等
肝機能等	GOT(AST)
	GPT(ALT)
	γ GTP
脂質	HDLコレステロール
	LDLコレステロール
糖代謝	中性脂肪
	空腹時血糖
服薬	HbA1c(NGSP)
	服薬(血圧)
	服薬(血糖)
	服薬(脂質)
	既往歴(脳血管疾患)
既往歴	既往歴(心血管疾患)
	既往歴(腎不全・人工透析)
	既往歴(貧血)

糖尿病発症確率 (例)



2型糖尿病のPRS 3群における
HbA1cの経時的推移 (模式図)



PRS

予測精度の検討

3. 情報解析

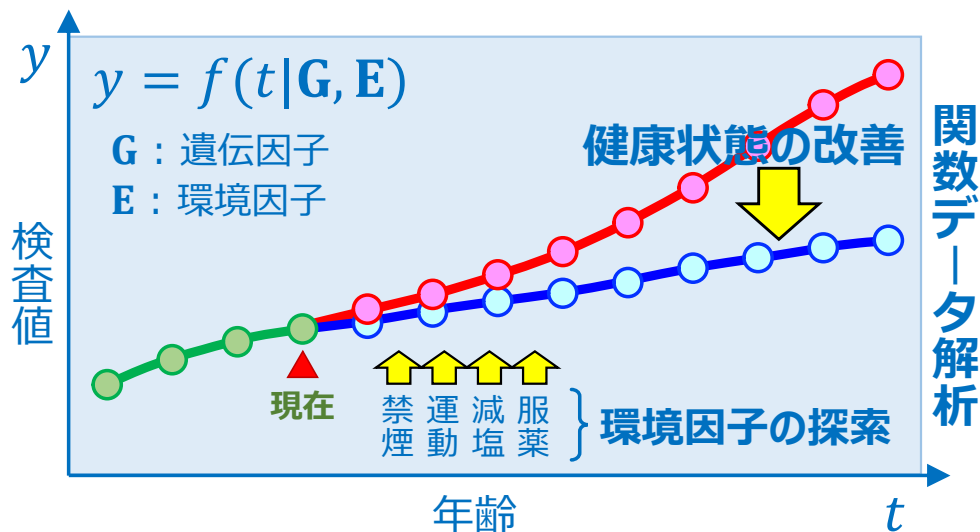
健診・ゲノム・オミクス統合データを用いた、 将来の疾患リスク、健康状態の予測モデルの構築

将来の健康状態のシミュレーション技術の構築

- ・遺伝因子Gと環境因子Eで検査値の経時的変化をモデル化
- ・将来の健康状態を制御する環境因子を実データ内で探索

大規模言語モデル
高次相互作用モデル
解釈可能AIモデル

検査値 y の経時的変化を年齢 t の関数 f で表現



(数値)データの解釈と注釈づけ
行動変容のための知識の言語化

38疾患の解析 → より単純な指標へ

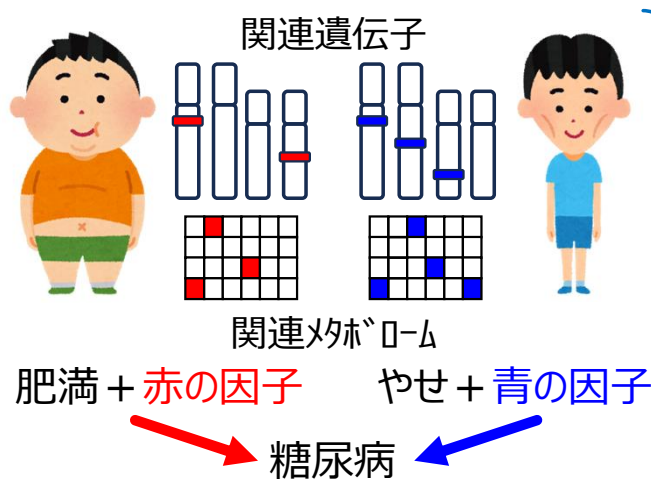
(東大中谷、東京電機大佐藤ら)



解釈可能な機械学習法を用いた高次相互作用モデルの構築

疾患リスク因子の組み合わせの探索手法と可視化手法の開発

同じ糖尿病でも肥満とやせではリスク因子が異なる



因子間の高次相互作用

組み合わせ最適化処理

分類条件 + 因子

階層的クラスタリング

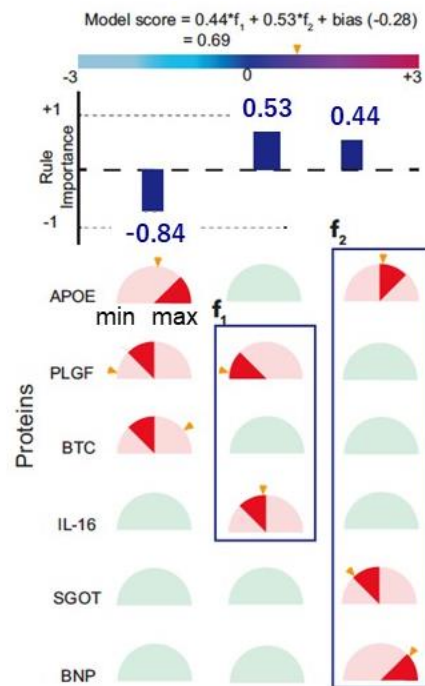
検体群のサブタイプ

データ内を探索

計算量が爆発

効率的な手法の開発
(背景知識も活用した計算方法)

因子の組み合わせと寄与率の可視化



因子間の高次相互作用（複数因子の組み合わせ効果）
と疾患リスクとの相関性の評価

4.
予防・診断

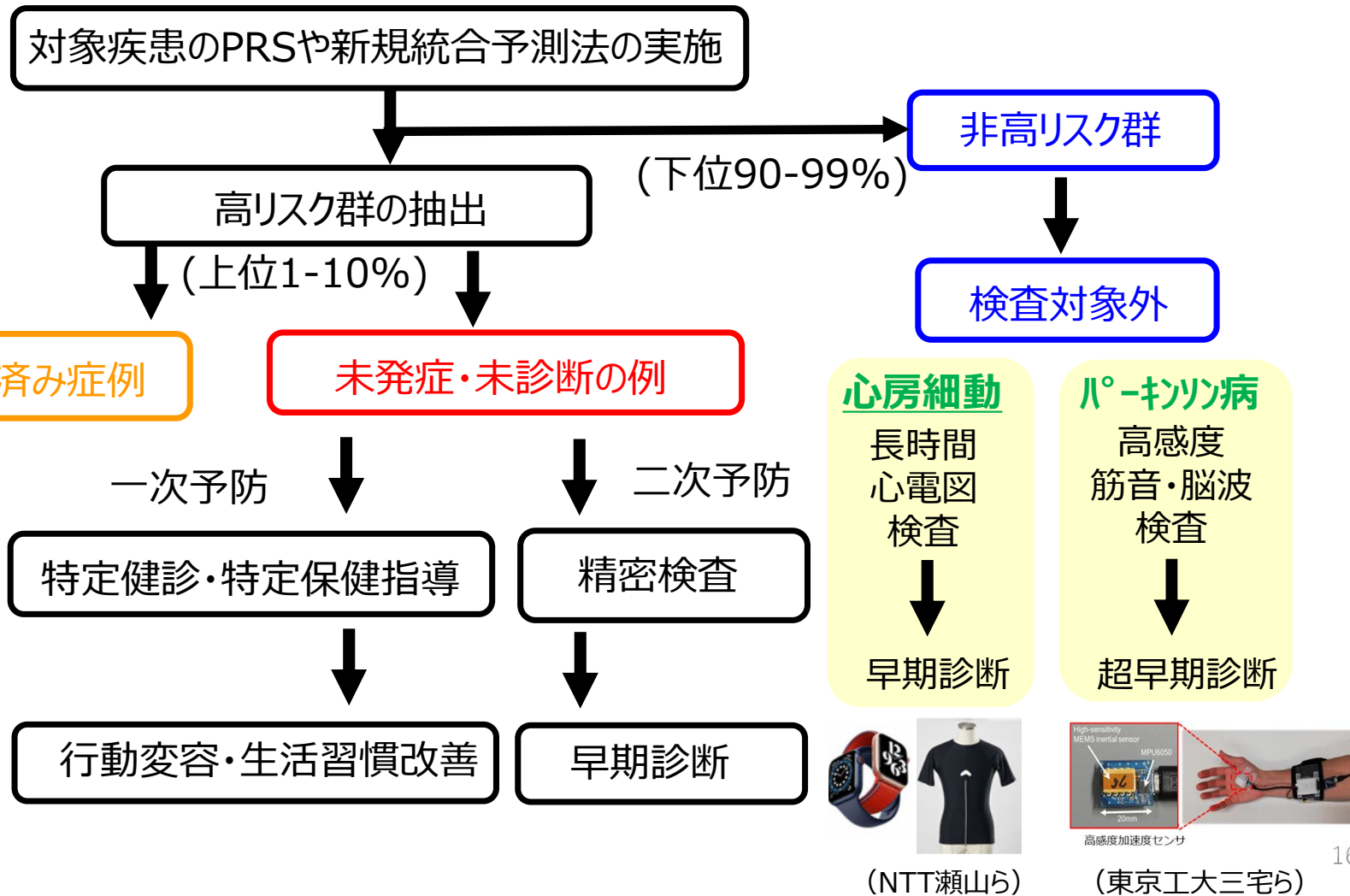
疾患リスク予測に基づく高リスク群の同定による 次世代健診法、超早期疾患診断法の構築



日本独自

行動変容による生活習慣病予防

- ・Web結果回付文章の検討
- ・行動変容プログラムの実施
- ・特定保健指導の枠組みの利用
- ・医療機関の相談窓口設置



5. まとめ

デジタルツインの構築による疾患予防：2030年の挑戦

課題1. コホート構築とデータ生成

生まれてから現在までの
個人データの活用

課題2. データ解析と デジタルツイン構築

AI 解析による
デジタルツインの構築

課題3. 次世代個別化健診と 予防・早期診断

リスク予測、シミュレーション
に基づく個別化疾患予防

