

第4回 協会けんぽ調査研究フォーラム 2017年5月31日

パネルディスカッション

# データ活用に必要な疫学・統計学 および医学的視点



国立保健医療科学院

生涯健康研究部長


横山 徹爾

# はじめに

- 近年、生活習慣病対策を科学的根拠に基づいて推進することの重要度が増している。
- 特定健診・特定保健指導、データヘルス計画等においては、各種データを効果的に活用し、根拠に基づいた計画の策定と実施、評価を行うことが不可欠。
- 活用可能なデータは増大しているが、十分にデータを活用している保険者はまだ少ない。

データ活用に必要な要件は？

# データ活用に必要なもの

 ① 大規模データを“加工集計”する

➤ ICT技術（含、便利なデータベースシステム）

② 最適な“解析”を行う

➤ 統計学・疫学理論

③ 解析結果を“解釈（分析）”する

➤ 医学的知識、疫学・統計学理論

# ①大規模データの“加工集計”

- かつては、**保健師等の専門職**が自分でエクセル等を使って、健診データ等を**こつこつと集計作業**することが多かった。
- それにより**保健活動に専念**する時間が奪われては本末転倒。
  - 事務職との**役割分担**
    - 例) **集計作業**は事務職⇒専門職はその**活用**
  - 便利な**データベースシステムの導入**
    - 例) 国保データベース(**KDB**)システム

# ①大規模データの“加工集計”（例1）

The screenshot shows the '国保データベース(KDB)システム メニュー' (KDB System Menu) interface. It features a grid of buttons for various data processing tasks. The '生活習慣病' (Lifestyle-related diseases) button is circled in red. The interface is organized into several functional groups:

- 健康課題の把握** (Understanding health issues): Includes buttons for regional overview, health issues from data, city-level data, and comparison of insured persons.
- 特定健診・保健指導集計** (Specific health checkups and health guidance summary): Includes buttons for specific health checkup output, elderly person output, and management of health guidance targets.
- 疾病別医療費分析** (Disease-specific medical cost analysis): Includes buttons for medical cost analysis (1) and (2), major/mid/minor categories, and detailed classification.
- 健診・医療・介護の突合** (Health checkups, medical care, and nursing coordination): Includes buttons for coordination status and comparison of insured persons.
- その他** (Others): Includes CSV data output, account output, basic course, option setting, and user management.
- 被保険者の突合** (Coordination of insured persons): Includes a button for late-term insured person coordination status.

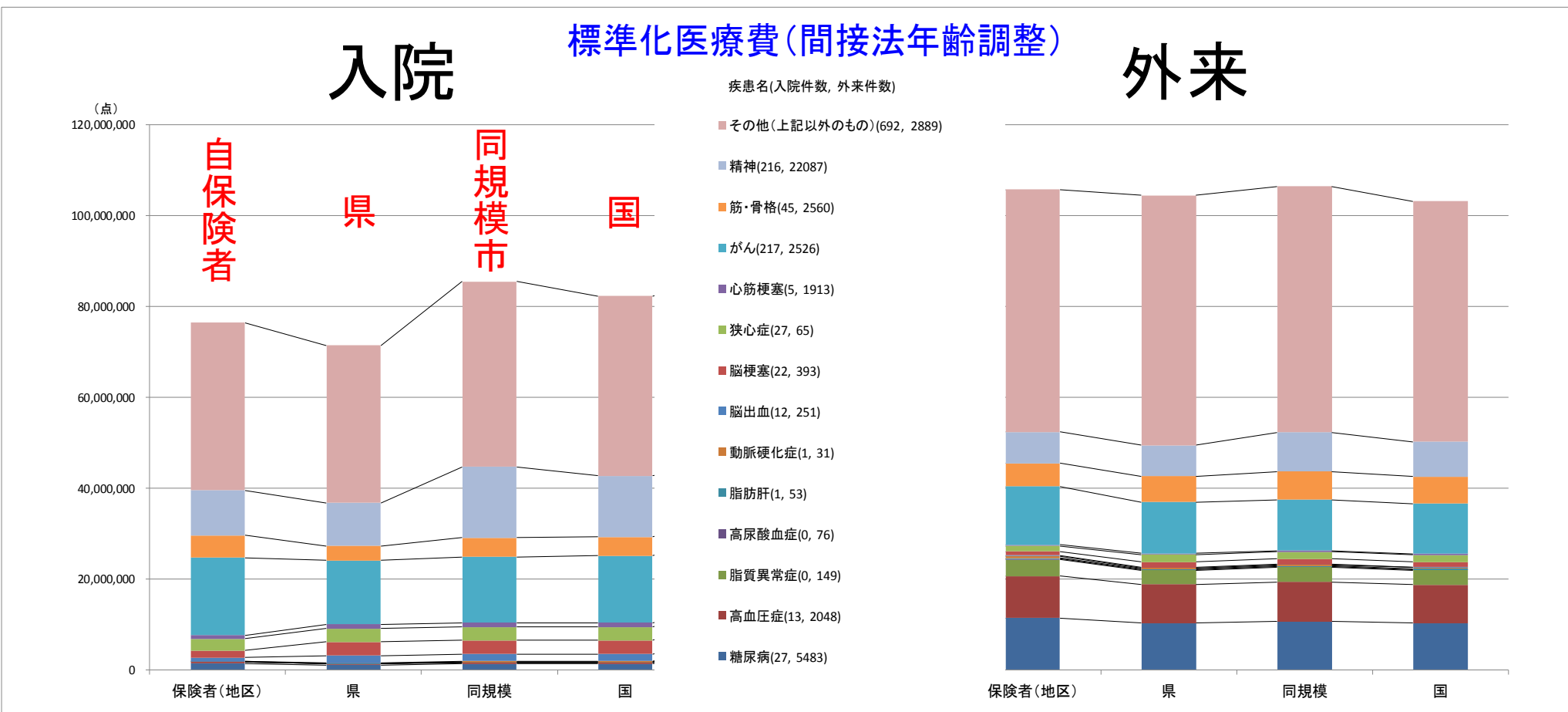
The '生活習慣病' button is located in the '疾病別医療費分析' section, under the '細小分類' (Detailed classification) category.

国保データベース(KDB)システム

# KDBのCSV出力を国立保健医療科学院の年齢調整ツールで加工

## 疾病別医療費分析(生活習慣病) サンプル市2 H26年度(累計) 男性

2016年1月22日



# データ活用に必要なもの

① 大規模データを“加工集計”する

➤ ICT技術（含、便利なデータベースシステム）

 ② 最適な“解析”を行う

➤ 統計学・疫学理論

③ 解析結果を“解釈（分析）”する

➤ 医学的知識、疫学・統計学理論

## ②最適な“解析”を行う

- 単純集計だけでは解釈が困難
  - 例) 特定健診での高血圧の頻度
    - A地域50% > B地域40%
      - 平均年齢: A地域 > B地域
      - 人数: A地域5/10人 > B地域40/100人
- 最低限、知っておきたい疫学・統計学的概念
  - 年齢調整
  - 偶然変動(検定、信頼区間)
  - 対照群との比較(平均への回帰)
  - 見える化
  - など。



# 年齢調整

- 死亡率、リスク因子（高血圧等）、医療費等は、高年齢ほど高くてあたりまえ。
- 地域間の比較および経時的な推移をモニタリングする際に、年齢調整した指標を用いる必要性。
  - 死亡率⇒標準化死亡比（SMR）
  - リスク因子⇒年齢調整割合、標準化該当比
  - 医療費⇒年齢調整平均、地域差指数など
- 病気の「なりやすさ」を知りたい場合（地域の健康状態の指標）
  - 年齢調整した割合・平均で比較する。
  - （または年齢別に比較する）
- 病気の「総量」を知りたい場合（必要な医療資源の量に関係）
  - 年齢調整しない（粗）割合・平均・人数等で比較。

# 偶然変動の考慮

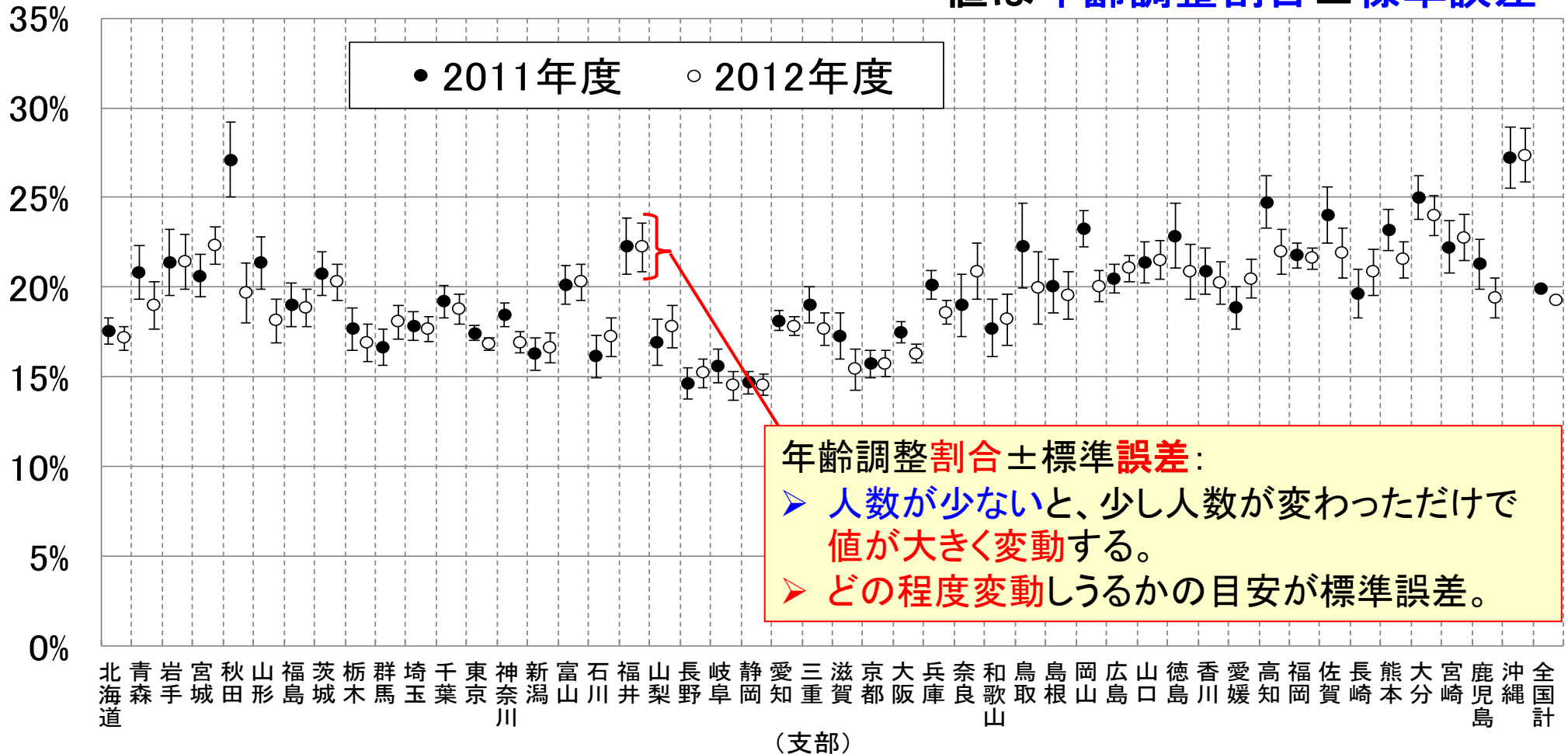
- 人口が少ないと、年次ごとの該当人数（死亡数、リスク因子保有者数、受療人数等）のわずかな増減によって指標（死亡率、該当割合、医療費等）の値が大きく変動することがあるため、地域間の比較および経時的な推移をモニタリングする際には、偶然変動を考慮する。
  - ①単年で評価せず経年的にデータを見ることによって長期的な傾向を把握する
  - ②複数年のデータをプールして指標を計算する
  - ③95%信頼区間、検定結果を併記する、など。

# 年齢調整・偶然変動の考慮(例1)

腹囲 $\geq 90$  cm(女)の割合の支部間・年度間比較(女性, 65-74歳)

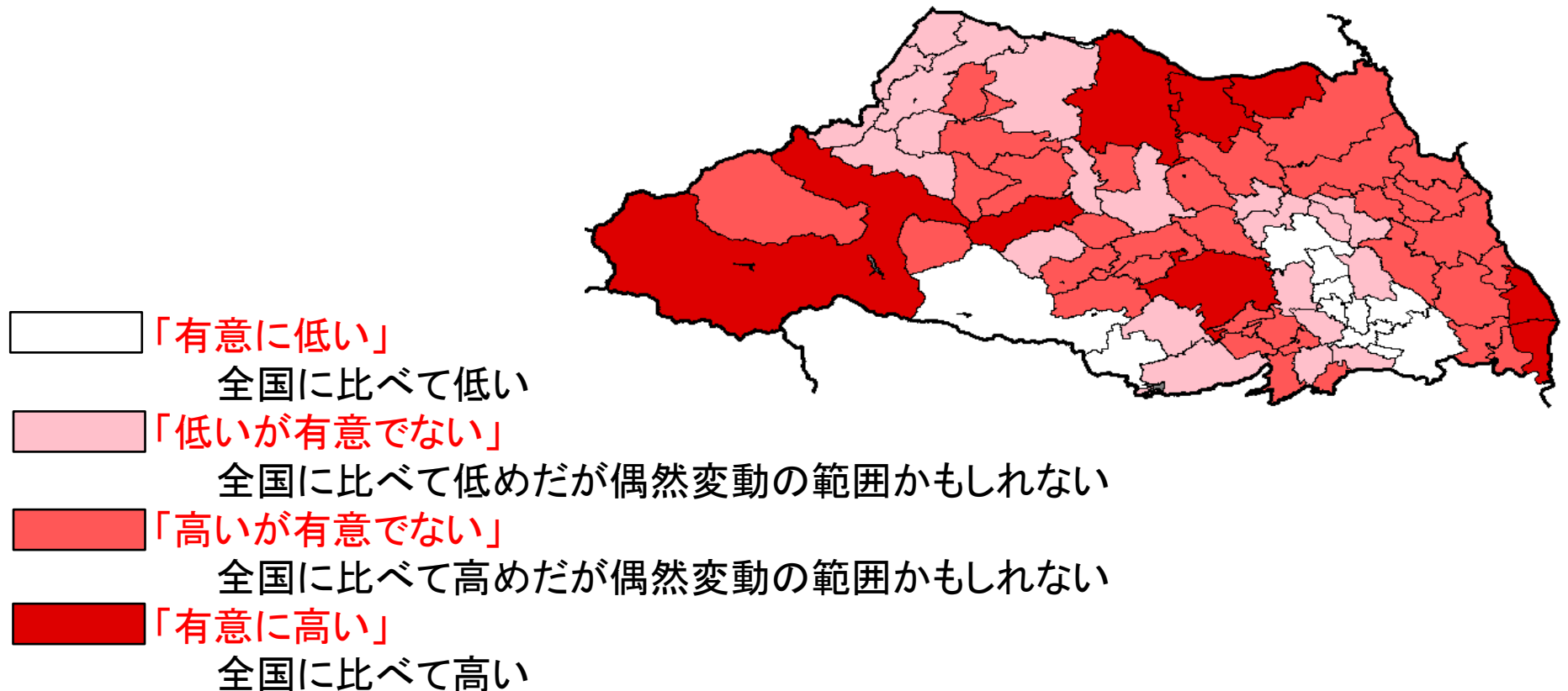
(割合)

値は年齢調整割合 $\pm$ 標準誤差



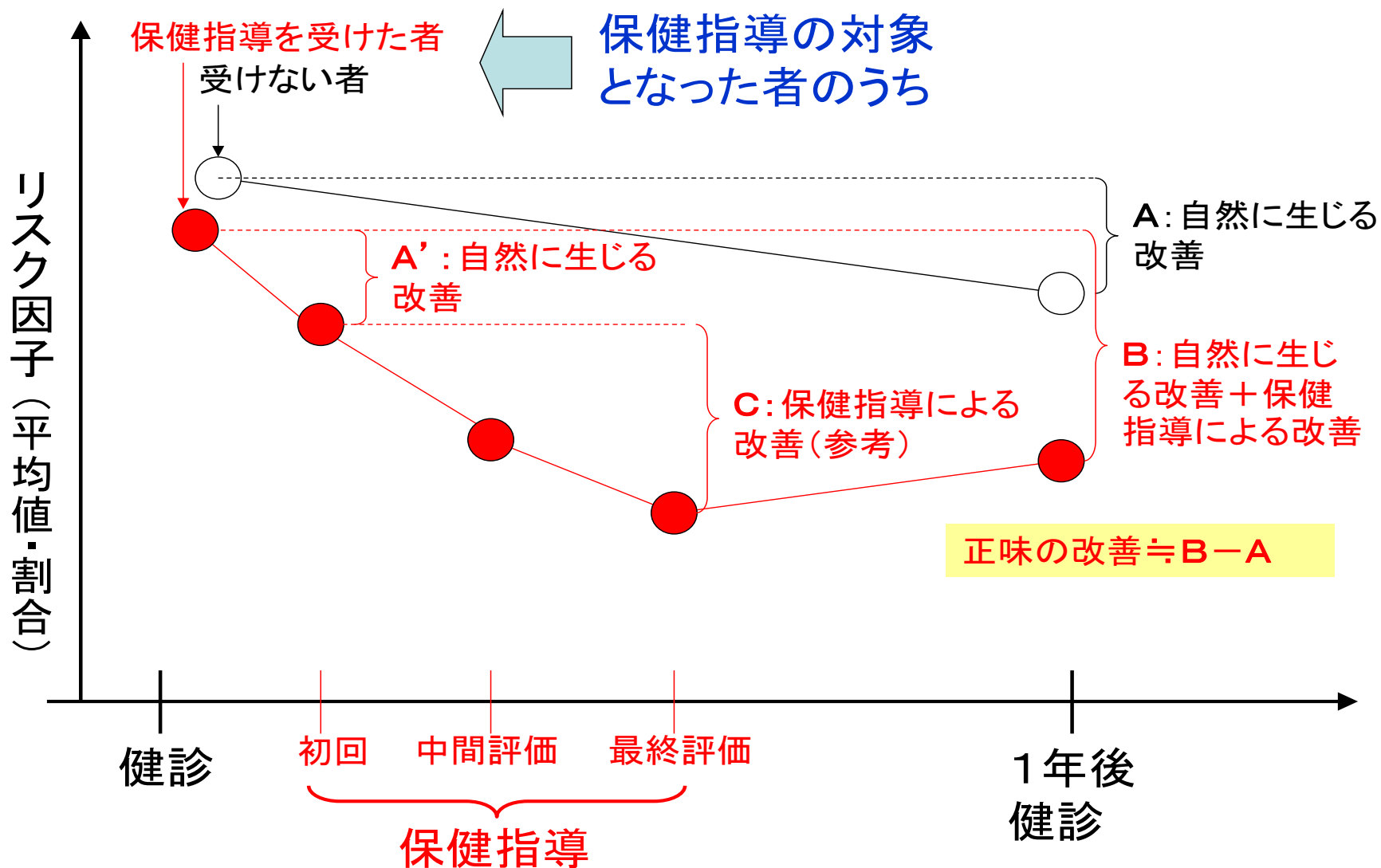
# 年齢調整・偶然変動の考慮(例2)

収縮期血圧 $\geq 130\text{mmHg}$ の地域比較  
男性・年齢調整値(標準化該当比)



# 対照群との比較

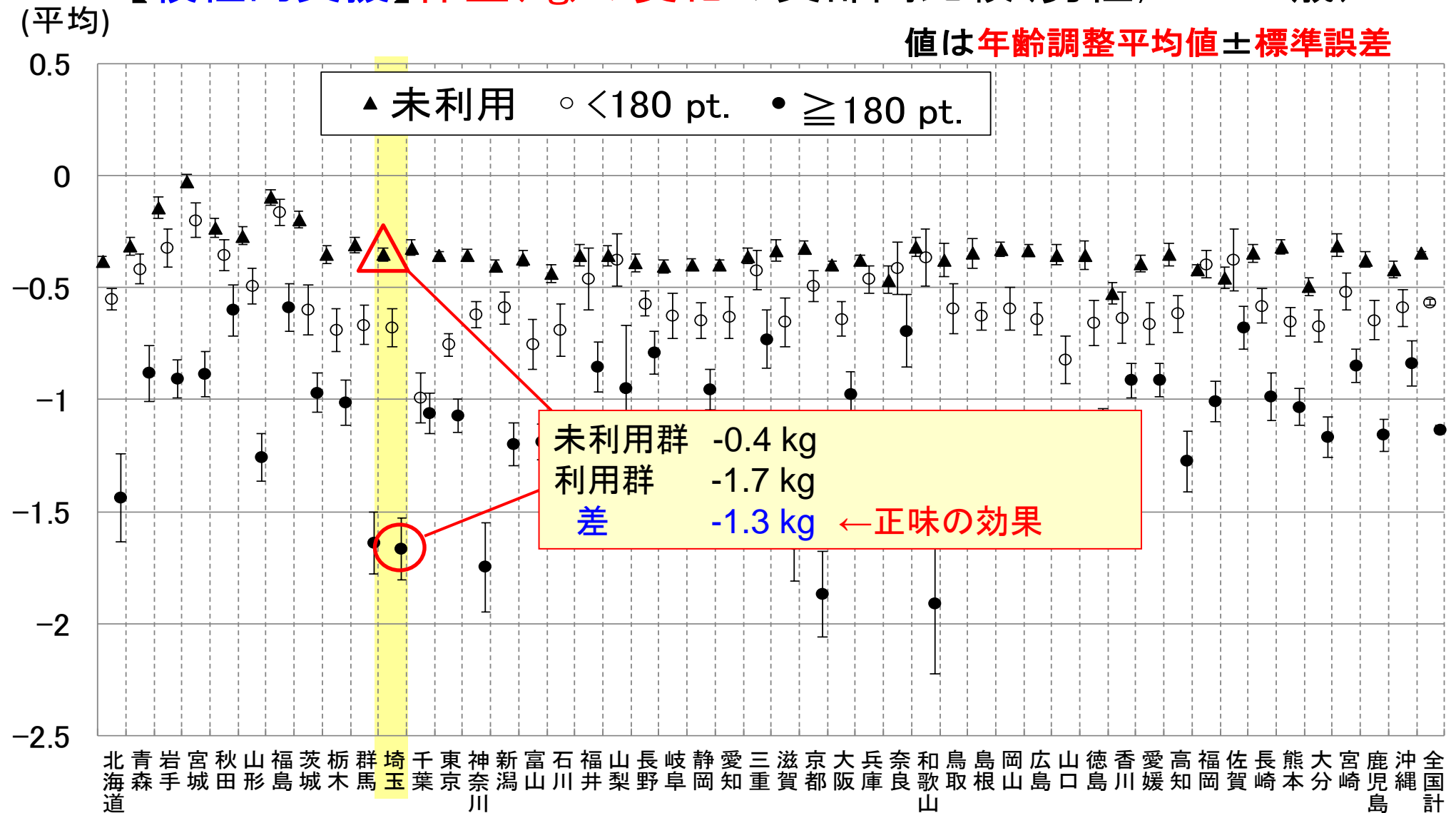
図1. 保健指導によるリスク因子等への「効果」の測り方の概念



# 対照群との比較(例1)

【積極的支援】体重(kg)の変化の支部間比較(男性, 40-64歳)

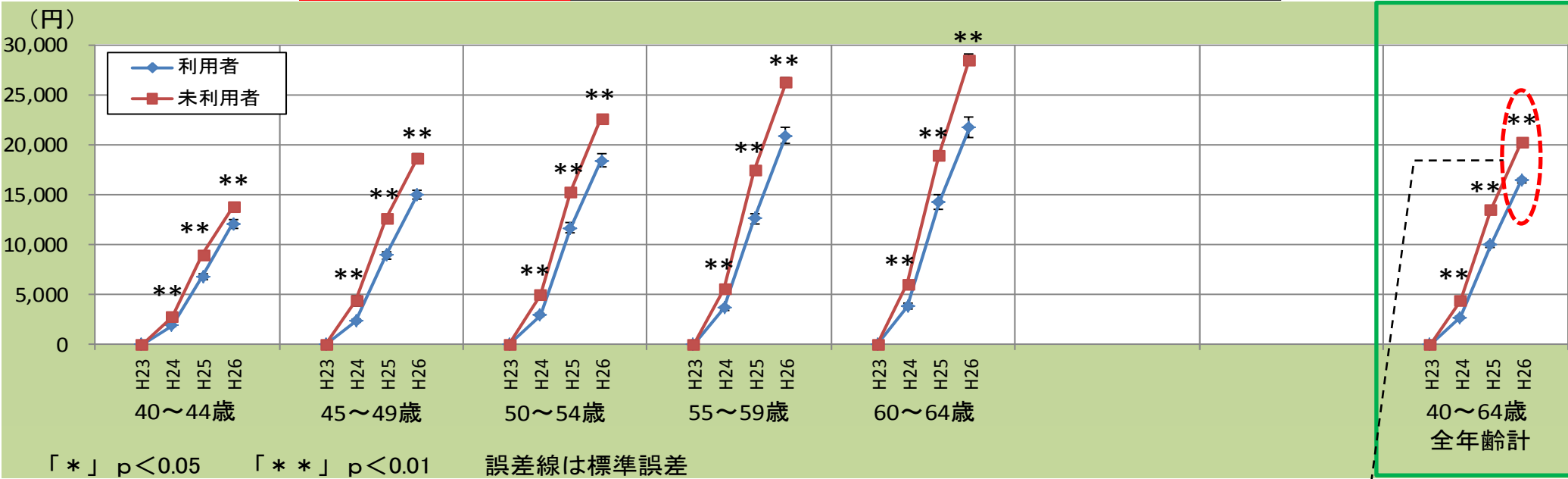
値は年齢調整平均値±標準誤差



# 対照群との比較(例2)

【積極的支援】利用者と未利用者の

その後3年間の入院外一人当たり医療費(男性)



- 全ての年齢階級で、未利用者より利用者の方が有意に低かった。
- 国のWG分析と同様の傾向が見られた。

26年度の差は約3,821円で、  
年齢調整後は  
**約3,790円**

## ★平成26年度医療費適正化効果額の試算

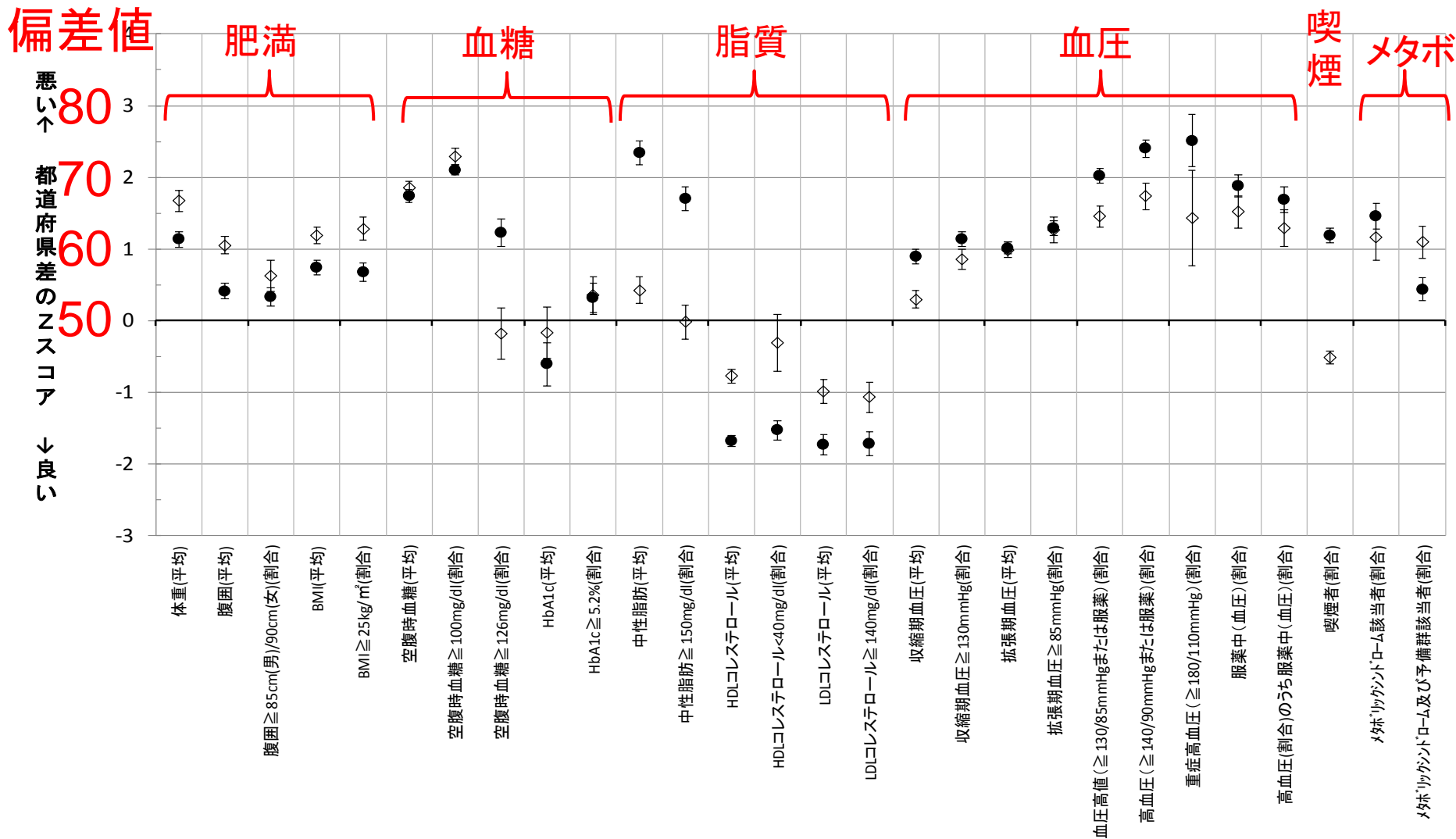
$$44,099(\text{利用者数}) \times 3,790(\text{26年度の差}) = \text{約1.7億円}$$

# 見える化(例1)

## 2012年度特定健診の支部別特徴の比較(年齢調整値) 秋田県

【40-74歳】

● 男性 ◇ 女性





# データ活用に必要なもの

① 大規模データを“加工集計”する

➤ ICT技術（含、便利なデータベースシステム）

② 最適な“解析”を行う

➤ 統計学・疫学理論

 ③ 解析結果を“解釈（分析）”する

➤ 医学的知識、疫学・統計学理論

# ③解析結果を“**解釈(分析)**”する

生活習慣⇒リスク因子⇒医療⇒介護⇒健康寿命  
の**上下関係(原因～結果)**を理解する

**【医学的知識が必要】**

平均寿命

—

健康寿命

=

不健康期間

死因別年齢調整死亡率・SMR

重症化

生活習慣病

介護(介護給付費、認定者数)

医療(入院・外来受療率、医療費)

保健指導(実施率・効果)

健診(受診率)

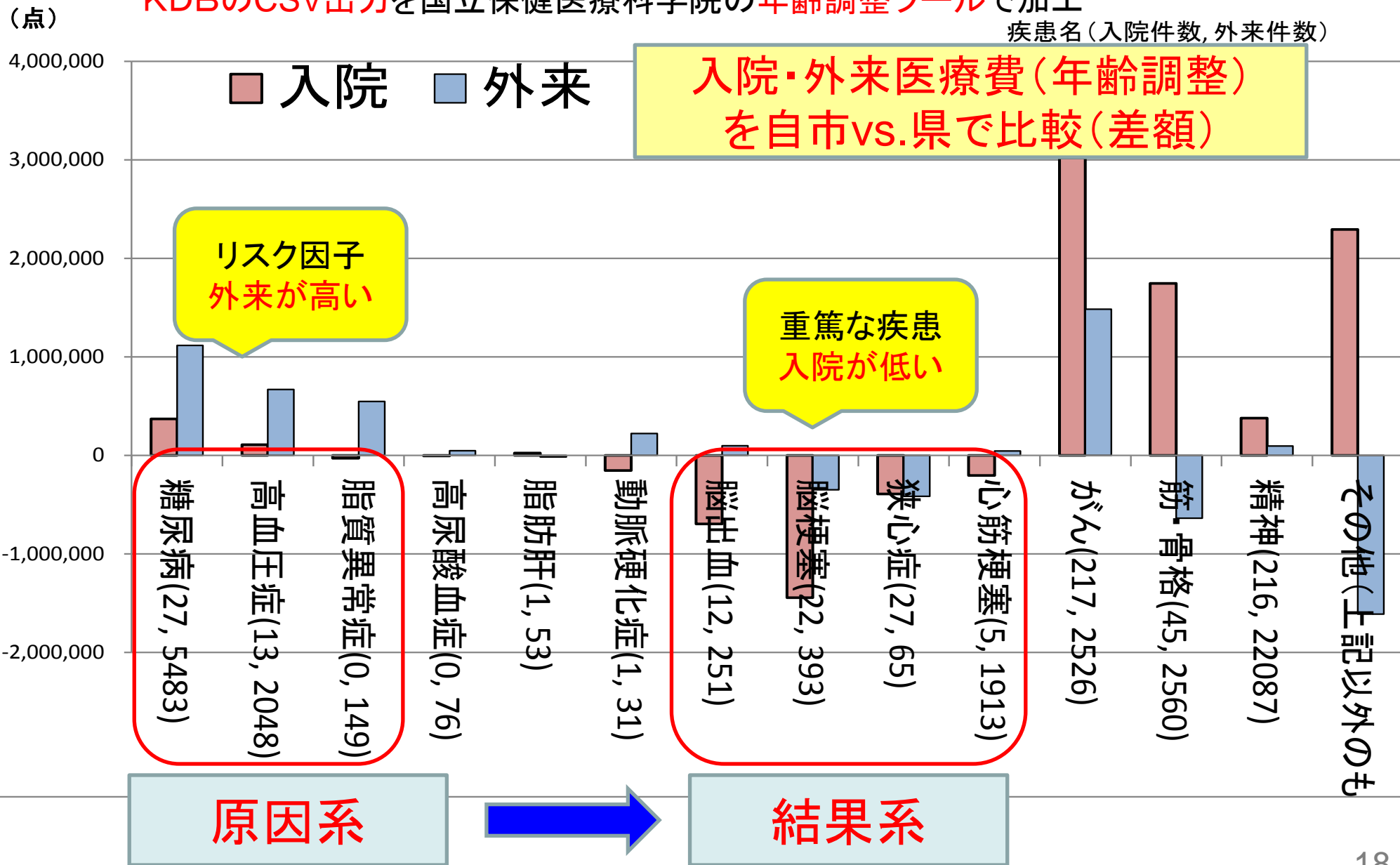
リスク因子

生活習慣・環境

# 解析結果を“**解釈(分析)**”する(例1)

KDBのCSV出力を国立保健医療科学院の年齢調整ツールで加工

疾患名(入院件数, 外来件数)

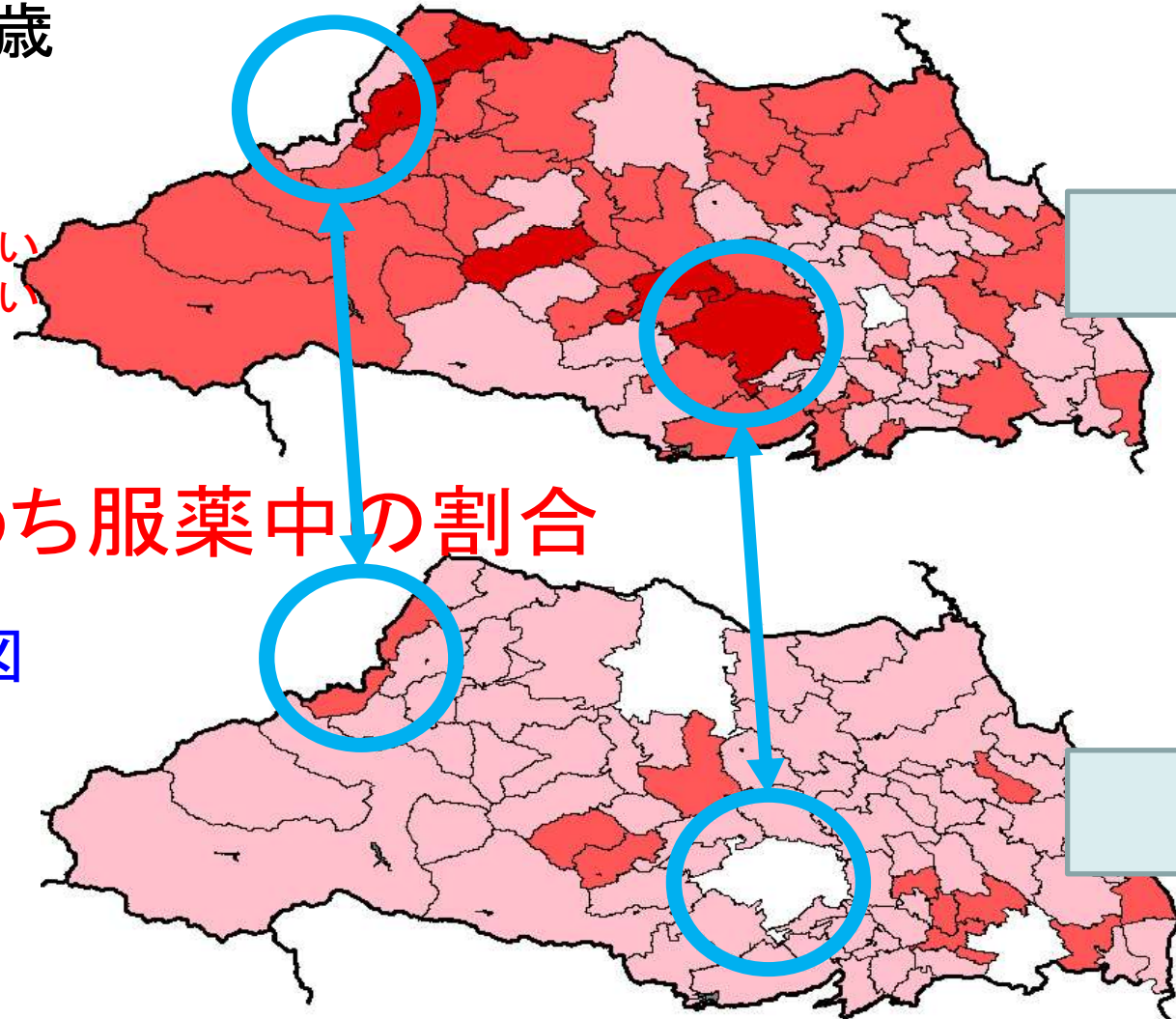


# 解析結果を“**解釈(分析)**”する(例2)

## 重症高血圧(Ⅲ度)の標準化該当比

男性, 40-74歳

- 有意に低い
- 低い有意でない
- 高い有意でない
- 有意に高い



結果系

## 高血圧のうち服薬中の割合

上図 vs. 下図

$r = -0.69$

$p < 0.0001$

原因系

# まとめ

## データ活用に必要なもの

- ① 大規模データを“**加工集計**”する
  - ICT技術(含、便利なデータベースシステム)
- ② 最適な“**解析**”を行う
  - 統計学・疫学理論
- ③ 解析結果を“**解釈(分析)**”する
  - 医学的知識、疫学・統計学理論

### ◆ 多分野からなるチームで役割分担して協働作業

- ICT技術を持つ事務職
- 医学知識を持つ医療専門職
- 疫学・統計学の専門家

### ◆ これらの能力を持つ人材を育成

- 保健医療データ分析専攻科(国立保健医療科学院)