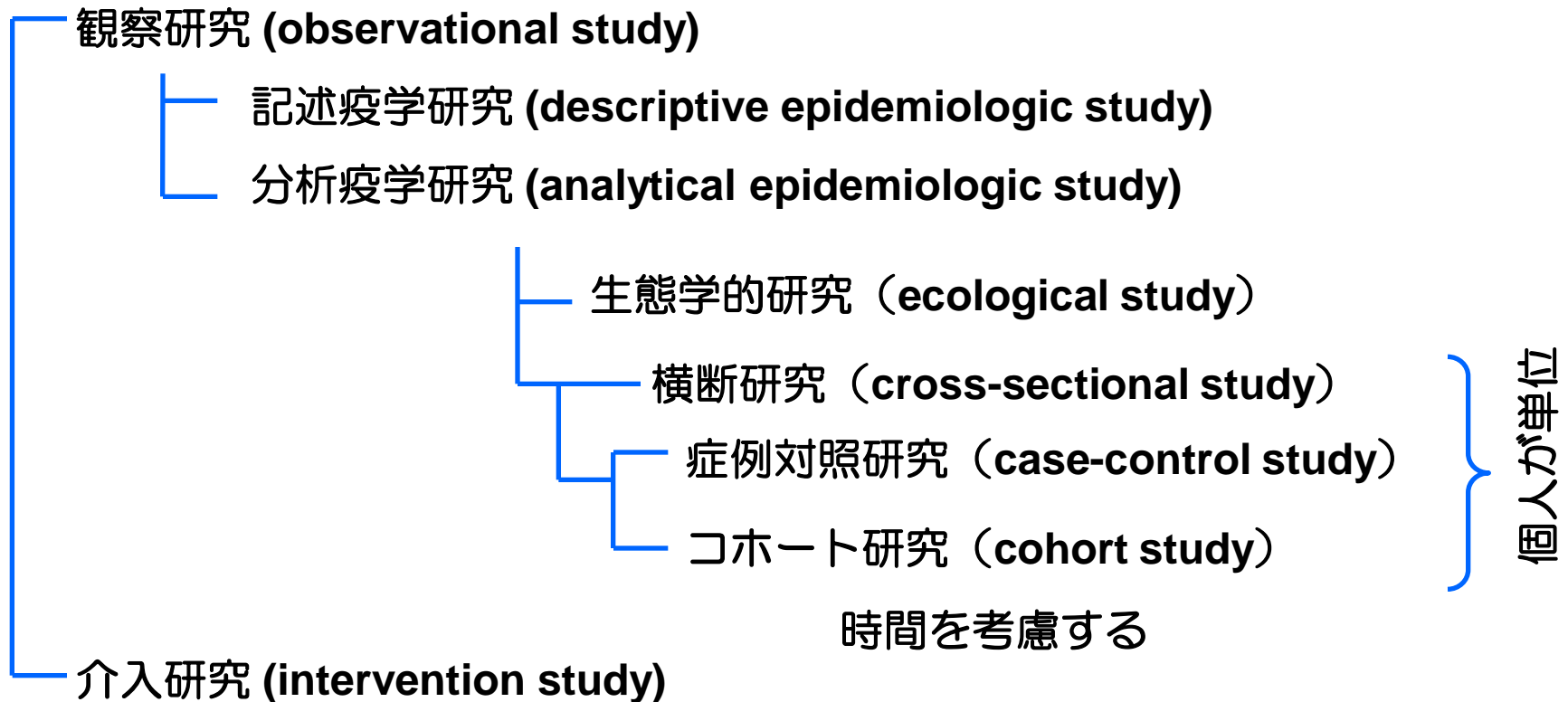
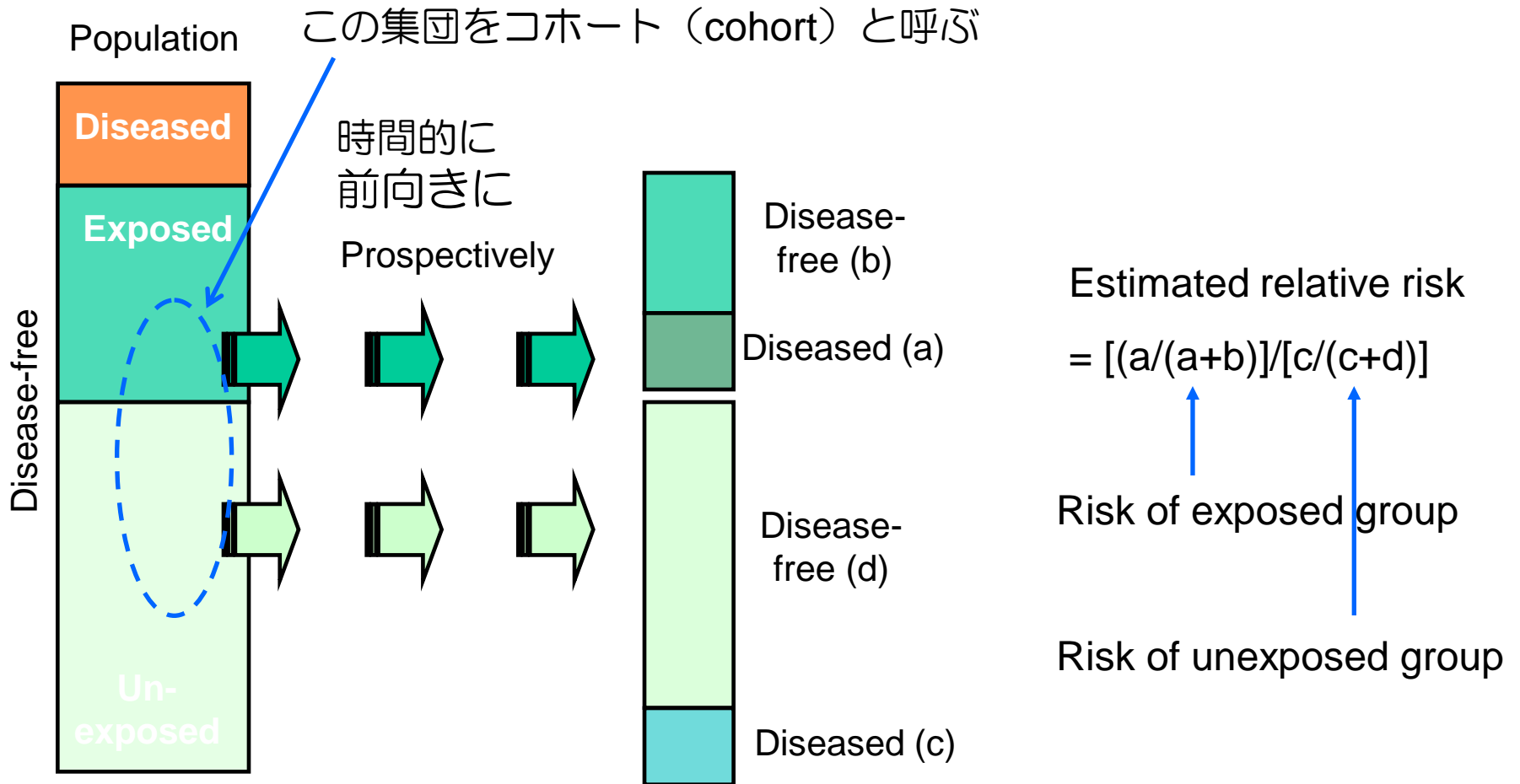


# コホート研究 (Cohort study)

原因を知りたい問題 (疾病) をもたない集団を (放置) 追跡し、問題が発生した集団としなかった集団について、それまでの危険因子 (曝露要因) の有無・程度を比較検討する疫学研究



# Cohort study: marching towards outcomes

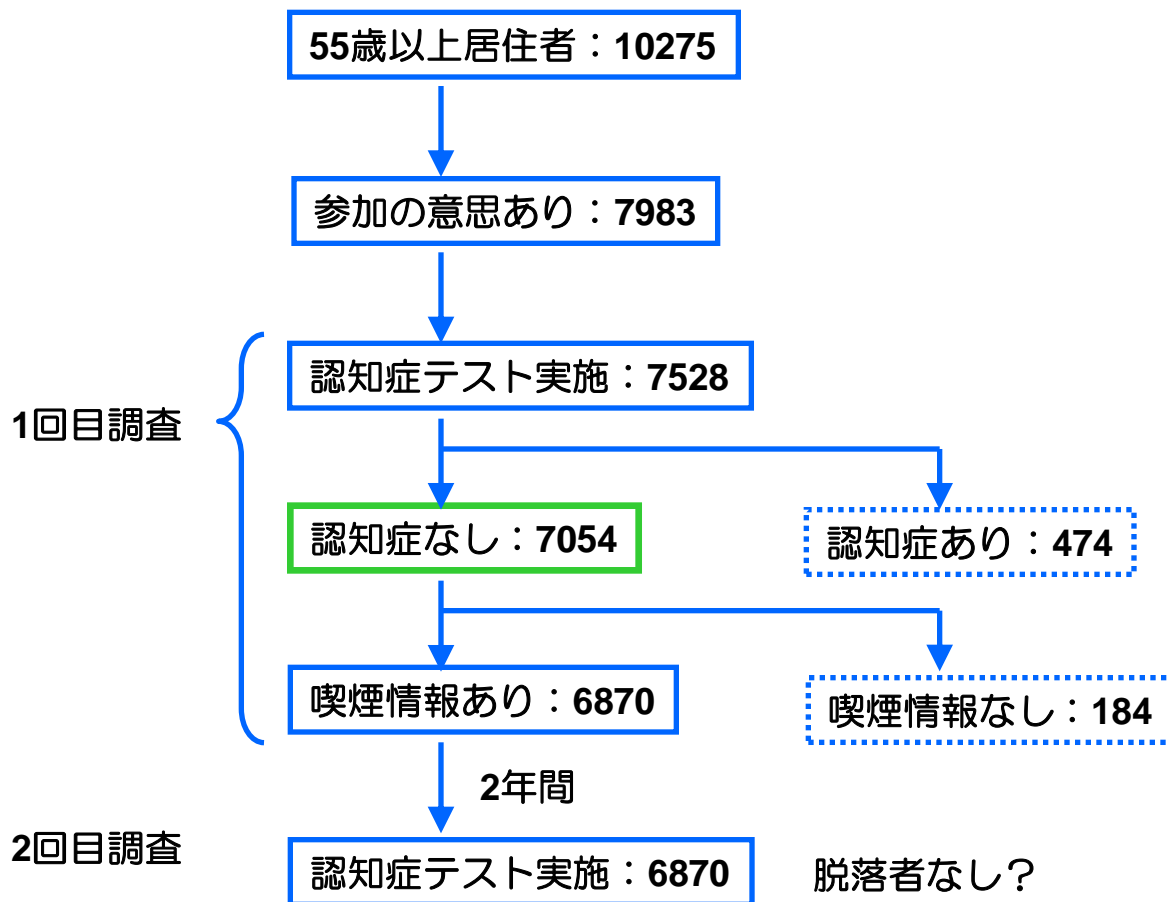
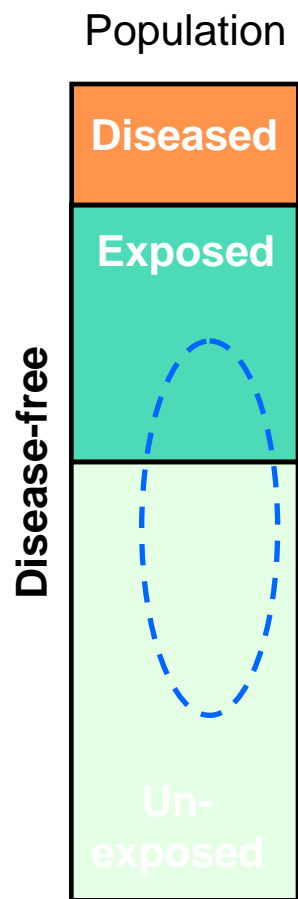


こんなにきれいにいくはずがない

# 対象者

## The Rotterdam Study: オランダの高齢者コホート

#3349. Ott A, et al. Lancet 1998; 351: 1840-3.

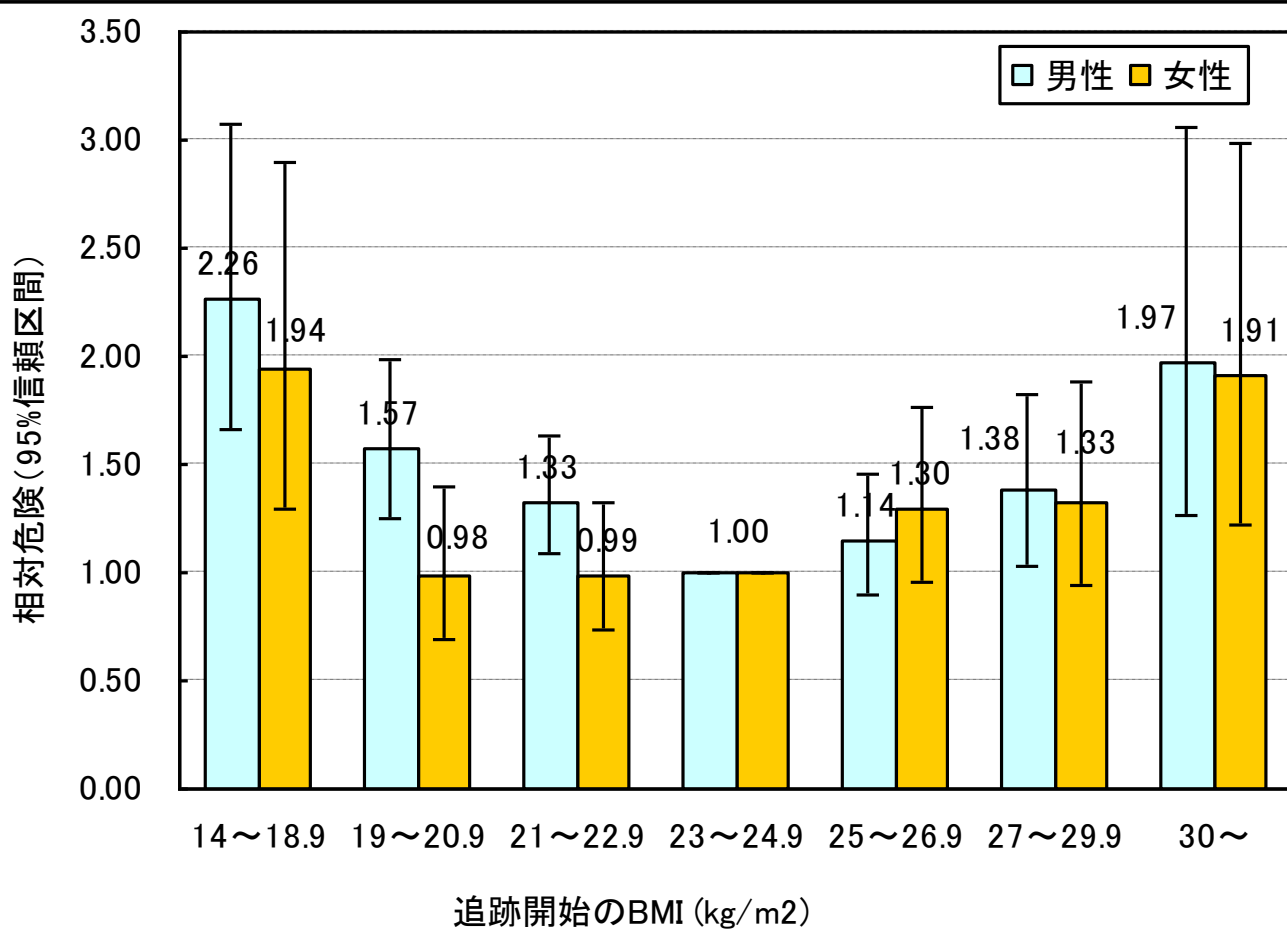


「とにかくたくさん集めればよい」というわけではない

参加の意思があり、対象疾患にかかっていなくて、1回目（ベースライン）の情報が全部ある人

# BMIと総死亡率の関連

#4917. Tsugane S, et al. Int J Obes 2002;26:529-37



身長と体重は自己申告

BMIが23~24.9の群に比べた相対危険 (95%信頼区間)

日本人男女 (19500、21315人) を10年間追跡した結果

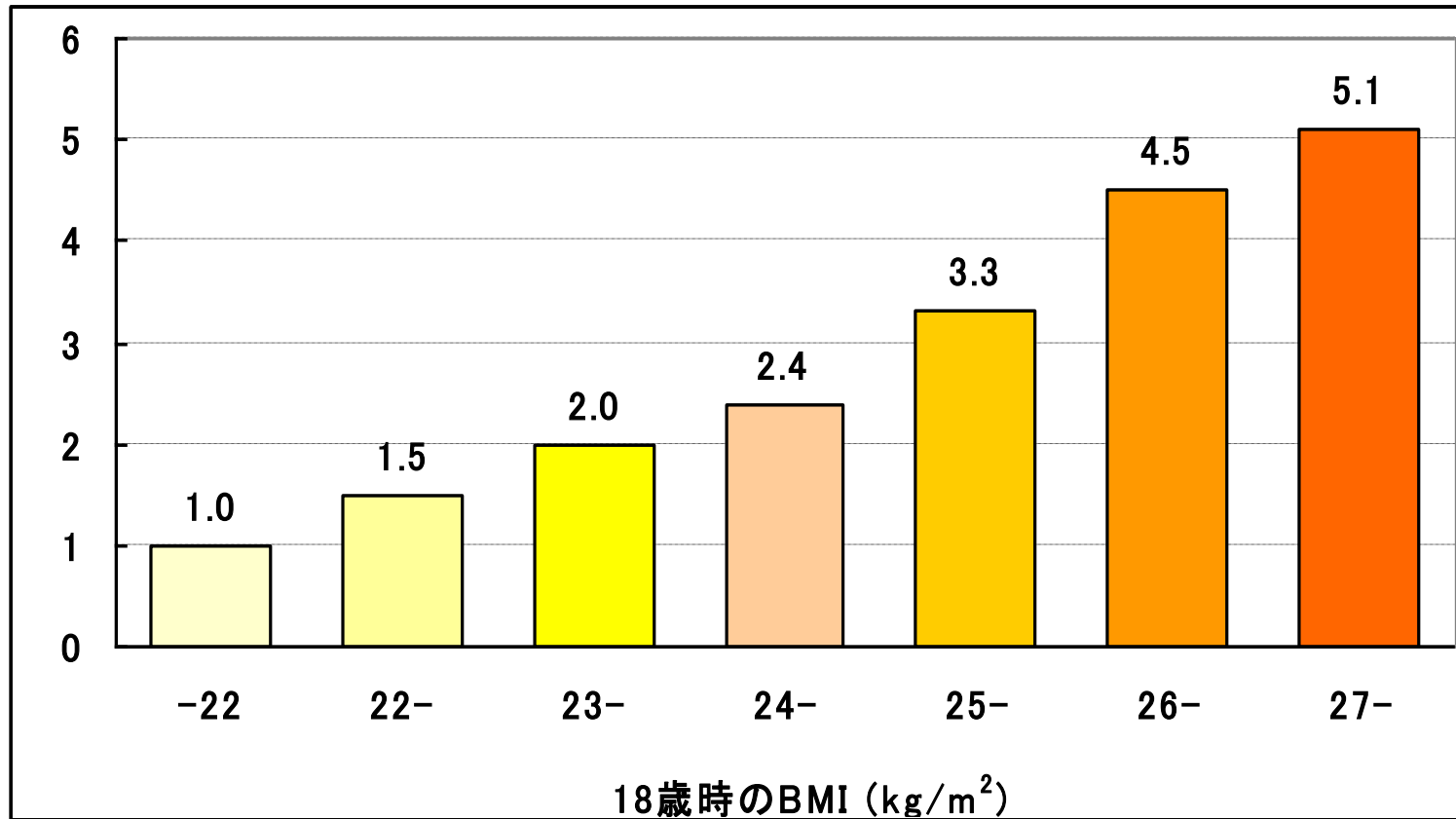
地域、年齢、喫煙習慣、飲酒習慣、教育歴、運動習慣、20歳以後の体重変化の影響を調整

厚生労働省多目的コホート研究 Japan Public Health Center-based Prospective Study on Cancer and Cardiovascular diseases (JPHC Study)

Outcomeをevent (死亡) でみた例

- ・・・コホート研究は発生がはっきりする疾患に向いている。

## 体格（BMI）と糖尿病家族歴の糖尿病発症への影響

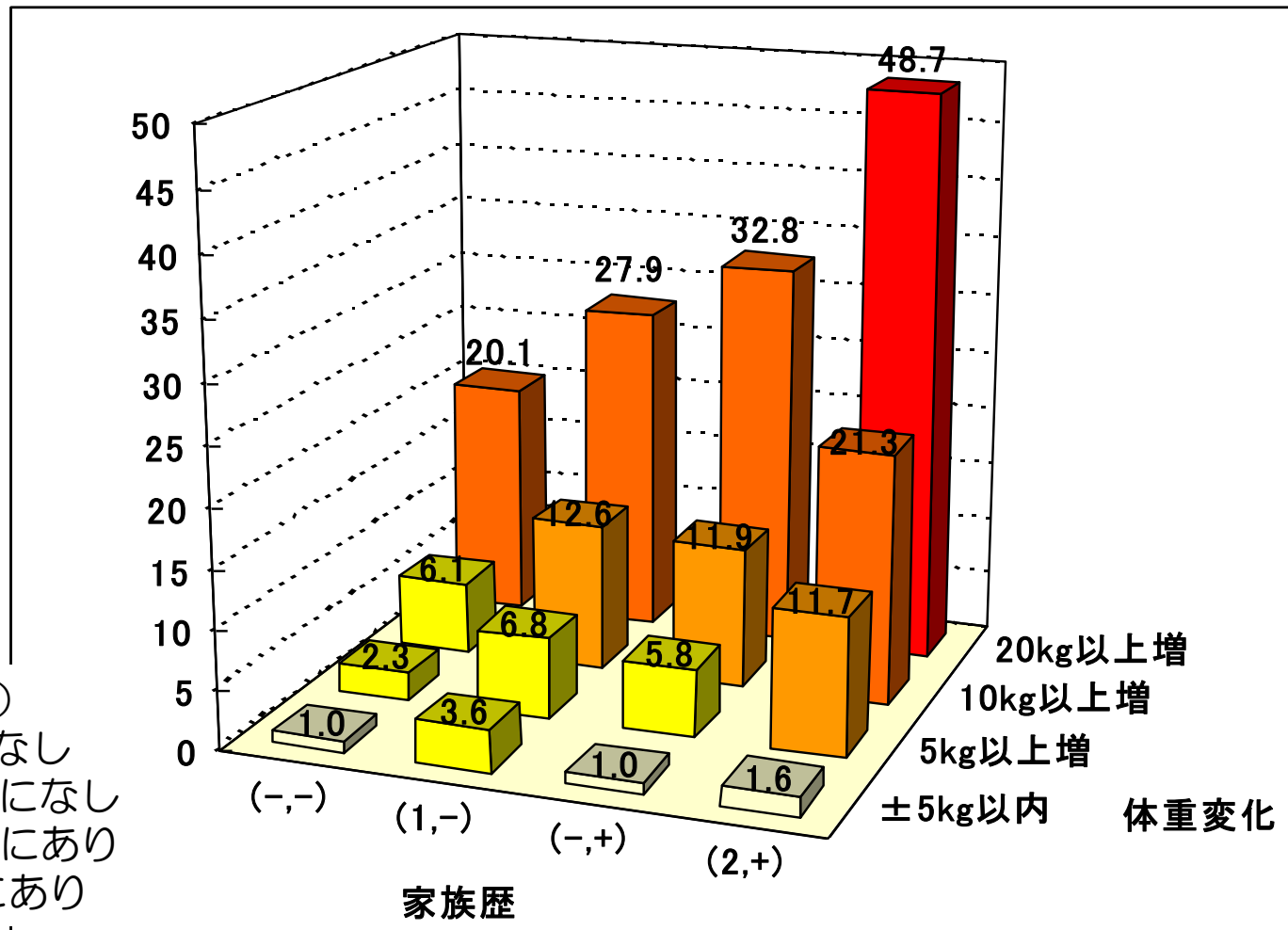


18歳時の体重と30歳以後の糖尿病発症の関係。BMIが22未満の群に比べた相対危険（アメリカ人女性、114281人を14年間追跡。糖尿病発症数=2204）

1976年に開始（30-55歳）。1980に「18歳のときの体重を尋ねた」 1976年から1990/06/01まで追跡した。...18歳のときに体重を測ったわけではない。後ろ向きに（retrospectively）なっている。けっこう、姑息？

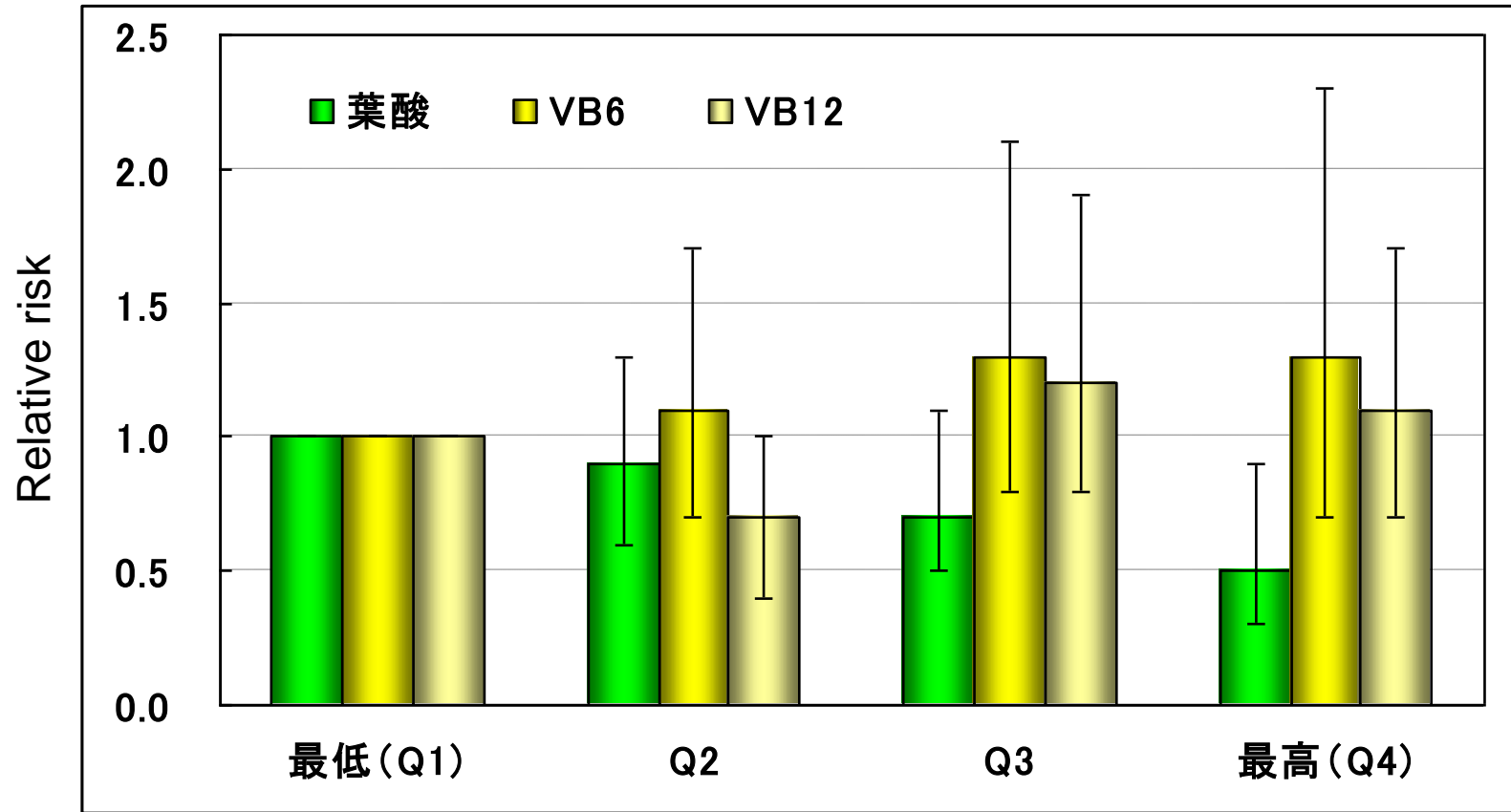
# 肥満と糖尿病家族歴の糖尿病発症への影響

家族歴（糖尿病の有無）  
 (-,-) 両親、兄弟ともになし  
 (1,-) 片親にあり、兄弟になし  
 (-,+) 両親はなし、兄弟にあり  
 (2,+) 両親、兄弟ともにあり



アメリカ人女性、114281人を14年間追跡した結果  
 (糖尿病発症数=2204)

# ホモシステイン関連ビタミン摂取量とアルツハイマー病発症との関連



摂取量＝食品+サプリメント

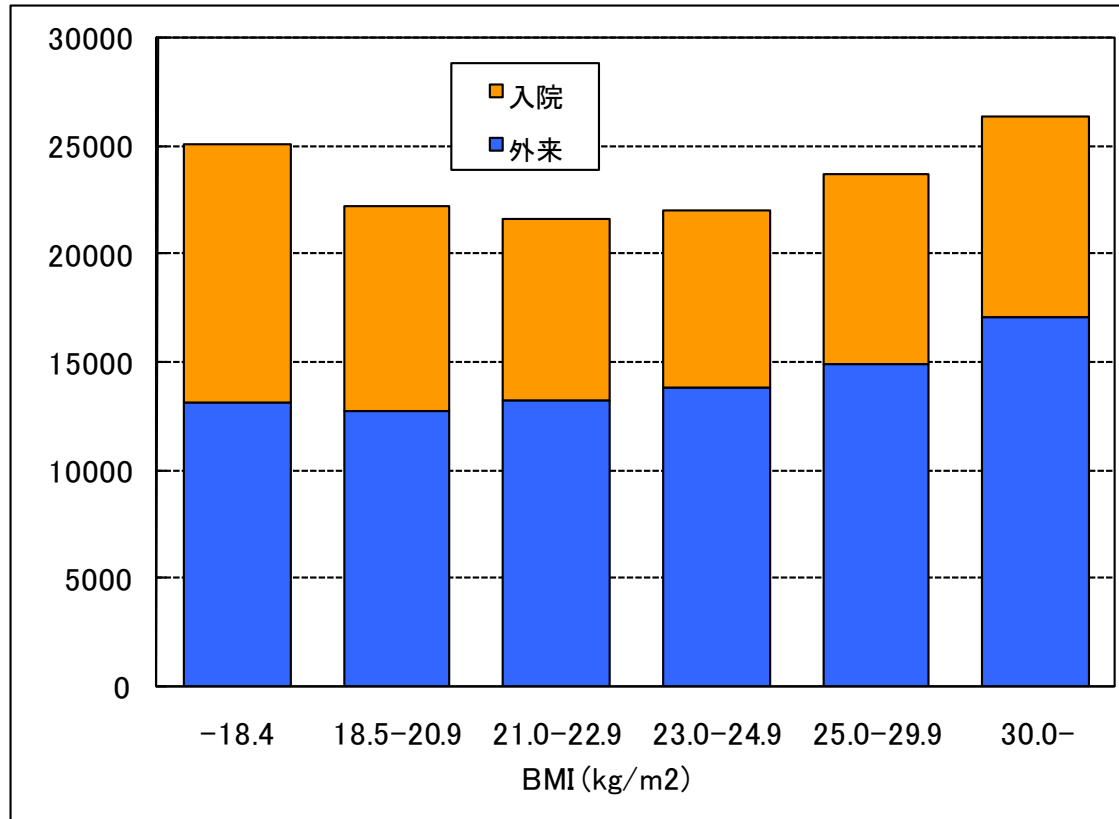
65歳以上、965人、追跡期間＝6.3年、（アメリカ）

調整因子：年齢、性、教育歴、アポリポ蛋白E4アリル有無、人種、他

#10432. Luchsinger, et al. Arch Neurol 2007; 64: 86-92.

Outcomeを変化でみた例（調査を2回している）

## 体格（BMI）と医療費の関連



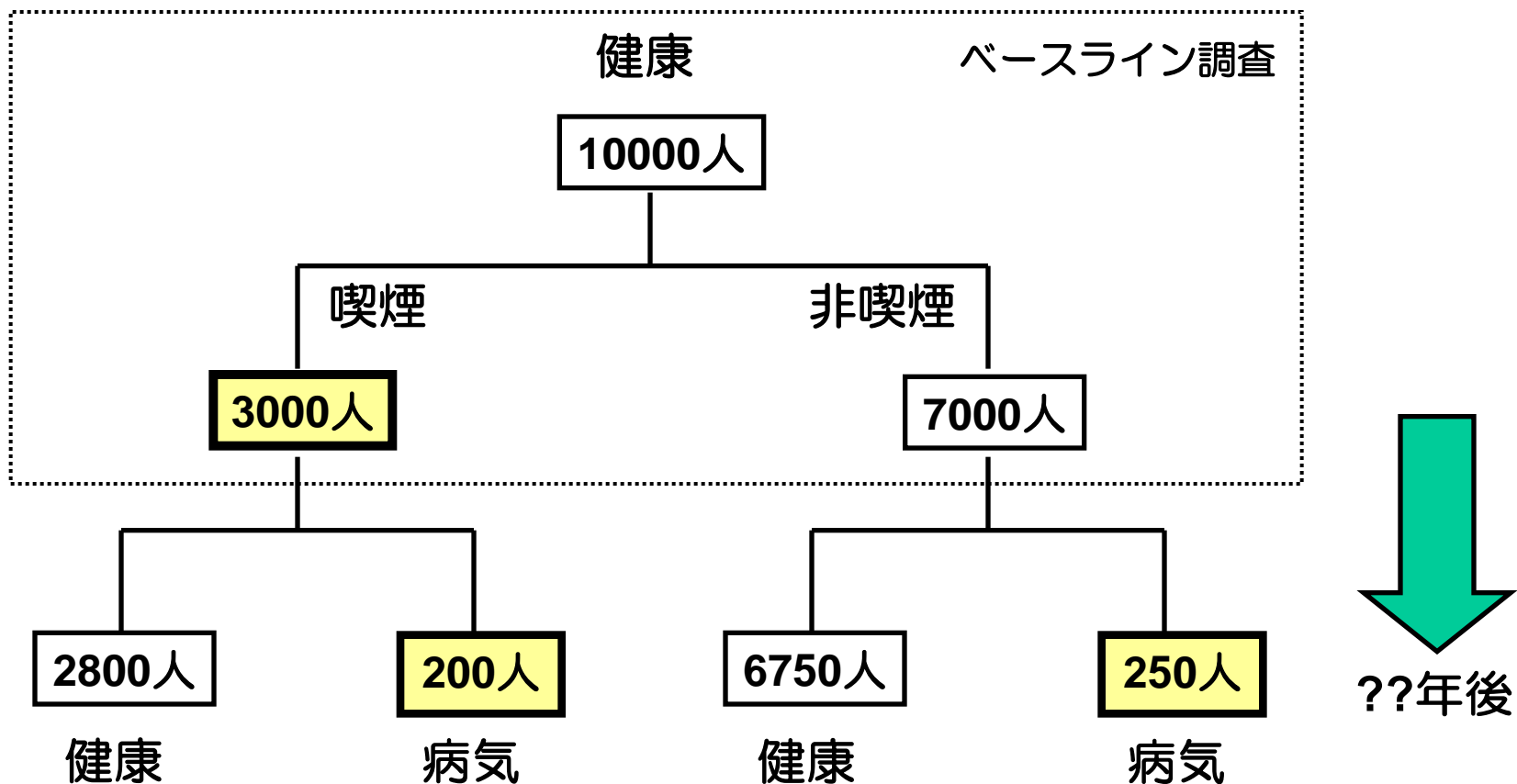
日本人中高年（40～79歳）における肥満度（BMI）とその後の1年間に支出した医療費（円/年）との関連：およそ4万人を1年間追跡した結果

#7063. Kuriyama S, et al. Int J Obes 2002; 26: 1069-74.

結果因子は継続して収集される場合もある。



# 喫煙と病気の発症との関連を調べるためのコホート研究（仮想データ）



① 非喫煙者に比べて喫煙者は何倍、病気にかかりやすいか？（相対危険: relative risk）

② 7000人の非喫煙者が喫煙していたら何人が病気にかかっていたか？（寄与危険: attributable risk）

この危険因子を排除できたら、どれくらいの人がこの病気から免れるか  
...が計算できる

## 相対危険

$$(200/3000) / (250/7000)$$

喫煙者の危険 (リスク)

非喫煙者の危険 (リスク)

喫煙者がもし非喫煙者だったら、  
 $3000 \times (250/7000)$  の病気が発生したはず。

実際の発生数は  $200 = 3000 \times (200/3000)$  だから、

喫煙者が喫煙していなかったら、

$$200 - 3000 \times (250/7000) = 3000 \times [(200/3000) - (250/7000)] = 93$$

だけ病気が減ったはず。

Population attributable risk

人口寄与危険

寄与危険 Attributable risk

「1人の喫煙者が喫煙していなかったら、病気が何人減るか？」の指標

「相対危険」と「寄与危険」 計算方法よりも用途のちがいに注意！

## 相対危険と寄与危険の例：肺がんと心筋梗塞のリスク（仮想データ）

	相対危険	非喫煙者の 死亡率 (人口10万 対)	喫煙者の 死亡率 (人口10万 対)	寄与危険
肺がん	32.4	0.07	2.27	2.20
心筋梗塞	1.4	7.32	9.93	2.61

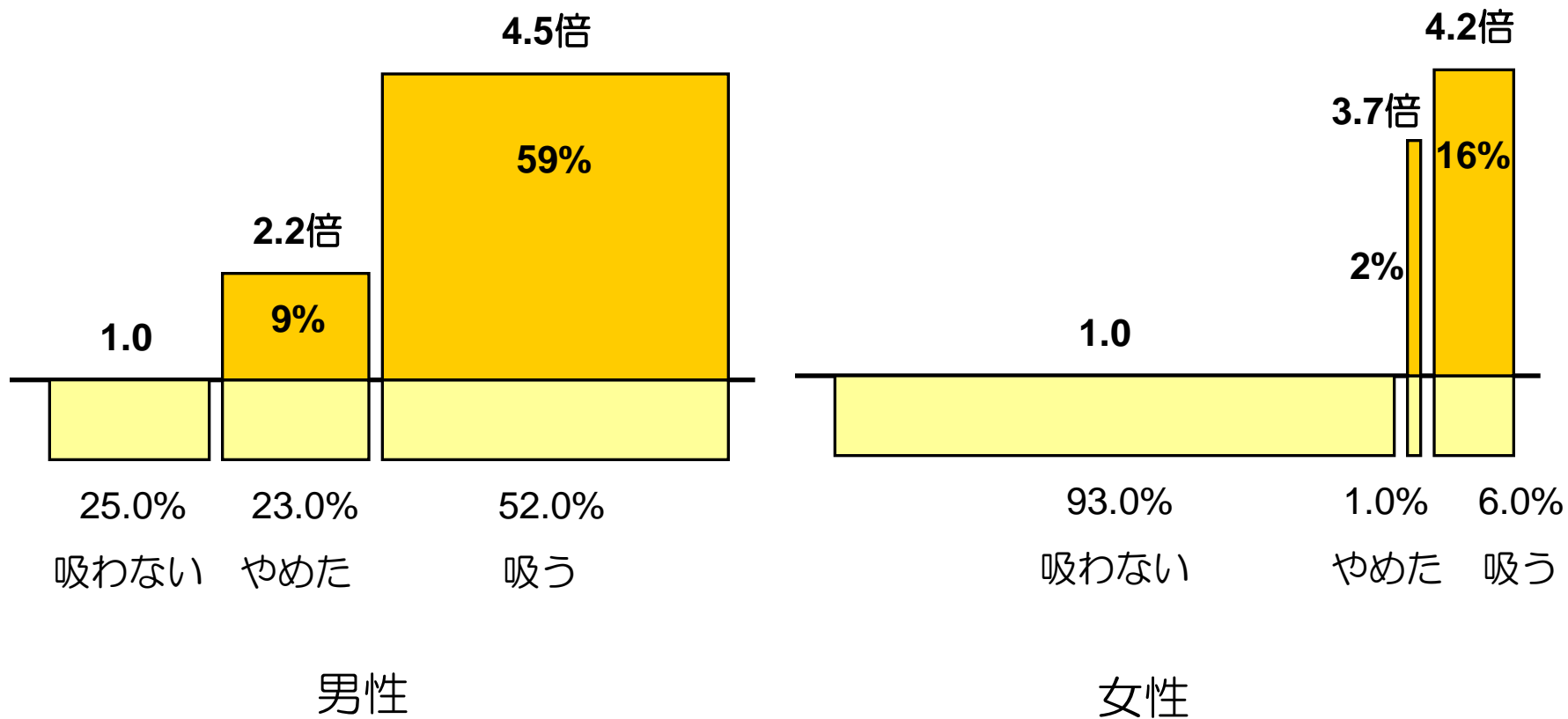
相対危険は、心筋梗塞に比べて肺がんで著しく高いが、心筋梗塞は全体として死亡率が肺がんに比べて高いため、寄与危険は心筋梗塞のほうが肺がんよりも高い。

喫煙者が禁煙した場合のリスクの低下は心筋梗塞よりも肺がんのほうが大きい。

一方、もし、集団全体が禁煙に成功したら、それが死亡リスクの改善に及ぼす効果は心筋梗塞のほうが肺がんよりも大きい。

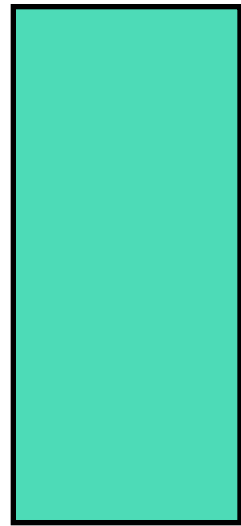
# 日本人における喫煙と肺癌の関係 (厚生労働省多目的コホート研究)

## 相対危険と集団寄与危険

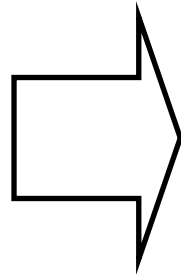


# Follow-up (追跡) Loss to follow-up (dropout: 脱落) がこわい

Eventを待つタイプ (死亡・循環器疾患/がん発症など)



追跡



健康

発症

死亡

転出

消息不明

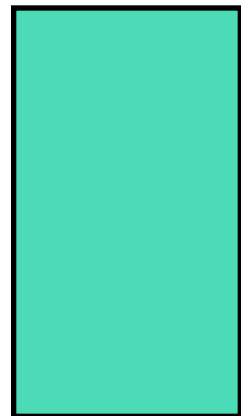
多くの場合

発症をeventとする場合  
のコホート集団

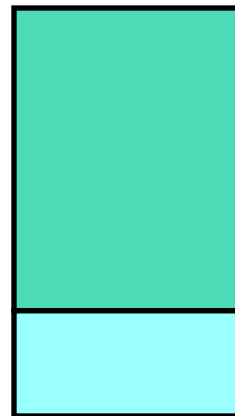
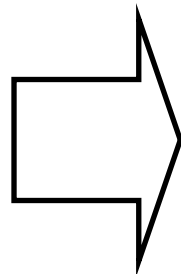
死亡をeventとする場合  
のコホート集団

ベースライン

変化を見るタイプ (高血圧・糖尿病・認知症など)



追跡



再調査

参加

非参加

原則的にベースライン  
と同じ調査であること

コホート集団

# Eventの見つけ方（死亡）

追跡用のデータベース

氏名、性別、生年月日、現住所

 Death

住民票

保健所の死亡小票

厚生労働省大臣官房統計  
情報部のデータベース

## 長所

全国どこでもできる

死亡日時がある

生年月日がある

## 短所

氏名と現住所が必要

お金がかかる

拒否例がある

5年で消える

死因は不明

死亡した市町村がわか  
らない

## 長所

氏名がわかる

死亡場所がある

死因がある

死亡した市町村がある

生年月日がある

## 短所

研究に参加している保健所  
からしか得られない

死因は最終的かつ標準化し  
たものではない

死因の詳細は不明

## 長所

全国のデータがある

最終的な標準化した死因  
がある

死亡した市町村がある

死亡した日時がある

生年月日がある

## 短所

利用は困難（可能）

氏名はない

（注意）佐々木のやや古い知識によります。その後変わった可能性があります。

# Eventの見つけ方（発症）

追跡用のデータベース

 Incidence

氏名、性別、生年月日、現住所

対象者への調査

診療記録の閲覧

## 長所

対象者以外への調査が不要

## 短所

対象者への調査が必要

非回答者

信頼度が低い

## 長所

疾患の詳細がわかる

## 短所

対象者以外も調べて、対象者を抽出しないといけない

非協力医療機関

協力必要医療機関が多く実際には無理

診療記録閲覧からのみ

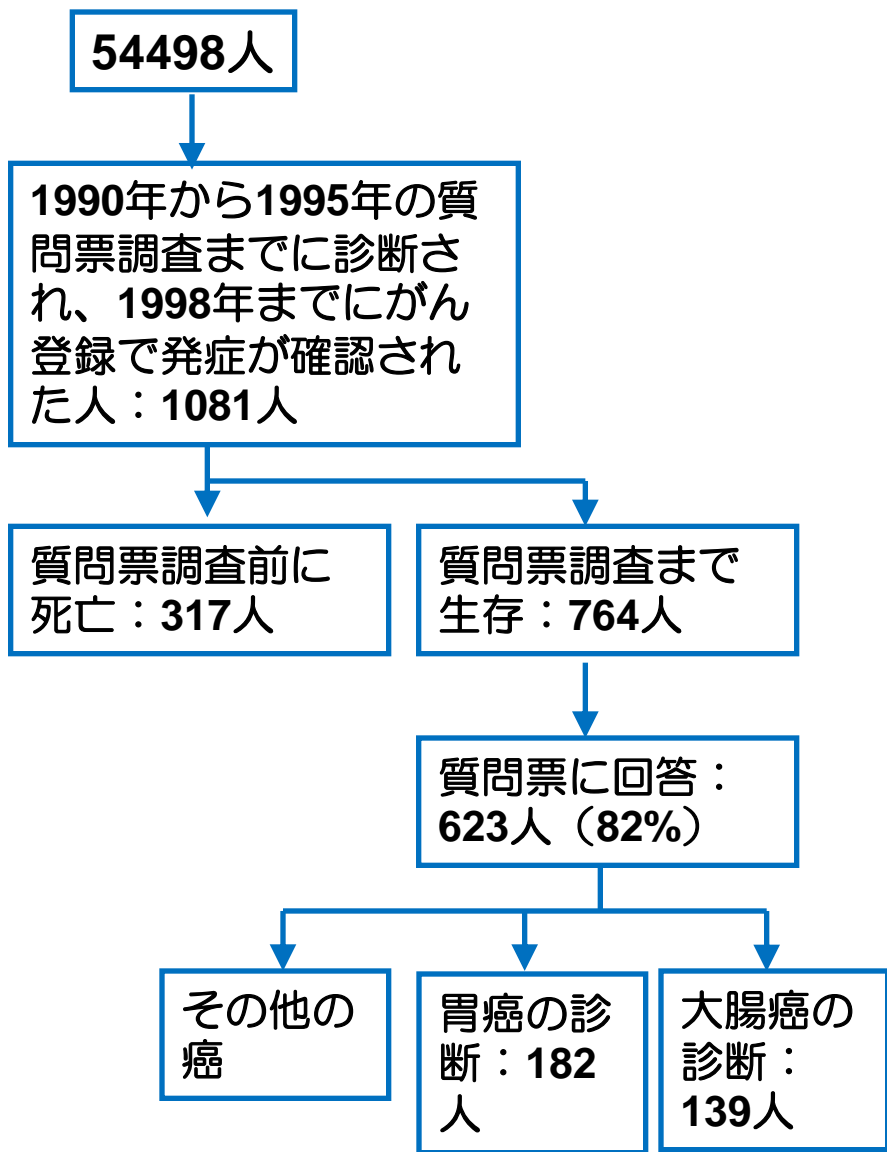
未受診

重複受診

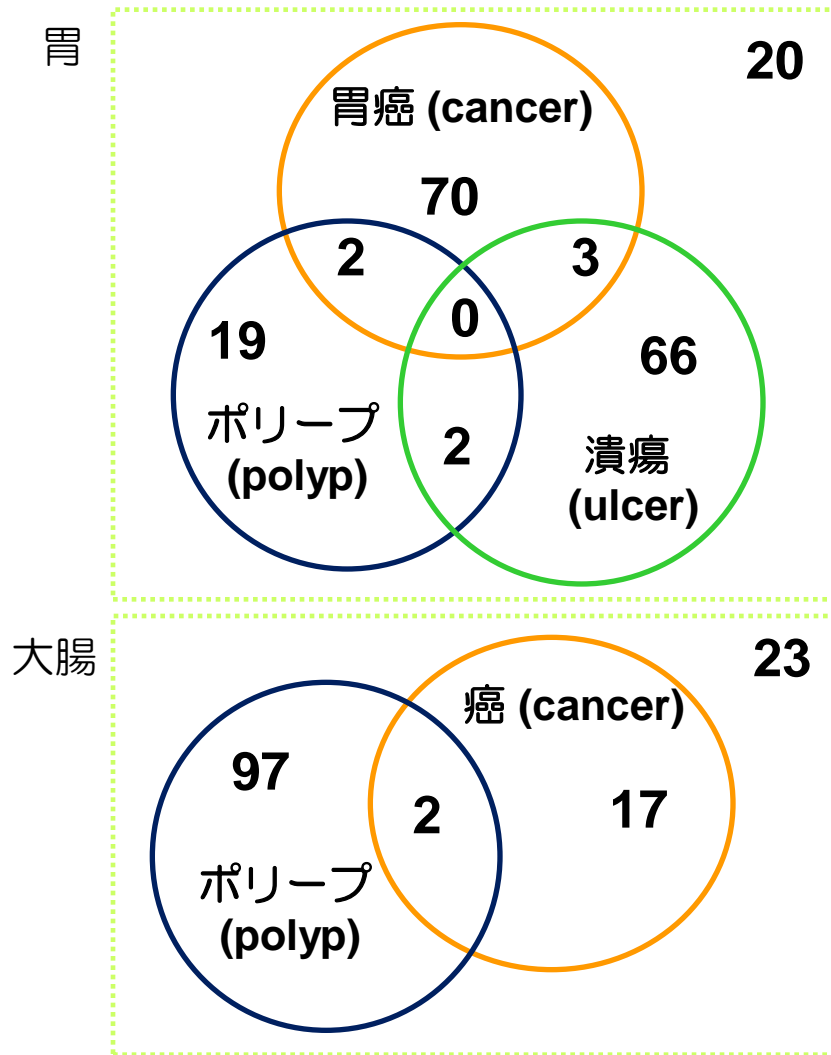
	協力病院	非協力病院
対象者	入手可能	得られない
対象者以外	不要だが調べてしまう（倫理的に問題）	

# 対象者にEventを尋ねることの信頼性

**Incidence**



がん登録で癌の発症が確認された人における自己申告の病名別の人数





# 同じ質問を用いること（標準化）がいかに難しいか？

#4851. Sobue, et al. Int J Cancer 2002; 99: 245-51.

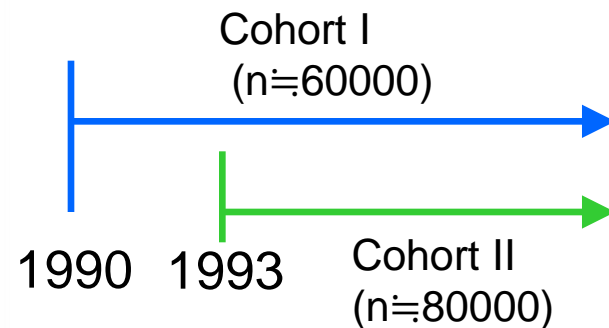
## QUESTIONS FOR CIGARETTE SMOKING

### Cohort I

Have you ever smoked cigarettes? yes/no  
 If yes, how old were you when you began smoking? \_\_\_years old  
 Are you currently smoking cigarettes? no/yes/sometimes  
 If no, how old were you when you quit smoking? \_\_\_years old  
 How many cigarettes per day do (did) you smoke? \_\_\_cigarettes/day

### Cohort II

Are you currently smoking cigarettes? yes/no  
 If yes, how many cigarettes per day do you smoke? \_\_\_cigarettes/day  
 How old were you when you began smoking? \_\_\_years old  
 If you no longer smoke, but smoked before,  
 How old were you when you quit smoking? \_\_\_years old  
 How many cigarettes per day did you smoke before? \_\_\_cigarettes/day  
 How old were you when you began smoking? \_\_\_years old



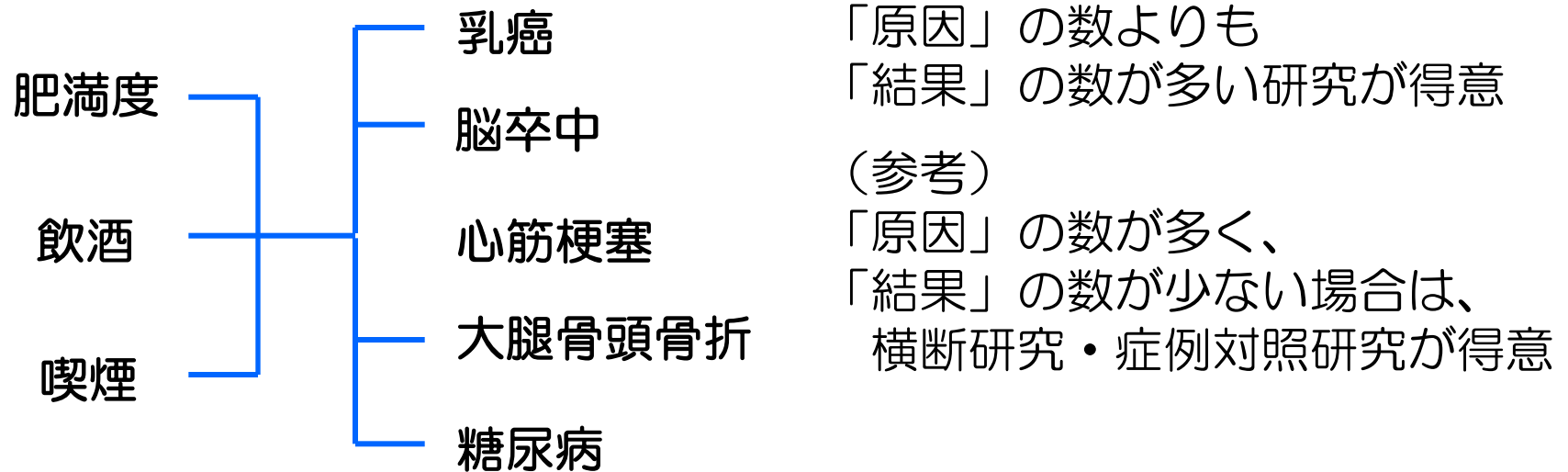
2つのコホートで質問文が微妙にちがう。

データを統合するときに苦労した。

はじめの質問だけに答え、残りが欠損（未回答）だと

...	現在喫煙	現在喫煙
	過去喫煙	過去喫煙
	喫煙歴なし	喫煙歴なし

## コホート研究の魅力：Multiple exposures & multiple outcomes



必要人数が結果因子によって異なる

結果因子の調べ方が結果因子によって異なる（実施可能性を調べておくこと）

結果因子によって交絡因子が異なる

結果因子が多いという魅力に目を奪われてはいけない。難しい点がたくさんある。

原因として何を調べるべきかを決めるのが難しい（結果の出現が未来だから）。

コホート研究の腕の見せ所は、「原因の予想」である。

なのに、原因の予想は不十分なままに、急いで始めたがる。

## コホート研究の成立条件

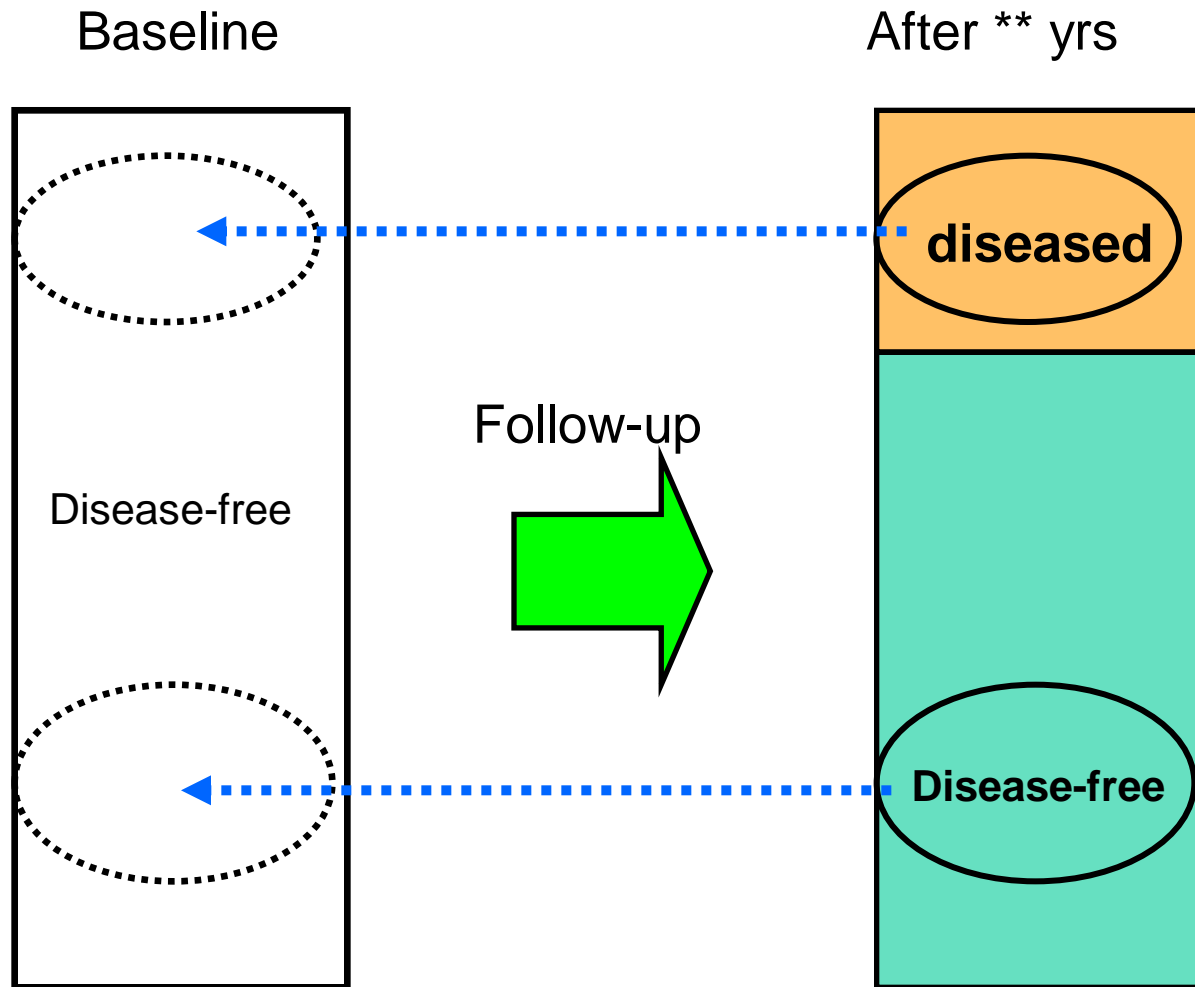
---

- 必要にして十分な数と内容の原因（曝露要因：exposures）を（対象者特性・交絡因子も含めて）調べられるか？（その測定の信頼度は保障されているか？）
- 結果要因（outcomes）を調べる（見つける）方法は確立しているか？（妥当性は？）
- 追跡体制は確立しているか？（その精度は明らかになっているか？）
- 必要な対象者数とイベント数は確保できる見込みはあるか？（必要な対象者数とイベント数は明らかになっているか？）

---

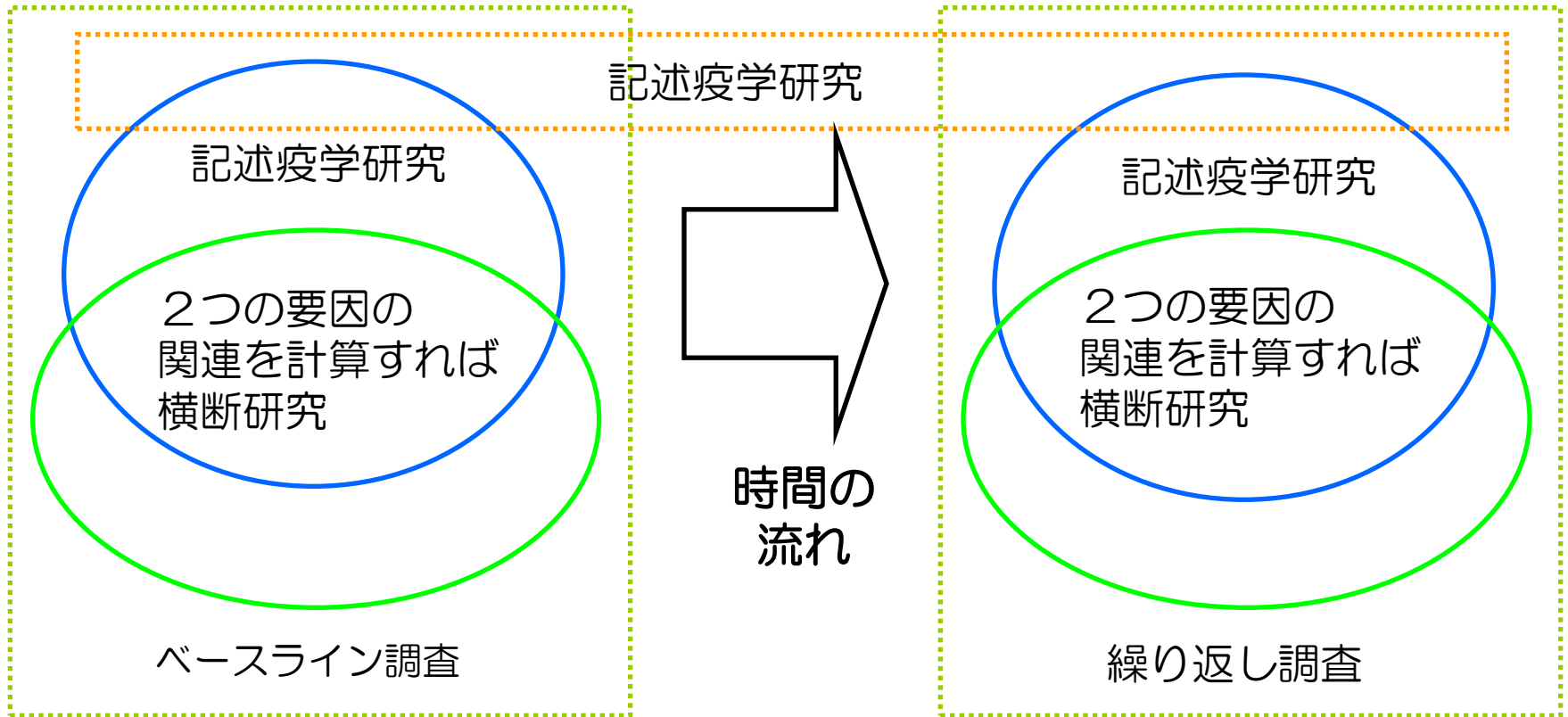
これらが満たされないままでの実施は極めて危険である。

# Nested case-control study (コホート内症例対照研究)



ベースラインで収集した試料（血清など）を有効に利用できる。

- ある変数の頻度、分布を観察すれば、記述疫学研究。
- 一時点のデータを取り出して、原因と結果の関連を検討すれば横断研究。

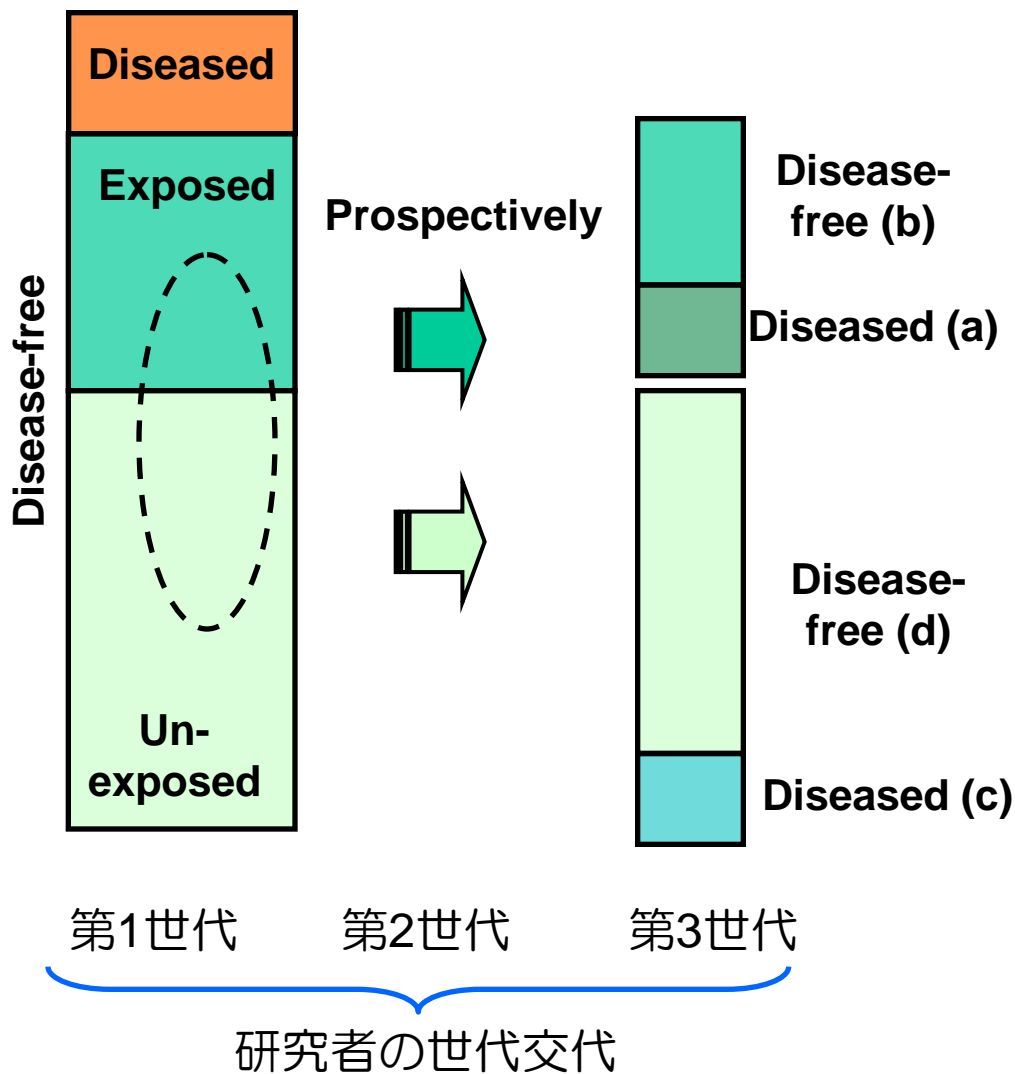


しかし、欲張らないこと。ほとんどのコホート集団は特殊な集団である。

しかも、コホート研究でできる調査内容は横断研究のものより粗いのがふつう。

さらに、横断研究より因果の仮説設定が甘くなりがち（未来予測だからしかたがない）。

# Cohort study: Marching towards outcomes



あくまでも exposure と outcome の関連を検討する研究である。等しくたいせつ。

結果は、ベースライン調査の質と、結果収集調査の質（追跡など）に依存する。

「論文を書く人には、ベースラインデータの質を上げるべきがない」ことに注意。

コホート研究の質は、研究計画者・研究のルールを敷いた人によって決まる。+ 追跡をしている人の熱意。

研究者の世代を超えた連携である

## Framingham Heart Study started in 1948.

---

Framingham: a town 30km west of Boston with population of 28,000 (10,000 for 30-59 yrs)

6507 randomly sampled from the residential registry (30-62 yrs)

4469 agreed to participate + 740 volunteers

= 5209 total original cohort

Medical and lifestyle checkup every 2 years (including subjects out of Framingham area)

Fund: 6,198,599 USD (2002)

Publications: 3906 (accessed 2008/06/26)

なぜフラミンガムだったのか？

1920年代に結核の実験的コミュニティ研究が実施されていた。住民に疫学研究参加の経験があり、理解があった。人口動態が安定している。保守的な町。低い失業率。ほとんどすべての職種があった（アメリカの典型的な町）。豊富な経験と情報をもった医療従事者がいる。ボストンに近い。1944年クッシング総合病院開院（傷痍軍人用）で住民はボランティアをしていた。

1961 「危険因子（risk factor）」ということばが登場

**Dr. Dawber (the 1st director) was always saying,**

「フラミンガム研究のような研究に従事するときはあなたたち医師が参加者に何かをしてあげているのではなく、参加者たちこそが医師に何かをしてきているのだ。

君たちは参加者の皆さんをすばらしく偉大な人々であるという認識で健診に当たらなくてはならない。

君たち医師は彼らが提供してくれている自己犠牲に対して感謝し報いなければならぬ。」

**One of the participants told,**

「（休暇中にフロリダで病気になって）地域の病院に行ったときに、『フラミンガムから来た？』と告げたところ、その医師が『えっ！あのフラミンガム研究の？』と驚いたそうです。

その医師は彼女がフラミンガム研究の初代参加者だと聞いてとても感激したそうです。

そのあまりの感激ぶりに彼女は『まるで映画スターにでもなったような気分だった』と言っていました。とても誇らしげでしたよ。同じようなことを別の参加者もブラジルで体験したそうです。」

---



## コホート研究でしばしばみられる問題

---

- ベースライン調査の項目が少ない。対象者数が多い、対象者の動機付けがじゅうぶんでないなどの理由により、簡単な調査項目が好まれる傾向がある。
  - ベースライン調査内容が結果を見るころになるとふさわしいものでなかったことが明らかになってくることがある。
  - 脱落を防止し、高い追跡率を保つための方法が作られないうちに始めてしまう研究がある。
  - 結果因子に興味をもっている研究者が中心になって組織される傾向が強い（結果よりも原因を先に調べるにもかかわらず！）。・・・大問題！
    - 曝露因子が未測定。測定精度が低い。
    - 中間因子（ある意味結果）を曝露因子の代理因子として使うことが多い。
  - 担当者が異動してしまう場合がある。
  - 長期的な予算は確約されない場合が多い。
-

成果を公表するのは結果を測る人であるが、  
成果の半分以上は立ち上げた人の手腕（研究デザイン、曝露測定）に依存する。

- 原因と結果の時間的関係を考慮できる魅力的な研究方法。
- 一見完璧のようにみえる。しかし、困難がいっぱい。
- 曝露因子は信頼度できるか？  
結果因子は信頼度できるか？
- 信頼度を上げるために考えうる対策の多くは実現可能性が乏しい。
- 過去にとられたデータは変えられない、という悲しさがある。
- 使うとき・結果を理解するときの注意：  
対象者が多いことよりも、研究の質が重要。  
集団特性に注意。

---

本日の宿題： コホート研究

cohort, prospective, follow-up, followed