

「集団」と「測定方法」

疫学 (epidemiology)

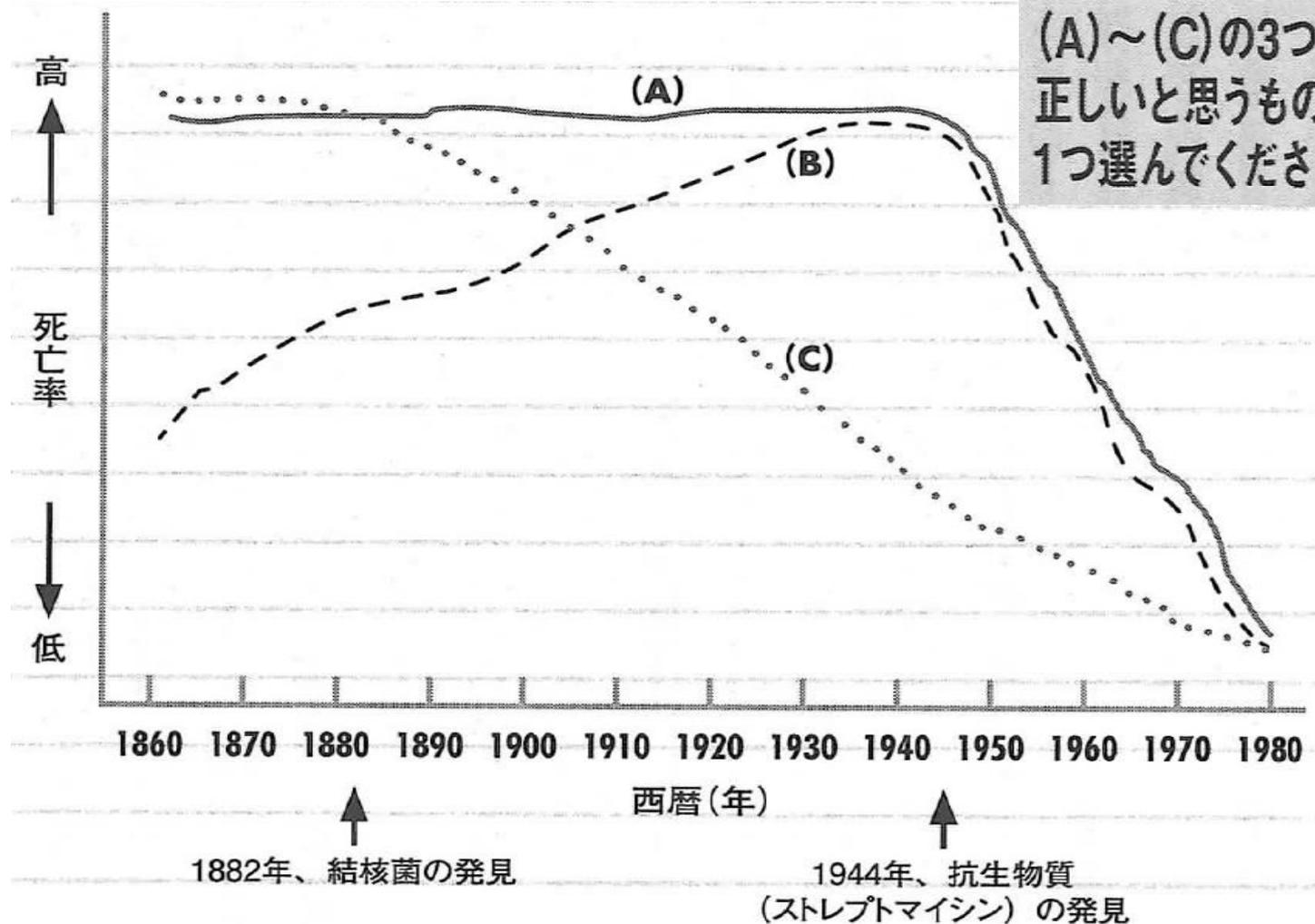
明確に規定された人間集団に出現する健康関連のいろいろな事象の頻度と分布およびそれらに影響を与える要因を明らかにして、健康関連の諸問題に対する有効な対策樹立に役立てるための科学

The study of the distribution and determinants of health-related status or events in **specified populations**, and an application of this study to control of health problems

結核死亡率の推移

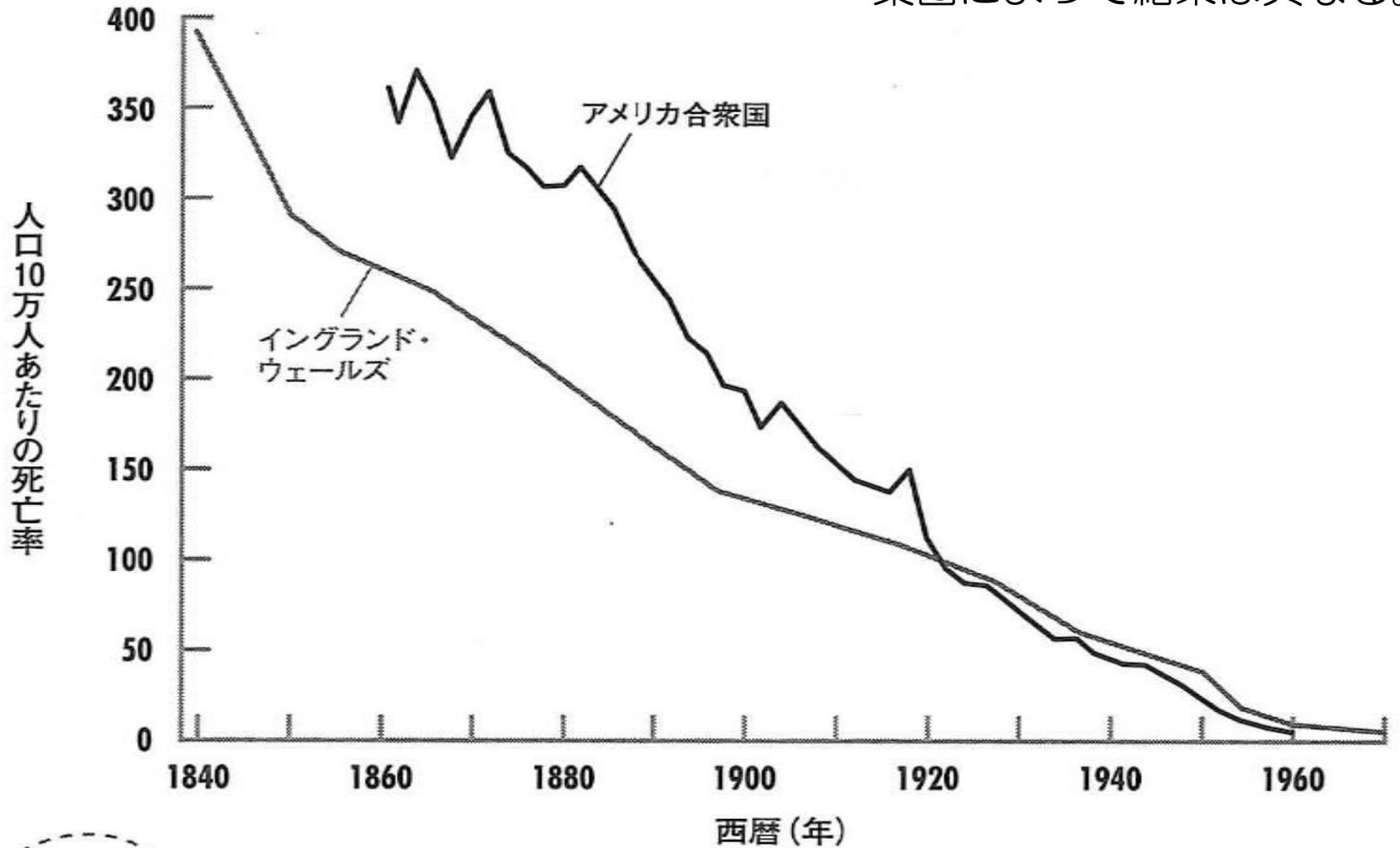
診断精度の問題はあるものの...

下図はアメリカ合衆国の結核死亡率（人口10万人あたり）の推移の概形を示したものです。(A)~(C)の3つの曲線から正しいと思うものを1つ選んでください。

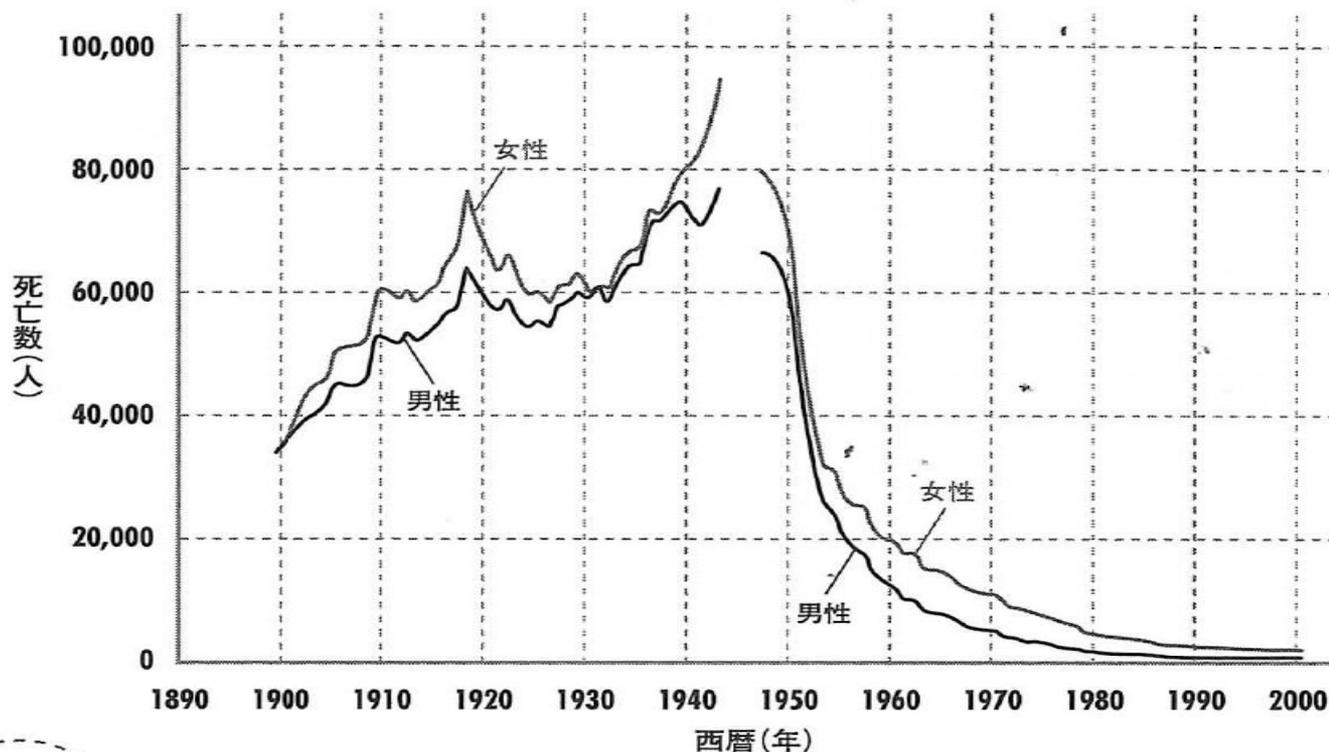


結核死亡率の推移（イギリスとアメリカ）

集団によって結果は異なる。



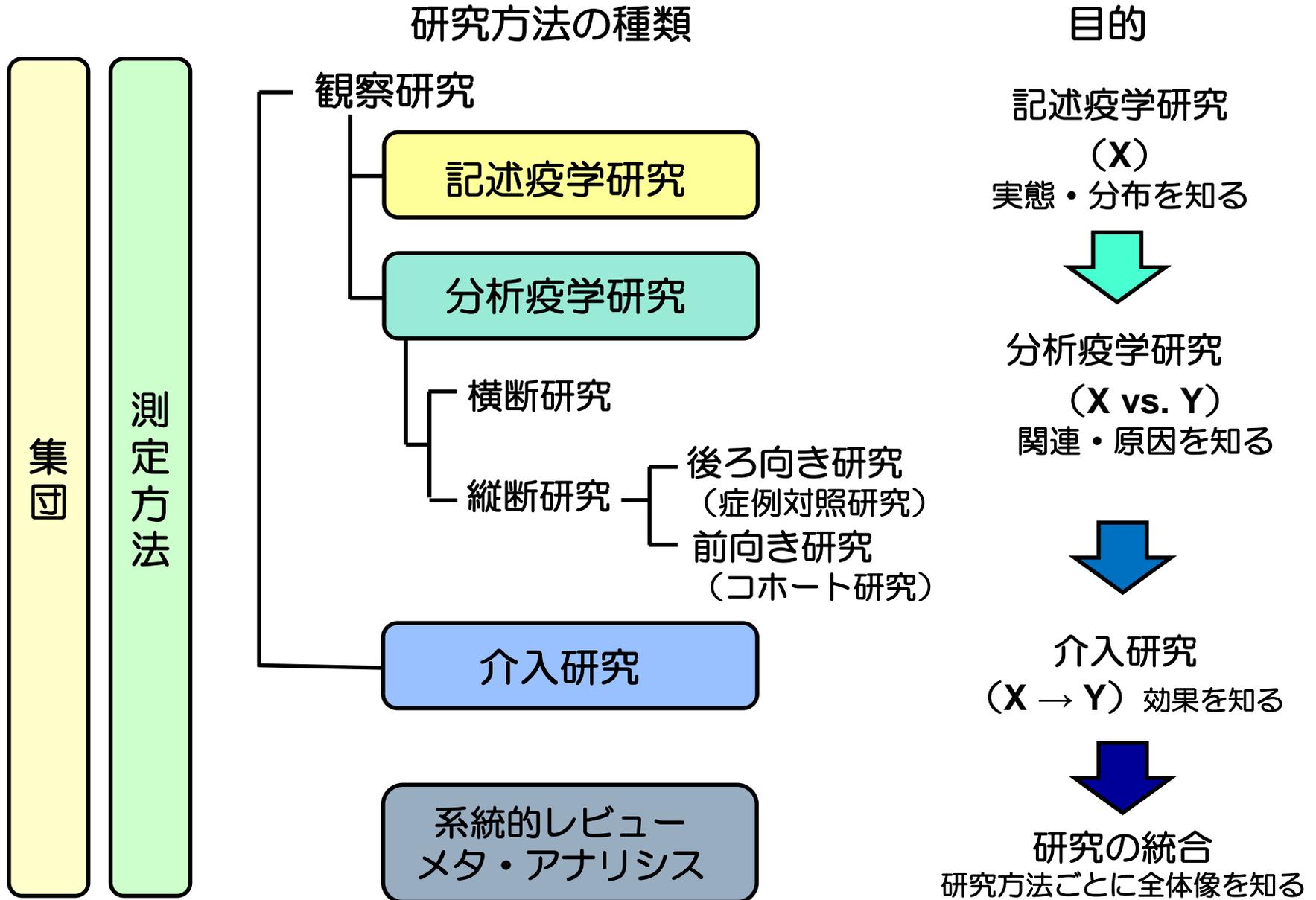
(注：1944～47年はデータが存在しない)



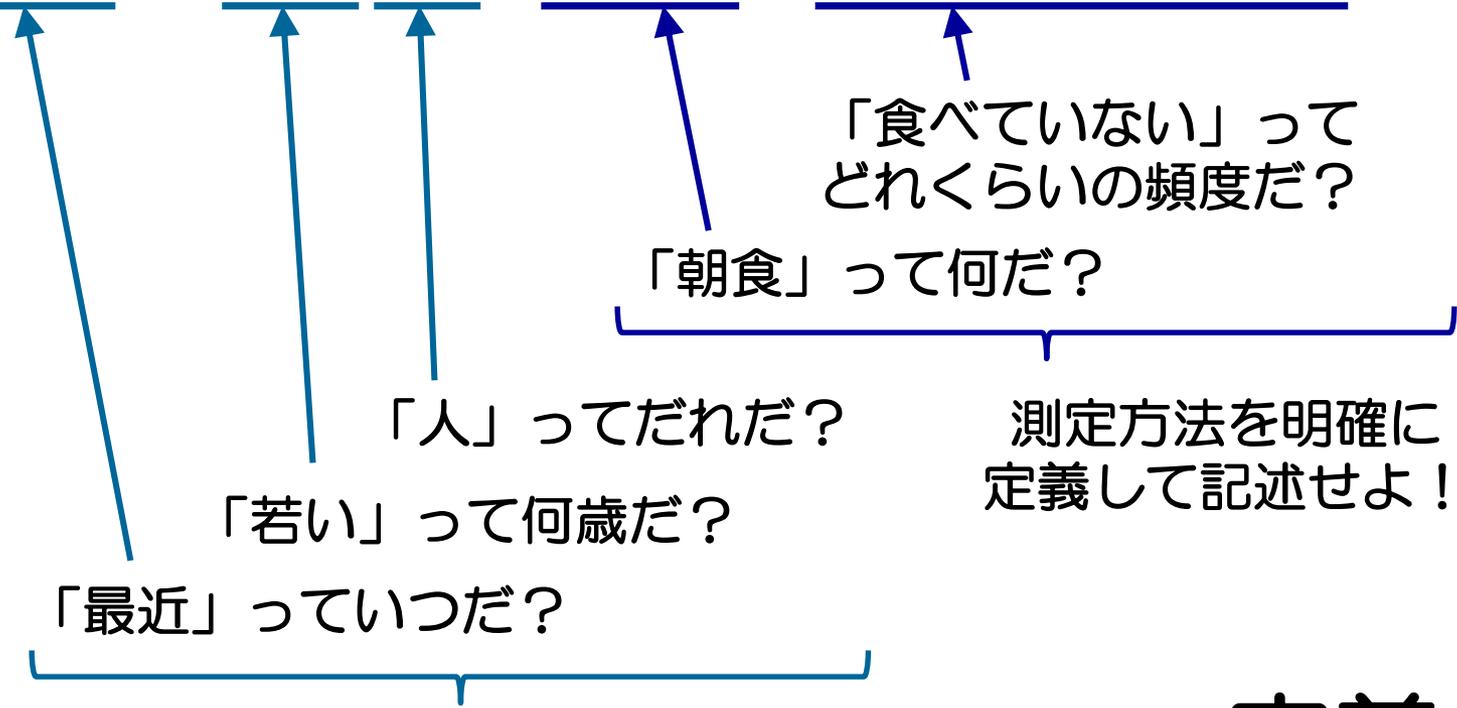
結核は結核菌で起こる。これは世界共通。しかし、その死亡率の推移は国（集団）によって大きく異なる。その対策も国（集団）によって大きく異なる。

疫学研究では、結果の前に必ず「どの集団か？」を見よう。

死亡率を集団間で横断的に比較すると「遺伝素因」で解釈されがちだが、推移を見ると、その解釈は変わってくることが多い。



最近の若い人は朝食を食べていない



定義！

論文や研究発表の序論で使われることが多い。
ここで誤るとその後の研究や発表はすべて意味がなくてなってしまう。

喫煙率 = 45%

集団は？

測定方法は？

こども？ / 成人？
男性？ / 女性？
健康者？ / 有病者？
だれが答え、だれが拒否したのか？

「喫煙」とは、
毎日1本以上吸っていることか？
それとも、1年以上休みなく吸い続けていたことか？
それとも、...
... これらのちがいで解釈と使い方は大きく異なる。

結果の数字だけでは何も言えない。
何も言ってはならない。

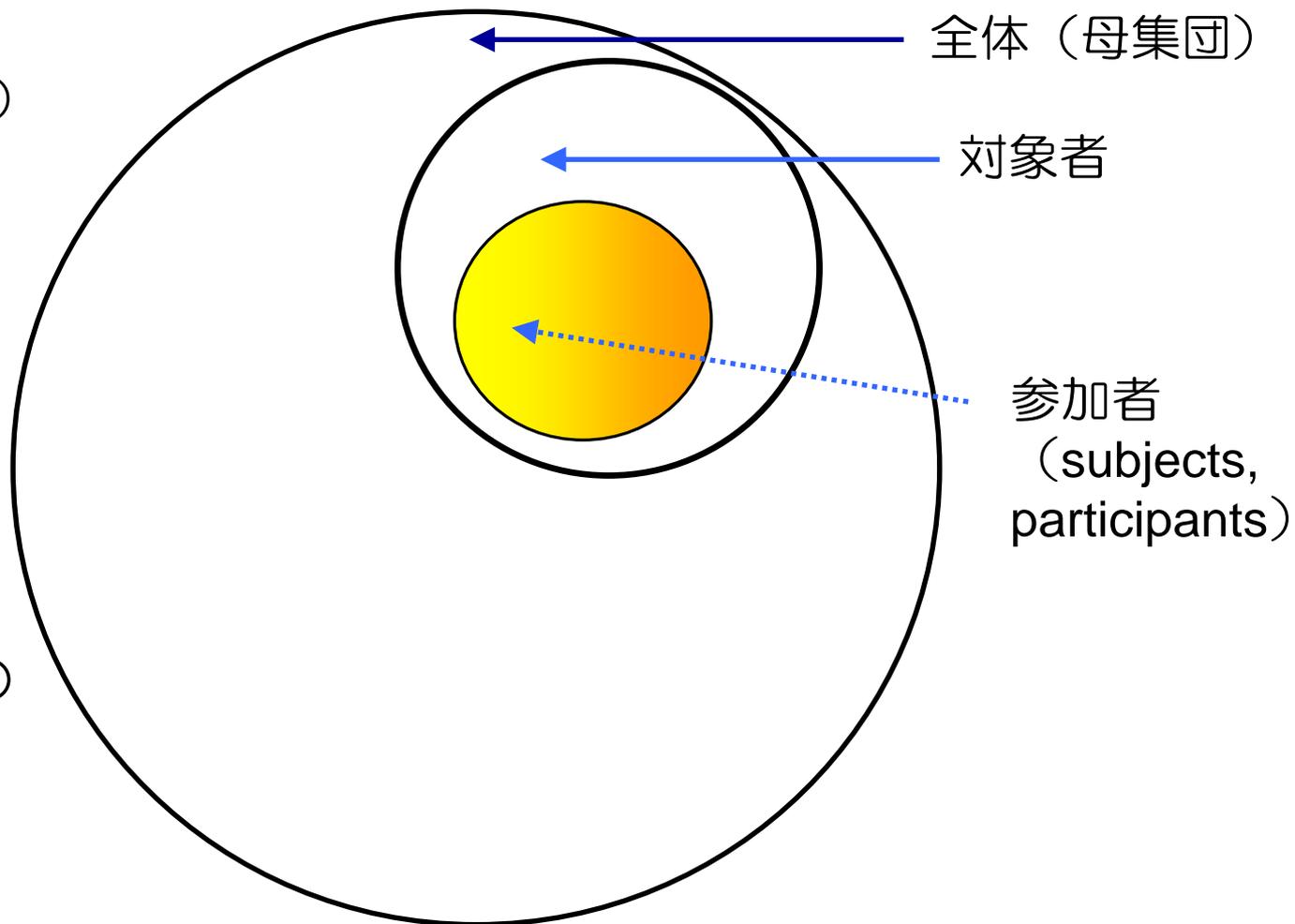
将来、
だれかが外部比較（この研究結果と他の研究の結果の比較）をするだろう。
そのとき、この結果を使ってもらうために、対象者特性と測定方法は必須。

対象集団 (target population)

対象者の抽出
(サンプリング)

被験者 ×
対象者 ○
参加者 ○

Patients ?
Subjects ○
Participants ○



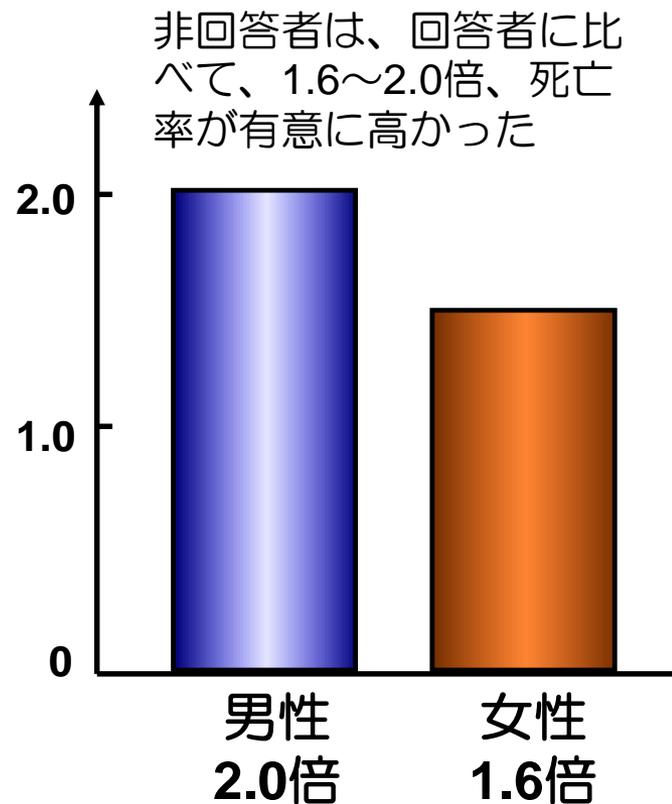
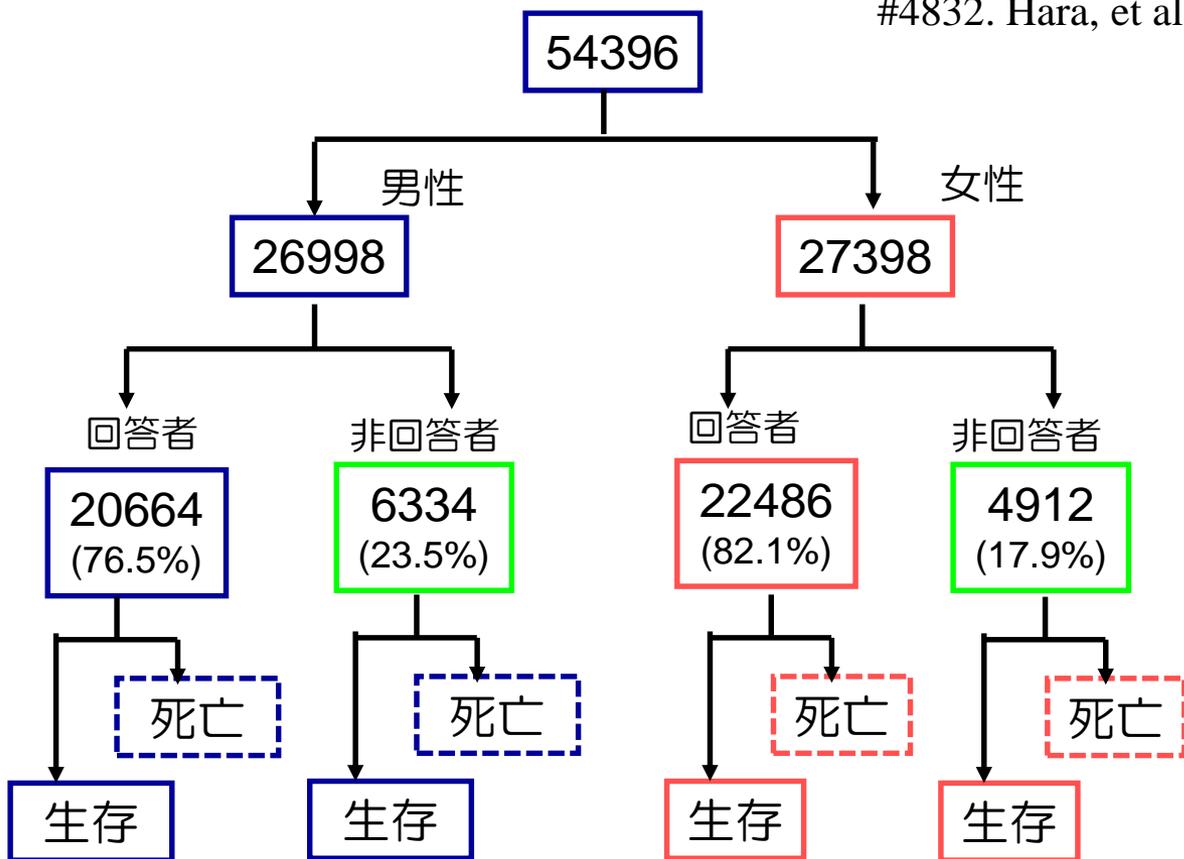
対象集団 (target population) とは調べたいと考えている集団のこと。
調べた (調査に応じた) 集団 (参加者) のことではない。

質問票調査を中年（40～59歳）5.4万人にした。

回答者4.3万人と、非回答者1.1万人について、調査後10年間の死亡率を比較した。

Comparison of cause-specific mortality between respondents and nonrespondents in a population-based prospective study: ten-year follow-up of JPHC Study Cohort I.

#4832. Hara, et al. J Clin Epidemiol 2002; 55: 150～6.



（解釈）病気がない人が質問票に回答した。健康的な生活の人が質問票に回答した。

（疑問）質問票の結果と疾病発症の関連を調べた。その結果を社会全体で使ってよいか？

先行研究・他の研究との結果の比較

序論（Introduction）・考察（Discussion）で必ず行う。

集団が類似 ⇔ 集団特性が類似

- 年齢・性
- 結果因子を大きく左右させたり、結果に関連していることが明らかになっている因子

「集団特性が類似した先行研究をいくつ発見できるか」は、論文の科学性を決める重要な要素である。

「集団代表性：population representativeness」をどこまで確保できるか？

集団代表性： 「サンプルで集団を代表できる」ということ。

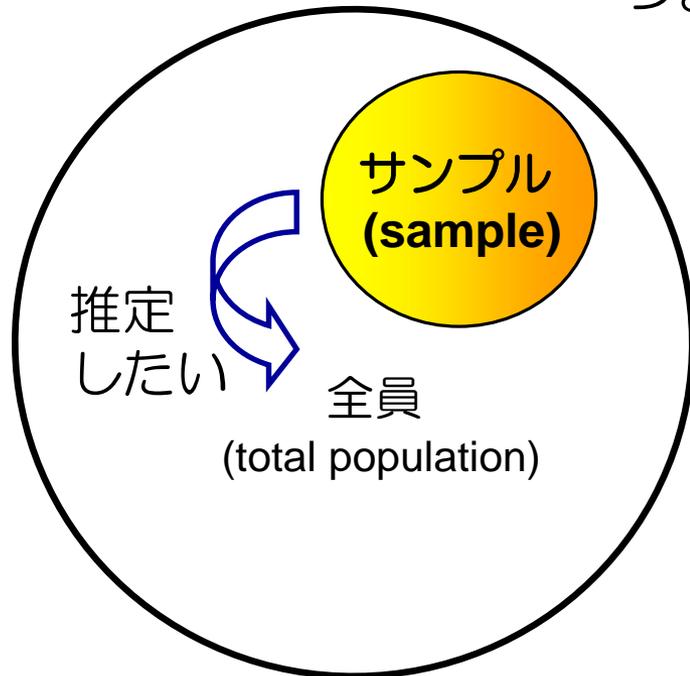
なぜ、集団代表性が求められるのか？

全員を調べられないから、その一部（サンプル）を調べる。

得られた結果を、集団全体から得られたものとして使いたいから。

得られた結果を、集団全体に適用したいから。

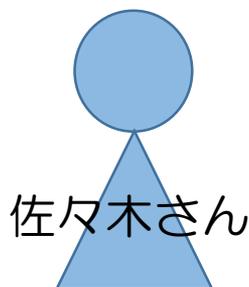
つまり、一般化（generalize）したいから。



「一般化」
...どこまでしたいのか？

あなたの頭は疫学的思考法に向いているか？

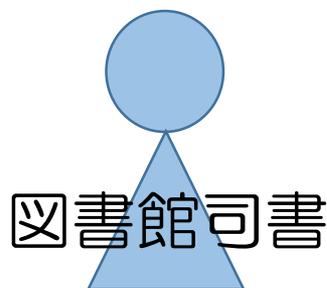
佐々木さんは、
とても内気で引っ込み思案だ。いつも頼りにはなる
が、基本的に他人には関心がなく、現実の世界にも
興味がないらしい。もの静かでやさしく、秩序や整
理整頓を好み、こまかいことにこだわる。



佐々木さんは、図書館司書でしょうか？
それとも農家の人でしょうか？

ダニエル・カーネマン。ファスト&スロー：あなたの意思はどのように決まるか？
ハヤカワ・ノンフィクション文庫。早川書房。P.20. (改変)

佐々木さんは、
とても内気で引っ込み思案だ。いつも頼りにはなる
が、基本的に他人には関心がなく、現実の世界にも
興味がないらしい。もの静かでやさしく、秩序や整
理整頓を好み、こまかいことにこだわる。

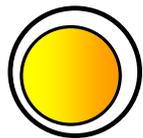


ここに図書館司書と農家の人が一ひとりずついます。
どちらかが佐々木さんです。どちらでしょうか？

(参考) ダニエル・カーネマン。ファスト&スロー：あなたの意思はどのように決まるか？
ハヤカワ・ノンフィクション文庫。早川書房。P.20. (改変)

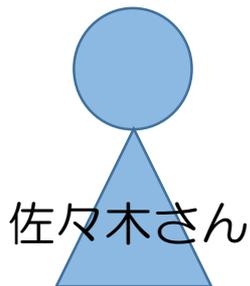
あなたの頭は 疫学的思考法に向いているか？

図書館司書



0.6万人

(注) 正確な数値ではありません。



農家
192.2万人

とても内気で引っ込み思案だ。いつも頼りにはなるが、基本的に他人には関心がなく、現実の世界にも興味がないらしい。もの静かでやさしく、秩序や整理整頓を好み、こまかいことにこだわる。

(参考) ダニエル・カーネマン。ファスト&スロー：あなたの意思はどのように決まるか？
ハヤカワ・ノンフィクション文庫。早川書房。P.20。(改変)

公共図書館経年変化 (日本図書館協会) 専門・兼任図書館司書・司書補：5574人 (2016年)

<http://www.jla.or.jp/library/statistics/tabid/94/Default.aspx>

農業就業人口 (農林水産省) 192.2万人 (2016年)

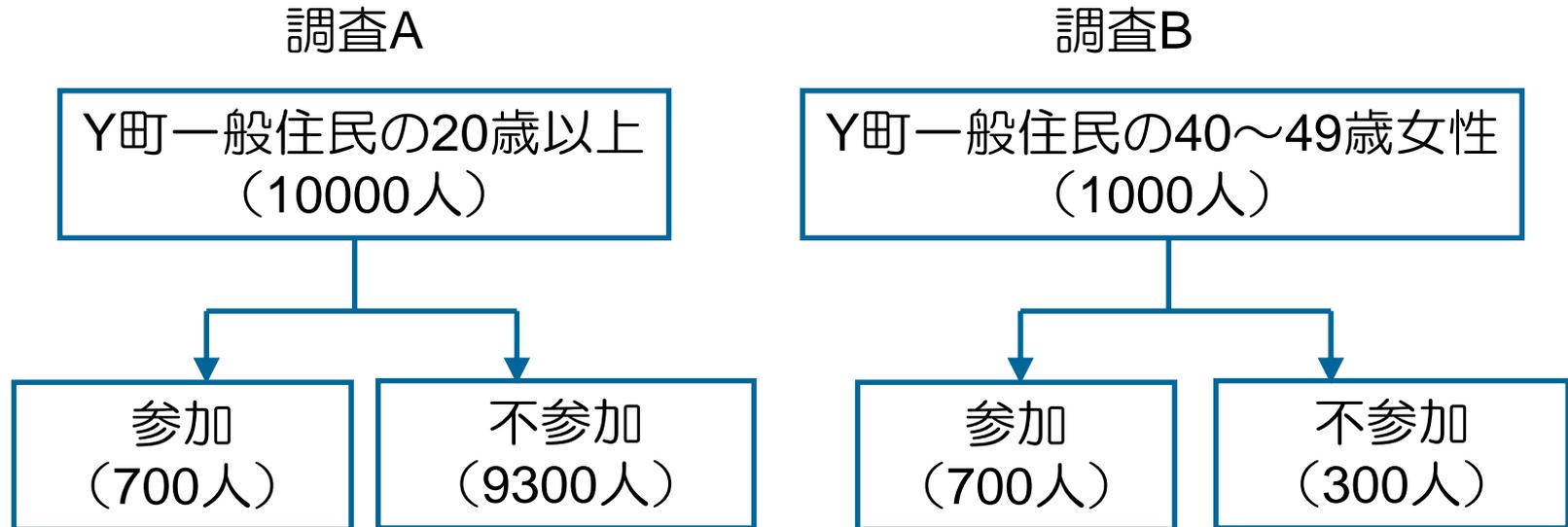
<http://www.maff.go.jp/j/tokei/sihyo/data/08.html> 2017/04/27 アクセス

対象集団の特性を広くとる・狭くとることの長所と短所

集団特性の広さ	広くとる	狭くとる
例	いろいろな県、20歳以上全員、男女両方。	A市在住、40～49歳、男性、治療を受けていない。
長所	結果を適用できる範囲が広い。	（類似の集団特性を有する集団ならば）外部集団との結果の比較可能性は高い。 交絡要因が少ない（真実を見やすい、解釈が容易）。
短所	（逆）	（逆）
注意点	<ul style="list-style-type: none"> • 人数が大きいことや一般化可能範囲が大きいことに目を奪われないこと。 • 参加率が特性によって異なる可能性を考えておくこと。 	<ul style="list-style-type: none"> • 一般化のむずかしさを自覚しておくこと。

「どちらを選ぶか」を計画時によく考えること。
シロかクロではなく、どちらを優先するかといった程度に考えるのが現実的。

「集団代表性の高さ」か「参加率の高さ」か？



両方とも参加者は700人。どちらの研究がよいか？

Aの結果は参加意欲が高い人にしか使えない。

Bの結果は40歳代女性にしか使えない。 ...こちらのほうが役に立ちそう。

なんとなくたくさんの人を集めようとせずに、目的を定めて、本当に必要な人にていねいをお願いするほうがよい。

しかし、「ていねいなお願い」ができない場合も多い。

「集団特性を広くとる」ことによる災い

「20歳以上（成人）700人を調べた結果、減塩を心がけている人は25%」
という結果が得られた。

（考えるべきこと）この結果は、この集団の役に立つか？
他の集団の役に立つか？

（疑問）

「減塩を心がける」という行動は血圧の影響を受けるのではないか？
すると、高血圧の有無（認知？）別にも解析しないとイケない。
血圧は年齢によっても大きくちがう。

すると、「年齢階級×高血圧の有無」別に、解析しなくてはならないか
もしれない。☞ おそらく、700人では足りない。

高血圧の有無	20～39歳	40～49歳	50～59歳	60～69歳	70歳以上
なし					
あり					

全体の数よりも、
層別解析したときのひとつの小集団の人数に注目する。

何をどのように解析したいのかをあらかじめ決めなければならないことに気付く。

何人必要ですか？

「解析最小単位の最少人数の集団」に注目する。
「調査対象集団全体の人数」ではない。

	20～29歳	30～39歳	40～49歳	50～59歳	60～69歳	70～79歳	合計
女性							
男性							
合計							

なのか？ それとも

	合計
合計	

(補足)

「70歳以上（上限なし）」は
あまり好ましくない。

なのか？ で、必要人数は異なる。

どのくらい年齢幅を広げて解析してよいのか？

男女は区別しなくてよいのか？

と、

どれくらいの人数を確保できるか？

の双方から調査人数を決める。（他にも考慮すべき要因はある：後日）

回答率・協力率・応諾率（response rate）

平成14年 国民栄養調査結果の概要について

<http://www.mhlw.go.jp/houdou/2003/12/h1224~4a.html>

I 調査の概要

調査の目的

この調査は、栄養改善法（昭和27年法律第248号）に基づき、国民の食品の摂取量、栄養素等摂取量の実態を把握すると同時に栄養と健康との関連を明らかにし、広く健康増進対策等に必要な基礎資料を得ることを目的とする。

調査客体

平成14年国民生活基礎調査により設定された単位区から無作為に抽出した300単位区内の世帯及び世帯員を調査客体とした。

調査実施世帯数は、4,246世帯であり、集計対象者数は、下記のとおりである。

	総数
総数	11491
男性	5377
女性	6114

この文章のもっとも大きな問題点は何か？

平成17年 国民健康栄養調査 調査の概要

国民健康・栄養の現状 平成17年厚生労働省国民健康・栄養調査報告より 第一出版、2008.

調査の目的	調査の対象及び客体	調査客体の概要
健康増進法（平成14年法律第103号）に基づき、国民の身体の状況、栄養素等摂取量及び生活習慣の状況を明らかにし、国民の健康の増進の総合的な推進を図るための基礎資料を得ることを目的とする。	対象は、平成17年国民生活基礎調査において設定された単位区内の世帯の世帯員で、平成17年11月1日現在で満1歳以上の者とした。客体は、平成17年国民生活基礎調査において設定された単位区から、層化無作為抽出した300単位区内の世帯（約5000世帯）及び世帯員（約15000人）とした。	無作為抽出された300単位区のうち調査の協力が得られた世帯数は、3608世帯である。

調査内容概略	対象年齢	対象者数
身体調査	1歳以上	7278
血液検査	20歳以上	3874
栄養摂取状況調査	1歳以上	8895
生活習慣調査	1歳以上	9137

多くの調査では、母集団の特性はおろか、母集団の人数さえあまりわからない。否、知りえないことが多い。

対象者の参加状況のゆがみが、結果の解釈に影響を及ぼさないか？

平成17年 国民健康栄養調査 調査の概要

国民健康・栄養の現状 平成17年厚生労働省国民健康・栄養調査報告より 第一出版、2008.

	世帯数	対象者数
客体	5000	15000
協力	3608	8895
割合(%)	72	59

世帯数に対して対象者数のほうが協力率が低い。

協力率が高い世帯の世帯員数が少ない傾向にある可能性がある。

全国平均は世帯員の少ない世帯の結果に偏るのではないか？

どの程度の影響（問題）なのだろうか？

都市部の協力度が低いと仮定すれば、逆のはずのようにも感じる。なぜだろう？ >>>

考えるための仮想データ

世帯内 人数	対象		参加		
	世帯	参加者	参加率	世帯	参加者
5	100	500	0.8	80	400
4	100	400	0.7	70	280
3	100	300	0.6	60	180
2	100	200	0.5	50	100
1	100	100	0.4	40	40
合計	500	1500		300	1000
参加率				60%	67%

なぜだろう？

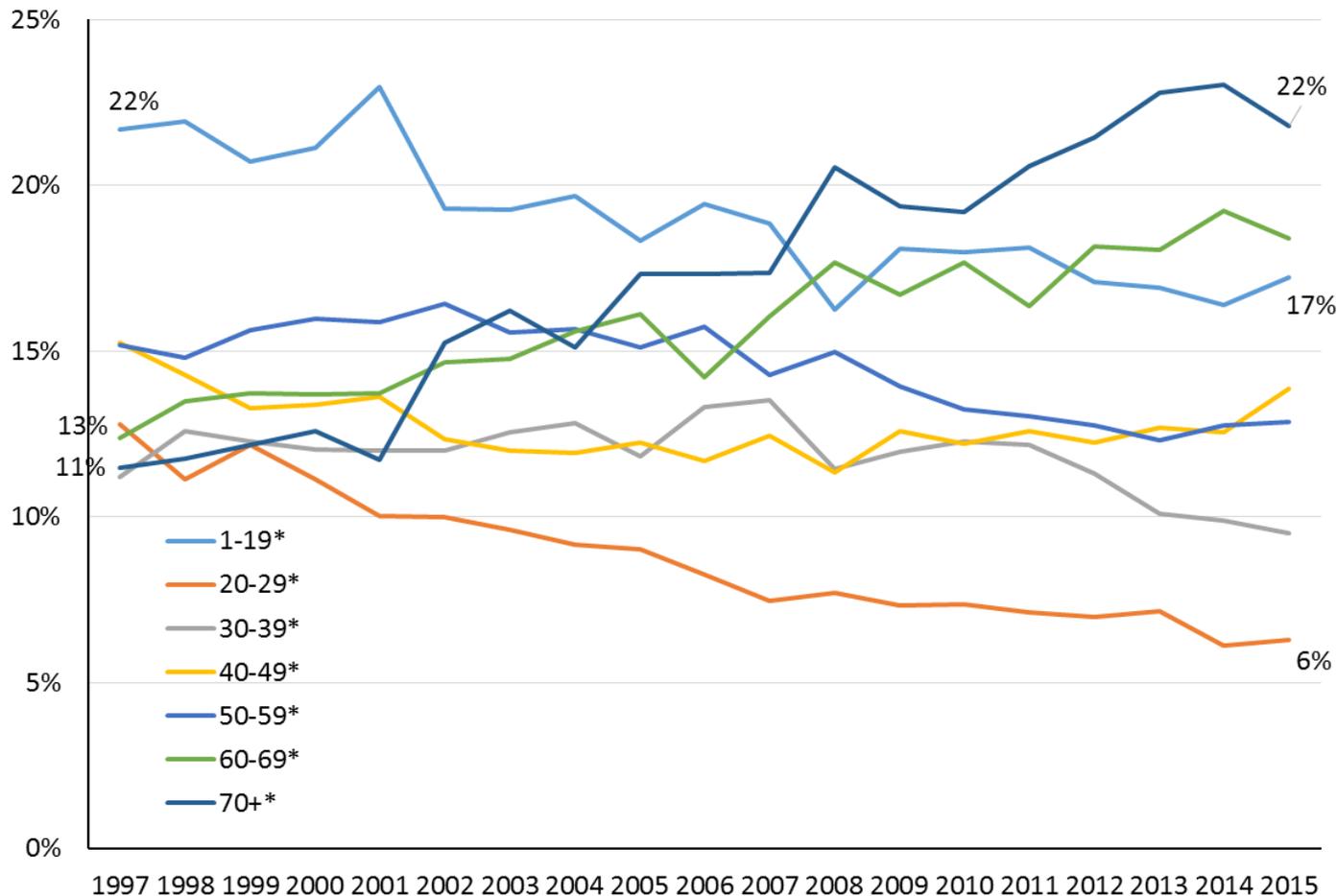
これは、粗探し（あらさがし）ではない。

調査を完全否定してはいけない。非難してもいけない。事情がある...はず。

疑問をもつこと、疑問をもって結果を解釈することがたいせつ。

調査を行った人に敬意を表すること。

報告者数の年齢階級別割合（％）



- 1～19歳、20～29歳、40～49歳が減って60～69歳と70歳以上が増えている。この間の結果（エネルギー摂取量の平均値など）を単純に比較してよいのか？
 - 参加率の変化か？
 - 人口構成の変化？
 - • • 年齢調整は不要か？

コホート研究における参加者数の推移

参加者数の推移（母子ペア数）：大阪母子保健研究

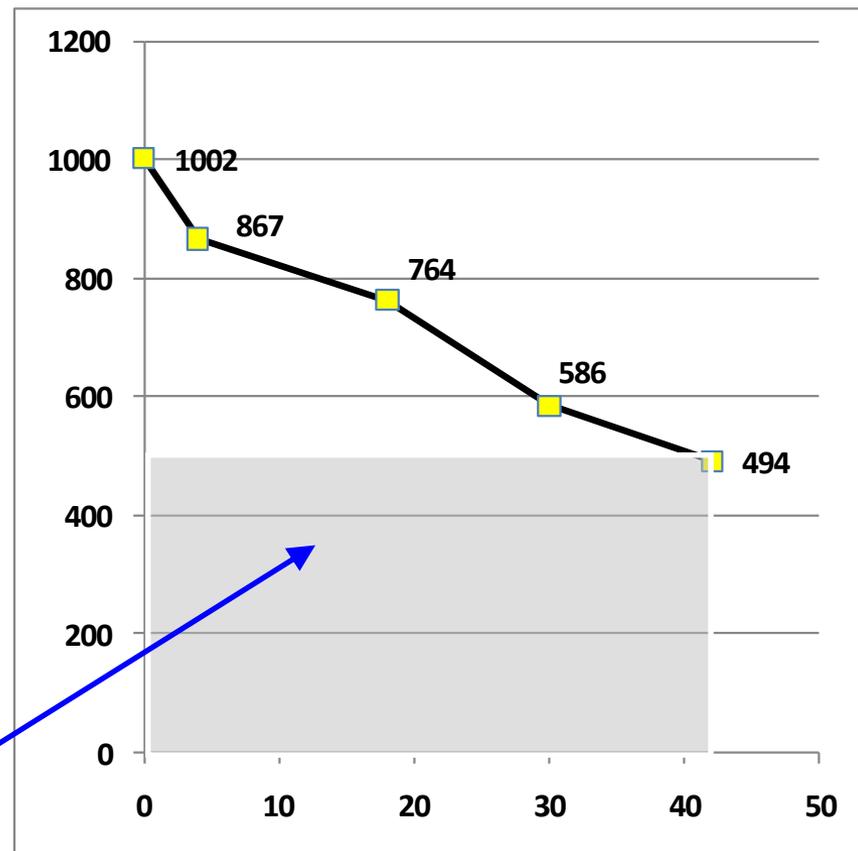
追跡研究において脱落は避けられない。

脱落による参加者数の減少は、「集団
代表性」の観点からも怖い。

どのような人が残り、どのような人が
脱落したのか？（それを知る方法はあ
まり存在しない）

どのような脱落なら問題は少ないか？
どのような脱落は問題が大きいか？

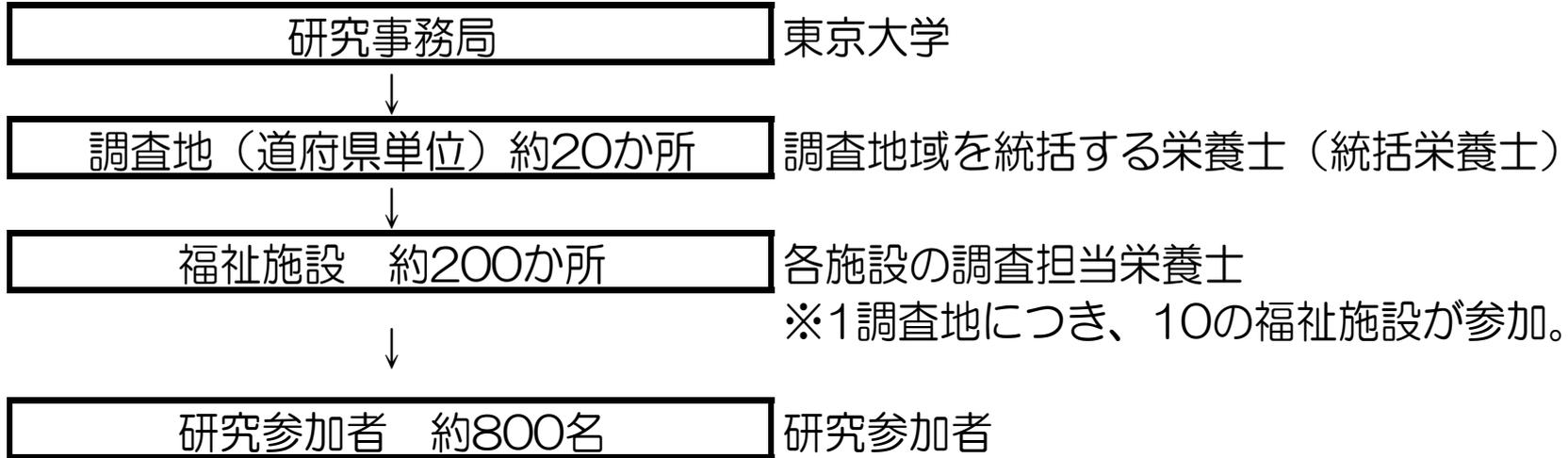
使えるデータはここだけ



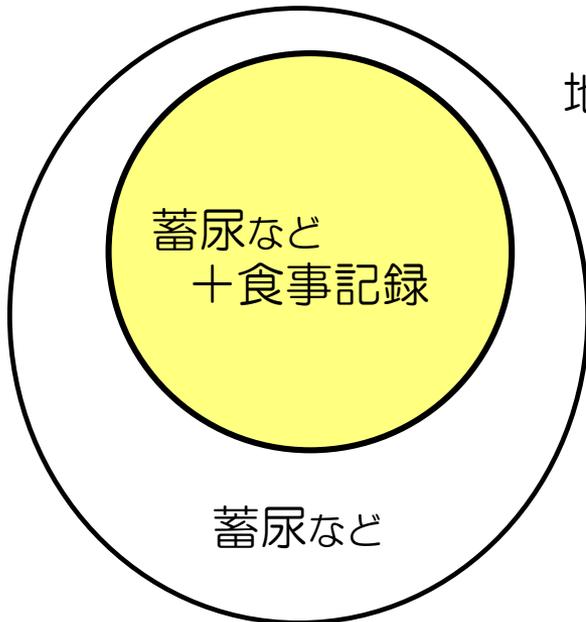
追跡期間 (月)

脱落した人と参加し続けた人のちがいが「原因と結果の関連」に影響しなければよいが、その保証はない（わからない）...場合が多い。
少なくとも、原因は比較が可能（必ず比較しておくこと）。

■ 塩研究（2012/02～03実施）：対象施設・対象者



※1参加施設から、原則4人の研究参加者が参加。
研究参加者数は1調査地域の中で40人になるよう調整し、施設ごとの参加者数は多少の増減あり。



地区ごとに

全体的人数



歳	20代	30代	40代	50代	60代
女性	4	4	4	4	4
男性	4	4	4	4	4

黄身的人数



歳	20代	30代	40代	50代	60代
女性	2	2	2	2	2
男性	2	2	2	2	2

■ 塩研究：参加施設、参加者数（結果）

		研究参加 施設数	研究参加 者全数	研究参加 男性数	研究参加 女性数	食事記録 無	食事記録 有
1	北海道	10	39	19	20	20	19
2	青森・岩手	10	40	20	20	20	20
3	山形	10	40	20	20	20	20
4	茨城	10	39	20	19	20	19
5	群馬	10	40	20	20	20	20
6	埼玉	10	40	20	20	20	20
7	神奈川	10	40	20	20	20	20
8	新潟・富山	10	40	20	20	20	20
9	静岡	9	35	17	18	18	17
10	大阪	10	40	20	20	20	20
11	兵庫	10	40	20	20	20	20
12	奈良	10	40	20	20	20	20
13	岡山	10	39	19	20	20	19
14	広島	10	40	20	20	20	20
15	山口	10	40	20	20	20	20
16	徳島	10	40	20	20	20	20
17	福岡・佐賀	10	39	20	19	20	19
18	熊本	10	40	20	20	20	20
19	大分	10	40	20	20	20	20
20	沖縄	10	40	20	20	20	20
	合計	199	791	395	396	398	393

地域、性、年齢
すべて互いに独立。
(調整の必要がない)

■ 塩研究：論文として投稿しました。査読者（reviewer）から対象者について質問を受けました。何でしょうか？

参加地区 (全20地区)

BDHQ (1回)
DHQ (1回)
質問票 (1回)



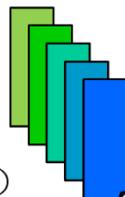
いろいろな質問票

24時間蓄尿
(2回)

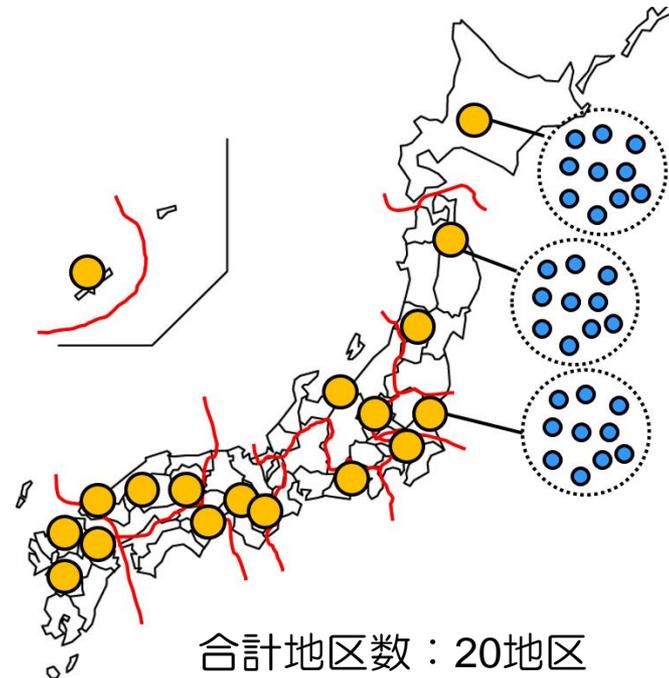


身長
体重
血圧
測定

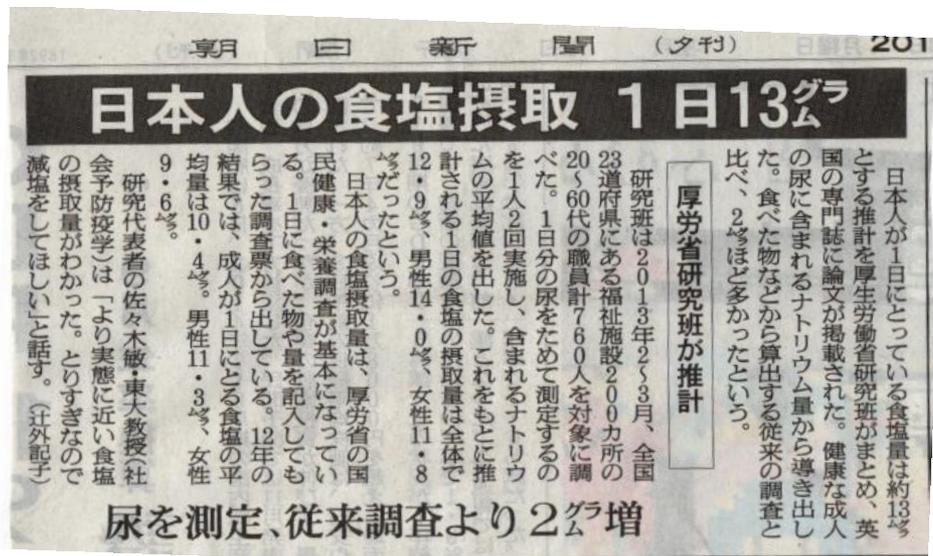
味覚閾値試験
試験紙法 (1回)



秤量式食事記録
(4日間)



合計地区数：20地区
合計道府県数：23道府県
合計施設数：200施設



J3679. 朝日新聞
夕刊 2014; 12月8日: 2.

「日本人の」と言うのはこんなにたいへん。

3) 研究にご参加いただける方の条件

研究にご参加いただく方は、いくつかの条件を満たしていなければなりません。どうしても一部の条件を満たせない、という場合には、統括栄養士にご相談ください。

●ご参加いただける方の条件

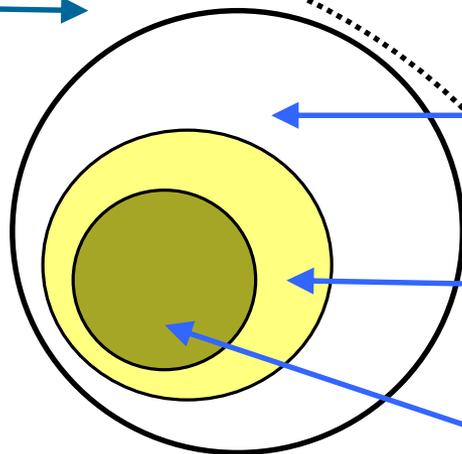
- 調査担当栄養士が勤務する福祉施設に勤務している方。
（常勤・非常勤・嘱託のいずれの勤務形態でも構いません。）
- 2013年2月1日現在で、20～69歳の方。
- 自分で質問票に答えられる方。

●ご参加いただけない方の条件

- 医療職（栄養士・管理栄養士を含む）の免許をもっている方。
（注：調査担当栄養士は表1の必要参加者数とは別に、希望によりご参加いただけます。）
- 明らかに特殊な食習慣（菜食主義など）の方。
- 2012年8月1日以後にその施設がある道府県、もしくは近隣の（境が接する）都道府県に引っ越ししてきた方。
- 現在ならびに過去1年間の間に、医師もしくは管理栄養士の指示のもと食事療法をしていた（いる）方。
- 調査開始時点で妊娠中、もしくは授乳中の女性。
- 過去に糖尿病の教育入院の経験がある方。

いろいろな「集団」

全体



調査対象者

調査協力者
(参加者)

解析対象者

人数 (n) が
どんどん減る

計画時に
考えるべき
順序

第2次栄養関連学科新入生調査 (2005/04)

#10163. Murakami, et al. Eur J Clin Nutr 2007; 61: 986-95.
#10948. Okubo, et al. Int J Obes 2008; 32: 541-9. 他

n=4679 (栄養士養成校: 54 大学/短大/専門学校)

この数が
わかっている。
すごい!

回答: 4394人

無回答: 393人

必要性・重要性を訴える
楽しみを作る (おみやげ)
負担軽減方法を考える。

女性: 4186人

男性: 226人

現実も
よく考慮する。

18~20歳: 4060人

他の年齢: 99人

少ない!
調査のていねいさが
ここに現れる。

調査が5月中に完了した。
食事調査データの信頼性が高い。
解析に必要なすべての変数がそろっている。

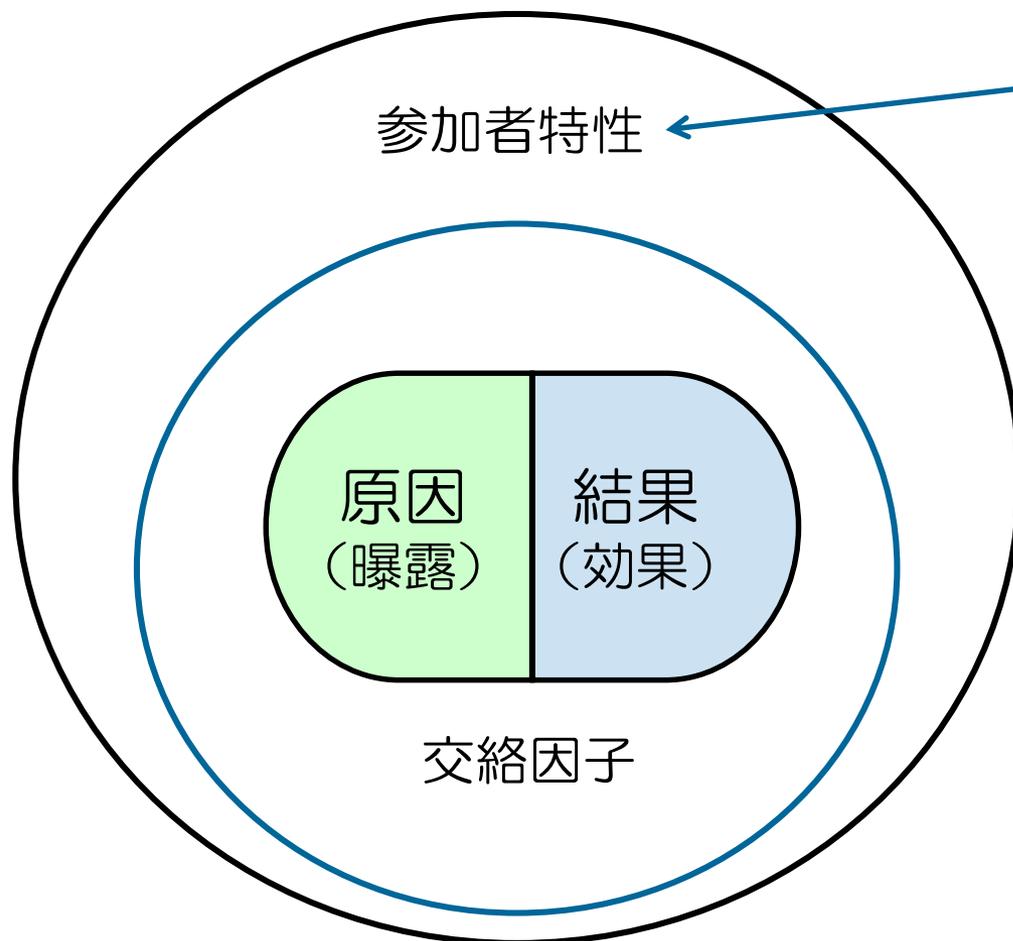
しなかった: 98人
その他: 23人
欠損あり: 12人

3931人 (4679人の84%, 4060人の97%)

調査人数はこの順序で決める!



相関・関連を調べる疫学研究（分析疫学研究）における4つの因子
介入研究にも通用する。



参加者の特性がよくわからないデータで、いくら「関連」を報告しても、その利用価値は低い。

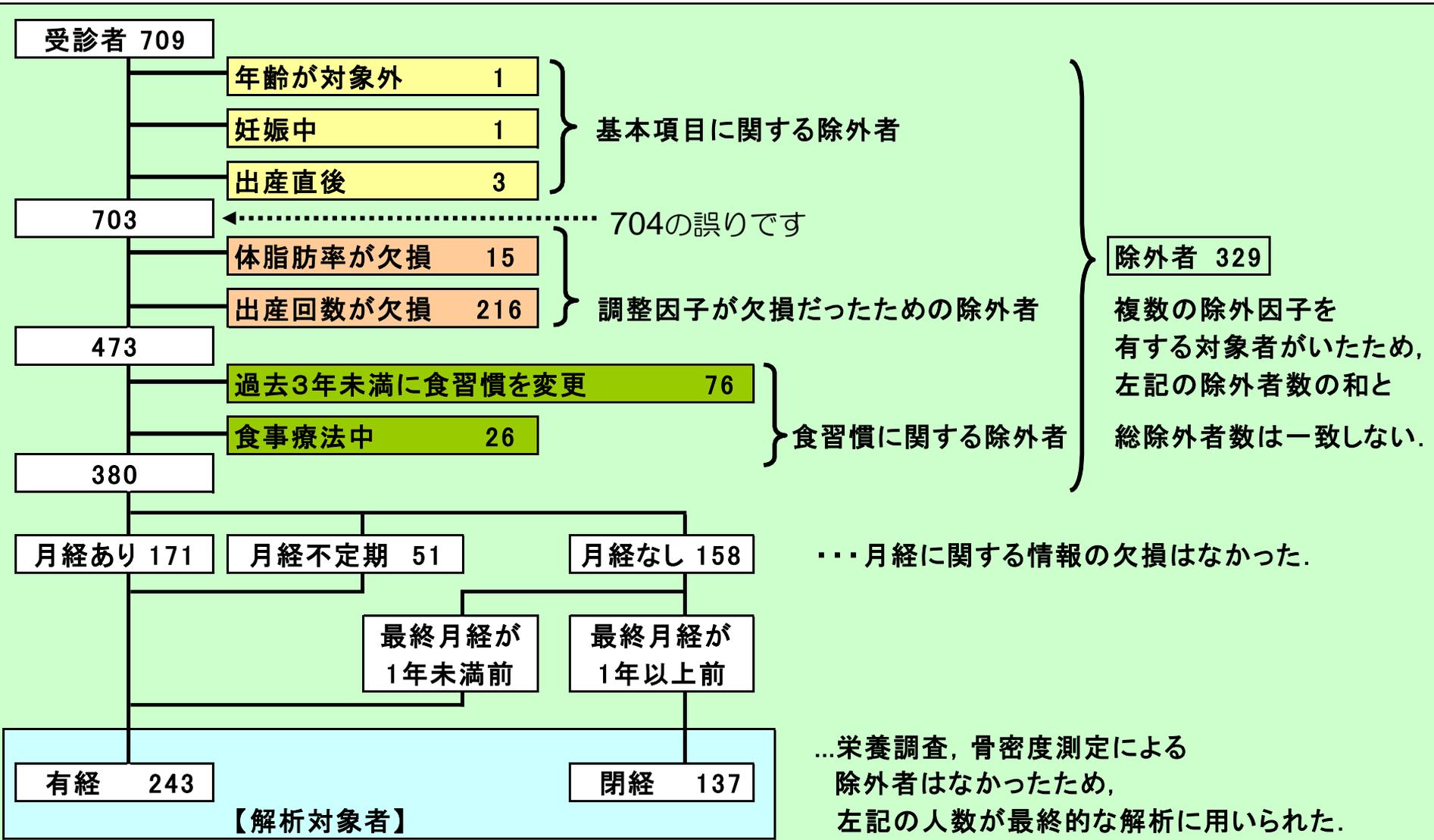
論文では、

表 1 (Table 1)

必要条件は、参加者特性。

十分条件は、原因と結果の関連。

調査完了から解析までの対象者人数の流れを示す図



#4696. Sasaki S, et al. J Nutr Sci Vitaminol 2001;47: 289-94.

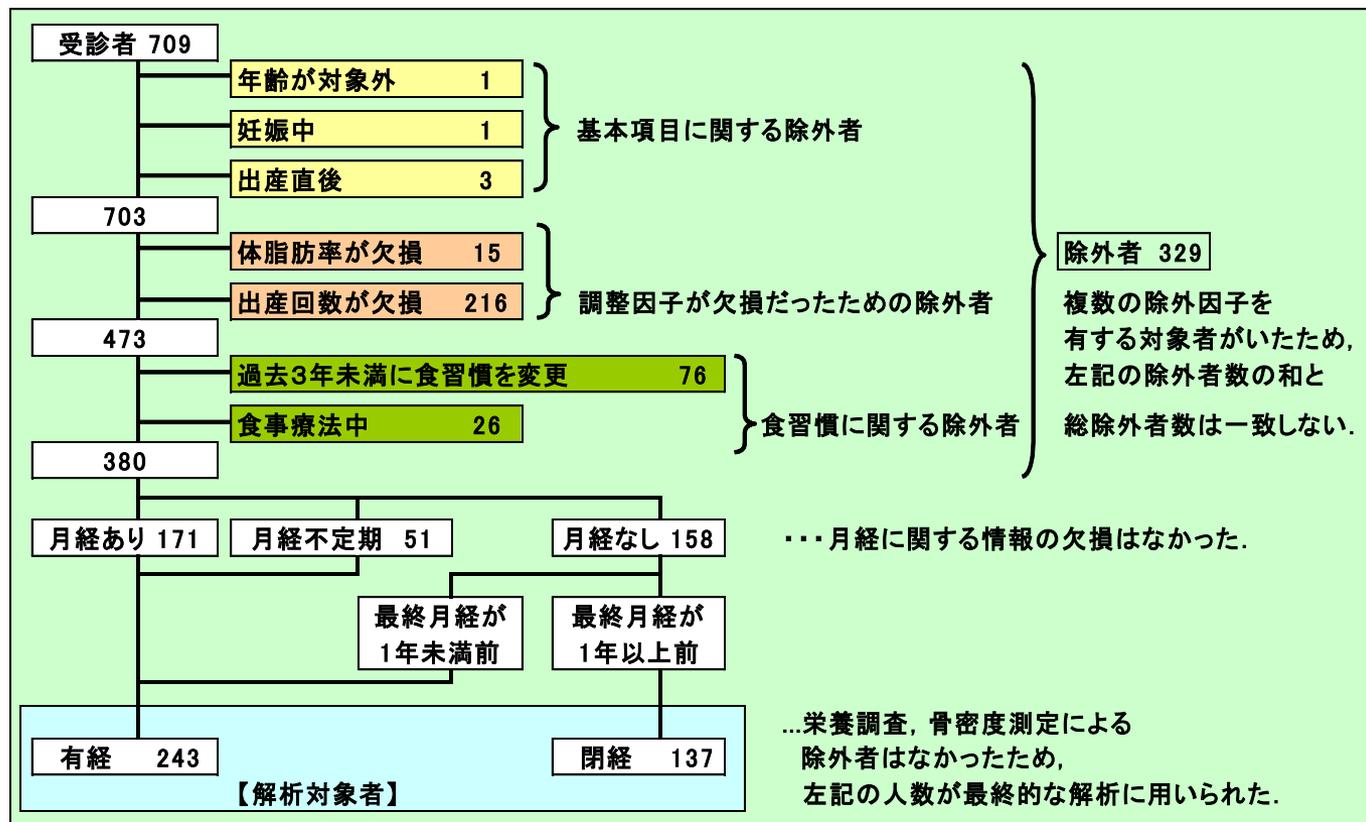
都合の悪いことも隠さずに発表・報告する姿勢がたいせつです。

掲載拒否 (reject) されました

Ca, Mg, Kなど骨代謝に影響する可能性が知られている栄養素の摂取量と骨密度との関連を、閉経前集団、閉経後集団に分けて解析し、その関連が両群間で異なることを示した。

ある国際誌から
掲載拒否 (reject)
されました。

なぜでしょうか？



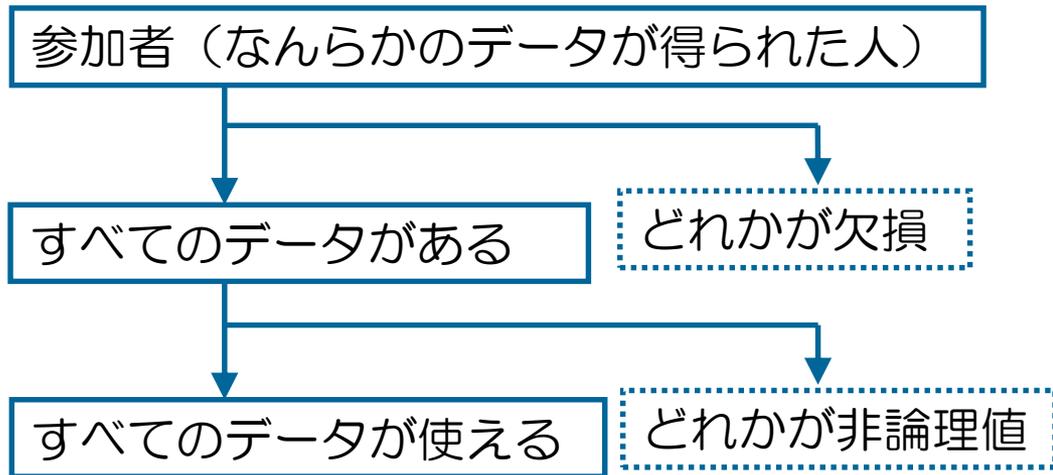
#4696. Sasaki S, et al. J Nutr Sci Vitaminol 2001;47: 289-94.

仮説を検証するために適した集団を用いること！

除外

調査ができたことと、解析できることは別

変数：a, b, c, d



集団代表性の崩れにつながることもある。

参加者側の問題
調査者側の問題

両方あるが、
調査者側の問題はなくすべき。

質の悪い調査（欠損データが多いなど）は除外が多い。

さらに、質の悪い調査は、除外すらしていない（気づいていない）。

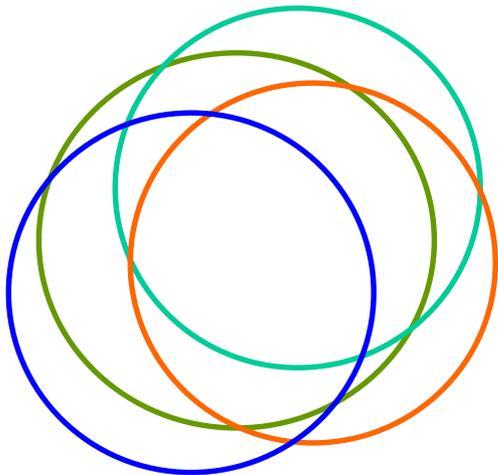
除外した人（人） × 調査に協力した時間（分/人） = 君が働いて返せ！

「調査に協力してもらったのに、データを使わないのは失礼」と考えたい。

👉 再調査ができない無記名調査でこの問題が特に大きいことに注意したい！

除外の方法・長所と短所

方法	長所	短所	お勧め度	分野
解析人数を固定する。	すべての解析で解析対象者が同じ (だれを解析したのかが明確) 結果を理解・説明しやすい。	人数の目減りが多い。	○	学術論文
変数ごとに除外者(解析対象者)を変える。	人数の目減りが少ない。	だれを解析したのかが明確でない。 (変数Xと変数Yで異なる人たちを対象としたことになる)。	×	(調査報告書)



(例) 国民健康・栄養調査では、体格の表と、食事調査結果の表では解析人数が異なる。

別々に用いるなら、これがよい
組み合わせて用いたいなら、よくない。

だれを解析したのか？
回答率(参加率)は？

欠損値・はずれ値

自分が調査者になった気持ちでみてほしい。

高齢者介護施設入所者を対象とした質問紙調査施設にある記録簿から転記をお願いした。

欠損と非論理値は最高6回までFAXを使って確認・再調査した。

人はまちがえる。

まちがえをとがめるのは目的ではない。

まちがえないような仕組みを作るべき。

費用対効果もしゅうぶんに考えること。

将来、疫学研究の仕組みを作ったり、指示を出したり、データを使ったりしたい人は、一度は現場を体験しておくこと。

およそ1万人のデータのなかから
見つかった非論理値

ID	身長 (cm)
7**1	0
195**32	14
101**1	14.05
68**2	15
65**7	318.9
45**6	1440
160**0	1531
202**7	9999
8**5	9999
45**0	9999
46**6	9999
116**3	15138

事件は会議室で起きてるんじゃない！ 現場で起きてるんだ！！

青島俊作 『踊る大捜査線』より

論文（報告）執筆時における重要注意

表1（table 1）は対象者特性（基本属性：basic characteristics）である

論文の表1は対象者特性（基本属性）を示す表

これがなくては、この後、どれだけ魅力的な結果を示されても、それがどのような集団から得られたものかがわからず、その結果の使い方もわからない（使えない）。

集団代表性が確保されているか否かの前に、対象者特性がきちんと書かれていないものがあまりにも多すぎる。

方法（methods）に脱落状況のフローチャートを入れる

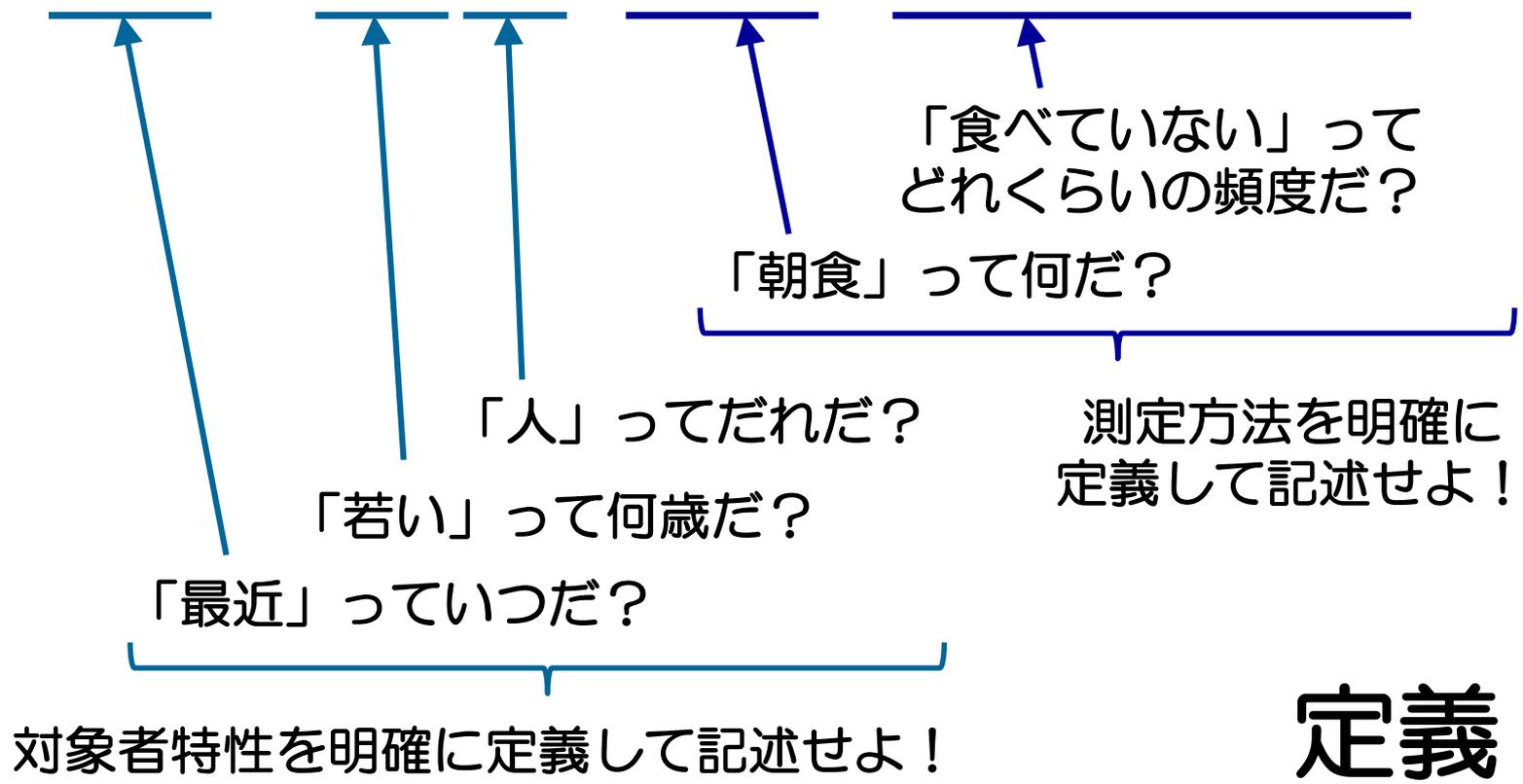
文章で記述する。または、図示する。

どれを分母にするかで、参加率などは異なる。

フローチャート（とそれぞれの値[実数]）を書いておけば、必要に応じて計算できて便利。

数値は、率よりも実数で記述しておくほうが親切。

最近の若い人は朝食を食べていない



論文や研究発表の序論で使われることが多い。
ここで誤るとその後の研究や発表はすべて意味がなくてなってしまう。

レポート

講義で扱った疫学方法を使った興味ある論文（英文）をPubMedで検索して、ひとつの論文について、和文（または英文）でまとめる。2ページ程度

- 氏名、学籍番号、提出期日、課題名
- 論文の書誌情報（著者名、論題、雑誌名、年・巻・ページ）
- 内容の概要（短めに）：抄録のコピーではなく、自分の文章で書くこと。図表は不要
- コメント（たいせつ!）：自分の感想、意見、批判、論理的なコメント
- 論文のコピー（全ページ）を添付
- 提出期日は原則として翌週の講義のとき（事後提出も認めます）

提出した翌週、いきなり、当てられて発表（5～10分）を命じられるので、心の準備をしておくこと。3人/日程度。発表内容や発表技術も評価対象です。

本日の結論 対象者

対象者の特性（属性）を調べ、記述すること！

実現可能性（最大限の努力をした上で）を考慮して行う。

集団代表性 ... 確保したいが、なかなかできない。

参加率・協力率 ... 高くしたいが、なかなかできない。

⇒ せめて、大切だということを知ってほしい。

} 社会の理解を得る
努力も必要

調査時にも、解釈の時・説明の時にも...

（実現可能性を考えて、計画し、実行し、解釈する姿勢が大切。）

対象の特性：疫学の論文ではじめに読むべき情報

「どんな人たちなのだろう？」と想像しながら読もう。

本日の課題：

対象者（対象集団）がおもしろい・すばらしい・問題だと感じた論文