

# 第6章

## 病気の原因を調べるための疫学研究2: コホート研究





## 本講義の内容

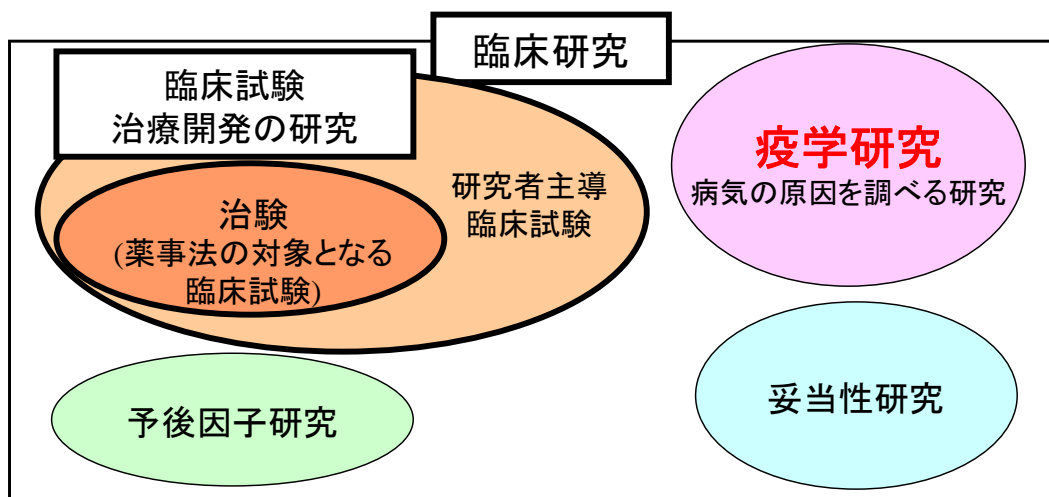
- コホート研究のデザインについて知る。
- コホート研究の調査方法を実例から学ぶ。
  - 調査項目
  - 追跡調査
  - 関連の指標
- コホート研究について、その利点と欠点を知る。
- ケース・コントロール研究とコホート研究がエビデンスレベルの中でどこに位置付けられているのかということを理解する。



前章、第5章の講義では、疫学研究デザインの1つである「ケース・コントロール研究」について説明しました。

本講義では、もう1つの代表的な疫学研究デザインである「コホート研究」について、「ケース・コントロール研究」と比較しながら解説します。

## 臨床研究とそれを支えるもの



**疫学研究とは病気の原因を調べる研究**



コホート研究は、臨床研究の中で疫学研究に分類されます。  
疫学は、病気の原因を調べる研究のことです。

## コホート (cohort) とは？

- 古代ローマ時代の兵隊の単位
  - 300-600人程度
- 一定期間追跡される人々の集団
  - 地域住民
  - 特定の職業集団: 医師、看護師、公務員など
  - 健康診査、検診受診者など



もともと、「コホート」という言葉は、古代ローマ時代の兵隊の単位でした。当時300人から600人程度の兵隊の集まりのことをコホートと呼んでいたようです。

現在、疫学で用いられる場合のコホートという言葉は「一定期間追跡される人々の集団」という意味で使われています。

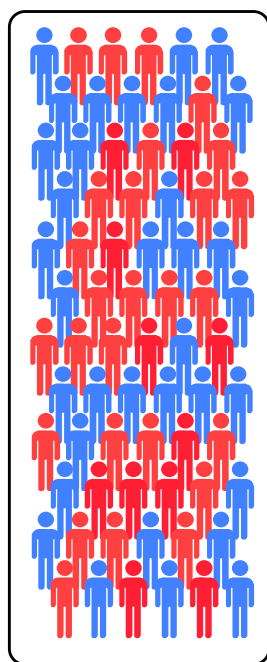
コホートとして追跡され、研究に用いられる集団の例として、地域の住民や特定の職業集団(代表的なものに医師・看護師・公務員など)、健康診査の受診者・検診受診者などがあります。

# コホート研究って何？

アンケート調査

追跡調査(10年以上)

データ解析

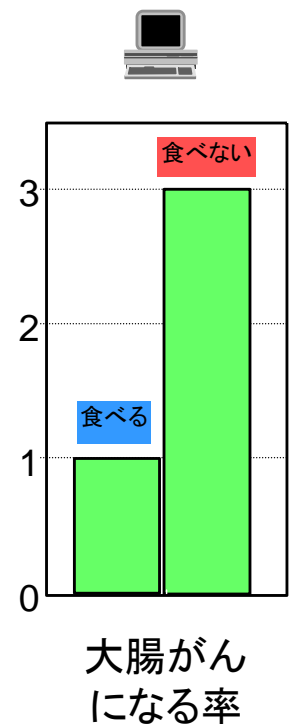
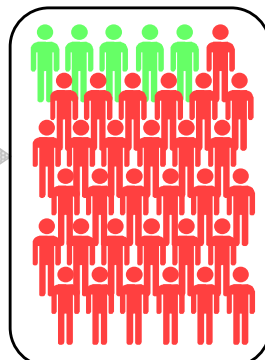
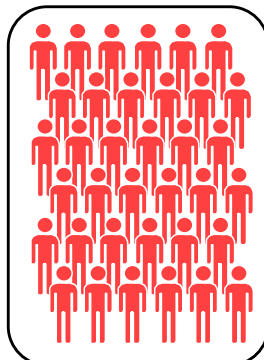
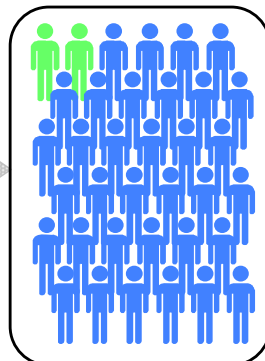
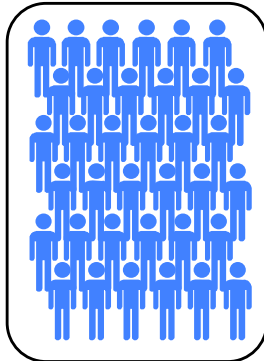


観察する

野菜を食べる



野菜を食べない



コホート研究は対象集団を定義することから始まります。

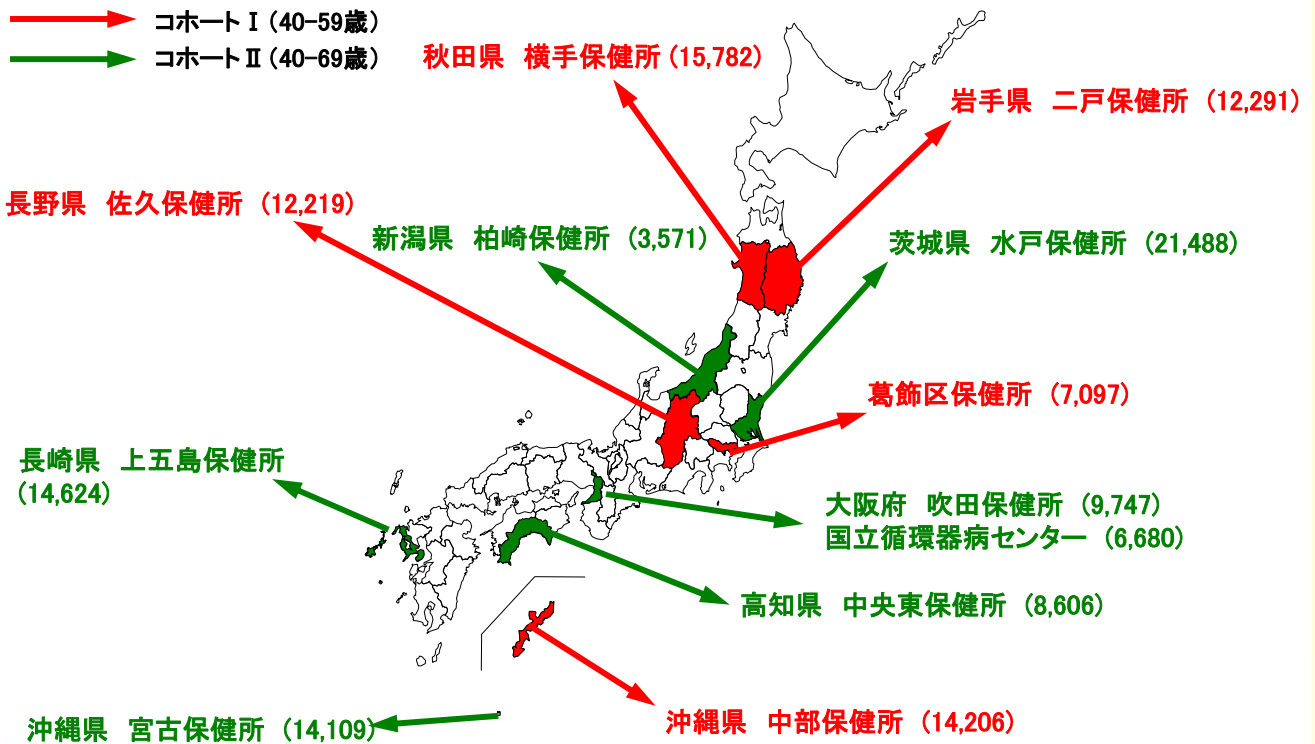
ケース・コントロール研究のときと同様に、野菜摂取と大腸がん罹患の関連を調べる研究を例に、コホート研究の説明をします。

まず対象者の集団(コホート)を定義し、その集団に対して一人ひとりの野菜摂取量を調べます。対象者の数が多い場合には、質問票などによって野菜摂取量を調べるのが一般的です。これらの集団を長期間、例えば10年間、その後の大腸がん発症を調べるために追跡を行います。

最初に把握した野菜摂取量で、コホート全体を野菜をよく食べるグループとあまり食べないグループに分け、それぞれのグループでその後の追跡期間中に発症した大腸がんの数を調べ、それぞれのグループでの大腸がん発症割合や年間発症率を計算します。

これらの発症率や発症割合を比べることにより、野菜をよく食べるグループでの大腸がんの発症があまり食べないグループに比べて低いのか、高いのか、あるいは変わらないのかを調べることによって、野菜摂取の大腸がんリスクに対する影響を調べる研究デザインをコホート研究と呼びます。

## コホート研究の例(JPHC研究)



<http://epi.ncc.go.jp/jphc/>



Introduction to Clinical Research

コホート研究の実例として、国立がんセンターの予防研究部が中心となって行っている厚生労働省研究班による「多目的コホート研究(JPHC研究)」を例に話を進めていきたいと思ひます。

JPHC研究の対象者は全国11保健所管内の市町村、約30市町村に在住の40歳から59歳、あるいは40歳から69歳の全住民(東京と大阪は一部住民)であり、合計約14万人のコホートです。

北は岩手県と秋田県から南は沖縄県の本島と宮古島まで、全国11か所が対象地域になっています。JPHC研究は、コホート I と呼ばれる1990年に調査を開始した地域と、コホート II と呼ばれる1993年に開始した地域からなります。

## コホート開始時調査項目

- **曝露**情報を収集する
  - 自記式質問票(アンケート調査票)
  - 採血
    - 健康診査の場を利用して提供を依頼
  - 健康診査の結果



JPHC研究では、開始時(ベースライン時)に大きく分けて3つの方法で情報を収集しました。

1つ目が自記式質問票

2つ目が血液

3つ目が健康診査の結果

です。これらの方法で調べた項目とその後の疾病の発症を調べるわけですが、これらの疾病の原因となりうる関心の要因を疫学では「曝露」と呼びます。

## ベースライン時の質問票に含まれる項目

- 14ページの カラー印刷版の自記式質問票
  - － 家族歴、既往歴
  - － 喫煙、飲酒習慣
  - － 身体活動
  - － ストレス、社会的支援
  - － 住居、職歴
  - － 性格
  - － 女性の出産歴
  - － 食事
    - 44食品(52)の食物摂取頻度調査



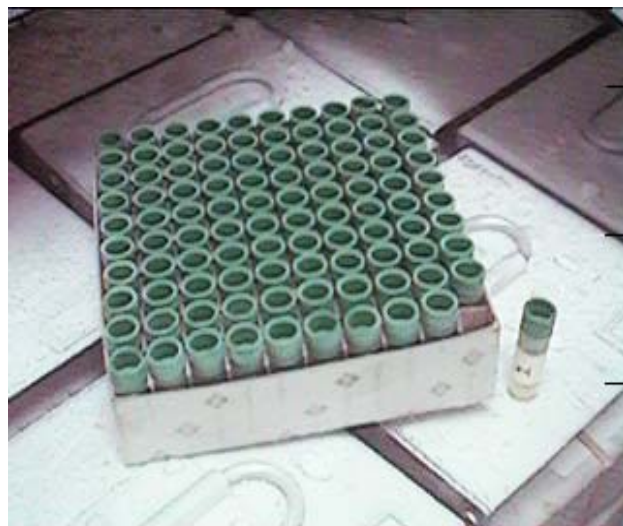
ベースライン調査では、カラー14ページの自記式の質問票を使用し、いろいろな曝露要因について調査を行いました。

研究内容や目的によって調査内容というのは異なりますが、JPHC研究では喫煙や飲酒、食事などの生活習慣とがんや循環器などの慢性疾患との関係を明らかにするというを目的としています。したがって、ベースライン時の質問票にはこれらの生活習慣に関する項目が含まれています。



## 生体試料(血液)

- 末梢血10 mL (ヘパリンチューブ)
- 12時間以内に遠心分離
- バッフィーコートと血漿(1mLずつ3本)に分注
- -80 °Cで保管

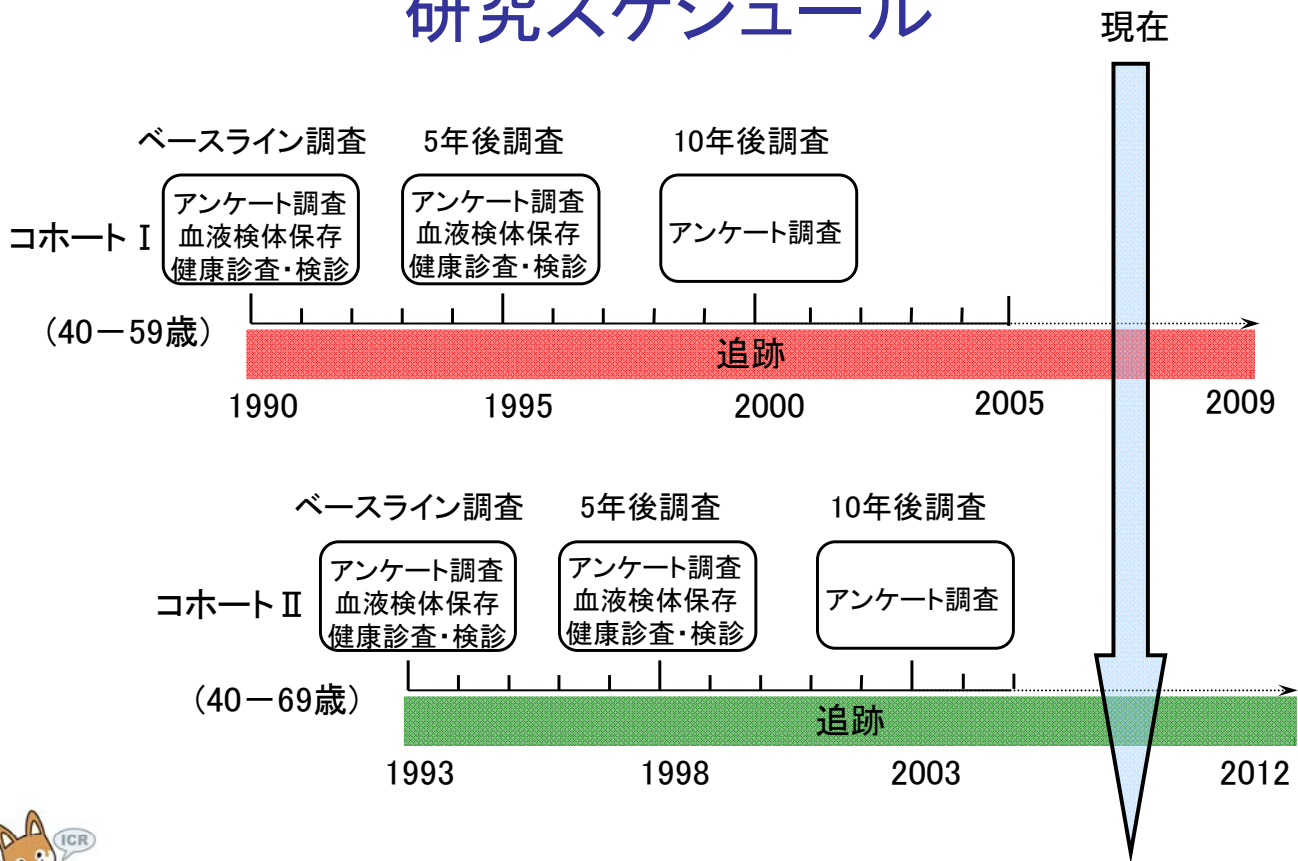


アンケート調査以外にも、血液、尿、髪の毛、爪という生体試料は、曝露評価の上で大事な試料になります。

JPHC研究では健康診断の場で研究のための血液提供をお願いし、採血して長期間冷凍保存しています。

長期間にわたるコホート研究では、採血した血液を冷凍保存し、将来、疾病に関する情報がたくさん集まって、いざ解析というときに関心のある物質について測定を行うことが一般的によく行われます。

## 研究スケジュール



JPHC研究の研究スケジュールは、研究開始から曝露についての定期的な調査と、継続的な追跡とからなっています。

1990年に調査を開始したコホートⅠ、1993年に開始したコホートⅡは、開始時点は違っても、その後は同じ間隔で調査を行っています。

ベースライン調査時には前述のようにアンケート調査と血液と健康診査データを集め、5年後にも同じ調査を繰り返しています。

ただし、食事についてはベースライン時よりも詳しいアンケート調査を行いました。

10年後にはアンケート調査だけを行う、という研究計画を立てています。

ベースライン調査が終わった時点、リクルートをした時点で追跡を開始し、10年後調査終了後もさらに追跡を続け、全体で20年間対象者を追跡するという研究スケジュールになっています。

## 追跡調査

- 異動と疾病の把握
  - 異動、死亡・死因
  - がん、脳卒中、心筋梗塞(疾病登録)
  - 糖尿病、白内障など(自己申告)
  - 健康状態確認を目的としたアンケート調査
  - ニュースレター送付



追跡調査の目的は、対象者の異動と疾病の把握です。

わかりやすくいうと、コホートの対象者が今どこに住んでいて、ご存命か、また、関心のある疾病にかかっているかどうかを把握する・確認するという作業です。

異動と疾病の把握にはいろいろな方法が考えられますが、JPHC研究の場合には、所定の手続きを経て、市町村から対象者の異動の情報をいただき、死亡者については保健所の死亡小票から死因を同定します。さらにはJPHC研究で行っている疾病登録のシステムからがんや脳卒中、心筋梗塞などの罹患の情報を集めます。対象者に直接これらの情報を尋ねる方法もありますし、都道府県がん登録など、既存の疾病登録を利用する方法もあります。

JPHC研究でも、糖尿病や白内障、その他の健康状態の確認については、アンケート調査を用いて対象者に尋ねるという方法をとっています。

また、対象者への結果の還元として、JPHC研究では研究成果などを含めたニュースレターを対象者に送付しています。ニュースレターを送付することによって、JPHC研究事務局が現在把握している住所に対象者が住んでいるのかということもわかりますので、ニュースレター送付も追跡調査の一環としてとらえることができます。

## 追跡調査における公的資料の利用

- 異動
  - 居住地、生死の確認
  - 住民基本台帳法:ここにある情報は基本的には公開情報
    - 何人に対しても閲覧・写しの交付請求を認めている。  
ただし「請求理由」を明らかにすることが要求され、  
不当な目的によることが明らかなきなどは拒否できる。
- 死亡
  - 人口動態調査の目的外使用の許可
- 罹患
  - 医療機関に協力をお願い
  - 地域がん登録の利用申請
- 追跡調査で死亡や罹患を把握することに加え、  
対象者の生存を確認することも大切。



追跡調査では、対象者の異動や罹患の情報などの個人情報を扱うことになります。ここでは、それらの情報を得るために我が国で利用できる公的な資料について説明します。

異動調査では居住地、生死の確認が目的になります。これらの情報は住民基本台帳から得ることができます。住民基本台帳は基本的に公開情報であり、正当な理由があれば誰でも閲覧できる情報です。

死亡情報や死因については、人口動態調査の目的外使用の利用許可申請をして、許可を得ることができれば把握することができます。許可をとるには通常、手間と時間がかかりますが、妥当な研究目的であれば許可を得ることは可能です。

罹患の把握については、調査対象地域の医療機関にお願いし、疾病の情報を集めるという方法があります。がんの場合には、地域がん登録がある地域では、正式な手続きを経てその情報を活用することができる場合もあります。

追跡調査に関しても、「疫学研究に関する倫理指針」に添った手順をとることが必要です。研究デザインによっては、必ずしも対象者からインフォームドコンセントを受けなくてもよい場合もありますが、その場合でも対象者に対して「異動調査で居住地や生死の確認をしています」「亡くなられたら人口動態調査から死因を把握します」などといった調査内容を広くお知らせすることが必要です。詳しくは「[疫学研究に関する倫理指針](#)」(リンクを飛ばしてください)をご覧ください。

## コホート研究の質を下げる要因

- **追跡脱落者の問題**
  - 例えば、大腸がん患者が選択的に追跡から脱落してしまうとき。
- **疾病把握に関する問題**
  - 例えば、野菜を食べる人と食べない人の間で疾病把握の精度が異なるとき。(受診医療機関の違いなどから)



コホート研究においてもその研究の質を下げる要因があります。  
ここでは2つ紹介します。

まず1つ目は、追跡ができなくなることによる問題です。

追跡期間が長くなることにより、だんだん途中で追跡できない人が出てきます。これを脱落(ドロップアウト)と呼びます。JPHC研究の場合、保健所管内を対象地区としているので、保健所管内から外に引っ越した場合には、がん登録のシステムから漏れてしまいます。対象地区外へ引っ越した場合にはその後の転帰がわからないので、その時点で脱落とし、解析のときには打ち切りデータとするなどといった処理を行います。その後の転帰と関係のない理由による脱落なら大きな問題ではないのですが、例えば大腸がんの患者が選択的に保健所の管内から管外へ引っ越し、追跡から脱落していつてしまうようなことが起こると、リスクが正しく推定できないといったことが起こります。

2つ目は、疾病把握に関する問題です。

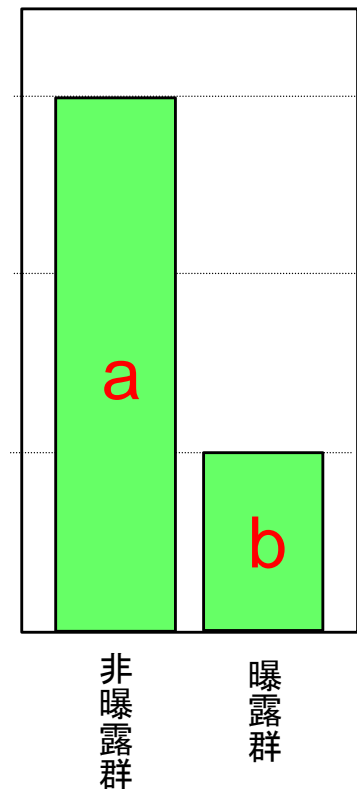
例えば、ある農村の人は野菜をよく食べ、ある都市の人は野菜をあまり食べないとします。その農村の近くにはがん診療の拠点病院があり、大腸がんになった人は多くがその病院を受診するために疾病把握の精度が高く、都市ではがん診療を行う医療機関がたくさんあるため、罹患情報をあちこちから集めるのが困難で、大腸がんの把握精度があまり高くないとします。このとき、野菜摂取と大腸がんの間に因果関係がなくても、見かけ上、野菜をよく食べるの方が大腸がんになりにくい、というような誤った結果が得られてしまいます。

## 関連の指標

- 相対リスク(Relative risk)
  - 疾病罹患と曝露との関連の強さを示す指標。1より大きいと曝露ありの方が疾病になりやすいことを示す。
- リスク比
  - 疾病発症頻度の比

$$\text{リスク比} = \frac{\text{曝露群の発症頻度}}{\text{非曝露群の発症頻度}} = \frac{b}{a}$$

疾病発症頻度



コホート研究において曝露と疾病(今回の例では野菜と大腸がん)の関連を調べる指標について説明します。

関連の指標には、差の指標や比の指標などいくつかあります。比の指標を相対リスクと呼びますが、ここではその中のリスク比について説明します。

リスクとは疾病発症の割合のことをいいます。つまり、ある対象集団の中の疾病の発症頻度です。

したがって、リスク比は2つの群の疾病発症頻度の比ということになります。

基準となる群の発症頻度を分母にし、リスクを見たい群の発症頻度を分子にします。

例えば野菜を食べない群の大腸がんの発症頻度を分母にし、野菜を食べる群の大腸がんの発症頻度を分子とすることによって計算できます。

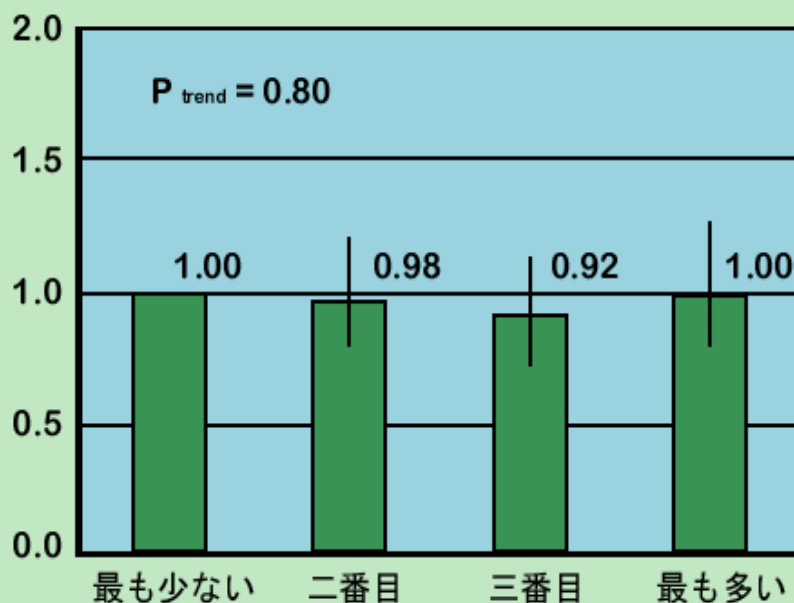
両者の頻度が同じ場合にはリスク比が1となり、野菜摂取と大腸がんの間には関係ないという結果になり、リスク比が1より小さければ、野菜を食べる方が大腸がんリスクが低い、ということになります。

ちなみに、単位時間(例えば1年)当たりの疾病発症頻度を発症率(ハザード)といい、その比を発症率比(ハザード比)といいます。ハザード比も相対リスクの一種です。

数学的な取り扱いのしやすさなどから「オッズ比」で結果が提示されることもありますが、「オッズ比」は直感的に解釈しづらいため、あまりお勧めできません。

## 野菜摂取と大腸がん罹患

—40-69歳の男女約90,000名を7-10年間追跡—



調整因子：性別、年齢、地域、大腸がん家族歴、BMI、身体活動、喫煙、  
飲酒、ビタミン剤の使用、エネルギー摂取、穀物・肉・魚摂取

Tsubono Y, et al. Br J Cancer 2005;92:1782-4.



これは、JPHC研究で調べた野菜摂取と大腸がんの関係です。

グラフは横軸に野菜摂取を「摂取量が少ない群」から「多い群」まで4つの群に分け、それぞれリスク比を計算しています。

最も少ない群を基準にリスク比を計算していますが、摂取頻度が多くなるにしたがって、リスク比はそれぞれ「0.98」「0.92」「1.00」となります。摂取量によって大腸がん発症頻度に差がないので、この研究からは野菜摂取と大腸がんの間には関連がみられない、と推論することになります。

## コホート研究の特徴

- 利点
  - － 観察の時間的順序と因果の順序が同じ
    - 先に野菜摂取量(原因)の測定し、その後の大腸がん罹患(結果)との関連を見る
  - － いろいろな疾病について調べられる
    - 野菜摂取量と胃がん罹患
- 欠点
  - － まれな疾病には適さない
    - 例として、約2万人を10年間追跡して179人の乳がん罹患
  - － 費用・時間がかかる
    - 年間数千万、かつ10年以上追跡しないと精度の高いデータが得られない



コホート研究の特徴についてまとめます。

コホート研究の利点は、観察の時間的順序と因果の順序が同じであることに由来しています。原因とする野菜の摂取量を先に測定し、その後結果とする大腸がんの罹患情報を把握し、その関連を調べるというデザインになっており、原因と結果という因果の順序と情報を得る時間の流れが一致しています。そのため、病気という結果の状態から遡って昔の原因を調べるケース・コントロール研究で生じてしまうリコールバイアス(思い出しバイアス)を避けることができます。また、コホート研究は同じ集団を追いかけるので、セレクションバイアスの影響を受けません。

次に、コホート研究では、いろいろな曝露と疾病について調べられることが利点です。

今回の例は野菜摂取量と大腸がんの話でしたが、タバコと胃がんや大豆と乳がんの関連も調べることができます。

その一方、欠点として、稀な疾病のリスクを調べるには適さないということが挙げられます。

日本人女性が罹るがんの中でもっとも罹患数の多い乳がんでも、約2万人のコホートを10年間追跡して把握した罹患は200人以下でした。実際に詳しい解析をしたいと思った場合には、200人という数字は十分な情報を持っていません。ましてや稀ながんや稀な疾患では、もっと大きなコホートをもっと長い期間追跡することが必要となり、現実問題として実施不可能な手間と時間がかかります。

コホート研究は膨大な費用と時間がかかります。乳がんのように多いがんでも、数万人のコホートを10年以上追跡することが必要であり、精度の高いデータを得るためには年間数千万円の費用が必要となるでしょう。

トータルで考えて、コホート研究はケース・コントロールより、エビデンスレベルが上といえます。ただし、野菜摂取といった曝露は本人の意思であり、実験的介入ではないため、統計解析で他の要因を完全にコントロールすることは不可能です。これは観察研究の限界であり、エビデンスレベルでは、ランダム化比較試験の下となってしまいます。これについては「第8章 生物統計学2:交絡・ランダム化と因果推論」で詳しく述べます。



## 日本のがんの大規模コホート研究

	開始年	人数	年齢	調査集団
環境省コホート	1983	121,610	40歳以上	宮城・愛知・大阪の4市5町住民
文部科学省コホート	1988	110,792	40-79歳	全国34市町村の住民と企業職員
厚生労働省コホート				
コホートI	1990	48,943	40-59歳	5保健所管内の住民
コホートII	1993	61,593	40-69歳	6保健所管内の住民
東北大コホート				
宮城県民コホート	1990	47,605	40-64歳	宮城県14町村の住民
大崎コホート	1994	52,029	40-79歳	大崎保健所管内の国民健康保険加入者
岐阜大コホート	1992	33,399	35歳以上	高山市の住民



コホート研究には多額のお金と時間がかかると説明しました。

日本全国を見渡しても大規模コホート研究は日本でも実際にそれほどたくさん行われているわけではありません。

がんを例にすると、10万人規模で日本で行われている大規模コホート研究は3つ程度しかありません。

JPHC研究は「厚生労働省コホート」と呼ばれているものです。がんに関してマスコミ報道されるコホート研究の結果の多くはここに挙げたものから得られたものです。



## ケース・コントロール研究とコホート研究の比較

ケース・コントロール研究		コホート研究
百から千	対象者	数万から数十万
ケースとコントロールで異なる場合が多い (セレクションバイアス)	母集団	同じ
発症後の情報 リコールバイアスの影響 最新の仮説に基づく 詳細な情報	曝露情報	発症前の情報 経時的に情報収集可能 解析時点では既にある情報に限られる
1つ 稀な疾患に有利	対象疾患	複数 稀な疾患は検討困難
結果→原因(後向き)	観察の方向	原因→結果(前向き)
短期間、数年以内	調査期間	長期間、10年以上
数百から数千万円	費用	億単位



最後に、Take Home Messagesの代わりにケース・コントロール研究とコホート研究の比較を表にまとめました。

表中の数字は、がんを対象とした研究を例とした場合の目安とお考えください。

なお最近では、コホート研究の中で質の高いケース・コントロール研究を行う(ネステッド・ケース・コントロール研究)という新しいデザインも提案されています。

疾病の原因を探求するにはコホート研究を行うほうがよいのですが、費用や時間などの実施可能性の点から実際にはケース・コントロール研究しかできないことが多いといえます。コホート研究、ケースコントロール研究の利点欠点をよく知って結果を解釈することが重要です。