

臨床研究と一般統計学 の基礎

大阪市立大学医学部 環境衛生学

林 朝茂

Tomoshige Hayashi 1

第49回日本臨床腎移植学会
＜第10回半日教育セミナー＞
COI開示

筆頭発表者名： 林 朝茂

本発表は臨床研究には含まれず、
開示すべきCOIはありません。

Tomoshige Hayashi 2

Evidenceとは？

- **人を対象とした信頼性の高い臨床研究による実証報告です。**
- EBMという言葉は、「**臨床疫学**」・「**根拠に基づく実践**」にあたります。
- ここでいう、「**実証的エビデンス**」と呼ばれるものは、**診療に役立つか役に立たないか**についての根拠のことをいうものである。

Tomoshige Hayashi 3

疾患の危険因子を検討する研究の歴史

肥満、糖尿病、高血圧、高脂血症、喫煙、性格などは、冠動脈疾患の危険因子である。



Framingham heart Studyの結果が根拠である。

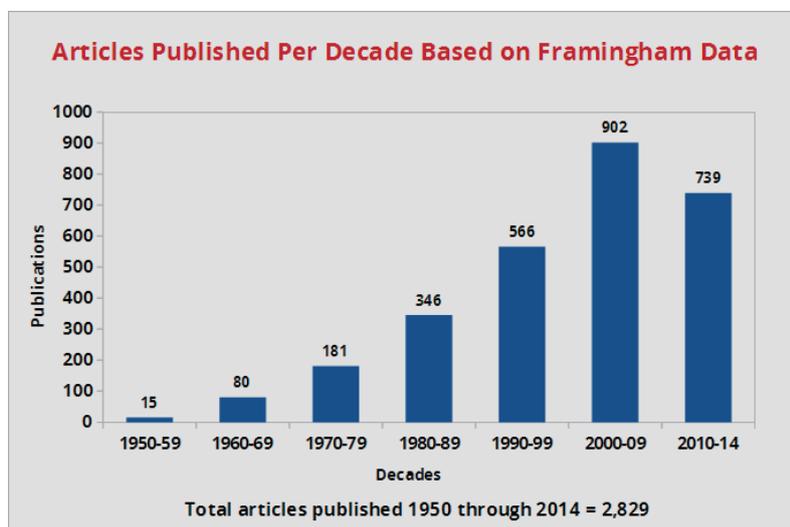
➤ 1948年に開始された**前向きコホート研究**

<http://www.framinghamheartstudy.org/>

Tomoshige Hayashi 4

フラミンガム研究からの発表論文数

<https://www.framinghamheartstudy.org/index.php>



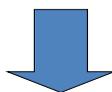
Tomoshige Hayashi ⁵

Drにとって役立つ 疫学の実践的知識

Tomoshige Hayashi ⁶

Epidemiologyって、なんだろう？

- 皆さんは、どんな印象をもっていますか？
- 医療に携わっていく上で、疫学の知識は、必要と認識されているでしょうか？



疫学の知識は、実際に、臨床の現場において、受診者に対応する上で、極めて重要なスキルと考えられます。

Tomoshige Hayashi 7



- 50歳の男性
- 空腹時の血糖値 130mg/dL
- HbA_{1c} 7.2%
- SBP/DBP
=150~160/90~100 mmHg
- Smokers
- BMI = 27 kg/m²

Tomoshige Hayashi 8

病気・将来・治療を話す上で、その根拠は、どこに置けばよいのでしょうか？

- 病気  病態生理
記述疫学・Diagnostic Test
- 将来(予防)  Prospective Cohort Study
(前向きコホート研究)
- 治療  Randomized controlled trial
(RCT): 無作為化比較試験

Tomoshige Hayashi

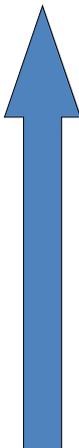
知っておくべき研究方法

- **Descriptive study: 記述疫学**
- **Observational Study: 観察研究**
 - ① Cohort Study
 - Prospective cohort study
 - Historical cohort study
 - ② Case-control study
 - Case-control study
 - Nested case-control study
 - Case-cohort study
 - ③ Cross-sectional study
- **Diagnostic test study: 診断法を評価する研究**
- **Randomized controlled trial: 無作為化比較対照試験**

Tomoshige Hayashi 10

EBMのHierarchyに一言
これって、絶対なのではないですか？

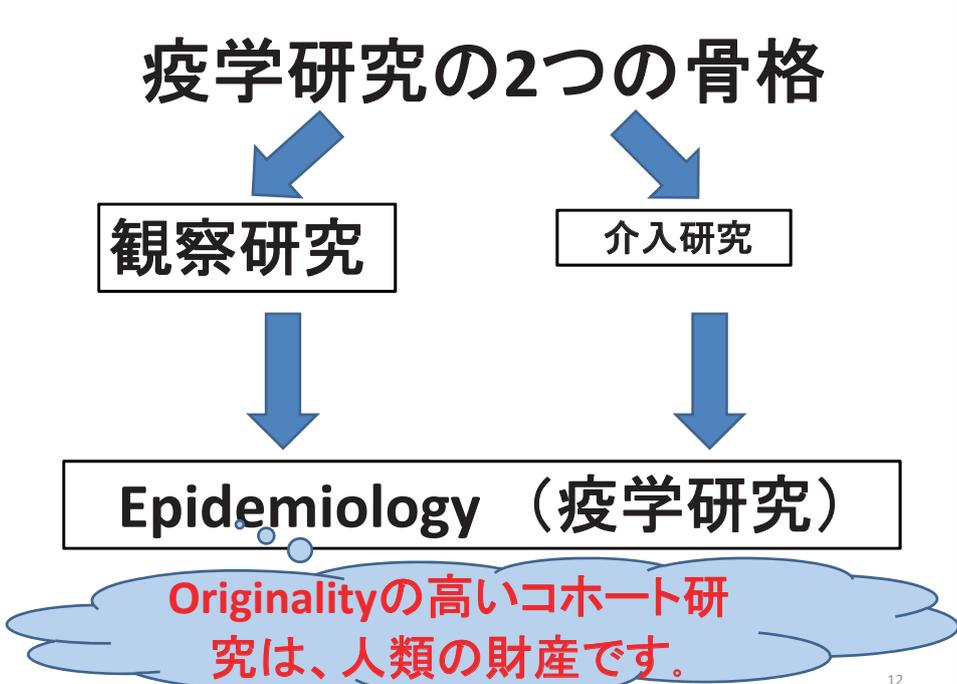
Evidenceの種類と水準



- 無作為化比較試験(RCT)・そのメタ分析
- Cohort Study
- Case-control Study
- 症例の集積
- 専門医の意見

Tomoshige Hayashi 11

疫学研究の2つの骨格



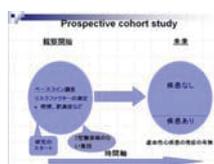
観察研究

介入研究

Epidemiology (疫学研究)

Originalityの高いコホート研究は、人類の財産です。

Tomoshige Hayashi 12



- 対象は、どこから選ぶか？
- サンプルサイズはどれくらい？
- 測定項目は、血液・問診・遺伝子・問診票・画像データ
 - ◆ 多変量解析の際に第3の因子を検討する上で十分か？
- 問診票は、その測定に再現性・妥当性があるか？
- 画像データや新しい検査機器での臨床データ
 - ◆ 再現性や妥当性はあるか？
- 参加基準・除外基準を明確に定義できるか？
- アウトカムの設定は、明確か？
- 観察期間中の定期Followは行うか？
- 行うならば、どのくらいの間隔で行うか？
- 多目的にコホート研究を設定しているか？
- 研究のテーマごとの統計方法はどうか？
- 結果のTableを想定してデータ収集をしているか？

13

疾病頻度の測定

Tomoshige Hayashi¹⁴

曝露とアウトカム: アウトカム

アウトカム: 疫学研究では、アウトカムとして、

① 疾病の発症

- 虚血性心疾患の発症
- 腎移植患者の透析再導入
- 腎癌の発症
- 腎移植患者の感染症発症

② あらゆる健康状態

- 要介護になる。
- 80歳まで自立して生きる。
- 透析患者で長期生存のFactorは？



比較可能な明確な定義をすることが重要。

Tomoshige Hayashi¹⁵

曝露とアウトカム: 曝露

曝露: アウトカム発症の以前に存在する状態を曝露と呼ぶ。

➤ 曝露の例

- 喫煙習慣
- 運動習慣
- 飲酒習慣
- 年齢
- 性別
- 遺伝子要因



- 曝露のうち、疾患の発症の原因とみなすことができる曝露を（曝露の有無や曝露の程度の違いで、疾患の発症確率が異なる曝露を）危険因子 (Risk factor) と呼ぶ。
- 危険因子 (Risk factor) であるためには、曝露→発症の時間軸が必要である。

Tomoshige Hayashi¹⁶

疫学指標

疫学の指標は通常

① 割合 **proportion**

② 比 **ratio**

③ 率 **rate**

のいずれかで表される。

Tomoshige Hayashi ¹⁷

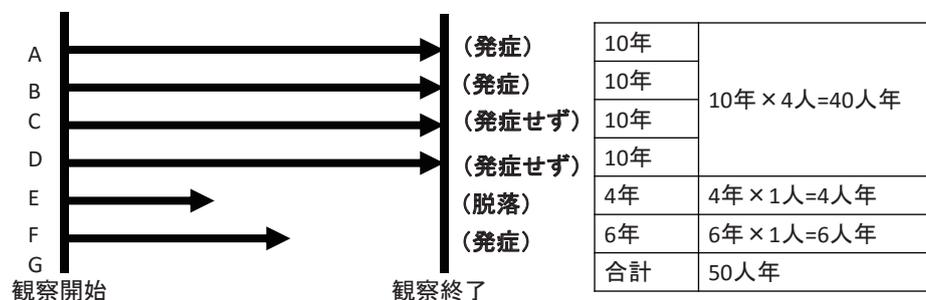
罹患率 Incidence rate

➤ 集団における疾患の発症の**rate**である。

$$\text{rate} = \frac{a}{b}$$

a: 疾病発生数。

b: 観察対象集団の各人の観察時間の合計。

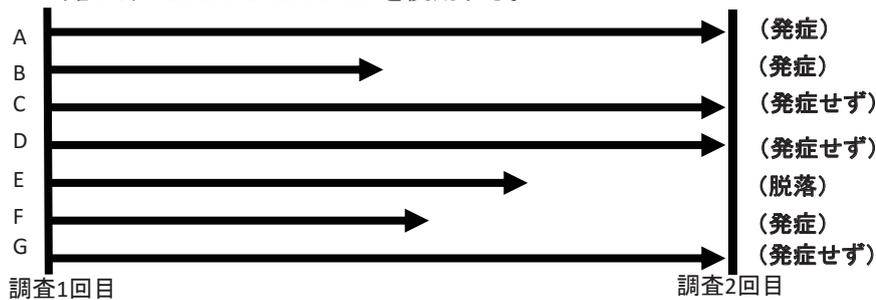


$$\text{Incidence rate} = \frac{3 \text{ person}}{50 \text{ Person-year}} = 6 \text{ 人 (/100人年)}$$

Tomoshige Hayashi ¹⁸

累積罹患率 Cumulative incidence

- ◆ 観察開始時の対象のうち、観察終了までに罹患した者の割合
- ◆ 観察開始から疾病発症までの時間を考慮しない。
- ◆ 中途脱落者(転出、他疾患での死亡)の扱いが難しい。
- ◆ 疾病の発症時期を特定することが困難な疾患では、罹患率を求めることが難しく、Cumulative incidenceを使用する。



$$\text{Cumulative incidence} = \frac{\text{追跡期間中に新規に発症した患者数}}{\text{全追跡対象者数(追跡不能例は除く)}}$$

19

Tomoshige Hayashi

Study Population(について)

- 研究を始める際に、どういった人たちを対象にしますか？
- Study populationは、誰でも、いいのでしょうか？
- Study populationをよく検討するのはなぜでしょうか？

Tomoshige Hayashi 20

対象の設定のポイントは？

- 研究結果のGeneralizability
- Settingはどうあるべきか？
- At riskな集団
- Follow-up率を初めから意識する。
- 要因の有無で対象を限定しない。
- サンプルサイズは？

Tomoshige Hayashi 21

Cohort studyの対象集団

疾患の発症リスクを検討したい？



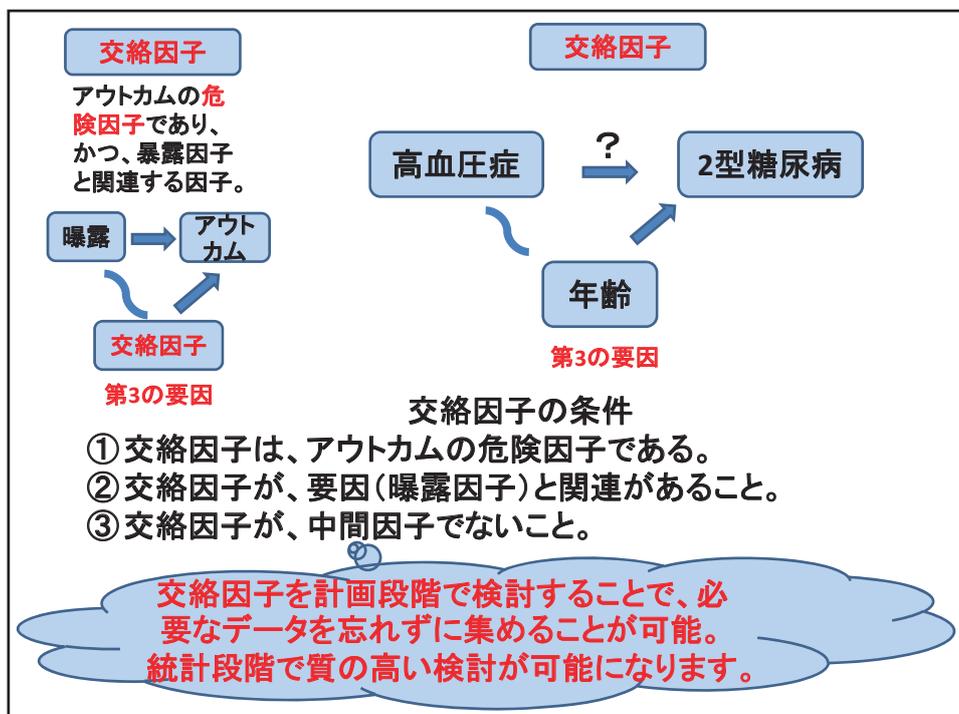
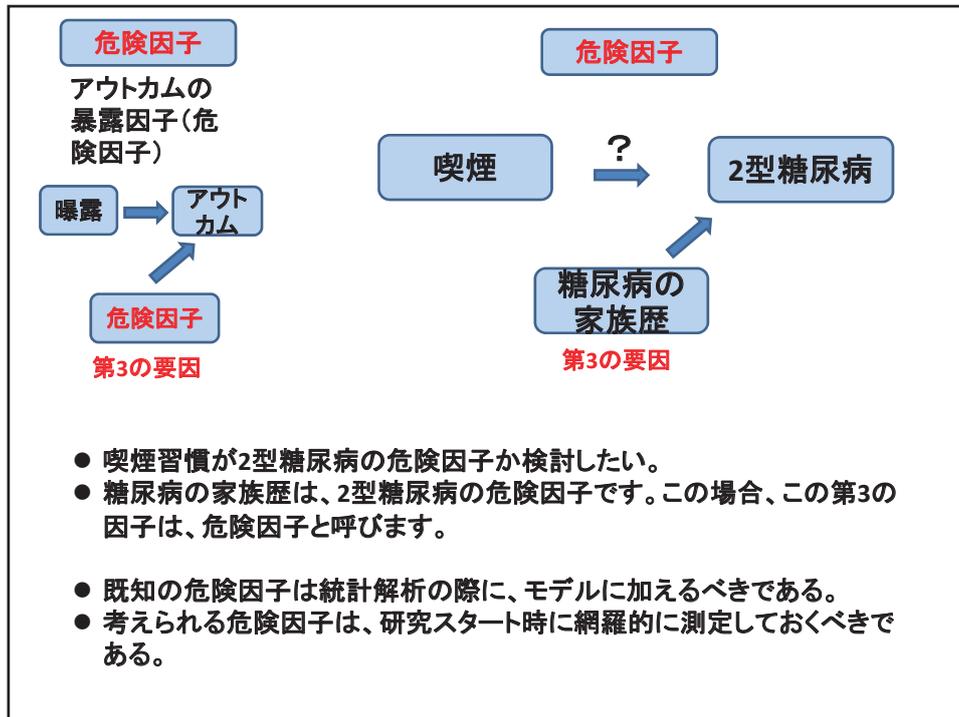
病気を持たない集団を対象とする。

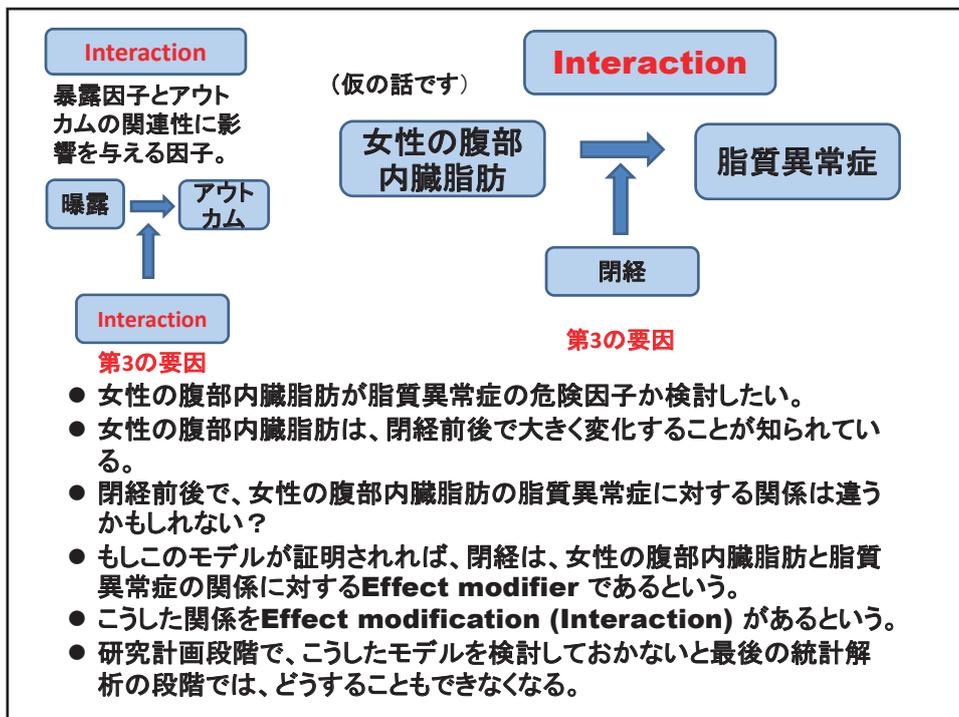
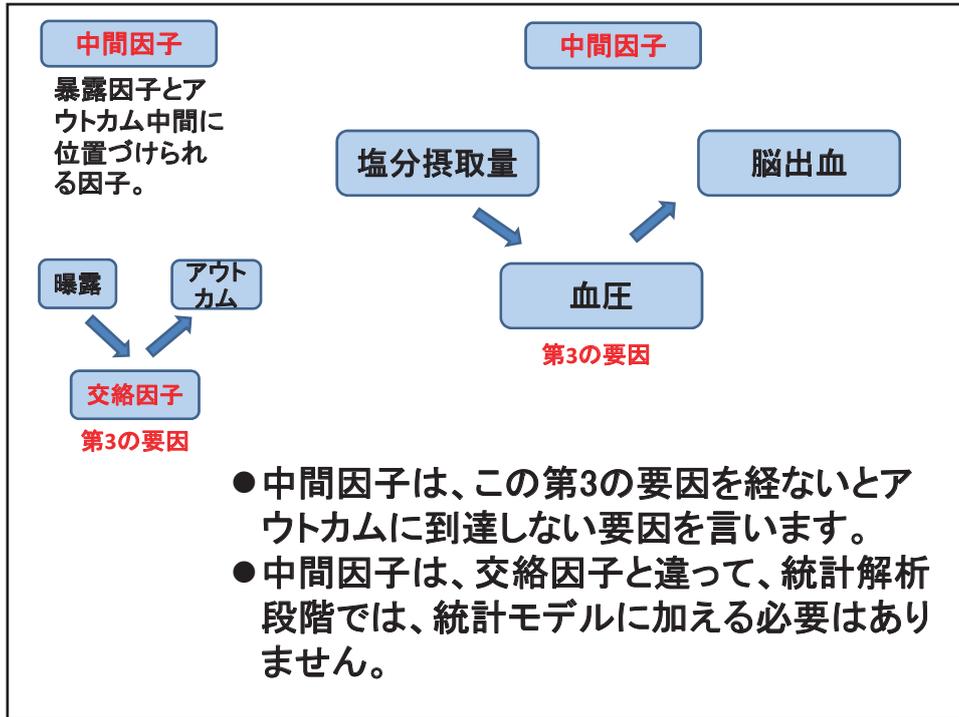
特定の疾患の合併症や疾患の予後のリスクを検討したい？

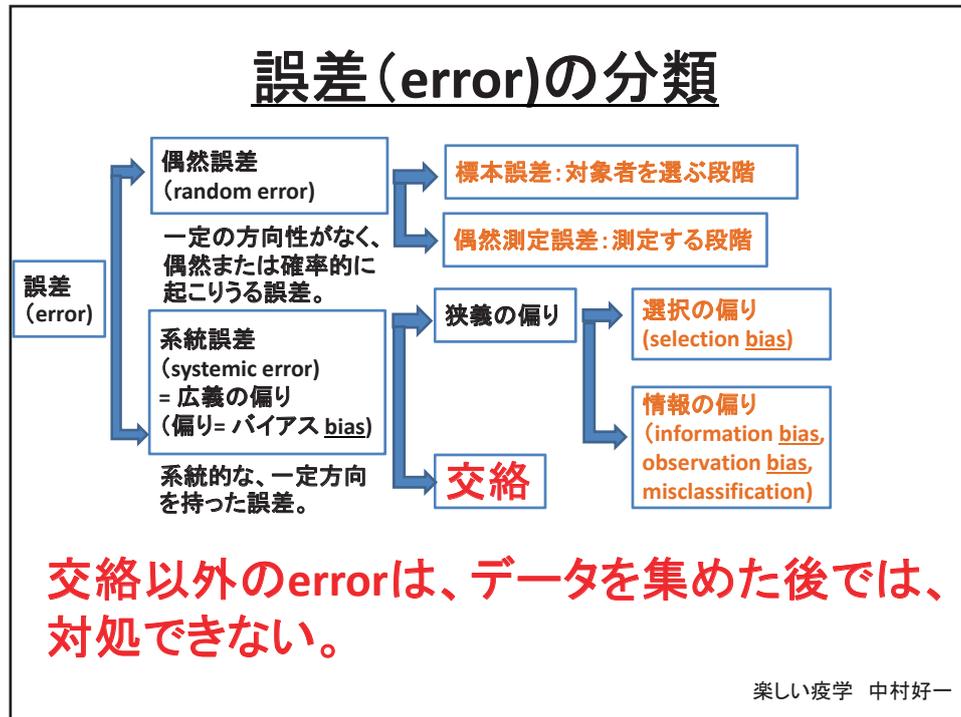


特定の患者集団を対象とする。

Tomoshige Hayashi 22







妥当性 (validity)

- 妥当性: 目的とする真の値(現象)に、測定値がどれほど近いのか、その程度を表す言葉で、正しい結論を導く上で重要。
- 測定値の妥当性は、ゴールドスタンダード(真の値を測定できる標準方法)との比較で評価するのが理想。



結論の妥当性を高める。

Tomoshige Hayashi

信頼性 (reliability)・再現性 (reproducibility)、 consistency (一致度)

➤ 信頼性: 測定値が測定ごとに安定した値を示す(再現性が高い)度合い。

➤ 信頼性を向上させる方法

- ① 測定方法の標準化
- ② 測定者のトレーニングと技能チェック
- ③ 測定方法の改善
- ④ 測定方法の自動化
- ⑤ 測定の反復

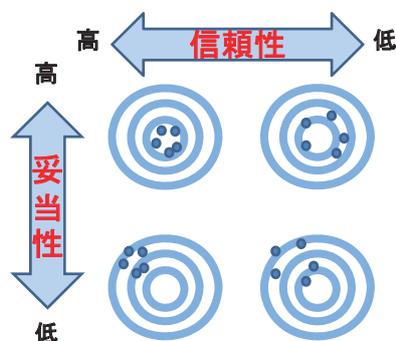


効果判定のパワーを高める。

Tomoshige Hayashi²⁹

測定: 信頼性と妥当性

- **信頼性・精度 (reliability)** ⇒ 偶然に左右されず安定して測っているか?
- **妥当性 (validity)** ⇒ 測りたいものを測っているか?



Tomoshige Hayashi³⁰

交絡とは？

観察研究には、必ず、多変量解析
が必要である。

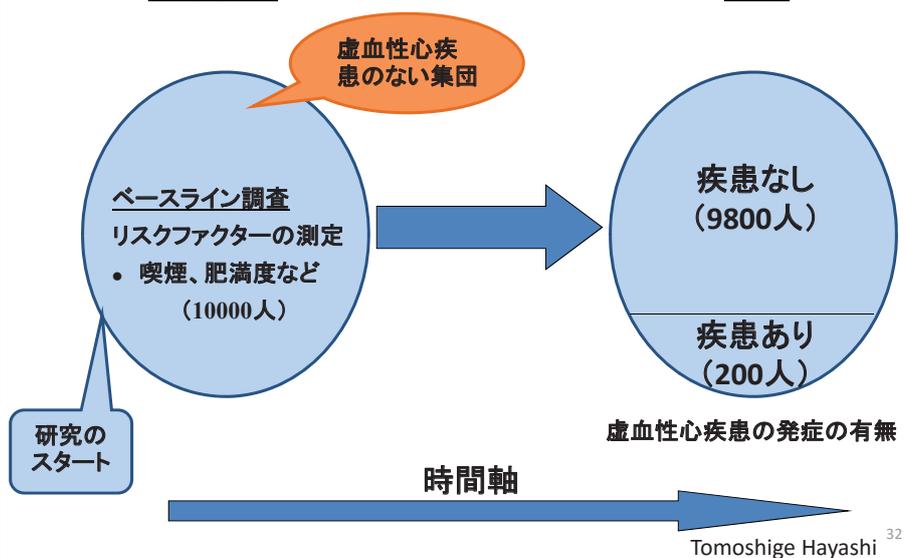
Why?

Tomoshige Hayashi 31

Prospective cohort study

観察開始

未来



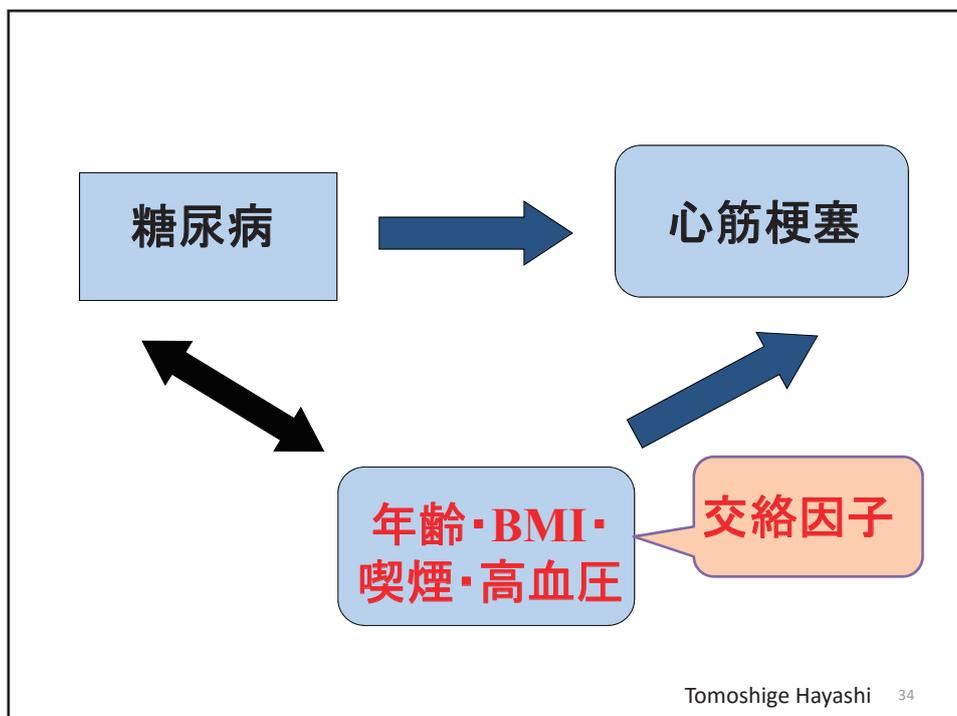
- 例) 1万人の冠動脈疾患のない人々を10年間観察したとする。10年間で、200人の心筋梗塞が発症したとする。その参加者の初年度のデータを以下に示す。

初年度のデータ	心筋梗塞発症なし (n=9800)	心筋梗塞発症 (n=200)	P
年齢	45	55	<0.001
BMI	22	26	<0.001
総コレステロール	200	250	<0.001
喫煙 (%)	40	50	<0.001
糖尿病 (%)	10	15	<0.001
高血圧 (%)	10	15	<0.001

(統計は、 χ^2 検定と対応のないt検定である。)

質問: このデータから、糖尿病のある人は、心筋梗塞を発症しやすいと結論づけることができるだろうか。

Tomoshige Hayashi 33



Tomoshige Hayashi 34

交絡因子の調節は、 どう調整すれば良いか。

Tomoshige Hayashi 35

交絡因子の調節は、どう調整すれば良いか。 解析時に行う方法。

- 層別化：交絡因子が、一つか、二つなら可能だが、多くなると、不可能。
- **多変量解析。**
(統計学は、疫学研究のtoolです。)



Risk Factorを検討する場合、**多変量解析**は、必須です。

Tomoshige Hayashi 36

疾患の危険因子を検討する研究の条件

- 研究デザイン
 - Prospective cohort study.
 - Historical cohort study.
 - 時間軸のある Case-control study.

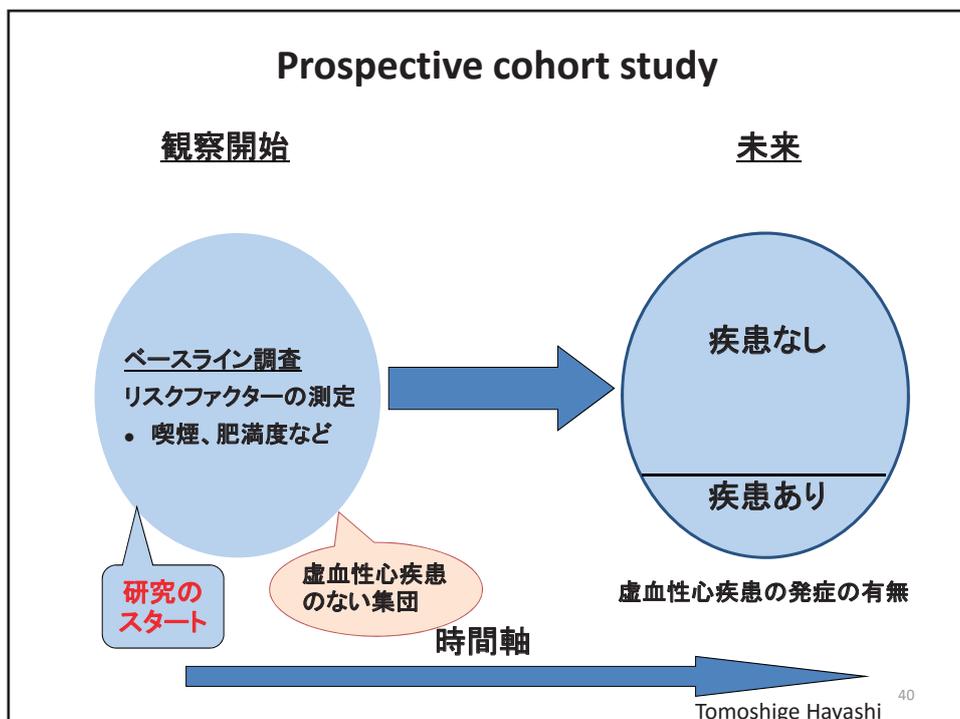
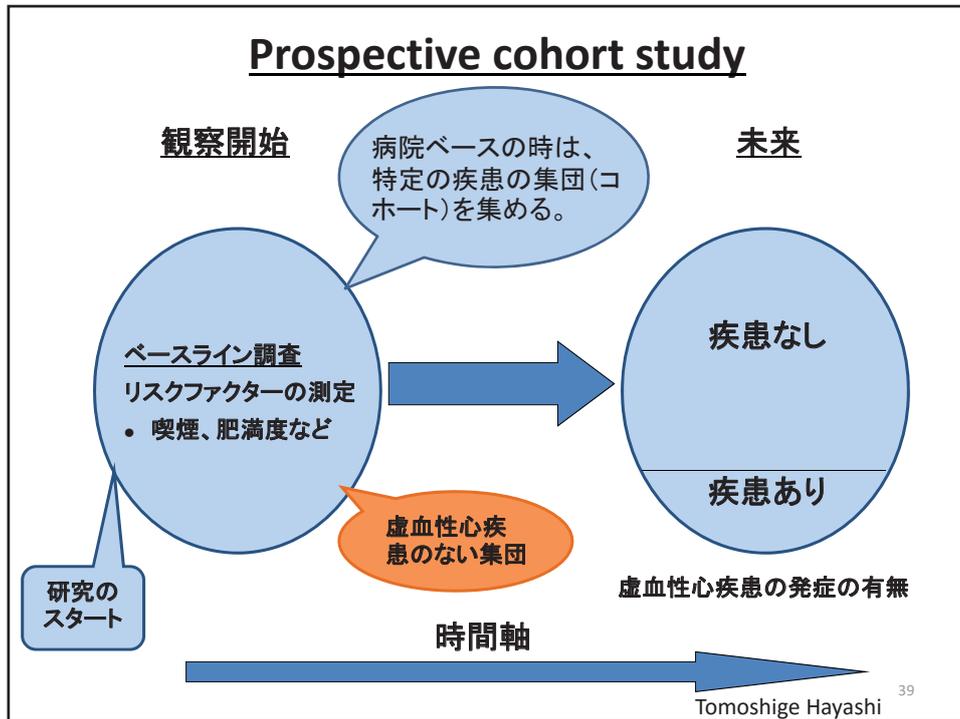
- 解析方法 — 多変量解析で、既知の危険因子が、考慮されていること。

Tomoshige Hayashi 37

知っておくべき研究方法

- Observational Study
 - ① Cohort Study
 - Prospective cohort study
 - Historical cohort study
 - ② Case-control study
 - Case-control study
 - Nested case-control study
 - Case-cohort study
 - ③ Cross-sectional study
- Randomized controlled trial
- Diagnostic test

Tomoshige Hayashi 38

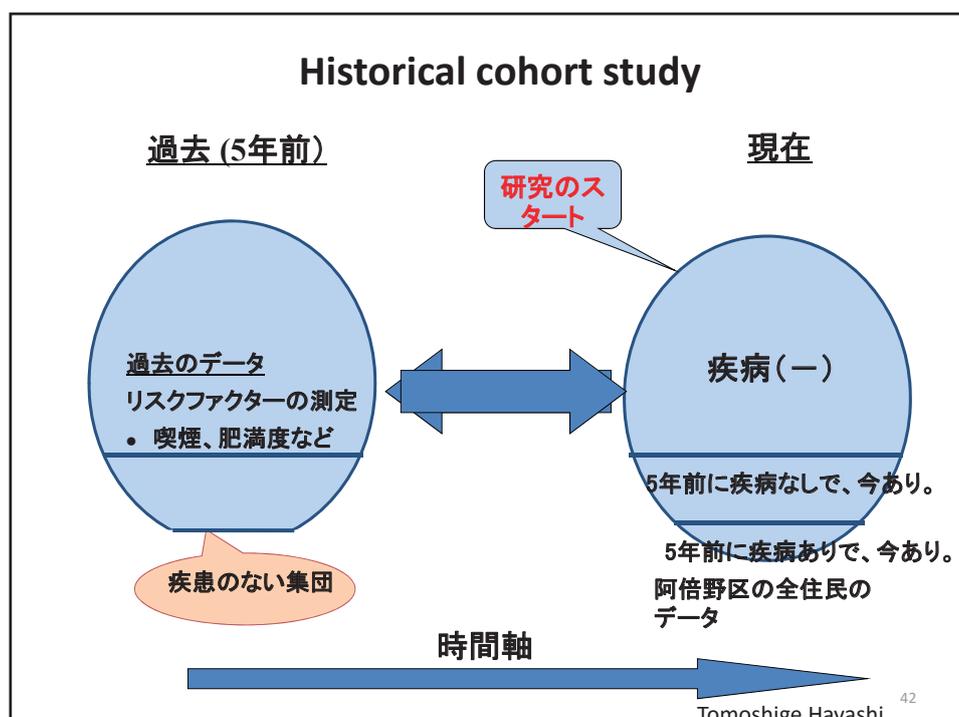


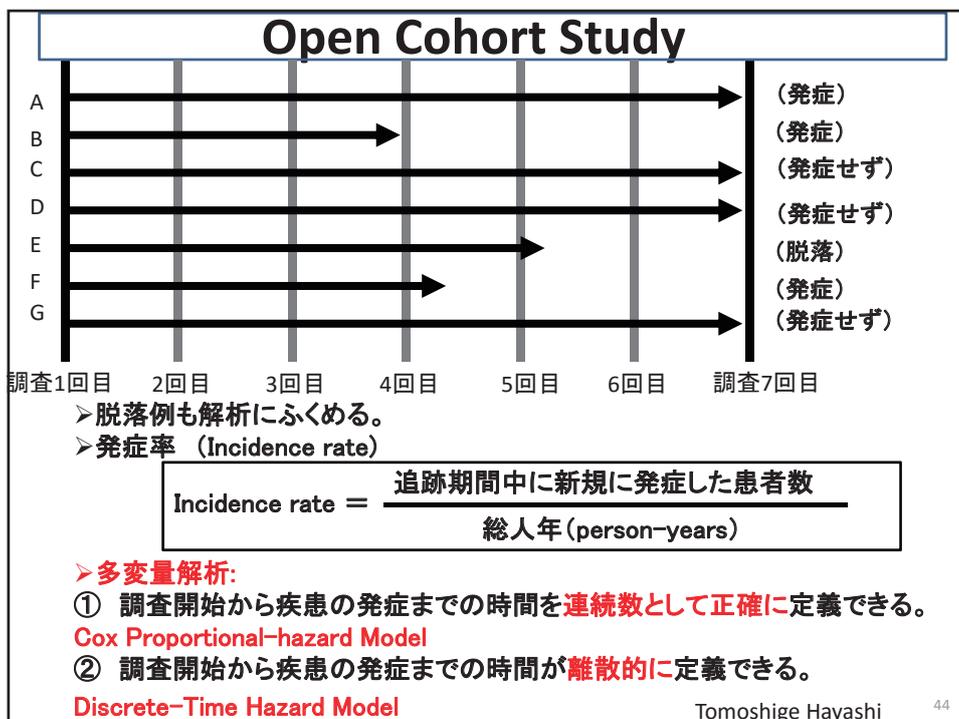
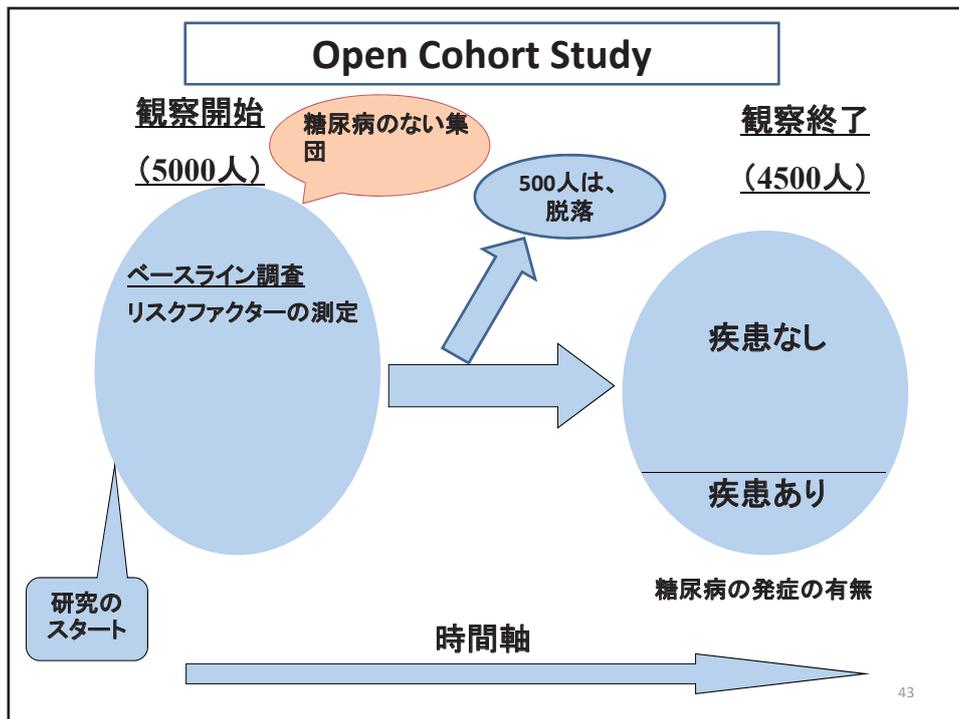
Cohort Studyの利点・欠点

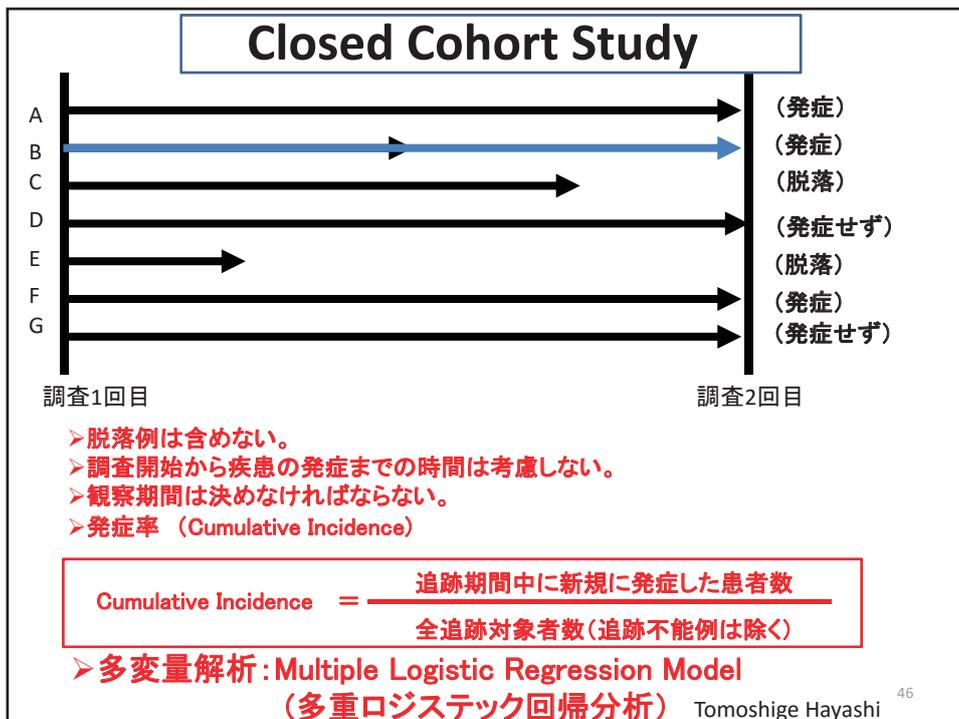
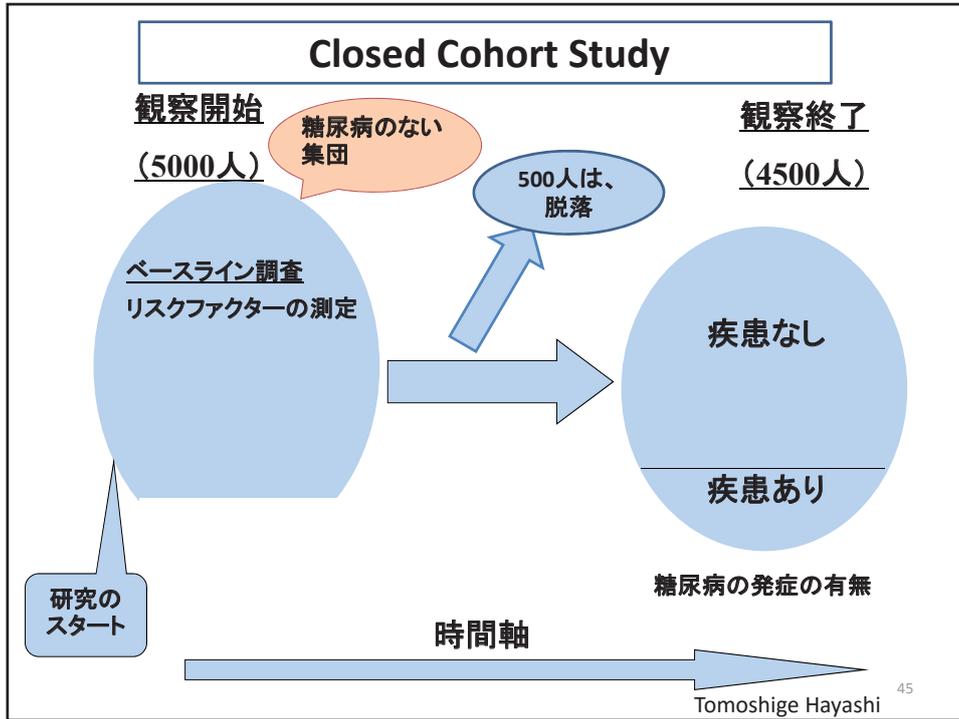
- 利点
 - 追跡観察を行うので**事象の発症順序が明確**である。
 - 要因とアウトカムの両方を測定できる。
 - 第3の因子、特に交絡因子を研究開始時に測定しておくことができる。
 - **要因に対して複数のアウトカムを検討できる。**
 - アウトカムの発症数が時間とともに増加する。
 - リスク比・発生率だけでなく、リスク差・発生率差などが計算可能である。
- 欠点
 - ランダムに発生しない脱落によるバイアスが生じやすい。
 - 診断バイアス・発見兆候バイアスが生じやすい。
 - 測定できない交絡因子の対処ができない。
 - お金や時間がかかる。
 - **発症がまれなアウトカムには不適。**

Tomoshige Hayashi 41

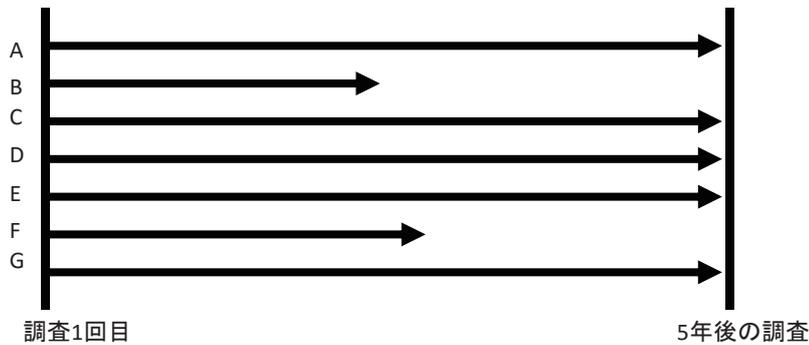
Historical cohort study







Closed Cohort Studyの例 5年後の腰椎の骨塩量の減少のリスクを 検討したい？



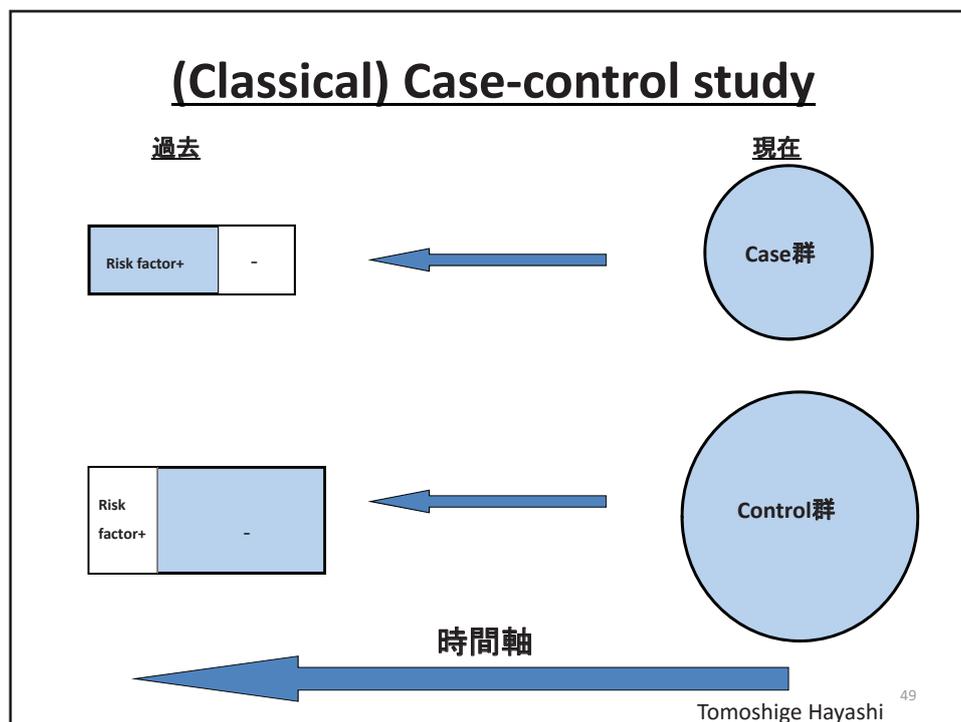
- 脱落例は含めない。
(調査1回目と2回目にも参加した人が対象。)
- 観察期間は決めなければならない。
- 多変量解析: Multiple Linear Regression Model

Tomoshige Hayashi ⁴⁷

知っておくべき研究方法

- Observation Study
 - ① Cohort Study
 - Prospective cohort study
 - Historical cohort study
 - ② Case-control study
 - Case-control study
 - Nested case-control study
 - Case-cohort study
 - ③ Cross-sectional study
- Randomized controlled trial
- Diagnostic test

Tomoshige Hayashi ⁴⁸



確率とオッズの関係

確率	確率%	オッズ $P/(1-P)$
0.01	1%	$1/99=0.0101$
0.1	10%	$10/90=0.101$
0.2	20%	$20/80=0.25$
0.3	30%	$30/70=0.428$
0.4	40%	$60/40=1.5$
0.5	50%	$50/50=1$
0.6	60%	$60/40=1.5$
0.7	70%	$70/30=2.33$
0.8	80%	$80/20=4.0$
0.9	90%	$90/10=9.0$
1.0	100%	∞

Tomoshige Hayashi ⁵⁰

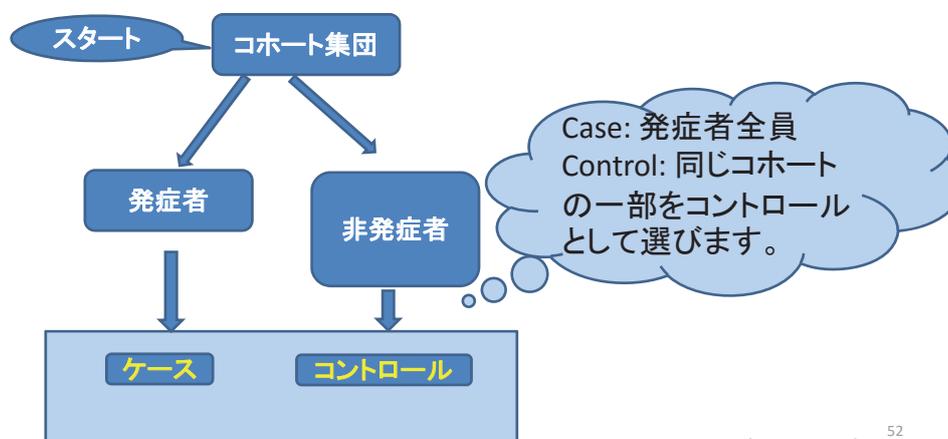
知っておくべき研究方法

- Observation Study
 - ① Cohort Study
 - Prospective cohort study
 - Historical cohort study
 - ② Case-control study
 - Case-control study
 - Nested case-control study
 - Case-cohort study
 - ③ Cross-sectional study
- Randomized controlled trial
- Diagnostic test

Tomoshige Hayashi 51

コホートを利用したCase-control Study

- コホート研究の中に、ケース・コントロール研究を組み込んだデザイン。



Tomoshige Hayashi 52

コホートを利用したCase-control Study

コホート研究の中に、ケース・コントロール研究を組み込んだデザイン。

コントロールの選択により2つに、分類される。

➤ **Nested case-control study**

➤ **Case-cohort Study**

Tomoshige Hayashi ⁵³

時間軸に沿った研究デザインの分類

➤ 時間軸がある場合。

① **Cohort Study**

- **Prospective cohort study**
- **Historical cohort study**

② **Case-control study**

- **Case-control study**
- **Nested case-control study**
- **Case-cohort study**

➤ 時間軸がない場合

① **Cross-sectional Study**

Tomoshige Hayashi ⁵⁴

Cross-sectional study

母集団

Risk factor+ -

Risk factor+ -

- 横断研究は、すべての測定が一時点で行われる。
- 有病率 (prevalence) が得られる。
- 疾患に関連する因子を持つ群と持たない群でのOdds ratio (オッズ比) を出すことができる。

Tomoshige Hayashi ⁵⁵

Cross-sectional study (横断研究)で、Risk factorの検討はできるのだろうか？

- 疾患に関連する因子と対象疾患の因果関係は、言及できない。

例) NHANES IIIから、小児の肥満とテレビ視聴時間に関連があることがわかりました。

- 肥満 → テレビ視聴時間延長
- テレビ視聴時間延長 → 肥満

Tomoshige Hayashi ⁵⁶

疫学研究のデザインと統計解析。
臨床医は、どのような解析を知っている必要があるのだろうか。

Tomoshige Hayashi ⁵⁷

臨床医が熟知すべき統計手法

多変量解析	Tableの表示	エンドポイント
Multiple Logistic Regression Model (多重ロジステック回帰分析)	Odds Ratio	2値で表せる
Cox Proportional-Hazard Model (コックス比例ハザードモデル)	Hazard Ratio	2値で表せる
Multiple Linear Regression Model (重回帰分析)		連続数

Tomoshige Hayashi ⁵⁸

問 多変量解析の目的は

- リスクを検討すること。
- 疾患の発症リスクを検討するモデル式を作ること。

Tomoshige Hayashi

問 統計学的に有意かどうかをどう読むか？

Cox proportional-hazard model・ Multiple logistic regression modelの場合

- Tableのどこを見ても、統計学的有意差を示すP値が記載されていないことがあります。しかし、95%信頼区間(95%CI)の記載があります。



統計学的に有意かどうかは、この95%CIが、その区間に1を挟んでいるかどうかです。1を挟んでいると統計学的に有意ではありません。

Tomoshige Hayashi

(例) 以下の表は、Cox比例ハザードモデルが用いられたProspective Cohort Studyの一例です。

Ann Intern Med. 1999 Jul 6;131(1):21-6.

Variable	Person-Years of Follow-up	Cases of Hypertension, n	Multivariate Relative Risk (95% CI)*	Further Multivariate Relative Risk (95% CI)†
Walk to work‡				
0-10 minutes	30 796	337	1.00 (reference)	1.00 (reference)
11-20 minutes	23 266	242	0.91 (0.77-1.08)	0.88 (0.75-1.04)
≥21 minutes	5722	47	0.70 (0.59-0.95)	0.71 (0.57-0.97)
Walk to work as a continuous variable (per 10 minutes)			0.88 (0.78-0.98)	0.88 (0.79-0.98)

* Adjusted for age, body mass index, alcohol consumption, leisure-time physical activity (regular physical exercise at least once weekly or less than once weekly), smoking status (current smoker, past smoker, or nonsmoker), and fasting plasma glucose level.

† Adjusted for age, body mass index, alcohol consumption, leisure-time physical activity (regular physical exercise at least once weekly or less than once weekly), smoking status (current smoker, past smoker, or nonsmoker), fasting plasma glucose level, systolic blood pressure, and diastolic blood pressure.

‡ P for trend = 0.02 for multivariate relative risk and further multivariate relative risk.

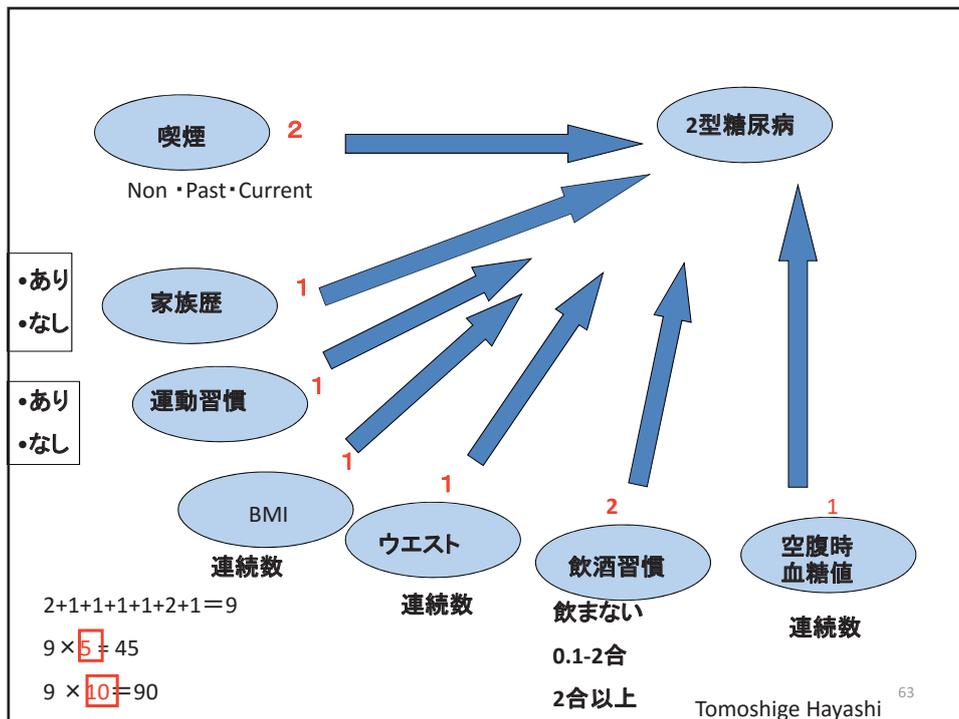
Tomoshige Hayashi 61

問 Prospective Cohort Studyをするには、どのくらいの規模の集団が必要でしょうか？
解析には、Cox proportional-hazard modelを用います。

- ① 10万人
- ② 1万人
- ③ 1000人
- ④ 500人
- ⑤ 100人

どの規模が一番すぐれた研究だと思えますか？

Tomoshige Hayashi 62



問: LinearかNon-linearかのチェックをしていますか?

Proportional Cox Hazard Model・Multiple Logistic Regression Model・Multiple Linear Regression Model共通

問: 多変量解析において、どのようにしてモデルに加える変数を選ぶべきか?

Proportional Cox Hazard Model・Multiple Logistic Regression Model・Multiple Linear Regression Model共通

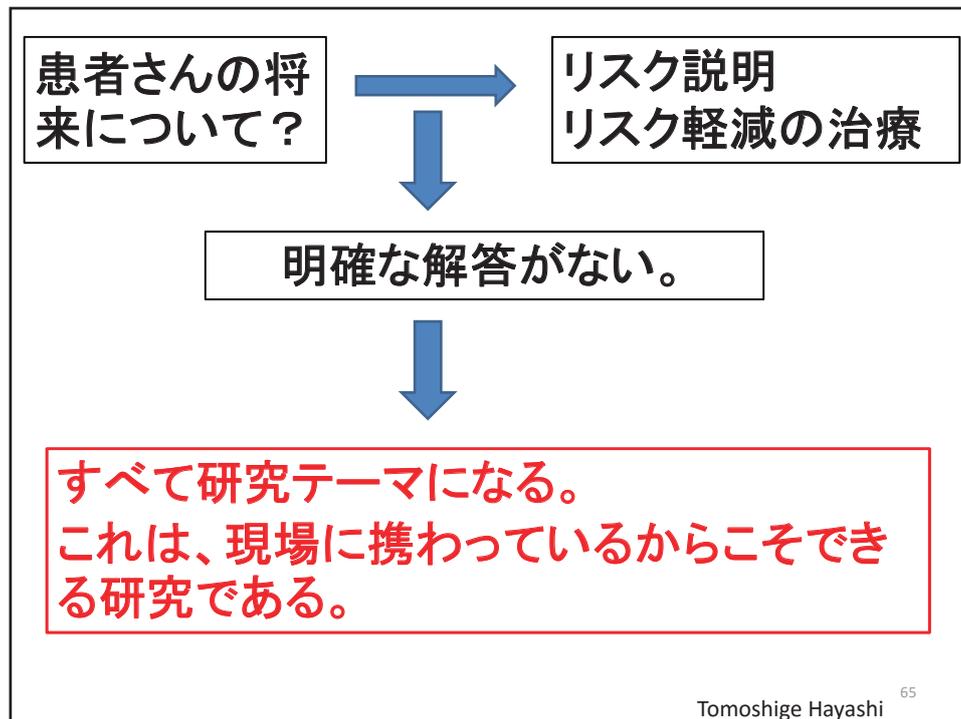
問: Effect Modification (Interaction: 交互作用) をチェックしていますか?

Proportional Cox Hazard Model・Multiple Logistic Regression Model・Multiple Linear Regression Model共通

問: Multicolliniarity (多重共線性)のチェックはしていますか?

問: 比例ハザード性のチェックはしていますか?
Proportional Cox Hazard Modelの場合

問: モデル式のチェックはしていますか?



疫学研究の注意事項!!

- ① 研究デザインを明確にしていない。
- ② 研究目的が明確でない。
- ③ 対象: Who, Where, Whenを明示しない。
- ④ 要因やアウトカムの定義が曖昧。
- ⑤ 変数の測定やアウトカムの測定に対して信頼性と妥当性を検討していない。
- ⑥ 解析方法を事前に決めない。
- ⑦ 統計解析は統計ソフトが全部してくれると思っている？(ガラガラポンではない)。
- ⑧ 研究結果の臨床的・社会的な適応を考えていない。

Tomoshige Hayashi

66

