

疫学・生物統計学（公衆衛生学第2回配布資料）

中澤 港

2017年10月9日

1 参考文献

- 柳川洋 (2012) 『疫学ノート—基礎から実践まで—』日本公衆衛生協会
- Rothman KJ (2012) Epidemiology: An Introduction 2nd Ed. Oxford Univ. Press (邦訳『ロスマンの疫学』)
- 『保健師・保健師をめざす学生のための なぜ? どうして? 4: 疫学・保健統計』メディックメディア
- 厚生労働省「治験」ホームページ (<http://www.mhlw.go.jp/topics/bukyoku/isei/chiken/index.html>)
- ゲルト・ギーゲレンツァー (2010) 『リスク・リテラシーが身につく統計的思考法』ハヤカワ文庫
- デイヴィッド・ザルツブルグ (2010) 『統計学を拓いた異才たち』日経ビジネス人文庫
- 佐藤俊哉 (2005) 『宇宙怪人しまりす 医療統計を学ぶ』岩波書店

2 概論

2.1 疫学 (epidemiology) の定義と目的

特定された集団*1で健康事象*2の頻度 (frequency) と分布 (distribution)*3を調べ、原因との因果関係 (causal relationship) を探り、病気や死亡を減らすことへの貢献を目指す学問

2.2 疫学研究のフレームワーク

疫学研究の手順は、“5W-Bridge” に集約される。いつ (When), どこで (Where), 誰が (Who), どんな病気に (What), 何故 (Why) 罹ったかを明らかにできれば、原因も突き止められるということ。

2.3 疾病分類 (classification of diseases)

第10回修正国際疾病分類。時空が異なっても共通の分類体系に依拠して定義された病気が同定できないと、病気の頻度や分布を調べることができない。そのため提案された、国際的に共通する疾病分類が、ICD (International Classification of Diseases)。1900年制定で、現行のものは1995年から使われている第10回修正版のためICD-10と呼ばれる。ICD-11へのreviseが進行中。

3 研究デザイン

研究デザインには観察研究（記述研究と分析研究）と介入研究があり、目的によって、最適なデザインの研究をすべきである。

- 疫学研究は大きく分けると観察研究と介入研究に分かれる
- 観察研究では、研究者自身が対象集団に対して意図的に介入し、疾病に関する状態を能動的に変えることはない
- 介入研究では、研究者自身が集団に対して意図的に介入し、能動的に割付けを行って、介入の結果によって疾病改善効果が見られるかどうかを検討する
- 介入研究は、観察研究によってかなりの確からしきで因果関係がありそうなことがわかってからでないと倫理的に実施できないので、アプローチの違いというよりも、段階の違いと考えるべき

*1 specified population

*2 概ね病気だが、病気に限らない

*3 時間分布と空間分布がある

3.1 観察的疫学研究のいろいろ

- 記述疫学：健康事象発生の頻度と分布のデータを収集し、図示や記述統計量を示すなどにより性状をつかむ。
- 分析疫学
 - － 自然実験：John Snow がロンドンのコレラで水道供給会社別に死亡率を比較したように、偶然比較可能な状況が整っていると強力な分析が可能だが、そういうケースは稀。
 - － 生態学的研究（地域相関研究）：国単位で一人当たり GDP と合計出生率の関連をプロットしてみるなど。誤謬に注意。
 - － 横断的研究：ある時点で調査対象が疾病をもっているかどうか、疾病の原因をもっているかどうか
 - － 症例対照研究：ある時点での患者群 (case) に対し、その時点でその疾病をもっていない対照群 (control) を選択し、リスク因子への過去の曝露状況を、患者群と対照群の間で比較
 - － コホート研究：あるリスク因子に曝露した集団を、その後、曝露コホートとして追跡調査 (followup study) し、疾病の発生率を観察。そのリスク因子に曝露していない点だけが異なる非曝露群との比較するのが理想
 - － 他に、ケースコホート研究、コホート内症例対照研究など

3.2 介入研究

- 介入研究では、研究者が要因曝露をセッティングすることにより、要因曝露の有無だけが異なる対照群を作り出すことができる。
- 因果関係の入口を決めて出口を評価するという点ではコホート研究に似ているが、曝露の有無をランダムに研究者が割付ける点と、想定される出口が1つである点の特異
- 薬を開発する際の臨床試験 (= clinical trial, 治験ともいう) で盛んに行われる。
- 臨床試験には第1相から第4相までである。
- 中でも RCT (Randomized Controlled Trial; ランダム化統制試験) は、最も科学的に厳密な仮説検定の方法とみなされている。
- 第3相臨床試験では、通常 RCT が行われる。

4 標本抽出法とサンプルサイズ

母集団を正しく代表する標本を抽出することが肝要。最も考え方が単純なのはランダムサンプリング (Random sampling, 単純無作為抽出法)。大集団からサンプリングの偏りをなくすためには、多段抽出法、とくに層別無作為抽出法*4を用いる。途上国の調査ではクラスターサンプリング (Cluster sampling, 集落抽出法) がよく使われる。

記述研究の最適サンプルサイズは、母集団のサイズ、許容標本誤差、予想頻度から計算できる。分析研究のサンプルサイズは使う統計的検定手法ごとに、予想される測定値のばらつき、臨床的あるいは科学的に意味のある差や相関のレベル、有意水準、検出力を決めて、それぞれの計算式に従って計算できる。ソフトウェアで計算可能だが、費用や時間などの制約により、最適サンプルサイズの研究ができるとは限らない。

5 疾病量の把握

疾病量は、横断的研究では有病割合または疾病オッズ、コホート研究では罹患率や累積罹患率 (リスク) として得られる。症例対照研究では疾病量そのものは原理的に把握できない*5。

5.1 疾病頻度の指標

横断的研究では有病割合と疾病オッズが得られる。症例対照研究では症例群、対照群それぞれにおける曝露オッズが得られる。コホート研究では罹患率や累積罹患率 (リスク) が得られる。以下3つの区別は重要。

有病割合 (prevalence) ある時点で調べた人のうち病気だった人の割合。無次元。

*4 Stratified random sampling. 層化無作為抽出法ともいう

*5 敢えて言えば、研究対象施設における罹患数は疾病量といえるが。

罹患率 (incidence rate) 観察した人・時間の合計で罹患数を割った値。(／時間) という次元をもつ*6

累積罹患率 (cumulative incidence rate) リスクともいう。観察開始時にいた人のうち、観察期間内に罹患した人の割合。無次元。

- 疾病負荷を示すのは有病割合、発生頻度を示すのは罹患率やリスク。
- 断面研究 (横断的研究) で得られるのは有病割合 (と疾病オッズ) のみ。罹患率や累積罹患率を調べるにはコホート研究が必要。記述的研究では、必ず得られるのは罹患数や死亡数 (の報告数) のみ

5.2 致命割合 (case-fatality ratio)

ある疾病に罹患し確定診断がついた人のうち、その疾病で死亡に至った人の割合 (通常、% で表す)*7 疾病の重篤度を示す (狂犬病は 100%、エボラウイルス病は 50~90%、高病原性鳥インフルエンザは 60% 程度、1918 年にパンデミックを起こしたスペイン風邪インフルエンザが 2.5%、季節性インフルエンザは 0.05~0.1%)

5.3 年齢による標準化

疾病罹患や死亡は年齢による影響を受けるので、年齢構成が異なる集団間で比較するためには (例えば年次推移を見たり国際比較するためには)、年齢による標準化を行う必要がある。標準化の方法には直接法と間接法がある。直接法の考え方は簡単だが、年齢別の率が必要なのでデータを得にくい途上国では使いにくい欠点がある。

- 対象集団の年齢構成が基準集団と同じだった場合に対象集団の年齢別死亡率に従って死亡が起こったら全体としての死亡率はどうかと考える
- 基準集団の年齢構成を重みとする、対象集団の年齢別死亡率の重み付き平均ともいえる
- 年齢構成を揃えて考えるために使う基準集団の年齢別人口を i 歳について P_i 、対象集団の年齢別死亡率を m_i と書くと下式の通り。

$$\text{直接法年齢調整死亡率} = \frac{\sum^i m_i \cdot P_i}{\sum^i P_i}$$

- 対象集団の年齢別死亡率の情報が必要
- 対象集団が基準集団の年齢別死亡率に従って死んだ場合に期待される死亡数で、実際の対象集団の死亡数を割った値を、標準化死亡比 (SMR) と呼ぶ。対象集団の死亡総数を d 、対象集団の年齢 i 歳の人口を p_i 、基準集団の年齢 i 歳の死亡率を M_i と書くことにすると、下式で得られる。つまり、対象集団の粗死亡率を、対象集団の年齢別人口を重みとする基準集団の年齢別死亡率の重み付き平均で割った値になる。

$$\text{SMR} = \frac{d}{\sum^i M_i \cdot p_i} = \frac{d / \sum^i p_i}{\sum^i M_i \cdot p_i / \sum^i p_i}$$

- SMR を基準集団の粗死亡率に掛けた値が間接法年齢調整死亡率となる。上記記号を使うと下式の通り。

$$\text{間接法年齢調整死亡率} = \text{SMR} \times \frac{\sum^i M_i \cdot P_i}{\sum^i P_i}$$

- 対象集団についての情報としては、年齢別人口と総死亡数だけで計算可能

6 危険因子とその影響 (効果) の指標

危険因子曝露の効果を示す指標としては、横断的研究と症例対照研究ではオッズ比、コホート研究や介入研究では罹患率比やリスク比、死亡率比を用いることができる (ただし介入研究の場合はコックス回帰でハザード比を計算する方が普通)。

6.1 影響 (効果) の指標

危険因子がある群と危険因子がない群の間で疾病量の差や比 (または変化率) をとるとというのが基本的な考え方。

*6 単位時間当たりの発生速度を意味する。

*7 西浦らは、2009 年パンデミックインフルエンザの流行時に、当初この定義に従って 0.4% など高めの値が報告されていたけれども、夏頃から受診しない人や確定診断がつかないままに治療を受けた人を考えると 0.045%、0.005% など低い値が相次いで報告され、それが CFR と混同されたために当初報告値が過大評価という批判を受けたことから、確定診断がついた症例中の致命割合 (confirmed case fatality ratio; cCFR) と症状から罹患したと考えられる人の中での死に至った割合 (symptomatic case fatality ratio; sCFR) を区別することを提唱している。

超過危険 (excess risk) 寄与危険 (attributable risk) ともいう。曝露群のリスクや罹患率から非曝露群のリスクや罹患率を引いた差。コホート研究と介入研究で得られる。リスクの差はリスク差、罹患率の差は罹患率差、死亡率の差は死亡率差。公衆衛生的なインパクトを見るのに適している。統計学的に0と有意に異なれば曝露は罹患に効果ありといえる。

相対危険 (relative risk) 危険因子への曝露群の非曝露群に対するリスクの比 (リスク比), 罹患率の比 (罹患率比), 死亡率の比 (死亡率比) の総称。コホート研究と介入研究で得られる。因果関係を検討するのに適している。公衆衛生的なインパクト評価としては過大な可能性。統計学的に1と有意に異なれば曝露は罹患に効果ありといえる。

オッズ比 (odds ratio) 断面研究では要因あり群の疾病オッズの要因なし群の疾病オッズに対する比 (疾病オッズ比), 症例対照研究では患者群の曝露オッズの対照群の曝露オッズに対する比 (曝露オッズ比) をいう。下記の通り両者は数学的に一致する。稀な疾患では相対危険の近似値となる。症例対照研究では, 対照群を密度依存サンプリングするなど工夫すれば, オッズ比が理論的に率比に一致する。

クロス集計表を作るとわかりやすい。

| | 病気あり | 病気なし |
|------|------|------|
| 曝露あり | a 人 | b 人 |
| 曝露なし | c 人 | d 人 |

このとき, 断面研究における^{*8}疾病オッズ比 (disease odds ratio) は,

$$\frac{(a/b)}{(c/d)} = \frac{ad}{bc}$$

患者対照研究における曝露オッズ比 (exposure odds ratio) は,

$$\frac{(a/c)}{(b/d)} = \frac{ad}{bc}$$

6.2 寄与割合・人口寄与割合

寄与割合 (Attributable Proportion) 曝露群の罹患率のうちその曝露が原因となっている割合。つまり罹患率差を曝露群の罹患率で割った値。罹患率比から1を引いて罹患率比で割った値とも等しい。

人口寄与割合 (Attributable Population=Attributable Fraction) 母集団の罹患率のうちその曝露が原因となっているものを取り除くとどれくらいの割合, 罹患率を下げられるか? という値。

6.3 相対危険と超過危険の関係

相対危険が1で超過危険が0という状態を除けば, 相対危険が経時的に不変, 即ち曝露群のリスクと非曝露群のリスクの比が年齢によらず一定 (>1) であるとしたら, 加齢に伴ってリスクそのものが大きくなることによって曝露群と非曝露群のリスクの差が拡大していくので, 超過危険が加齢とともに増加する。逆に, 超過危険が経時的に不変, 即ち曝露群と非曝露群のリスクの差が年齢によらず一定であるとしたら, やはり加齢に伴ってリスクそのものが大きくなることによって, 相対危険は減少していく。

^{*8} コホート研究でも考えることは可能だが, コホート研究ではリスク比や率比を計算できるので, 敢えてオッズ比を計算する理由はないだろう。