

2001年1月26日土曜講座研究発表より 素人のための疫学入門

人々に共通するリスクの
有無や程度は
どのように推し量ることが
できるのか？

上田昌文 + 小牧史枝

この講座の前半では、疫学分析の道具となる統計学の基礎事項のうち、「標本から母集団の平均・標準偏差を推定し、母平均の95%信頼区間を求めること」、そして「2つの標本を比較して差があるかどうかを検定すること(主にt検定の場合について)」を学びました(発表は主として小牧史枝が担当)。

講座の後半では、私たちが抱えている健康に関わるリスクをできるかぎり正確に定量的にとらえ、しかも、健康被害の生物メカニズムを知り得なくてもある程度まで原因にせまるための方法として、疫学という学問を紹介します。疫学は、統計学の手法を使いながら常識的な論理の積み重ねでリスクをとらえようとするものですから、その基礎事項を学ぶことは誰でもできますし、様々なリスクを適切に判断するためにはその考え方に慣れ親しんでおくことが大切なのです。私たちが一つの有効な知的ツールとして疫学的考え方が生かせるようになることが目標です。

疫学とは

- 語源 epidemiology “人間集団に関する学問”
epi = upon, among / demos = people / logos = science
- 定義: 人間集団内の健康、疾病、異常事象の発現頻度に関する法則性を見出し、発生要因を探ること、疾病予防や健康増進に役立てる学問。
- 典型事例:
 - ・1854年
ロンドンのコレラ発生をつきとめたJ.Snow
 - ・1890年代
脚気の原因を推測し予防法を見つけた高木兼寛

疫学とは何でしょうか? 「疾病の分布状態を調査し、その原因を探り、疾病を予防することを目的とする学問」と簡単に述べておくことはできますが、その語源を探ると“人間集団に関する学問”であることがわかります。集団をとおして見られる、疾病の主に発生頻度の法則性から、その原因を探っていくというのが疫学の特徴です。

すでにJ.Snowはコレラ菌が発見される30年も前に、そして高木兼寛はビタミンBの発見に先立つこと10数年の時点で疫学的手法を用いて病気の流行を防ぎました。

疾病を集団的に分析するときの 指標(1)発生頻度

- 比(ratio) $x:y$ または x/y (0 比)
- 割合・比率(proportion) 全体の中で占める比 (0 割合 1)
- 率(rate) 単位時間における事象の変化分 (0 率)

例: 3年間で平均10万人の女性集団で乳ガン患者が120人発生
・1年における発患者数 = 40人/年 (ただし10万人の中で)
・1年における発生率 = $40/10,0000$ [年]

では実際に疾病を集団的に分析するとは何をすればよいのでしょうか? それにはまず疾病の発生の頻度を正確にとらえることです。そのために、比、割合、率という言葉を使い分けます。

人年法(person-year method)

- 1人年 = 1人の対象者を1年間観察した場合の観察期間を1単位として表す。
例: 10人の人を2年、それとは別30人の人を4年観察した場合、
 $10 \times 2 + 30 \times 4 = 130$ 人年
・ただし疾患が観察期間のはじめと終わりで発生頻度が大きく変わることもある。
・観察から脱落する人の経過の確認も必要になることがある。

疾病の起こり具合を見ると、観察期間を決めたらその最初から最後まで対象とする人々をすべて観察することが理想です。しかし現実にはそれは難しい。したがって、固定された期間に限定せず、個々人についてそれぞれ観察可能な期間を足し合わせることで、全体の状況をつかもうとします。それが「人年」の考え方です。

罹患(発症、発生)率 (incidence rate)

$$\text{罹患率} = \frac{\text{観察期間内の新患者発生数}}{\text{観察集団内の全員の人年の総和}}$$

- 疾患のない人が特定の疾患にかかる速度(期間は特定しない)
- 対象者全員について病気にかかり得る状態にある期間の合計人年をとる
- 実際には観察開始時と終了時の対象人口の平均で考えることが多い

まず、その人年を用いて把握することができるのが「罹患率」です。

たとえば「この冬の土曜講座メンバーのインフルエンザ罹患率」を考えてみましょう。必ずしも全員について、12月1日から2月28日まで観察できるわけではありません。メンバーの中に途中で南半球(夏です!)へ旅行に行く者があれば、どう扱いますか? また12月1日以前にすでにインフルエンザにかかっていた人はどうしましょうか? 知りたいのは、「この冬にいったいどれくらいのスピードでインフルエンザが広がるのか」ということです。つまり「ある一定の時間がたてば、どれくらいの人が新たにその病気にかかるか」です。それを算出するのが、「罹患率」です。

罹患率で注意すべき点

- 観察期間とは、対象者が疾病に罹患する危険にさらされている時間である。したがって
- 該当する疾病にすでに罹患している者は観察対象集団に含まない。
- 観察中に罹患した者はそれ以降の観察では、対象から除外する。
- 疾病が発症して調査の間に治癒しても有病者と見なす。
- 観察対象の個人個人について、その観察開始時期が異なってもよい。
- 途中で観察できなくなっても観察できた期間は観察人年に加える。
- 年の途中で転入、転出、死亡、罹患が発生した場合には、0.5年と計算する。

罹患率の定義がわかれば、疫学調査を厳密にすすめるとき、このような条件を設けなければならないことは、理解していただけるでしょう。

累積罹患(発症、発生)率 (cumulative incidence)

$$\text{累積罹患率} = \frac{\text{一定観察期間内の新患者発生数}}{\text{観察開始時点での人口集団の人数}}$$

- いわゆるリスク
- 対象疾患にかかっていない人が特定期間中にそれにかかる割合

次に、「疾病の広がる速さ」がわかれば、当然今注目する集団において、「その観察期間でどれくらいの割合の人がその病気になるのだろうか」ということを知りたくなります。つまり「土曜講座の何割の人がこの冬にインフルエンザにかかったのか」ということです。この割合が高ければ、「インフルエンザという病気のリスクが大きい」ということとなりますから、この「累積罹患率」はいわゆる「リスク」というのにかなり近い概念なのです。

たとえば土曜講座のメンバーが100人いて、先の12月のインフルエンザ罹患率を求めるには、その12月に新たにインフルエンザにかかった人の人数を把握すればよいのです。それが20人のとき、累積罹患率は0.2ということになります。

気をつけて欲しいのは「率」と言っていますが、先の定義に従うなら、ここでは「累積罹患割合」という言うべきである、ということです。

有病割合(prevalence)

$$\text{有病割合} = \frac{\text{ある時点で対象とする病気にかかっている人の数}}{\text{ある時点でのすべての観察対象者数}}$$

もう一つ、単純だけれど別の「病気の頻度」をとらえる方法があります。「病気の広がる速度」や「新たにどれくらいの割合で罹るか」ではなく、ただ「今どれくらいの人が罹っているか」を示す「有病割合」です。

今日この日に土曜講座メンバーではどれくらいインフルエンザにかかっている人がいるのでしょうか? この場合は、一度罹ってはいるがもう直った人は、カウントできません。したがって、罹患期間の短い病気は、「有病割合」が低くなる傾向があります。

「有病割合」は病気の新発症を表すものではないので、罹患率と異なり、疾病発症の要因を明らかにする指標としてはあまり有用ではありません。しかし有病割合を知ることが、行政的な施策としてその病気に罹患している人を治療し援助する場合に、必要な医療資源や社会福祉資源を算定する上で、極めて有用な指標となります。

死亡率、累積死亡率、致死率

$$\text{死亡率} = \frac{\text{観察期間内でのその病気による死亡者の数}}{\text{観察集団内の全員の人数の総和}}$$

$$\text{累積死亡率} = \frac{\text{一定期間内に新たにその病気死亡した者の数}}{\text{観察開始時点での人口集団の人数}}$$

$$\text{致死率} = \frac{\text{その疾患で死亡した人数}}{\text{ある特定の疾患にかかった人の数}}$$

疾病の特殊な結果場合として「死亡」を考えれば、死亡についても、以上で示した指標をまったく同様に考えることができます。すなわち、「死亡率」とは、「ある一定期間の死亡数が観察集団の人口に占める割合」であり、「死亡という事象の発生率」です。

ここでたぶん一番身近なのは「致死率」でしょう。これは分母が「その疾患に罹った人」ですから、「死亡率」とは違います。いわ

ゆる不治の病は「致死率100%」であって、「死亡率100%」という言い方はおかしいことになります。

疫学 演習問題1

- (1) ある医学検査で、5年間に270例の十二指腸潰瘍がある町の男性に起こった。その町の男性人口は、観察開始時18,500人、終了時には21,500人であった。どのような指標が測定できるか。
- (2) 65歳の男性1000人の集団検診を行なった。100人がある病気にかかっていることが分かった。続く10年間を経過観察したところ、160人がこの病気にかかった。病気の発生に関してどのような指標が測定できるか。

では、下の解答を見ないで、この問題を解いてみてください(文献1より)。

疫学 解答1

- (1) 観察期間の間の平均人口は20,000人。
対象集団の年人は100,000人年。
わかるのは罹患率で、 $270 / 100,000 = 0.0027 / \text{年}$
- (2) わかるのは有病割合と累積罹患率で、罹患率は求められない。
有病割合 = $100 / 1000 = 0.1$
累積罹患率 = $160 / (1000 - 100) = 0.17$
(10年間において)

いかがでしたか？ 罹患率と累積罹患率の違いを実感していただけただけでしょうか？ ではもう1題(文献4より)。

疫学 演習問題2

ある町の2000年1月1日の人口は80,000人、2000年12月31日の人口は90,000人であった。この1年間の肺ガン死亡数は40人であった。また、1月1日現在の結核患者数は45人、1年間の新発生結核患者数は30人であった。1月にインフルエンザが流行があり、1,000人の患者が発生し、そのうち8人がインフルエンザが原因で死亡した。

- 1年間の観察年人は？(平均をとって考える)
- 年間の肺ガン死亡率(10万人年対)は？
- 1月1日現在の結核有病割合(人口10万対)は？
- 年間の結核罹患率(10万人年対)は？
- 年間の結核の期間有病割合(人口10万対)は？(平均して考える)
- 年間のインフルエンザ死亡率(10万人年対)は？
- インフルエンザの致死率(%)は？

疫学 解答2

$$(80,000+90,000) / 2=85,000 \text{人年}$$

$$40 / 85,000 \times 100,000 = 47.1$$

$$45 / 80,000 \times 100,000 = 56.3$$

$$30 / 85,000 \times 10,000 = 35.3$$

$$(45+30) / 85,000 \times 100,000 = 88.2$$

$$8 / 85,000 \times 10,000 = 9.4$$

$$8 / 1,000 \times 100 = 0.8$$

すべて簡単な計算で求めることができますね。

疾病を集団的に分析するときの指標(2 - 1)リスクの差

- 暴露効果の指標
暴露群(リスク(危険)因子にさらされた集団)と非暴露群との違いを測る
(1)リスクの差を見る
発生率差
= 暴露群の発生率(罹患率) - 非暴露群の発生率
単位は[数/人年](- から まで)
- 寄与危険度
= 暴露群の累積発生率 - 非暴露群の累積発生率
単位は無単位 (- 1から1まで)

さて、今度は「病気がどれくらい起きているか」ということから一歩進んで、「いったいその病気は何によってもたらされているか」を推測することへと目を向けてみます。

疫学では、今着目しているその病気、障害、異常などの原因となりそうなことさらに目星をつけ、その原因にさらされた人(「暴露した人」とさらされていない人(「暴露していない人」)を対照させることで、その原因を探っていくとします。そのときの一つの方法は、この2集団について「リスクの差」の大きさを調べることで、上記2種類の指標に応じて、「発生率差」と「寄与危険度」が定義できます。

疾病を集団的に分析するときの指標(2-2)リスクの比

■ (2)リスクの比を見る

$$\text{発生率比} = \frac{\text{曝露群の発生率}}{\text{非曝露群の発生率}}$$

$$\text{相対危険度(RR)} = \frac{\text{曝露群の累積発生率}}{\text{非曝露群の累積発生率}}$$

$$\text{寄与割合} = \frac{\text{曝露群の発生率} - \text{非曝露群の発生率}}{\text{曝露群の発生率}}$$

さらに、「リスクの比」を見ることもできます。疫学ではこちらをよく用います。つまり、「曝露群は非曝露群に比べて、何倍早く病気になるやすいか」という発生率比、「その病気になる人が何倍多くなるのか」という相対危険度、そして、「その原因だけを取上げてみると、その病気にかかりやすくしていることに、どれくらい寄与しているといえるのか」を表す寄与割合、です。

これらのうち、相対危険度(Relative Risk, RR)はいわゆる「リスク比」ともいわれるもので、この指標をもとにして、たとえば「小児白血病のリスクが3.5倍に高まった」と言うわけですね。

後述しますが、相対危険度RRは、計算の際に曝露の有無で群を分ける必要があるため、もっぱらコホート研究で算出される指標です。症例対照研究ではオッズ比をもって相対危険度とみなすこととなります。

では、問題です。相対危険度を2つ求めます(文献4より)。

疫学 演習問題3

幼少時に甲状腺腫のために放射線治療を受けた2,872人の人に、24例の甲状腺癌が、52例の良性甲状腺腫が発生した。対照群として、5,055人の上記のような治療を受けなかった子どもが選ばれた(先の子どもの兄弟姉妹から)。追跡期間中に対照群には甲状腺癌は発生せず、6例の良性甲状腺腫が発生した。

甲状腺癌と良性甲状腺腫について相対危険度を求め、結果を考察せよ。

疫学 解答3

- 曝露群の甲状腺癌の累積発生率 = $24 / 2,872 = 0.00836$
- 非曝露群の甲状腺癌の累積発生率 = 0
- よって甲状腺癌の相対危険度 = $0.00836 / 0 =$
- 曝露群の良性甲状腺腫の累積発生率 = $52 / 2,872 = 0.01811$
- 非曝露群の良性甲状腺腫の累積発生率 = $6 / 5,055 = 0.00119$
- よって甲状腺腫の相対危険度 = $0.01811 / 0.00119 = 15$
- 甲状腺癌については、おそらく非曝露群の人数が少なすぎるか追跡期間が短かすぎて、甲状腺癌が検出できなかったため、相対危険度は推し量れない。
- 良性甲状腺腫については曝露群が非曝露群に比べて15倍リスクが高まる。

いかがでしたか。相対危険度が無限大になるということは、現実にはあり得ないはずですから、そのような不自然な結果をもたらしてしまったわけを考える必要があります。疫学についてはいつも、導き出され数字を見るだけでなく、それがどういう条件のもとでの観察で導き出されたのかを確かめておく必要があります。

疫学研究のデザイン

観察研究

横断研究: 特定の一時点で研究する。

継続研究: 時間経過を考慮したもの。

症例対照研究 case control study

疾病の有無で集団を設定する。

コホート研究 cohort study

曝露の有無で集団を設定する。

前向きコホート研究 / 後向きコホート研究

介入研究: 実験条件を人為的に作成する。

臨床試験 野外試験 地域介入研究

疫学の研究は、観察研究と介入研究に大別されます。観察研究の一番の基本は「記述研究」ですが、これは先のスノーや高木の研究にあるように、ていねいに緻密にデータを記述し、そこから推し量れることがらを次第につかみ出していくというものです。

疫学で最もよく使われているのが、コホート研究とケース・コントロール研究です。

コホート研究(cohort study)は後に詳しく取上げますが、前向き(prospective これから先)と後ろ向き(retrospective 過去に遡る)の2方向の研究があります。

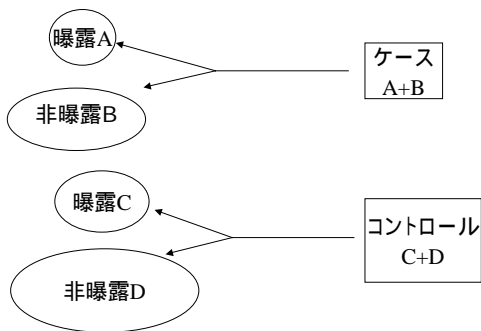
患者対照研究(case-control study)も後に詳しく取上げます。この2つの違い、つまり「疾病の有無」か「曝露の有無」のどちらに着目しているか、が大事です。

横断的研究は、仮説としてみている疾病の要因と疾病とを同じ時点で(断面で)調査するものです。原因と結果を同時に調査しているため、両者に統計的な関係が認められたとしても、それが因果関係を反映しているのか否かを区別することが難しいという問題点があります。

例えば、カロチンの摂取が多い群ほど慢性胃炎の割合が低いという関係が断面研究で認められたとしても、カロチンを多く摂取したから慢性胃炎の割合が低いのか、反対に、慢性胃炎の割合が低く健康な集団だからカロチンを含む野菜をたくさん食べられるのかを区別することはできないのです。

介入研究は、すべて「前向き」になります。臨床試験は患者を対象とし、たとえばエイズ新薬の開発がそうです。野外試験は、一般住民が対象でたとえばポリオワクチンの効果をそれで見ます。地域介入研究は、たとえばある地域の水道水のフッ素を入れた場合の虫歯に対する効果、などがあります。

ケース・コントロール研究とは？



ケース・コントロール研究とは何を意味するのでしょうか？

ケースとは「症例」すなわち、ある疾患なり障害なりをもった「患者」のことです。コントロールとはそれと「対照」させる「非患者」です。すなわちその病気にかかっていない者です。この図をみて、ケース・コントロール研究の意図を汲み取ることができますか？

そうです。この研究によって、ある特定の曝露要因がどれくらい疾患の発生に寄与しているかを推測することができるのです。

ケース・コントロール研究の特徴

- 疾病の有無によって患者群と対照群を設定し、過去に遡ってその中での曝露者の割合を比較する。(コホート研究は原因の有無でコホートを定義するのに対して、症例対象研究では結果の有無で対象を分類)。
- 症例対照研究の利点は、非常に稀な結果をもたらす要因のリスクを評価できることや、コホート研究に比べてコストが安いこと。
一方、様々なバイアスが混入しやすいという欠点を持つ。

ケース・コントロール研究は「後向き研究(retrospective study)」とも呼ばれます。最初に、すでに疾病にかかった人を「症例(case)」として選び出します。次に、この「症例」と性別や年齢などの条件が似た人を「対照(control)」として選びます。「対照」は、「症例」と同じ地域に住む健康な住民から選ばれる場合(「住民対照」と呼ばれる)と、「症例」と同じ病因に入院している患者から選ばれる場合(「病院対照」)などがあります。「症例」と「対照」の双方に対して、疾病の原因と考えられる要因(例えば食生活など)を、過去にさかのぼって調査し、両者で比較するわけです。

簡単に言えば、「症例」とは患者のことであり、「コントロール：対照」とは健康な人のことであり、この研究では「患者」と「健康な人」を選んで比較するわけです。例えば、サルモネラ菌で食中毒に罹った患者が散発的に現われたとします。患者たちからの聴取調査により「いか菓子」が原因として浮かんできました。そこで、「健康な人」をランダムに選び、過去に「いか菓子」を食べたことがあるか調査し、発症の有無と「いか菓子」の関係を究明する、

といった具合です。ケース・コントロール研究は、すでに疾病に罹患した者を対象にするため、コホート研究よりも手間や労力が少なく済みます。その反面、仮説要因を過去にさかのぼって調査するため、例えば肺がん患者が過去の喫煙量を実際以上に過大に思い出して申告するなどして、仮説要因と疾病との関連を過大評価したり、反対に過小評価したりする危険性があります(情報バイアス)。また、「住民対照」を選択する際に、一般住民の中でも特に調査に協力的で生活習慣も健康的な者を偏って選んだり、「病院対照」を選択する際に、一般の住民より生活習慣の不健康な者を選んだりして、仮説要因と疾病との関連を正しく評価できない危険性もあります(選択バイアス)。ケース・コントロール研究の結果を解釈する際には、これらの問題点を十分に考慮する必要があります。

ケース・コントロール研究を実施する上でのポイント

- 【選択バイアス】を排除すること
患者と対照群の設定が曝露経験の有無によって影響を受けばバイアスを生じる。
- 【マッチング】によって精度を高めること
比較の精度を高めるために(少数例でも精度を上げるために)ケースとコントロールで、曝露要因以外のいくつかの条件について条件が揃った「組」を作って、その組の数で比較することがある。

ケース・コントロール研究で有名なものには「能動喫煙と肺癌」(Hill & Doll 1950年、1952年)や「サリドマイド事件」(Lentz 1961年)、『奪われし未来』でも詳しく紹介されている「DES(ジエチルスチルベストロール、合成エストロゲン)と膵腺癌」(Herbst 1971年)などがあります。これらはすべて、患者ならびに対照に選ばれた者にインタビューを行い、仮説要因のリスクを計量して、その原因を推定することに成功した事例です。

こうした研究では、バイアスをいかに慎重に排除するかが大事です。たとえば、「喫煙と肺ガンの関連を調べる」場合、「喫煙者」を優先的に診断するような状況があれば、肺癌患者に必然的に「喫煙者」が多くなってしまおうし、「妊娠中に服用した薬が原因で奇形の子どもが生まれたかどうか」を調べるとき、奇形の子どもを産んだ母親は奇形でない子を産んだ母親より、妊娠中の服用歴を正確に思い出すことが多いものです。

研究の精度を上げるために、「マッチング」も用いられます。たとえば、「生理用品(タンポン)の継続使用と毒性性ショック症候群との関連」では、「女性でしかも同じ年齢」という条件を揃えて1:1の「ケース:コントロール」のペアを作って調べられました(Shands 1980年)。

ケース・コントロール研究のデータ解析(1) [マッチングのない場合]

- 曝露した患者(ケース)の割合 = $a/a+c$
- ケースにおける曝露オッズ = (曝露した割合)/(曝露していない割合) = $(a/a+c)/(c/a+c) = a/c$
- コントロールの曝露オッズ = b/d

曝露	患者	対照	計
あり	a	b	a + b
なし	c	d	c + d
計	a+c	b+d	a+b+c+d

オッズ比

$$= \frac{\text{ケースにおける曝露オッズ}}{\text{コントロールにおける曝露オッズ}}$$

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

ケース・コントロール研究では、仮説となる疾病要因と疾病の関連性をあらわす指標として、オッズ比(odds ratio)が使われます。この比を求めるために、「2×2の表」を作らなければならないのですが、収集したいいろいろな情報を、この表を作成して分析と評価に持ち込むという点が、疫学の核心をなしています。

食中毒の例で言うと、
 疾病の有無(食中毒のある、なしの2つの場合)
 曝露の有無(食べた、食べないの2つの場合)

	有症	健康
食べた	a	b
食べない	c	d

で合計2×2の4つ指標の類別をつけ、それぞれの指標の人数を確定することで、原因推定のための出発点が与えられます。そしてこの場合、オッズ比(OR)の値によって、

- OR = 1の場合: 食べた人と食べない人の発症割合は同じ(両群に差はない)
- OR > 1の場合: 食べた人の発症割合の方が大きい(食べた人の方がより発症している)
- OR < 1の場合: 食べない人の発症割合の方が大きい(食べない人の方がより発症している)

が示されることとなります。

タバコの例で言うと、喫煙者の非喫煙者に対する肺がん罹患のオッズ比が4であれば、喫煙者は非喫煙者と比べて4倍肺がんに罹りやすいことを意味します。

ケース・コントロール研究のデータ解析(2) [マッチングのある場合]

- この場合は、各欄に入るのはペア数であり、wとzはケースとコントロールの曝露の有無がまったく同じ(結果に何も情報を与えない)。よって、この場合のオッズ比は x/y で与えられる。

	対照群		計
	曝露あり	曝露なし	
患者群			
曝露あり	w	x	w+x
曝露なし	y	z	y+z
計	w+y	x+z	w+x+y+z

これは、詳しく説明しませんが、先のマッチングしない場合と、表の個々の項目の意味合いが違っていることに注意してください。

相対危険度とオッズ比

	疾病あり	疾病なし	計
曝露あり	a	b	a + b
曝露なし	c	d	c + d

コホート研究
 曝露群のリスク = $a / (a+b)$
 非曝露群のリスク = $c / (c+d)$
 相対危険度 = $\frac{a/(a+b)}{c/(c+d)}$

ケース・コントロール研究
 ケースの曝露オッズ = a / c
 コントロールの曝露オッズ = b / d
 オッズ比 = $(a / c) / (b / d)$
 ここで疾患が稀なとき(a, cが大変小さく)、2つの比は一致する。

曝露	患者	対照
あり	a	b
なし	c	d
計	a+c	b+d

事象がどのくらい確実に起こるかの度合いを表現するには確率とオッズの2通りの方法があります(競馬に詳しい方は「オッズ」と聞いてピンときますか?)

ある事象の確率 probability とは、何回か施行したうちでその事象が起こり得る回数の中で、通常は0から1の数字で表されます。

ある事象のオッズ odds とは、その事象が起こりそうもないと思われる回数に対する起こりそうと思われる回数の比と定義され、通常は比や分数で表されます。

ある事象が起こる確率を P とすると、その事象が起こらない確率は $1 - P$ になります。その事象のオッズはこの2つの値の比 $P/(1 - P)$ になります。

もしその事象がめったに起こらないものだとすると、Pは非常に小さくなり $(1 - P)$ が1に近づくため、オッズと確率は近似してきます。

この理屈によって、発生状況が稀な疾患の場合は、オッズ比を相対危険度の近似として扱うことができます。注意すべきはコホート研究での相対危険度は曝露群と非曝露群の比較であるのに対し、ケース・コントロール研究でのオッズ比は症例群と対照群の比較であり、見ている方向が異なる、ということです。

オッズ比に伴う95%信頼区間

- マッチングしていない場合のオッズ比を とすると、その時の95%信頼区間は、次の式の±で与えられる値を、上下それぞれの信頼限界の近似とすることができる。

$$\times \exp\{\pm 1.96\sqrt{(1/a) + (1/b) + (1/c) + (1/d)}\}$$

曝露	患者	対照	計
あり	a	b	a + b
なし	c	d	c + d
計	a+c	b+d	a+b+c+d

この式の由来は、残念ながらここでは説明し切れません。統計学の本を見て確認していただくようお願いいたします。

疫学 演習問題4

- 喫煙と心筋梗塞の相関に関するケース・コントロール研究で以下のような結果が得られた。

	非喫煙者	1/2箱(1日)喫煙者	1箱(1日)喫煙者	2箱(1日)喫煙者
症例数	31	9	39	18
対照数	2,706	710	1,825	605

それぞれのオッズ比(相対危険度)を求めよ。
 (1) 1日1/2箱 (2) 1日1箱 (3) 1日2箱
 また、全喫煙者と非喫煙者を比較して、
 (4) 心筋梗塞のオッズ比(相対危険度)

それぞれのオッズ比を定義にしたがって求めましょう(文献1より)。

疫学 解答4

- (1) 1.1
- (2) 1.9
- (3) 2.6
- (4) 1.83

計算の過程は省いていますが、容易に答が出せたでしょう。

95%信頼区間(confidence interval)

- 信頼区間はデータの統計的な安定性の指標。
- たとえば「相対危険度4.0で95%信頼区間が2.5 - 5.5」の時、どのようなイメージでその安定性をとらえることができるか。

95%信頼区間はデータの統計的な安定性の指標になるのです。

多人数の調査で得られた結果の方が、小人数の調査の結果よ

りも、統計的に安定しています。この安定性の程度を定量的に示したものが信頼区間であるとみなすことができます。

例えば、相対危険度が4で95%信頼区間が2.5-5.5であれば、「同じ人数で100回調査を行えば少なくとも95回は2.5-5.5の範囲の相対危険度を観察する」ことを意味します。対象者の数が多くなるほど95%信頼区間の範囲はせまくなります。つまりそれだけデータの統計的な安定性は高くなるのです。

そして、「比」を問題にしている場合は、相対危険度の特定の値(例えば4)を見るだけではなく、その95%信頼区間(例えば2.5-5.5)を見てその安定性を評価することが必要になります。比の場合は、信頼区間が1.0をまたがると有意差があるとは言えなくなってしまう(RR=1.0あるいはOR=1.0で両群に差がないことを意味していましたから)。

統計的有意差(statistical significance) と p 値

- データの統計的な安定性の指標。通常、「統計的検定に関する確率値(P値)」が5%未満($P < 0.05$)の時に、「統計的有意差がある」と判断される。

データに「統計的有意差」がなければ、対象者が少なくデータが不安定なために、統計的なばらつきの影響でその結果になった可能性を否定できないのでした。

「統計的有意差」があれば、このような可能性は低く、統計的なばらつき以外の要因が影響している、たとえば検定の際に立てた対立仮説「仮説要因と疾病が実際に関連している」が成り立つ、などと解釈するわけです。

いま、喫煙者と非喫煙者の胃がん発生率を比較したところ、喫煙者の発生率は非喫煙者の2倍、つまり相対危険度は2だったとしましょう。もしもこの2という相対危険度に「統計的有意差」がなければ、実際には(母集団では)喫煙と胃がんに関係がないにもかかわらず、研究の対象者数が少なくデータが不安定なために、見かけ上2という相対危険度を観察した可能性を否定できません。

それに対して、もしも相対危険度に「統計的有意差」があれば、「対象者が少なくデータが不安定なためにこの結果を観察した」という可能性は低くなり、実際にも(母集団でも)喫煙と胃がんに関係がある関係がありそうだと解釈するわけです。

これまで紹介したオッズ比、信頼区間、p値はデータの数の多い少ないと次のような関係があることは、直感的に理解していただけたと思います。

	オッズ比	信頼区間	p 値
データ 多	ほぼ変わらない	狭い	小さい
データ 少	広い	広い	大きい

コホート研究 cohort study

■ コホートとは？

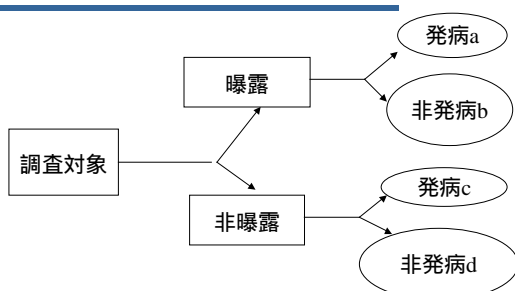
年齢・職業・人種等を同一にする集団、あるいは同一の外的要因に暴露された人間集団
コホート研究

ある要因に暴露された群と暴露されない群を設定し、それぞれの群における疾患発生率を比較する研究

疾病の要因と考えられている情報に基づいて調査集団を決定し、その後の疾病や死亡の起こり方が、要因の有無によってどのように異なるかを観察するのが、コホート研究です。

コホートとは易しく言うと「共通の性格をもつ集団」ですから、たとえば食中毒の調査の場合、「食べた」人と「食べない」人を選んで調査することが、コホート研究にあたります。具体的に言うと、例えばある講演会に出席した集団から、懇親会まで出席した人(食べた人)と出席しなかった人(食べなかった人)を同じ数選び、食中毒の発症の有無を調査するような場合のことです。

コホート研究



この図からコホート研究が「曝露の有無」によって集団を選定していることがわらうでしょう。

しかしたとえば、ある化学物質の影響をコホートで探ろうとするとき、

- ・その要因物質への曝露が明確であるか
- ・その曝露量の評価が明確であるか
- ・観察期間内に疾病の生態を知ることができるか
- ・曝露とそれに関連しているだろうと予測できる結果に関して、用量反応関係(曝露が多くなればそれだけ疾病の頻度が上がるという量的な比例関係)が得られるか

などいろいろな課題をかかえていることがわらうでしょう。

コホート研究の特徴

- コストが高い 大規模な集団を設定し長期にわたって追跡する必要があり、特に測定しようとする疾病が非常に稀な場合にはさらに規模を大きくする必要があります。
- 調査開始時に設定された曝露要因についてしかその影響を測定できない
- 分類 : 調査を開始した時点で疾患が発生しているか否かによって
前向きコホート研究 prospective cohort study
現在の曝露情報をもとに将来の疾病発生状況を見ていく。
後ろ向きコホート研究 retrospective cohort study / historical cohort study 過去の時点の曝露情報をもとに、現在の疾病発生状況を見る。

前向きコホート研究(prospective cohort study)では、多数の健康人の集団を対象として、最初に、疾病の原因となる可能性のある要因(喫煙・食生活・血液データなど)を調査します。次に、この集団を追跡調査して、疾病にかかる者を確認します。その上で、最初に調査した要因とその後の疾病の発生との関連を分析するのです。例えば、喫煙者と非喫煙者で、その後の肺がんの発生率を比較します。血清コレステロール値と心筋梗塞との関連や、喫煙と肺がんの関連など、今日では常識的な知見も、この研究方法で明らかにされた古典的な成果といえます。

多人数の集団(数万 - から数十万人)を長期間(5 - 20年)にわたって追跡調査をする場合は、前向きコホート研究には多大な手間と費用がかかります。しかし最近では、世界中で大規模コホート研究が行われるようになってきました。日本では、文部省コホート、環境庁コホート、厚生省コホート(1990年開始、13万人)の三つが進行中です。また米国ではハーバード大学がすすめた12万人の看護婦さんを対象にした大規模女性コホート研究(1970年代から20年以上追跡)が有名です。その結果、たとえば経口避妊薬(ピル)の現在の使用者の間では乳がんのリスクが少し増加していることなどがわかってきています。

後ろ向きコホート研究(retrospective cohort study)では、すでに曝露がおこってしまった後で、研究者が事後的に(後ろ向きに)その状況を調べ、さらにその集団を追跡調査することで、疾病の発生を確認します。事故によって高濃度の化学物質や放射線などにさらされた産業労働者の曝露状況を事後的に調べ、その集団のがん発生を追跡調査によって明らかにする場合などにこの研究方法が用いられます。広島への原爆投下後、ABCC(原爆傷害調査委員会)が開始した被爆者の疫学調査は、世界最大規模のコホート調査です。この場合、すべて被曝線量にもとづいて追跡することになるので、死亡率の研究を除き、低線量被曝集団との比較が必要になります。

コホート研究のデータ解析

相対危険度(リスク比: Risk Ratio: RR)

曝露群のリスク = $a / (a+b)$

非曝露群のリスク = $c / (c+d)$

相対危険度 = $\{a/(a+b)\} / \{c/(c+d)\}$

相対危険度RRの近似的な信頼区間

$$(RR) \times \exp\left\{ \pm 1.96 \sqrt{\frac{1 - a/(a+b)}{a} + \frac{1 - c/(c+d)}{c}} \right\}$$

これも天下りの式で申し訳ありません。統計学のテキストをみていただくほかありません。

相対危険度は容易に理解してもらえますと思います。

選択バイアス(selection bias)

- **バイアス**の一つ。研究の対象者の選択方法に問題があるために、仮説要因と疾病の真の関連が正しく評価されずゆがめられてしまう(両者の関連を過大評価したり過小評価したりする)現象。

データの統計的なバラツキについては、統計的検定を行なうことでどれくらいの信頼度をもって「差は偶然ではない」とみなすことができるのか、を示すことができました。ところが、データを扱う際に生まれる誤差は、こうした性質のものではありません。調査する側、調査される側の意図(一方的な思いこみなど)により生じる誤差があり、これが分析に重要な影響を与えることがあるのです。こうした誤差のことをバイアスと呼んでいます。

例えば、あるお店で食中毒が起こったとの報道がなされた後に、「自分もその店を利用し、調子が悪い」などの電話が増えることはよく知られていますが、こうしたその情報を鵜呑みにすると、「食べた」「有症」の人のみが増えて(2×2表のaの数値が増え)、見かけ上オッズ比が高くなるわけです。

また例えば、がん検診の有効性を評価する時に、検診を受診した集団と受診しない集団でがん死亡率を比較することを考えます。この場合、検診を受診した集団はもともと生活習慣が健康的でがんのリスクが低ければ、仮に検診に全く効果がない場合でも、見かけ上、検診を受診した集団の死亡率は受診しない集団の死亡率よりも低くなり、検診の真の効果を過大評価してしまいます。このように、もともと特性の異なる集団を研究対象者として選択して比較すると、仮説要因(ここではがん検診の実施)の影響を過大評価あるいは過小評価してしまう危険性があるのです。

交絡(confounding)

交絡要因(confounding factor)

- **バイアス**の一種。仮説要因と疾病との真の関連が、第三の要因の影響によって、ゆがめられてしまう現象。

例えば、野菜の摂取が多い集団と少ない集団で、大腸がんの発生率を比較することを考えて見ましょう。

この場合、野菜の摂取が多い集団は、肉類の摂取が少ない傾向があるでしょう。肉類の多量摂取は、それ自体が大腸がんのリスクを上昇させると考えられています。このため、野菜繊維の多量摂取群と少量摂取群の大腸がん発生率を、両群の肉類の摂取量の違いを無視して単純に比較すると、野菜の予防効果を過小評価してしまう。つまり、野菜という「仮説要因」と大腸がんという「疾病」の真の関連が、肉類という「交絡要因」の影響でゆがめられてしまうことになるのです。

いま、喫煙者と非喫煙者で肺がんの発生率を比較したところ、喫煙者の発生率は非喫煙者の5倍だった、つまり相対危険度は5だったとしてみましょ。この時、喫煙者の方がより高齢で、緑黄色野菜の摂取量も少なかったとします。肺がんの発生率は高齢者や緑黄色野菜の摂取の少ない者でも高くなります。この場合、これらの要因を考慮しないで、単純に喫煙者と非喫煙者の発生率を比べると、喫煙の影響を実際以上に過大評価してしまうことになります。これが交絡です。

疫学 演習問題5

- 次の図はスウェーデン(カロリンスカ研究所)による、送電線周辺での極低周波電磁波被曝と小児白血病との関連を探った研究(1992年)の概要である。この研究からどのような結論がどのような信頼性を伴って導くことができるか、検討しなさい。

さて、最後の総合問題です。がんばってください。

疫学 解答5

相対危険度(オッズ比)の計算

$$\frac{\text{患者の曝露オッズ (3.0mG以上) } 7 \text{人}}{\text{対照の曝露オッズ (0~0.9mG) } 32 \text{人}} = \frac{27 \text{人}}{475 \text{人}} = 3.8$$

(= 3.8, a=7, b=32, c=27, d=47 を公式に代入して解く)

磁界曝露量が3mG以上の時、白血病の患者の割合が0~0.9mGの場合に比べて3.8倍となる。オッズ比3.8は関連性を示すが、95%信頼区間の幅が広い点をどう見るかが問題になりそう。

そのやり方の一つは、たとえば曝露量の異なる2つの階層を選んで、たとえば0~0.9mGを「非曝露群」、3.0mG以上を「曝露群」とみなし、オッズ比を求めることです。

患者の曝露オッズは7/27、対照の曝露オッズは32/475です。したがってこの場合のオッズ比は、 $(7/27) \div (32/475) = 3.8$

となり、相対的なリスクの比(相対危険度)は、3.8倍とみなすことができます。

これは言い換えると、「高圧線からの電磁波を3.0mG以上の強さで被曝している場合は、その強さが1.0mGの場合と比べて、小児白血病の増加は3.8倍になる」ということです。

ここで、先に示した別の計算式を用いて「95%信頼区間」を求めると、1.4~9.7となります。この結果は、「もし同じ人数で調査を繰り返せば、100回のうち少なくとも95回は、1.4~9.7のオッズ比を得る」ということを意味します。確かに被曝する電磁波の強さと小児白血病の増加とは関連があることが示されていますが、その時の95%信頼区間の幅がかなり広い点をどう見るかが問題になりそうです。

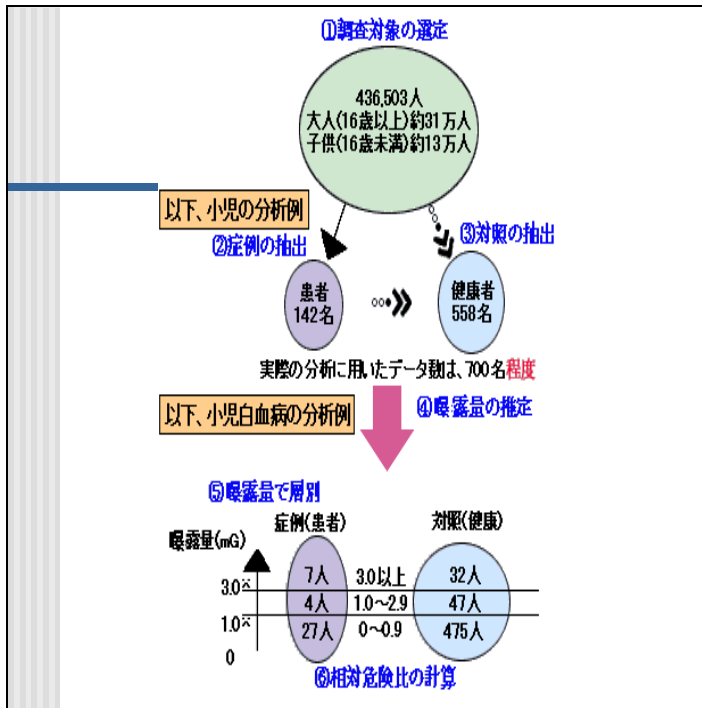
スウェーデン政府はこの結果を受けて、1993年から2~3mGを目安にして、子ども用施設などの移動や送電線の撤去などを開始して、国際的に注目されました。さて、先の公式を使って、信頼区間を求めることができました。

相関関係と因果関係

- 関連性(association)の2つの概念
- 相関関係(統計的な関連性)(correlation)
- ↓ × ↑
- 因果関係(原因と結果の対応関係)(causality)

疫学によって私たちは原因と結果の関係をどのようにとらえることができるでしょうか。

統計的な関連性を示す場合に「2つの事象に相関関係がある」



ここでは、実際に行なわれたスウェーデン(カロリンスカ研究所)研究による、送電線周辺での極低周波電磁波被曝と小児白血病との関連を探った研究(1992年)を見てみます。(商用周波数電磁波の磁場の強さの単位はmG:ミリガウス)

この研究では、スウェーデン登録所情報をもとに、症例(白血病、脳腫瘍、その他のガン患者)を抽出しています。対照の抽出は、患者1人に対し、4人の健康者を性、年齢、地域、送電線の種類などを揃えて無作為に抽出しています。

曝露量の推定を、各症例(患者)、対照(患者でない人)について年平均負荷時の送電量と、送電線と居住家屋との距離から計算により推定しています。家電製品からの曝露や屋外曝露は考慮していません。

抽出された症例、対照を推定された曝露量にもとづき分類(階層化)しています(0~0.9mG、1.0~2.9mG、3.0mG以上)。ただし、3.0mG以上の症例は7名と少ないです。これは研究の一番の欠点と言えるかもしれません。スウェーデンではそもそも住宅地の近くに送電線が少なく、そのため近隣の小児白血病の患者数もわずかであるからです。

ところでこの調査の場合、曝露の有無によって集団を選定したのではなく、症例の有無によって集団を選定していますから、ケース・コントロール研究に相当します。そのなかで、では「曝露群」と「非曝露群」をどう定義づけなければいでしょうか?そしてあなたはこのデータから定量的なリスクをどう導き出しますか?

という言い方をしますが、もちろんこれだけでは因果関係があることを意味しません。相関関係は因果関係の必要条件ではあっても十分条件ではありません。

ではどのようにして因果関係を確定できるのでしょうか？

疫学調査はデータにもとづいて、曝露群が非曝露群に比較して疾病の発生が増えているかどうかを定量的に示すことができます(相対危険度)。これは疾病の原因を探りあてていくために必要不可欠な分析指標であることは明らかだと思われます。ところがこれまで私たちは、「その現象が起こるメカニズムが解明されない限り、原因を示すことにはならない」「疫学研究のみの結果にもとづいて病気の原因を探ることはできない」といった考え方に影響されて、疫学が示す結果を軽視してきたとは言えないでしょうか。

疫学を無視あるいは軽視したために大変な悲劇となった事例があります。戦後最大の公害として誰もが知る水俣病です。

水俣病を仮に、魚介類を食べたためにある地域で起こった食中毒事件であるとみなすなら、当然行政は曝露地域とおぼしき地域で住民の調査を行い、患者集団と原因食品の摂取との関係を疫学手法を用いて明らかにしようとするでしょう。そしてその結果にもとづいて、その地域での魚介類の摂取にストップをかけ、被害を防止しようとするはずでしょう。

ところが現実には、この種の調査はなされず、“患者が本当に水俣病であるかどうかの診断”(四肢末端優位の感覚障害・運動失調・求心性視野狭窄、難聴などの兆候による患者「認定」)ばかりが問題にされてきたのです。申請者が申請しないと患者としてカウントされず、その患者も認定審査委員会が認定しないと“真の患者”とはみなされない、という異例の対応がとられ続けたわけです。この事件への対応として一番必要だったのは、魚介類の摂取もしくはチッソ水俣工場の廃水と、水俣病とみなすことのできるような症状との因果関係を明らかにすることだったのに、まさにその因果関係解明への科学的アプローチ(この場合は疫学分析)がまったく欠落していたのです。

水俣病に関して以上のことを実証的に明らかにしたのは岡山大学医学部衛生学教室の津田敏秀さんですが(「水俣病問題に関する意見書」『水俣病研究』第1巻、1999年)、津田さんは、“因果関係はメカニズムが解明されてはじめて示すことができる”という考え方が、むしろ因果関係を現実に即してまともに究明していくことを妨げてしまう、との観点から次のように批判しています。

「一見“判明”したメカニズムも、突き詰めるとその先にやはり未確定な部分を常に有しているのであり、いずれかのレベルで納得し合意を形成せざるを得ない。しかしこのようにあるレベルを定めたメカニズム議論は実際には希で、曖昧で具体性に欠ける“メカニズムが明らかでない”という表現が実際には多い。具体性に欠ける“メカニズム”が問い直されず、いかにも大多数の人々が納得する“メカニズム”が実在し観察可能であるかの如くに語られてしまうことの方に問題があったと考えられる。(中略)メカニズムはヒトにとって必要なレベルで論じるべきなのである。」(「我が国の社会医学における因果関係論の構築を目指して」『日本衛生学雑誌』55号462-473頁、2000年)

たとえば疫学のテキストには、因果関係が成り立つかどうかを判断する条件として次のようなものが掲げられています。

因果関係の判断条件

- 例：喫煙と肺ガンで考えると……
- (1) 普遍性(consistency): 時間、場所、対象者を異にするどの集団でも見出せる
- (2) 強度(strength): 曝露量が増えれば罹患率は上がる「量-反応関係」(dose-response relationship)
- (3) 特異性(specificity): 一対一の関係性がそれなりに強く示される
- (4) 時間的關係(temporal relationship): 原因が時間に必ず先行している
- (5) 論理的な一致(coherence): 動物実験など医学・生物学の知見と合致している

もちろんこれですべてが尽くされているわけではありませんし、もっと精緻に理屈を組み立てることもできると思いますが、たとえばこのような判断条件が生かされるためには、現実には次のような取り組みが必要になることが理解できるのではないのでしょうか。

健康リスクに関わるような問題では、因果関係を解明する作業は、現実には人々の間で生じていることがらを見落とさず、予断を交えずに観察し、集団に関わることがらで定量化できる部分ではきちんと定量化すること、そして結果の事象(病気)の頻度を増加させる要因が浮かびかかってきた場合は、それを原因とみなしながら、できるだけ速やかに対処すること。

疫学から見えてくるのは、実際的で常識的な確からしさに基づいて因果を推定し、適切な社会的対応を施していくことの大切さではないかと思うのです。

参考文献 おすすめできる疫学の入門書

- 1 津田敏秀『疫学講座資料集』化学物質問題市民研究会・頒布 2000年10月
- 2 加藤克己・高橋秀人(編)『疫学概論 理論と方法』朝倉書店 2000年
- 3 A.ウオーカー『疫学研究の考え方・進め方 観察から推測へ』新興医学出版社 1996年
- 4 柳川洋ほか『地域の保健活動のための疫学』日本公衆衛生協会 2000年
- 5 D.L.ストレイナー・G.R.ノーマン『論文が読める! 早わかり疫学』メディカル・サイエンス・インターナショナル 2000年
- 6 R.S.Greenberg 他『Medical Epidemiology』(3th ed.) McGraw Hill 2001年(第2版の邦訳は『医学が分かる疫学』新興医学出版社 1995年)
- 7 R.F.Morton 他『A Study Guide to Epidemiology and Biostatistics』(5th ed.) Aspen 2001年
- 8 B.B.Gerstman『Epidemiology Kept Simple』Wiley-Liss 2001年

日本金属学会 2002年春期大会での発表より