

## 解説

## 騒音と人々の健康に関わりを探る調査～疫学調査の手法の解説～\*

後 藤 恭 一\*\*

## 1. はじめに

空港周辺において住民を対象に騒音による健康影響の調べる唯一とも思われる研究方法が疫学研究である。疫学研究で騒音と健康事象との間に統計学的有意性が認められた場合には、騒音と健康関連性にエビデンスが認められたと称し、因果関係があることを疑う。2011年、WHO 欧州事務局が発行した「欧州のための夜間騒音ガイドライン」では、騒音と健康影響の関連性については、疫学研究に立脚してガイドライン値が定められている。そこで、疫学研究の仕組み、リスクの指標とその解釈、そして疫学研究の限界について解説する。

## 2. 疫学とは：個人ではなく、より多くの人に合う解決方法を提案する学問。

一言で言えば疫学とは、「人間集団における健康状態の頻度を測定して、予防する“実践的”学問」である。あえて“実践的”と記したのは疫学の原点を振り返ると分かる。疫学の草分けは1854年に英国で実施されたコレラの調査である。当時、病気は空気を伝わる悪臭「瘴気（しょうき）」であるとする説（miasma theory）が考えられていた。だが英国人医師のジョン・スノー（John Snow）は、同じ井戸の水を飲む人にコレラを発症する割合が高いことを突き止めた。そして、感染経路である井戸水の飲用を禁止することにより、コレラの予防を成功した。日本でも同様の例がある。東京慈恵会医科大学の創始者である高木兼寛（たかぎかねひろ）は、1882年（明治15年）に海軍医務局副長に就任し

た後、当時日本軍を悩ませていた脚気の対策に取り組んだ。高木は、脚気（かっけ）の原因が食物にあると考え、海軍の兵士たちに改善食（主として洋食）を与えた群と改善食を摂らなかった群を比較した。その結果、前者のグループで脚気が激減したというのは、良く知られているエピソードである。ちなみにこの時、陸軍の軍医だった森林太郎（森鷗外）は脚気病原菌説を唱えて、菌の同定に力を注いだ。が成果は得られず、陸軍の脚気罹患者が減ることもなかった。ロベルト・コッホによりコレラ菌が発見されたのは1884年、抗脚気のビタミンが発見されたのは1910年、二人の業績よりもずっと後になる。原因は分からないが疾病の予防法を見いだすというのは疫学の強みといえる。

その後、疫学は公害病や天災、生活習慣病などに研究対象を広げてきた。ただし、ウイルスでインフルエンザが引き起こされたり、病原性大腸菌で下痢や嘔吐等の胃腸症状を呈したりする等の発生要因が明確に特定できるものとは事情が異なり、騒音による身体影響は喫煙、飲酒、食事などの生活習慣や、その他の環境汚染など、多くの環境要因が関わって発生する。また、騒音曝露から影響が現れる時間の長さも関係し、長期に渡るほど騒音以外の様々な要因が複雑に関与するため、騒音曝露と疾病の綾を解きほぐすのは相当困難となる。さらに原因を特定することは難しい。疫学ではコントロール出来ない要因に接触したり、取り込むことを「ばく（曝）露（exposure）（暴ではない）」という。

\* Explanation of epidemiological research on the influence of noise on health

\*\*（前）航空環境研究センター 調査研究部 部長代理（現）いわき明星大学 看護学部 講師

なお、疫学ではある疾病を引き起こす因子を、「原因」とは言わず「要因」と呼ぶ。一般的には、「原因」は1つしかない場合や1つに特定できる場合に用い、「要因」は複数ある場合や1つに特定できない場合に用いる。また、ある事象を起こしたものを「原因」、ある事象に影響のあるものを「要因」と区別する見方もある。さらに、過去の特定の事象を起こしたものを「原因」、今後ある事象を起こす可能性のあるもの「要因」と区別することもある。「要因」と「原因」は、さまざまなシーンでよく使われる言葉であり、意味も似ているためによく混同されるが、このように整理すると疫学研究でなぜ「要因」という用語を用いるかが分かる。

### 3. 疫学研究の種類

#### 3.1 研究デザイン：実験か観察か

疫学では、研究デザインが重要となる。研究デザインとは、誰を対象に、どのような方法で、何を評価・測定方法として、どれくらいの期間をかけるかといった研究計画を組み立てることである。研究方法「型」を大別すると、要因を人為的に与えて影響を調べる「実験（介入）疫学」と、研究者は対象者を観察して曝露による影響を調べる「観察疫学」の2つがある。ジョン・スノーと高木兼寛の逸話に触れたが、ジョン・スノーが行ったのは観察疫学であり、高木兼寛が行ったのは介入疫学である。

介入疫学の場合には研究者が曝露量を人為（実験）的に変化させることによって反応の違いを観察することができる。一方、観察疫学の場合には研究者は曝露量を変化することが出来ないため、曝露と結果の両者を観察することになる。観察疫学の場合、この部分に不確実性（不確定性）が入り込む。騒音による健康影響に関する疫学研究には、この不確実性が壁となる。これについては後ほど触れる。

#### 3.2 観察疫学の方法「型」

騒音による健康影響に関する疫学研究は、かつては横断研究が多かったが、最近の欧米における報告はコホート研究 (cohort study) や、大規模調査として生態学的調査も行われている。最近ではメタ解析も多く行われている。そこで、先ず疫学方法を紹介する。

##### 1) 横断研究：疾病と要因の保有状況を同時に調べる方法

横断研究は、ある集団のある一時点における状況を観察して、要因と結果の関連を調べる方法である。疾病の評価としては、集団中における病気の者の割合（有病率）が用いられる。

この研究の利点は、調査が比較的容易で、短期間で多くの対象者を対象とできることである。そのため疫学の他、社会調査でもこの手法が良く用いられる。欠点としては、疾病と要因の時間的な前後関係が不明なため、因果関係の推定が困難なことである。例えば、運動不足と肥満に関連性が認められたとしても、「運動不足だから肥満になる」とも考えられるし、「肥満だから運動不足（嫌い）」になるとも考えられるため、横断研究では因果関係を明らかにすることは出来ない。そのため、現状を把握するとともに、その後の因果関係を追求するためのヒントを与える（研究仮説を立案）入口としての研究と位置づけられている。

##### 2) コホート研究：疾病の発生を観察する方法

そこで、健常者の集団を2回以上追跡して、曝露の有無、あるいは多少によって事象（結果）が生じるかを観察することにより、より確からしい因果関係を推定しようとするのがコホート研究である。研究に時間の流れが加わるのが特徴である。コホート研究は観察疫学研究の中心となる研究である。同義語として、追跡研究、縦断的研究がある。コホート (cohort) とは元々、古代ローマの軍隊の数百人程度の

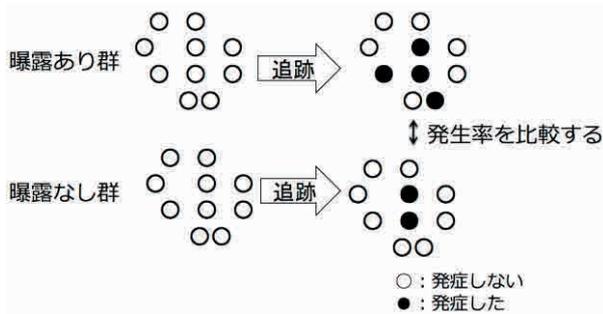


図-1 コホート研究のイメージ○・●の単位は個人

兵員単位を表すもので、疫学研究では転じて、ある共通の特徴を持つ集団の意味で使われる。例えば、同じ地域に住んでいる、同じ職業を持っている、同じ年に生まれたなど、共通の特性を持つ集団を追跡して、その集団からどのような疾病が発生し、また健康状態が変化したかなどを観察して、各種要因との関連を明らかにしようとするものである。仮説を検証する研究として位置づけられている。

コホート研究は時間の流れに沿って観察することが出来るので、環境の変化、症状の変化を正確に捉えることができるとのメリットがある。一方、時間や費用が多くかかることや、稀な疾病や、長い潜伏期間を持つ疾病には不向きである。また、追跡不能になるケースもあるとのデメリットがある。

最近、国際学会で発表される騒音による健康影響に関する疫学調査は、このコホート調査によるものが多い。

### 3) 症例対照研究：疾病の原因を過去にさかのぼって調べる方法

ある時点で疾病に罹患した集団と、その対照として罹患していない集団の過去に遡って曝露状況を調査して、両者を比較して要因と疾病の関連を評価する研究手法である。ケース・コントロール研究、患者対照研究、結果対照研究とも呼ばれる。例えば、特定の疾病の発生に騒音が関係すると考えられる場合、症例群として現在疾病に罹っている患者と、対照群として疾病ではない人を集めて、過去

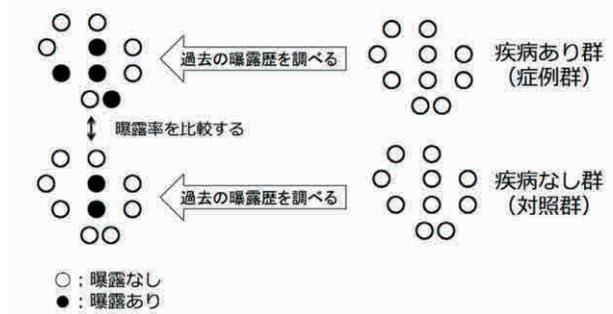


図-2 症例対照研究のイメージ○・●の単位は個人

の曝露量を調査し、騒音と疾病の関係について調べるものである。

この調査は、比較的簡単に調査が行え、追跡調査が不要という長所があるが、思い出しバイアス\*の影響を受けやすい。また症例と特性の揃った対照を選択することが困難である短所もある。

### 4) 生態学的研究：疾病と関連要因を地域または集団単位で検討する方法

2013年10月に医学関連で権威ある研究雑誌「The BMJ」に航空機の騒音によって心血管疾患のリスクが増加することを示す生態学的研究の結果が掲載された。英国のヒースロー空港付近の住民360万人のデータを用いて、ヒースロー空港近辺の21の地区における脳卒中と心臓疾患による入院件数および死亡率を航空機騒音のデータと照らし合わせた。研究対象の全域を12,110のエリアに分割し、

\* 症例と対照とは過去の事象に対する思い出しやすさが異なり、症例の方が過去の記憶が明確である。従って、本来は両者で差がない場合でも、症例群の方が曝露を受けているとの誤った結果が得られてしまう。

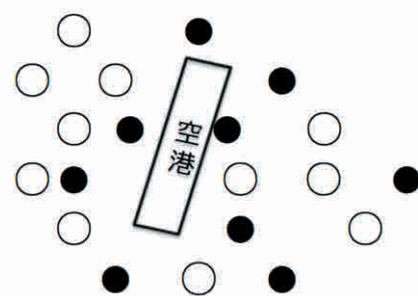


図-3 生態学的研究のイメージ○・●の単位は集団

各エリアの 2001～2005 年にかけての騒音レベルを調べた。結果、航空機騒音レベルの増加によって心血管疾患の入院件数も死亡率も増加する傾向が見られた。ただし、研究者は、今回見られた心血管疾患による入院・死亡のリスクの増加の原因が、航空機騒音以外の要因による可能性もあると報告している。

この研究の最大の欠点は、曝露や疾病発生を個人単位ではなく、集団としてしか把握していないため、「曝露と疾病発生に関連あり」と言うための根拠としては弱い。また、生態学的錯誤も生じる。例えば、集団レベルでの変数間に観測される関連が、必ずしも個人レベルで存在する関連を表すものではないことによるバイアスがある。

#### 5) メタ解析：複数の研究を収集・統合して、より高い見地から分析する方法

メタ解析（メタアナリシス meta-analysis、pooled analysis と呼ばれる）は、単独に実施された複数の疫学研究を統合して評価する手法である。メタ解析と聞くと統計手法の一つと思われるが、メタ解析は研究をデザインし、複数の調査結果を統合して、統計学的検定を行うという一連の手順である。

具体的にメタ解析の手順を示すと以下通りになる。統計解析が研究の初期段階で検討される。

- 1) 研究テーマの決定と仮説の設定
- 2) デザイン、エンドポイント、統計解析方法の決定
- 3) 研究資料の収集
- 4) 適格基準によるデータ抽出
- 5) 抽出データの統合
- 6) 統合データの検討（検定）

メタ解析において、3.の研究資料とは、研究の信頼性が高い“全ての”研究資料である。そうすることによりバイアスが排除される。一方で、別のバイアスが入り込む。出版バイアス（お蔵入りバイアス）である。「影響が有り」の論文の方がジャーナルに掲載されや

すく、「影響が無い」論文は掲載されにくい。また、研究中の段階で中止になることも多い。研究資料の収集がメタ解析の成果の鍵を握る。

#### 4. リスクの定量化

騒音の疫学調査では騒音曝露によるリスクは、寄与リスク（attributable risk, AR）、リスク比（risk ratio, RR）あるいはオッズ比（odds ratio, OR）としてリスクを定量化する。

先の図-1を集計すると分割表（contingency table）と呼ばれる表が得られる（表-1）。なお、公衆衛生学の分野では行方向には要因、列方向には結果を置く。また、表の縦線は引かないルールがある。

表-1 図-1の集計結果 1

		発症		横計
		有	無	
曝露	有	4	6	10
	無	2	8	10

実際にリスクを算出してみる。「発症した」数を横計の数で割って得られた「割合」がリスクである。なお、医学では病気にかかることを罹患（りかん）と呼ぶ。

#### 1) 寄与リスク

曝露「有」群と「無」群のそれぞれのリスクは群全体のうち何割が発症したか割合で示される。

表-1（図-1）の集計結果 2

		発症		横計	割合
		有	無		(リスク)
曝露	有	4	6	10	40%(4/10)
	無	2	8	10	20%(2/10)

表-1では曝露「有」群が40%、「無」群が20%となる。曝露が無くても20%発症することから、そこで曝露「有」群から「無」群の

リスクの差分を取る。これを寄与リスク（寄与危険度）と呼ぶ。

寄与リスクは20% (40% -20%) となる。寄与リスクとは、ある曝露により集団における発症がどのくらい増えるかを示す指標なので「曝露が有った群は無かった群によりもリスクが20%増し」とも表現することができる。

### 2) 相対リスク

寄与リスクがリスクの絶対量を示すのに対して、ある曝露により、集団における発症が何倍高いか、と相対的にリスクを示すものが相対リスクである。相対リスクは、曝露「有」群と「無」群の比を取ることに由来する。表-1の場合の相対リスクは2 (40% ÷ 20%) である。「曝露が有った群は無かった群に比べてリスクは2倍増し」と表現することができる。疫学で単にリスクといった場合には、この相対リスクを指すことが多い。

### 3) オッズ比

リスク20%増し、リスクは2倍といった表現は日常の報道で目（耳）にするが、オッズ比は馴染みのない用語である。オッズ (odds) とは「見込み」を表し、「起こる確率と起こらない確率の比」である。これを式にすると下式になる。

$$odds = \frac{p}{1-p}$$

「曝露有り」群のオッズは0.67 (4/6) となる。「曝露無し」群のオッズは0.25 (2/8) となる。

表-1 (図-1) の集計結果3

		発症		比
		有	無	(オッズ)
曝露	有	4	6	0.67 (4/6)
	無	2	8	0.25 (2/8)

曝露無しを分母に、曝露有りを分子としたものがオッズ比であり、2.68 (0.67/0.25) となる。

では、このオッズ比から、曝露群は非曝露群に比べて罹患するリスクは2.68倍であると言っているのだから、そういった解釈は論文も含めて時折見かけるが、オッズ比が相対リスクに近似するのは、稀な罹患、つまりイベント発症確率が十分小さい疾患の場合のみである。

具体例を示してこれを説明する。飲酒、喫煙、性別の発症を調べた結果を分割表として示したものが表-2である（仮想データ）。

表-2 飲酒、喫煙、性別の集計

		発症		横計	割合 (リスク)	比 (オッズ)
		有	無			
飲酒	有	4	6	10	40% (4/10)	0.67 (4/6)
	無	2	8	10	20% (2/10)	0.25 (2/8)

相対リスクとオッズ比を表-3に整理する。

表-3 相対リスクとオッズ比の整理

	発症率	相対リスク	オッズ比
飲酒	40%	2.00	2.68
喫煙	20%	2.00	2.27
性別	60%	2.00	3.49

相対リスクは全て2.00である。一方、オッズ比は大きく異なっている。オッズ比と相対リスクの関係は常に、オッズ比 > 相対リスクとなっているが、発症率が高いほどオッズ比と相対リスクの乖離が大きくなる。従って稀な罹患の場合、オッズ比は相対リスクに近似する。稀とは、人口100人あたり1人以下と定義されることが多い1。なお、オッズ比は0~∞の値を取り、オッズ比が1とは、事象の起こりやすさが両群で同じということである。オッズが1未満の場合には、曝露有り群の方が事象が起こりにくいことを示し、逆に、1より大きい場合には、事象がより起こりやすいことを意味する。本来オッズ比は、先に説明した、症例対照研究（ケースコントロール研究）のリスクを示す指標である。

では、疾病頻度が比較的高い場合や、コホート研究で、何故オッズ比がリスク比の近似値として用いられるのであろうか。薬 (A～C) と効果を集計したものを表-3に示す (仮想データ)。表-2と同様にオッズ比>相対リスクの関係になっている。

表-4 薬 (A～C) と効果の集計

薬	効果		横計	相対リスク	オッズ比
	有	無			
A薬	有	46	64	110	4.83
	無	9	95	109	

複数の曝露因子に対する要因の相対リスクとオッズ比の順番が一致していることがわかる。オッズ比の数値自体がリスクを  $x$  倍にするという意味を持たないが、複数の要因のリスクの順番が示されることからオッズ比が用いられる。

#### 4) ハザード比

オッズ比と似たようなものにハザード比 (hazard ratio: HR) がある。これは発症の有無だけではなく、「いつ発症したのか」という時間情報も加味して計算した発症率 (ハザード) がリスクの指標として用いられる。航空機騒音の疫学研究では用いられることが殆どないため、説明は良書に譲る。

### 5. 統計学的推定と検定

疫学研究の結果は偶然の影響にも左右されるので統計学的推定や検定を行い、精度評価を行う必要がある。

#### 5.1 統計学的推定

統計学的推定 (estimation) とは、サンプルから得られた平均値から母集団の平均値を推定することである。

##### 点推定

標本から得られた平均値を母集団の平均値

(母平均値) として推定することを点推定という。例えば、ある地点で7日間、航空機騒音を測定したところ、期間中の  $L_{den}$  平均が 61.0dB だったとする。この 61.0dB を母集団の推定値 (この例の場合、年間平均値が該当する) とするのが「点推定」の考え方である。

##### 区間推定

母平均はピッタリ 61.0dB とは限らないため、61.0dB とピンポイントで推定しまうと、推定が外れる危険性が生じる。そこで幅を持たせ「この調査結果から母集団の推定値は  $xx\%$  の確率で 60.8~62.1dB の範囲中にある」と推定するのが「区間推定」である。つまり、「標本から推定すると、母平均はこの値からこの値までの間にはいるのではないか」という推定を行うものである。区間は「信頼区間 (CI, Confidence Interval)」と呼ばれ確率的に決められる。疫学では通常「確率 95%」が用いられ、これを「95%信頼区間 (95% CI)」と呼ぶ。信頼区間が広いほど、推定を外す確率が小さくなる。

##### 標準誤差 (Standard error)

区間推定をするために必要な統計量が標準誤差である。標準誤差は、標本データの散らばり (標準偏差) が大きいほど母平均と標本平均の差 (標準誤差) は大きくなる。また、標本数が少ないほど標準誤差も多くなる。この理論を散らばり (標準偏差) を  $s$ 、標本数を  $n$  として式にすると、下式が得られる。標準誤差、標準誤差、標準偏差と似たような用語が続くが概念が異なるので注意したい。

SE 標準誤差は下式で算出される。

$$SE = s / \sqrt{n}$$

$s$ : 標準偏差      $n$ : 標本数

分子が大きくなるほど SE は大きくなり、分母が小さくなるほど SE が大きくなるのがわかる。

大標本の場合の 95%信頼区間は下式から求

めることが出来る。

$$95\% \text{ CI} = \text{標本平均} \pm 1.96 \times \text{SE}$$

小標本の場合の95%信頼区間の式は下記の通りとなる。

$$95\% \text{ CI} = \text{標本平均} \pm t \text{ 値} \times \text{SE}$$

例えば、先に例として掲げた航空機騒音を実例として計算してみる。7日間の航空機騒音の測定において、期間中のLden平均が61.0dB、標準偏差が0.8dBだったとする。

t分布表をもとに臨界値、t値を求める。t分布とは0を中心とした左右対称の分布である。図中のtが臨界値である。山の高さは自由度によって異なり、自由度が大きいほど山の高さは高くなる。

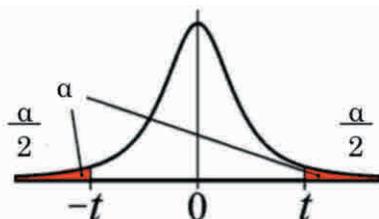


図-4 t分布

t分布表には、自由度がdf (degree of freedom) である場合の赤色部分の面積が $\alpha$  (両側) となるtの値が示されている。95%信頼区間の場合の $\alpha$ は、 $1-0.95=0.05$  (5%) である。下表から、df 6 と 0.05 が交差する値が臨界値であり、2.4469 と読み取ることができる。t値はt分布表から求めることができる。Excelでも関数TINV(両側確率P, 自由度df)を用いてt値を求めることができる。「=TINV(0.05,6)」と入力すると、2.446911851 とt値が返される。

表-4 t分布表

df	$\alpha$ (両側)			
	0.10	0.05	0.01	0.001
1	6.3138	12.706	63.657	636.62
2	2.9200	4.3027	9.9248	31.598
3	2.3534	3.1825	5.8409	12.941
4	2.1318	2.7764	4.6041	8.610
5	2.0150	2.5706	4.0321	6.859
6	1.9432	2.4469	3.7074	5.959

これらの値から信頼区間は、

$$95\% \text{ CI} = \text{標本平均} \pm t \text{ 値} \times \text{SE}$$

$$= 61.0 \pm 2.4469 \times 0.8 / \sqrt{10}$$

従って、下限値は60.381、上限値は61.619となり「母集団の推定値は95%の確率で60.4～61.6の範囲の中に収まる」というように示すことができる。

なお、自由度(df)と聞き慣れない用語が出てきたので簡単に説明する。自由度は独立に動くことができる数を意味する。例えば、平均値が2の時、1つ目のデータが「1」、2つ目のデータが「3」であれば、3つ個目のデータは必然的に「2」となる。つまり、3個の観測値のうち2個が決まれば最後の1個は値が拘束されてしまうので、自由に値を取れるデータの個数が1つ分減らした数、つまり、 $df=n-1$ となる。なお、統計量によって自由度の計算方法が異なることに留意されたい。

## 5.2 統計学的有意差検定

統計学的検定 (Statistical Significance Test) とは Test との文字が示すとおり○×式で差や関連があるか否かを統計学的に検証することである。

日本工業規格 (JIS) では、(有意差) 検定 (statistical test) とは「帰無仮説を棄却し対立仮説を支持するか、又は帰無仮説を棄却しないかを観測値に基づいて決めるための統計的手続き。その手続きは、帰無仮説が成立しているにもかかわらず棄却する確率が $\alpha$ 以下になるように決められる。この $\alpha$ を有意水準という。」と定義する。少々この説明では難しいので、ある人が「犯人である」ということを立証する手順を考えてみる。

1. 『Aが犯人であること』を立証したい。(立証したい仮説を「対立仮説」と呼ぶ)
2. 犯人であるかは神(とその本人)のみぞ知る真実である。そこで『Aは犯人とは無関係』から検証をスタートする(この仮説を「帰無仮説(きむかせつ)」と呼ぶ)

3. 「証拠」を収集する。偶然に証拠が得られることもあるが、証拠が多いほど犯人である確率が高い。なお、犯人である確率が高くなるにつれて「冤罪確率 ( $\alpha$ )」は小さくなる。冤罪確率は、誤った判断を犯す確率ともいえるので危険率とも呼ばれる。

4. 冤罪確率をもとに判断する。

1) 設定レベル一定以下の場合

証拠が一定以上収集されるのは、『A は犯人とは無関係』とする仮説と矛盾する。そこで帰無仮説を棄却して『A が犯人である』との対立仮説を支持する

2) 設定レベルを越える場合

帰無仮説である『A は犯人とは無関係』を保留する。しかしこれは「A が犯人ではない」ことを立証したわけではない。あくまで「A を犯人と認証するだけの証拠が見つからなかった」ので「犯人であるとは言えない」との消極的判断となる。

このように、ある主張Aを証明するのに、Aでないという前提からは矛盾が生ずると示すことで行う証明法を背理法（はいりほう）と言う。

これを元の有意差検定の手順に戻してみる。

1. 『差・関連が0とは異なる』を立証したい
2. しかし、真は不明なので『差・関連は0である』から検証を開始する
3. 証拠「検定統計量」から p 値 Probability を算出する。検定統計量が大きいほど危険確率 ( $\alpha$ ) は小さくなる。1-p =  $\alpha$  の関係がある。
4.  $\alpha$  = 5% 水準（ハードル）基準で判断する
  - 1)  $\alpha$  以下の場合『差・関連は0である』を棄却し「差・関連がある」と解釈する
  - 2)  $\alpha$  を越える場合『差・関連は0である』は棄却されず保留される。ただし「差・関連は0」ではなく誤差の範囲に留まり「差があるとはいえない」と解釈する。

このように置き換えることができる。すなわち、犯人が「差・関連」、証拠が「検定統計量」、冤罪確率が「 $\alpha$ （危険率）」に相当する。

先に示した例、7日間の航空機騒音の測定において、期間中の  $L_{den}$  平均が 61.0dB、標準偏差が 0.8dB に基づき計算してみる。ここでのテーマは「この地点の観察値は 61.0dB だが、母平均 62.0dB と差があると言って良いのか」ということになる。「差がある」と言えれば、観察値は「有意に低い」と判断することができる。集めたデータの平均と既知の母平均を比較したい場合には、下式の学生 t 検定 (Student) の t 統計量を求める。検定手法は「1 標本 t 検定」と呼ばれる。

$$t = \frac{\text{標本平均} - \text{母平均}}{\frac{\text{標準偏差}}{\sqrt{\text{標本数}}}}$$

式 Student-t 検定統計量

この式を用い t 値を求めると -3.3072 と計算される。

$$t = \frac{61.0 - 62.0}{0.8 / \sqrt{7}}$$

観測数が7のときの自由度は6である。自由度6、 $\alpha$  5% (両側) の t 値は 2.4469 である。証拠能力を表す統計量が臨界値である有意水準値を上回っていると、帰無仮説が棄却され「母集団と差がある」と判断することができる。従って、観察値は「有意に低い」と判断することができる。

ところで、先に推定において算出された 95% 信頼区間は 60.4 ~ 61.6 であった。62.0 は 95% 信頼区間の範囲の外にある。このことから、信頼区間の範囲の外に仮説値があれば統計学的に有意となり、逆に信頼区間の範囲の中にあれば統計学的に有意とは言えないと言える。両者は推論の方向が逆の関係にある。

## 6. 多変量解析

### 6.1 なぜ多変量解析を使うのか

医学・生物学における共通の狙いの1つは、例えば要因の曝露の有無と罹患の有無の関係といったように、疾病と疾病要因間に存

在する関係を明らかにすることにある。しかし、疾病の発症には様々な要因が関与するため、得られた結果がそのまま因果関係を示すものではない。

表-2および表-3（再掲）の集計結果が得られたとする。

表-2 飲酒、喫煙、性別の集計（再掲）

	発症 (名)	横		割合 (リスク)	比 (オッズ)	
		有	無			計
飲酒	有	4	6	10	40%(4/10)	0.67(4/6)
	無	2	8	10	20%(2/10)	0.25(2/8)
喫煙	有	2	8	10	20%(2/10)	0.25(2/8)
	無	1	9	10	10%(1/10)	0.11(1/9)
性別	男	6	4	10	60%(6/10)	1.50(6/4)
	女	3	7	10	30%(3/7)	0.43(3/7)

性別のオッズ比が最も高いことから、直ちに疾病発症には性差が関与することを表すと結論することはできない。男性は女性よりも飲酒や喫煙率が高いため、性差の背後にある生活習慣が関与した見かけ上の因果関係であることが考えられる。さらにお酒を飲む人はたばこを吸う、といったように飲酒と喫煙にも関係がある。そこでそれらの複雑な関係を取り除いて、飲酒、喫煙、性別と発症のダイレクトな関係を調べる必要がある。

多変量解析の「多」とは、「複数の要因（変量）を同時に使う」という意味である。「複数の要因を同じに使って」解析することにより、他の要因による影響を調整した（取り除いた）影響度合いを調べることができる。

代表的な多変量解析として、重回帰分析が良く知られている。例えば3つの要因があり、それらの結果 y への影響を見ようとした場合には、以下の3つの単回帰（式）分析を繰り返すのではなく、

$$y = \beta_1 * x_1 + \beta_0$$

$$y = \beta_2 * x_2 + \beta_0$$

$$y = \beta_3 * x_3 + \beta_0$$

下式を用いて分析すれば、各要因の y への単独した影響力が  $\beta$  として示される\*\*。

$$y = \beta_1 * x_1 + \beta_2 * x_2 + \beta_p * x_3 + \beta_0$$

騒音の疫学調査では**多重ロジスティック分析**が多く用いられる。両者の違いは式中の y に当たる結果のデータの尺度の違いである。血圧値や、血糖値のような連続尺度データの場合には重回帰分析を用いる。他方、y がカテゴリのような順序尺度データや、「病気になる / ならない」あるいは「死亡した / 生存している」のような名義尺度データの場合にはロジスティック回帰分析を用いる。ロジスティック (logistic) とは「物流」という意味ではなく「記号論理学の」という意味である。

$$\log \frac{p}{1-p} = \beta_1 x_1 + \beta_2 x_2 + \beta_p x_p + \beta_0$$

なお、多重ロジスティック分析の場合は、影響力の大きさを示す  $\beta$  はオッズ比として示される。さらにオッズ比は例えば「1.2 (95% CI: 0.5 ~ 1.5)」95%信頼区間も同時に示されることが多い。信頼区間の中に1が含まれていない場合には、統計的有意差が認められたことを意味する。

統計解析で取り扱うデータの尺度を少し詳しく説明する。尺度は、名義尺度、順序尺度、連続尺度の3つに大きく分類することができる。

#### ①名義尺度

単に区別するために用いられている尺度であり、例えば、血液型でA型・B型・O型・AB型を、それぞれ0・1・2・3と数値に対応させたもの。数値の大小関係はなく、平均値を求めることに意味は無い。

\*\* 回帰係数 (b) は重回帰分析することにより算出される。しかし、b は単位依存性があり、単位を変換するとそれに応じて値も異なる。また、単位が異なるもの同士の b を比較することは意味をなさない。そこで各データを平均値0、標準偏差を1と標準化して重回帰分析を行うことによって算出される標準化偏回帰係数 ( $\beta$ ) を影響力の指標として判断する。

## ②順序尺度

大小関係にのみ意味がある尺度。例えば、治療効果の判定において、悪化・不変・改善・著効を、それぞれ -1・0・1・2 と数値に対応させたもの。数値の差がまちまちなので、平均値は定義できないが中央値は定義できる。

## ③連続尺度

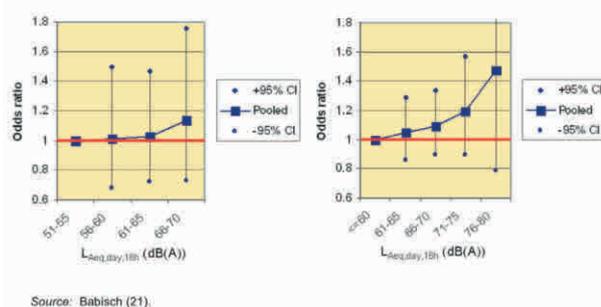
数値の差のみに意味がある尺度や、数値の差とともに数値の比にも意味がある尺度。前者の例としては、摂氏温度のように温度が 10℃ から 15℃ になったときに、50% の温度上昇があったとは言えないが、差分の 5℃ という数値には意味がある。後者の例としては、体重が 50kg から 60kg になったときと、100kg から 110kg になったときとは、同じ 10kg の増加であっても、前者は 20% 増、後者は 10% 増といえる。しかし、統計処理の際には両者は区別しない。比が定義できるということは絶対零点を持つことと同じことを表す。

## 6.2 騒音による疫学研究の実例

これまで説明した内容をもとに道路交通騒音と心臓疾患に関する疫学研究の論文の結果をしてみる。下図は、WHO 欧州事務局が 2011 年に公表した「Burden of disease from environmental noise」に掲載された、自動車騒音による心筋梗塞 (myocardial infarction) の関係である 3。

この結果から、WHO 欧州事務局は「ヨーロッパのための夜間騒音ガイドライン値 (NIGHT NOISE GUIDELINES FOR EUROPE)」を導き出している。

Fig. 2.1 & 2.2. Pooled effect estimates (meta-analysis) of the association between road traffic noise and the prevalence (Fig. 2.1, left) and incidence (Fig. 2.2, right) of myocardial infarction (odds ratio +/- 95% confidence interval)



グラフのタイトルを見ると Pooled effects estimates(meta-analysis) となっており、複数のコホート研究をまとめて評価するメタ解析による解析結果であることが分かる。左図は心筋梗塞の有病率 (疾病を有している者の割合: Prevalence)、右は罹患率 (疾病に罹った者の割合: incidence) と騒音レベルとの関係が示されている。統計学的手法として多重ロジスティック分析を用いて、年齢や性等の他要因を調整して、最も騒音レベルが低い LAeq,day,16h 51-55(dB(A)) を基準とし、オッズ比を 1 とした時の 56-60、61-65、66-70、・・・の相対オッズ比の値が図中 (■) に示している。■から上下に伸びている線が 95% 信頼区間である。

図を見ると、Prevalence(有病率)、incidence(罹患率)ともに騒音が高くなるにつれてオッズ比の上昇が見られる。では、この関係は統計学的に有意と言えるのであろうか。この場合の帰無仮説は「オッズ比は 1 である」。先に説明した通り「95%信頼区間に 1 が含まれて」いなければ「オッズ比は統計学的有意差」があるといえる。逆に 95% 信頼区間が 1 を含んでいれば「統計学的有意差は認められない」ことになる。つまりこのグラフからは、騒音増加に伴って相対リスクが上昇するとは言えないと読み取れる。しかし、WHO レポートでは「図 2.1 と図 2.2 からは 60 dB (A) を超えるノイズレベルの場合に心筋梗塞リスクが増加する (For noise levels greater than 60 dB(A), the myocardial infarction risk increases.)」と記述されている。95% 信頼区間、統計学的有意差は考慮されることなく、60dB(A) が越えたところからオッズ比が上昇することを視覚的 (ビジュアル) に捉えて解釈している。なお、原典を確認すると、有意差が得られていないことが明記されている。このあたりに WHO 欧州事務局の政治的思惑を読み止めることが出来る。WHO 欧州事務局は、55dB 以上で心疾患のリスクのあることから、子どもや高齢者、妊婦、交代制勤務

者などの高感受性群を含むほとんどの住民を健康影響から保護するためのガイドラインとして40dB（夜間帯23.00-07.00における屋外の等価騒音レベル：Lnight,outside）が定められた。また、様々な事由で短期間にガイドラインを達成できない場合の暫定目標値として55dBが提案している。

さらにこのメタ解析は、同一の研究者によるデータのみを統合して解析を行った結果である。メタ解析の手順に照らし合わせると、手順3と手順4に不備があり、手順6が適格に解釈されていないということになる。

## 7. 疫学の限界

冒頭で「原因は分からないものの予防法を見いだすというのは疫学の強みといえる」と述べたが、逆に「原因と因果が明確に分からないままの情報が提供される」のが疫学の弱みとも言える。騒音と疾病の関係が得られたとしても、それが真の関係ではない可能性を多々含んでいるとも言える。

さらに解析上で性、年齢、人種、社会経済的要因を調整したといっても全て拭い去れるわけでもない。また、地区ごとの騒音レベルとの関係の様に見えるが、騒音量で区分されたエリアの違いに騒音以外の地域特性が乗っかってくる。地域特性は、地域に基づく特性もあれば、そこに居住する者の特性、さらにそれらが相互に複雑に関連し合い、多変量解析で解き明かすこと困難となる。実験とは異なり、生活している人々を観察して、要因と結果の関連性を推定するのが疫学であるために、一定の限界はある。偏りや交絡因子による誤差を排除することが困難であるという問題をそもそも含んでいる。

騒音による疫学研究の場合に一番の課題が個人個人の曝露レベルの不確実性（不確定性）である。騒音と同様に環境影響を取扱う環境疫学の場合、食物や飲料水による影響では個人毎の曝露（摂取）量が調べられ健康事象と

の関係が検討される。水銀、鉛、ヒ素、カドミウム等の重金属の影響を調べる場合には、個人個人の毛髪に含まれる重量から曝露量を推定する。このように個人個人の曝露量と健康事象との関係を調べるのに対して、騒音による影響の場合には、曝露量はブロック毎の代表的な値を求めて、そこに居住する者は全て同じ曝露量とする、いわば個人個人を対象とした生態学的デザインによる調査となっている。では、自宅周囲の騒音レベルを正確に予測出来れば良いのであろうか。しかし、家屋の構造、主に生活する部屋の位置関係によって個人曝露レベルは相当異なるし、同居する居住者によっても、昼間は居住地を離れて学校、職場に通う者と、日中を居住地で過ごす主婦では曝露量は相当異なる。さらに、疫学における曝露量とは、どれくらいの音のレベルを聞いていたのか、どれくらいの時間聞いていたのか、の両方を考慮する。つまり、「曝露レベル×曝露時間（期間）」を考慮するが、これを考慮した研究は見当たらない。これらが騒音による疫学研究の最大の課題であり、限界となっている。

また、職場を対象とした研究では騒音により高血圧を誘発する報告例があるが、地域を対象とした環境疫学では騒音と健康被害（循環器障害）の関連は薄いとされている。例えば航空機や道路からの騒音によってアノイアンス（心理的な煩わしさ）への影響は良く知られるが、明確な健康被害に結びつくことはない。職場と地域では騒音レベルの違いが大きく異なることが考えられる。職場では比較的レベルが高い騒音が研究対象となるが、地域では日常レベルの騒音が対象となる。そのため、地域では、アノイアンス～ストレスを介した間接的な影響を観察することになる。そのため、個人個人の騒音の受け止め方が相当異なり、こうした個人差により集団としての健康影響が捉えにくいとも考えられる。

一方、こうした課題はあるにせよ、そもそ

も我が国における騒音による影響に関する疫学的知見が乏しい。生活習慣病には体質、遺伝的要素も関与するために、人種が異なる疫学調査の結果を我が国に当てはめることが出来ない。欧州や米国では大規模なコホート調査が計画、あるいは実施されているが、我が国も実施に向けた検討も必要である。

## 8. まとめ

騒音によって高血圧が生じる、ある食品を食べると長生きできる等々、多くの医学に関する報道は、疫学研究にもとづくものである。

疫学は「学問」というよりもむしろ「方法論」といった方が適切だと思われる。何の方法かといえば、「因果関係を推論する」ための方法であり、「疾病を防止あるいは予防する」ための方法である。そのどちらにも統計学、正式にいえば生物統計学が用いられる。疫学の三本柱は「疫学理論、生物統計学、因果推論」である。

疫学理論とは、簡単に言えば時間の経過の中で病気が「発生」するということを整理することである。少し乱暴な言い方をすれば「品質管理」と同様の考えである。品質管理は故障が発生することを知り、防止するためのものであるから、その考え方は疫学と似通っている。

生物統計学とは、医療・健康科学分野を対象とした応用統計学であり、疫学研究における研究デザイン（設計）と統計解析の両方を含む方法論のことである。

因果（関係の）推論は、どちらが原因でどちらが結果なのかを理論的に考えることである。「理論的に」といっても単に日常的な常識を突き詰めて考えているだけである。

従って、疫学は医学・医療の日常的常識を定式化した方法論と言える。「疫学の使命は医学研究の結果が示す数値の意味と、その背景となる方法論を分かりやすく説明することであり、それ以上の判断は社会が行うべきで

ある」という意見もあることから、疫学の結果は、聞き手（読み手）側も日常的常識の範囲で吟味して（批評的）評価する必要があるのではないだろうか。もう一つ加えれば、「あれがなければこれがなし B would be false if it were not for A.」を考えるのも重要である。

日常生活における騒音による影響は、産業職場における騒音レベルよりも低いいため、心理～ストレスを介した影響となる。そのため個人差の関与が大きいために騒音影響を見出すのが困難とも言える。今後の疫学研究の知見の積み重ねが望まれる。

以上、騒音による健康影響に関する疫学研究について、疫学にまつわる話題や、リスクや、リスクの指標等について解説するとともに、結果の実例を示して解析した。ご参考になるものがあれば、幸いである。

- 1 田中平三. 疫学入門演習 - 原理と方法 - 東京：南山堂、1989；45-46.
- 2 日本規格協会, JIS Z 8101-1:1999 統計 - 用語と記号 - 第1部：確率及び一般統計用語
- 3 World Health Organization /Europe, Burden of disease from environmental noise (2011)