

疫学エレメント

本書を読み進むにあたって、各項で適宜臨床研究のエレメントについて解説してあるが、ここでは初級編を読む前にどうしても知っておいてほしい基本的な内容について解説してある。

Exposure vs. Outcome

- 人は何らかの因子に暴露(exposure)されることによって疾病などの結果(outcome)を発生する。臨床研究の目的は、この2つの間に真の関係を証明することである。exposureというと放射線被爆のようなものを思い浮かべるが、治療であっても、運動であっても、性別であってもかまわない。一方、結果というと白血病などの病気発生を思い浮かべるが、病気の再発、薬の副作用、あるいは離婚などでもoutcome にし得る。そして、バイアス、交絡因子(confounding)、偶然(chance)の影響を除外できると仮定して、exposure とoutcome の関係について結論付けることができる。

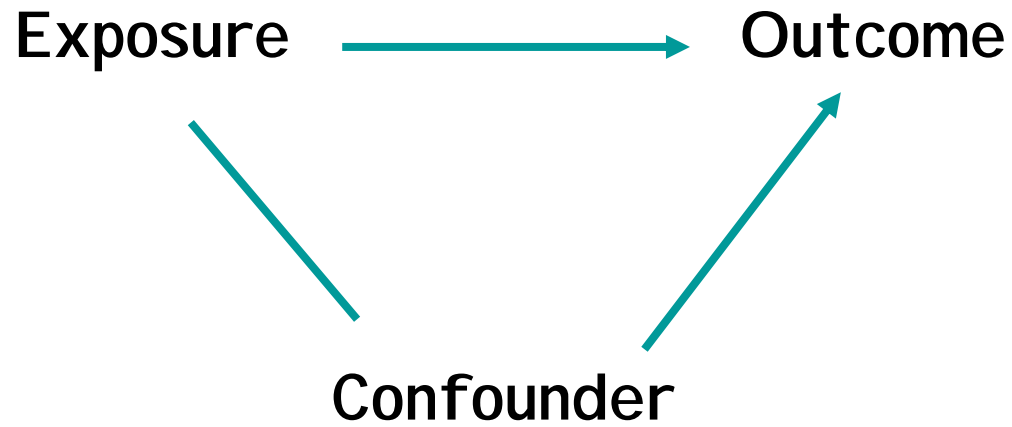


図1. 交絡因子(confounder): 暴露(exposure)と関係があり、結果(outcome)に対して暴露因子とは独立した危険因子を指す。

Bias

- 定義: バイアスとは、暴露と結果の間の関係を歪める結果となる研究デザイン上の不備: Systematic flaw in the design that results in an incorrect estimate of the association between exposure and outcome; である。バイアスが、一度データに混入してしまうと、これを除去することは困難である。これに対して後で説明する交絡因子 (confounder) はデータ回収後も調整可能である。
- 先のウーバンの例で言えば、インタビュアーは、「汚染された井戸の水を飲むと白血病、奇形、アレルギーなどの病気をもちやすい」という仮説を知っていた可能性が大きい。何故なら、白血病の母親もインタビュアーに含まれていたからである。「電話番号からウーバンのどこに住んでいたかは判らない」とはしているが、ウーバンの住人であれば、話の端々からどこに住んでいるかを憶測することは十分に可能であっただろう。汚染された井戸水の地域住人に対しては、無意識のうちにもしつこく疾病発生の有無や程度を誘導したかもしれない。

- 仮にウーバンの住人の多くが、井戸水の汚染があったことを地域新聞やうわさなどで知っていたかもしれない。そして、汚染井戸水地域の住人で、自分の子供が奇形であった場合、インタビュアーに妊娠中井戸水をのんだかなどを尋ねられれば真剣に思い出そうとして、本当の飲水量より過大に申告するかもしれない。逆に、自分の子供が健康であれば過小申告する可能性がある。そうすると、本当は汚染井戸水と病気発生の間に全く相関関係が無いのに、バイアスの混入により相関関係ありと結論してしまうかもしれない。
- また、「全部で8190家庭を電話調査対象とし、7134家庭(88%)から回答を得ることができた。残りの975の電話は平均3.9回かけたが通じなかった。1149人(18%)は質問を拒絶し、32人は英語が話せないために除外、28人は十分質問できておらず除外した。」とあるが、コントロールで拒絶した18%に、自分の子供が奇形をもっているためインタビューを拒否した人たちが、インタビューを受けた人たちより多く含まれていたかもしれない。更に1970年代当時、電話を持っていない家庭がどれ程あったか判らない。しかし、少なくとも白血病を発症したケースがどうであったかを示す必要がある。ケースの家庭に電話があればよいが、いくつかでも電話を持っていない場合、それだけでバイアスが混入する。つまり、電話を持っていない家庭は社会経済レベルが低い可能性があり、そのような家庭が汚染井戸水周辺に多かったかもしれない。
- 前者のように暴露あるいは結果を測定する際にデータに混入する場合を測定バイアス(measurement bias or observation bias), 後者のように対象を選別する際に発生するものを選択バイアス(selection bias)と呼ぶ。

Confounder & Effect Modification

- 交絡因子(confounder) vs. Effect modification
- 図1にあるように、暴露と関係があり、結果に対して暴露因子とは独立した危険因子を指す。具体的には、多変量解析などで調整し、その関係が10%以上変化するようであれば交絡因子と判断する。
- 一方、Effect modifier (適切な日本語訳がみつからない)は層化因子によって結果に及ぼす影響が異なる場合をいう。
- 両者は、exposure, outcome からみれば第三の因子となる(直接因果関係に関与しない)。理論だけをきいてもイメージがわからないと思うので、例を用いて説明したい。

chance

- 偶然(chance)
- 私達は一部の集団からデータを得て全体を推論する。その際、偶然偏った集団を選んでしまうかもしれない。その偏りの程度をp値あるいは信頼区間で表現する。

Longitudinal vs. cross sectional

- Longitudinal とは exposure と outcome の間に時間差があることを示している。Exposure がきっかけで outcome が発生するとしたら、必ず exposure は outcome に先行する。よって、因果関係を捉える場合には、この時間関係を知る必要がでてくる。通常、臨床試験、ケース・コントロール研究、コホート研究は longitudinal である。一方、アンケート調査のような cross sectional study からは時間的關係をつかみにくい。例えば、アンケート調査で、子供の日頃の運動量と肥満度を調べるとする(図2)。どうやら、肥満度と運動量の間には相関関係がありそうである。しかし、運動しないから肥満になるのか、肥満だから運動しないのかは、あるいは両者の相互作用かわからない。アンケートで代表される Cross sectional study は簡便であるが、推論できる根拠に乏しく、その点で longitudinal study より劣る。

Questionnaire at Elementary School

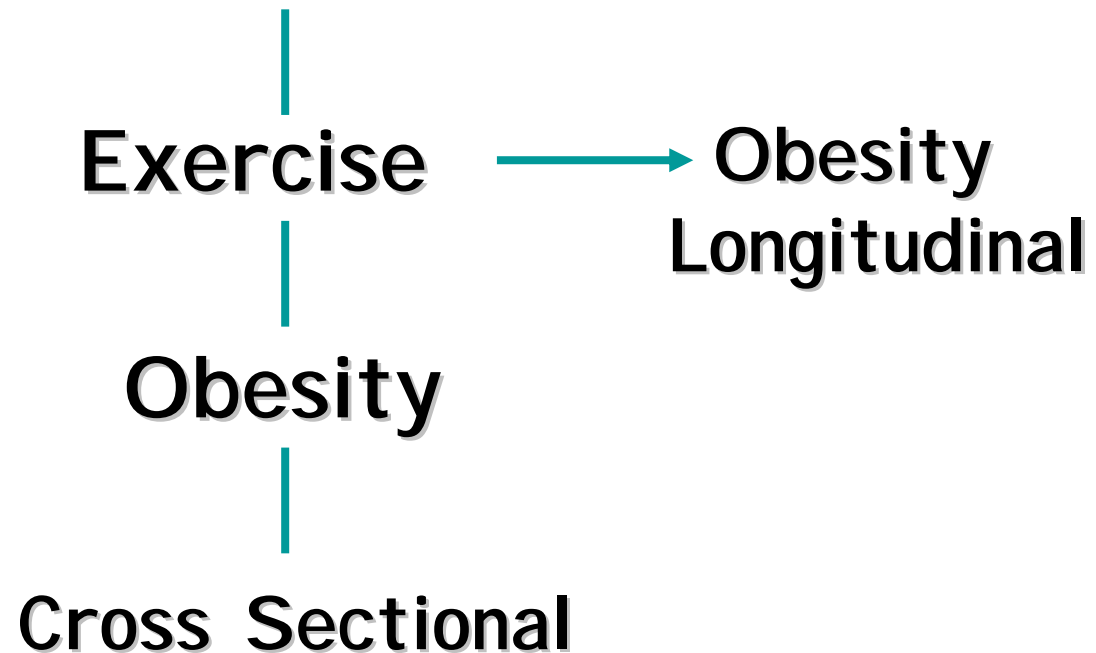


図2. Longitudinal と Cross Sectionalの関係：運動と肥満の因果関係を調査するのにCross Sectional ではどちらが先だか判らないので、相関関係を算出することはできても、因果関係について論ずることはできない。

Discrete vs. continuous

- Discrete とは Yes / No (or 1/0) で表現されるものである。例えば、男女の差、白血病である・ないなどである。一方、continuous は連続的な数値で表されるもので、白血球数や脈拍数などである。しかし、白血球数が12,000以上の時に1、それ未満のときに0とすれば、discrete data に変換することができる。臨床研究では結果がdiscrete であることが多く、このような場合には、odds ratio(OR)やrelative risk (RR)を用いる。一方、結果が連続変数である場合にはt-testなどで比較する。更に、白血病初診時の白血球数で、5万以上を3、2-5万を2、2万以下を1といった具合に分けるような場合はcategorical と呼ぶ。赤、青、黄から一番好きな色を選んでもらったとする。赤か否か、赤が1で表されれば、青と黄は0になる。このような変数をdummy variable と呼ぶ。

Prevalence vs. Cumulative Incidence

- Prevalence (罹病率)は、ある時点(期間ではない)におけるある疾患の割合(proportion)を指す。例えば「日本における1998年の新生児死亡率は0.05%だった。」のような表現がそうである。Prevalence はしばしば cross sectional study で用いられる。あるいは公衆衛生学で重要となる。一方、臨床研究におけるリスクとは、ある人がある一定期間にある病気(結果)になる確率を指し、cumulative incidence (発症率)とはある集団における平均のリスクを指す。例えば「3歳児が今後10年間に不慮の事故で死亡する0.1%である」「誰しも死亡するリスクは100%である。」「女性が生涯睾丸腫瘍になるリスクは0%である。」「ステージIVの悪性リンパ腫が5年間に再発するリスクは35%である。」などである。Prevalence と risk/cumulative incidence を比較すると前者はpoint で疾患の割合を見ているのに対して、後者は一定期間にどれくらいの人が病気に罹患したかを測定している(図3)。よって、Cumulative incidenceにおいて、測定開始時にはみな病気を持っていない点を銘記しておく必要がある。一部の臨床論文はこの重要なルールを無視している。そしてどれくらい観察したかも大切な点である。例えば生涯にわたってのリスクが1%であれば低いと言えるが、1年間のリスクが1%といえ、教室中100人のうち1年後には1人病気を発症しているわけだから、1%でもリスクは高いと言える。

- また異なる研究データを比較するときも注意が必要である。例えば研究Aで10年間のリスクが12%であるのに対して、研究Bでは生涯でのリスクが12%であれば、cumulative incidence は異なることになる。ある疾患になるリスク(or cumulative incidence: CI)はコホート研究により算出される(後述)。よって、コホート研究において研究開始時病気を持っている人は対象から除外する。また病気になりえない人も除外する。早期胃癌で胃を全摘した人や、子宮筋腫で子宮を全摘した人は、胃癌や子宮癌の臨床研究対象からは外される。さらに、結果を発生した時点で経過観察は中止される(センサー)。何故ならずでに@risk ではないからである。

インフルエンザ

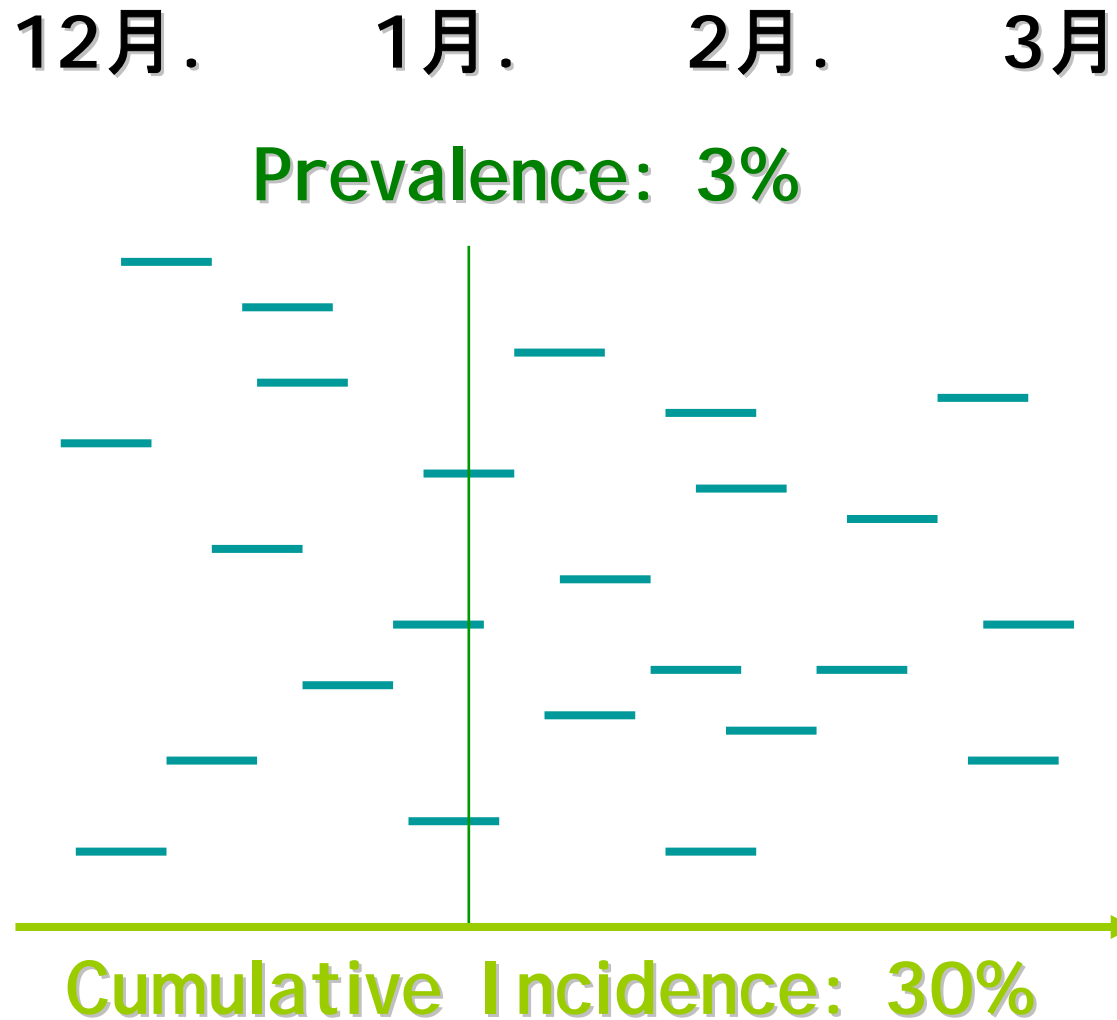
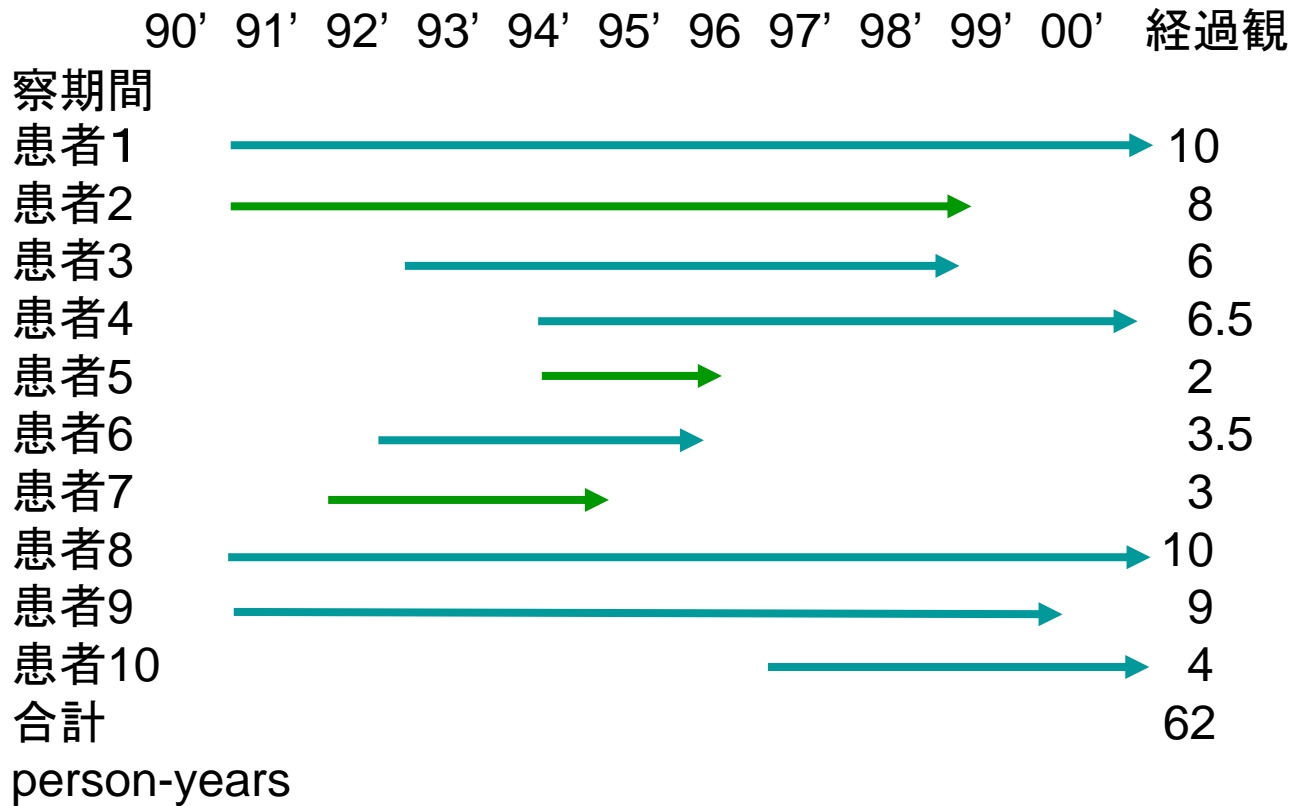


図3 . Prevalence と cumulative incidence の関係 : クラスで1月14日にインフルエンザで休んでいる生徒は3%であったが、シーズン中、インフルエンザに罹患した生徒はクラスの30%に及んだ。

Cumulative incidence vs. Incidence rate

- Cumulative incidence はある一定期間にoutcome を発生するリスクである。しかしながら、うまく測定できない場合がある。途中から研究に加える場合や(variable time at entry)、途中で経過を追えなくなる場合(loss to follow-up)など臨床研究の際しばしば経験する。このような場合下記のようにperson-time でとらえると問題は解決される。
- 3人の患児(5,7,10)が白血病を再発して死亡したとすると、incidence rate = $3 / 62$ person-years になる(図4)。分子が1 - 100 の間に来るよう分母を10,100,,,10,000などで示すのが慣例である。上の例は48.4/1,000 person-years と書く方が一般的である。このような場合においてのみrate を使用する。そしてrate はperson-time の単位をもつ。他の教科書でcumulative incidence rate, prevalence rate といった使い方も見受けられるが、本書では使用しない。またrate 以外のprevalence, cumulative incidence などは割合だから0 - 1 (0 - 100%)で示され、単位を持たない。一方、Rate は割合(proportion)と違って0から無限大までに及ぶ。

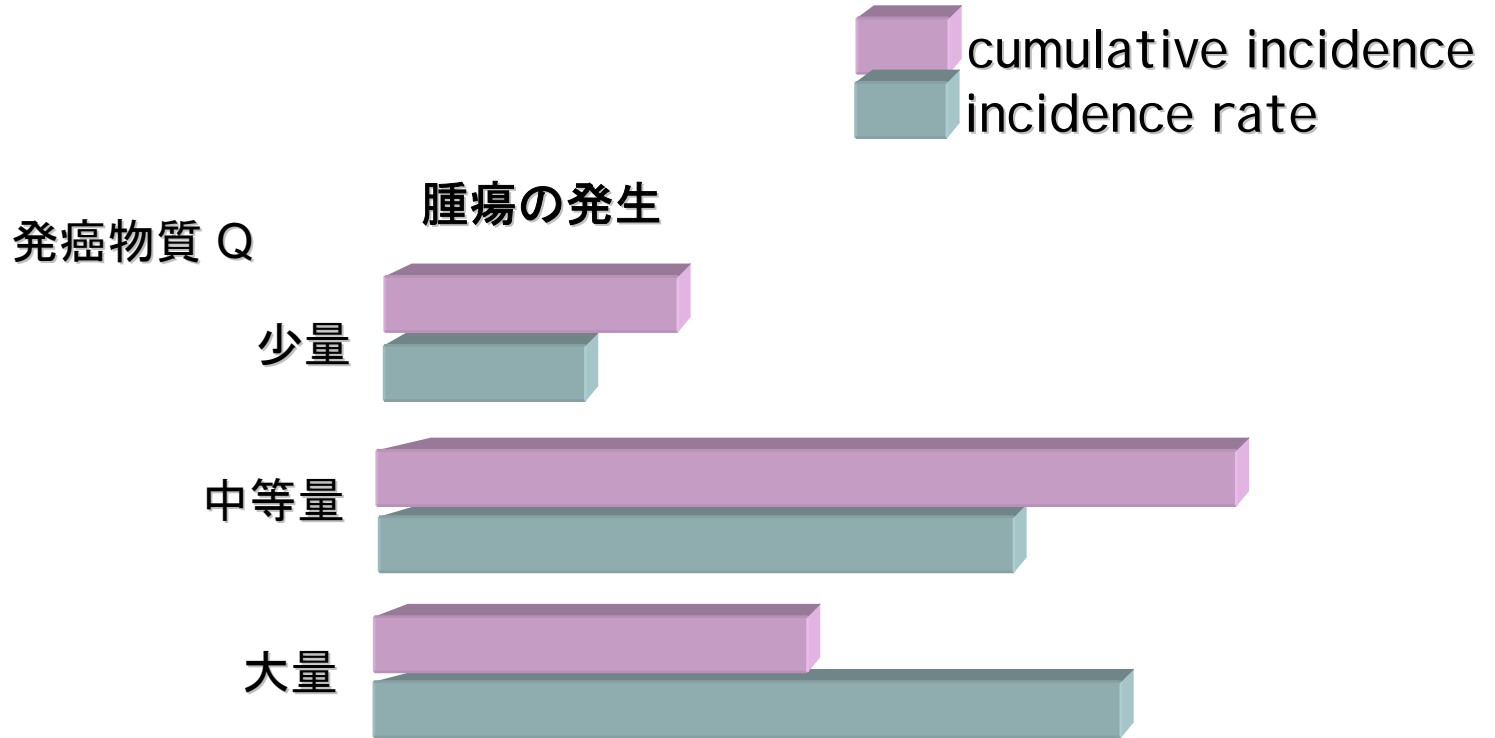
例1



$$3/62 = 0.048 = 48 / 1000$$

図4 Incidence rate: 観察期間の長短に応じてウエイトを変える。いわゆる人年法でウエイトをみていく。

- Competing risksもcumulative incidence やrisk において問題になる。例えば物質Qはマウスに腫瘍を発生させる。そこでマウスに物質Qを2年間与えて影響を比較した。すると大量投与では肝毒性が強く多くが半年未満で死亡してしまったためほとんど腫瘍の発生をみなかった。一方中等量与えたものでは2年後多くのマウスで肝臓腫瘍が観察された。この結果をどう考えたらよいだろうか(図5)。
- このような場合にもincidence rate が効果を発揮する。肝毒性で早期に死亡したマウスはperson-year (mouse year)で考えると、そのウエイトは小さくなる。よって肝臓腫瘍を発生したマウスのウエイトが大きくなり、competing risks の問題を解決することができる。



Overcome (1) variable time at entry, (2) loss to follow up, (3) competing risk

図5 Competing risks : cumulative incidence で考えたときcompeting risk により High dose より、middle dose 投与において腫瘍の発生を多くみているが、incidence rateを用いることにより、middle dose よりhigh dose の方が腫瘍の発生が多くなる。このようにincidence rate を用いることによって、competing risk の問題、途中から研究に加えたい場合、途中で経過を追えなくなる場合でも解析可能となる。

- 以上のようにvariable time at entry, loss of follow-up, competing risksの問題から臨床研究においてはincidence rate が可能な限り好んで用いられる。
- Incidence rate (I), Prevalence (P), Duration (D)の間には以下の公式が成り立つ。
- $P = ID / (1 + ID) \approx ID$ ($P < 0.1$) or $P/(1 - P) = ID$ (odds of disease)
- ここでもし病気の罹患率(prevalence)が0に近づけば(希な病気)、
- $P/(1 - P) \approx P$ と近似できる。よって疾患の罹患率が希であれば、incidence rate の平均に疾患の平均罹病期間をかけてやることによって算出される。 $P \approx ID$ ($P < 0.1$)
- Prevalenceはincidence rate とduration を一緒にしているので、ある意味で情報の一部を捨てていることになる。
- 慢性疾患はincidence rate が低くても罹病期間が長いため、急性疾患ではincidence rate が高くても罹病期間が短いためprevalence は同じになり得る。ある都市で、多発性硬化症と咽頭炎のアンケート調査を行ったとすると、ほぼ同じ数値を得るかもしれない(表1)。

表1 incidence rate, prevalence, duration の関係。多発性硬化症は稀な疾患であり、咽頭炎はよくある病気である。前者は罹病期間が長く後者は短いため、罹病率は両疾患であまり変わらなくなる。

	Incidence rate (I)	Duration (D)	Prevalence (P)
多発性硬化症	3/100,000 PY	25 years	75/100,000
咽頭炎	10,000/100,000 PY	3 days	82/100,000

コホート研究 vs. ケース・コントロール研究

- 観察研究にはexposure を軸にoutcome を評価するコホート研究と、outcome を軸にexposure を観察するケース・コントロール研究がある。
- ある地区で、北の住人はダイオキシンの汚染された井戸水を飲み、南の住人は汚染されていない井戸水を飲むとする。この地区の妊婦とコンタクトをとり、先天性奇形の出産する頻度を比較した。これは前向きコホート研究である(図6)。
- RR (relative risk, risk ratio) = $1.55/0.36 = 4.30$; 北に住む妊婦は南の4倍余り奇形を生むリスクが高い。 RD (risk difference = absolute risk) = $1.55 - 0.36 = 1.19$ (%); 北に住む妊婦は南に住む妊婦より奇形を生むリスクが1.2%高い。
- NNH (number needed to harm) = $1/RD = 84$; 84人に1人の妊婦は北に住むことによって奇形をもつ児を生むと計算される。
- $Attributable\ risk\ percent\ (AR\%) = (RR-1)/RR \times 100 = 76.7\%$; 北地区にみられた奇形の77%は、北地区に住んだことによる。北地区の妊婦が南地区に住んでいたら、奇形の発生は77%少なかった。
- $Population\ attributable\ risk\ percent\ (PAR\%) = [p(RR-1)]/[p(RR-1)+1]$
ここでは、北地区に住む人は $393/2332 = 17\%$, $p = 0.17$
- $PAR\% = 35\%$; もしも、汚染された水が原因であるとしたら、この水を飲まないことにより地区全体として35%奇形を減らすことができる。分母を person-time をもちいることによってriskをincident rate として示すこともできる。この場合のRR はrate ratio となり、同じrisk ratio と同じで混同しやすいが意味はほぼイコールである。

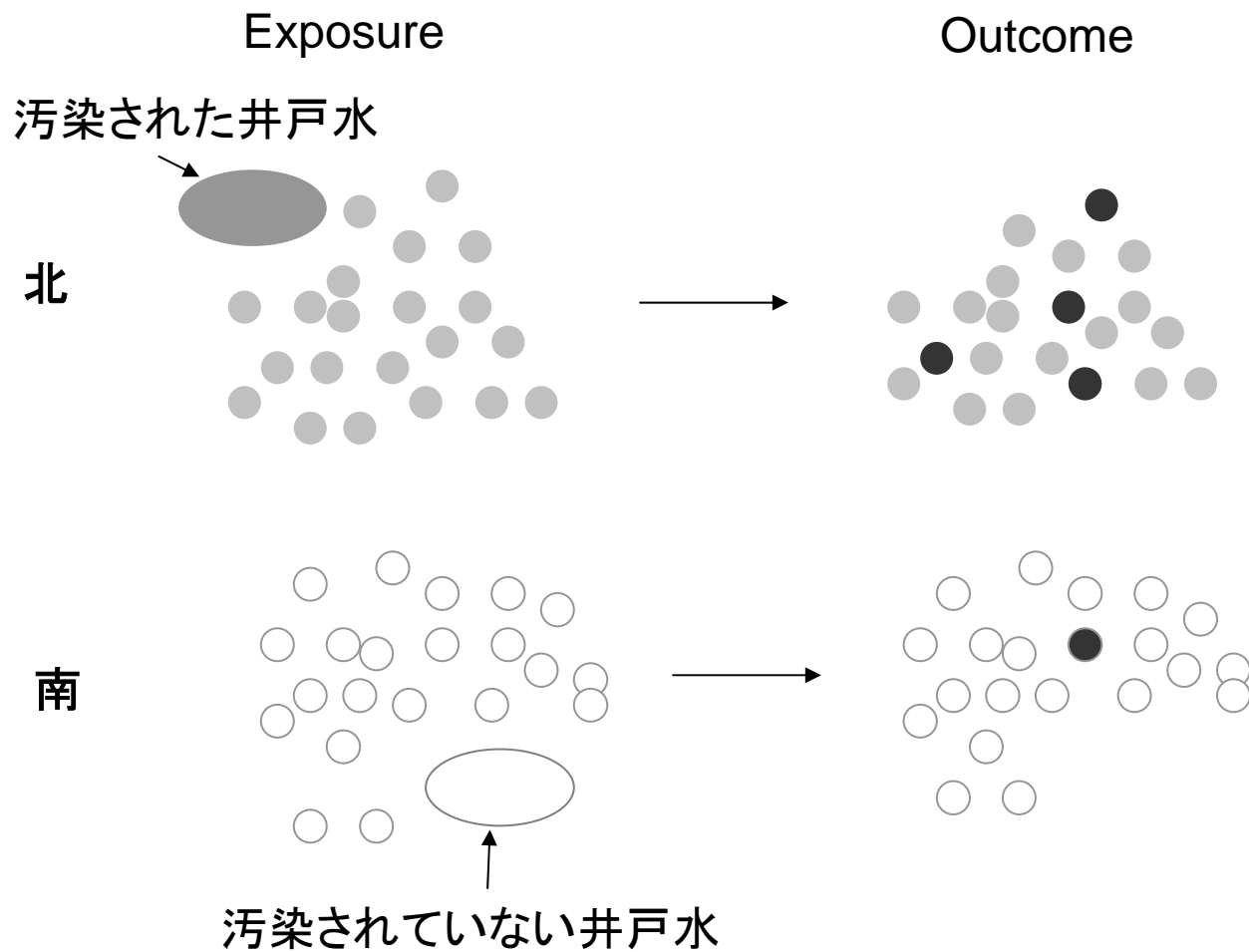


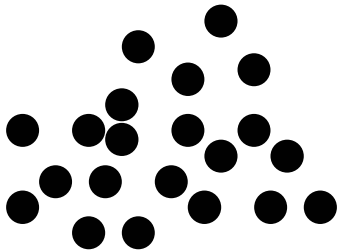
図6 コホート研究:汚染された井戸水を飲んで生活していた北に住む妊婦と、そうでない南に住む妊婦を対象にして、先天奇形の出産について評価した。

表2 コホート研究の結果

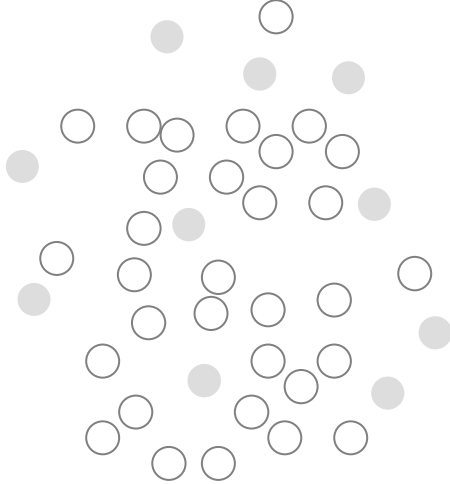
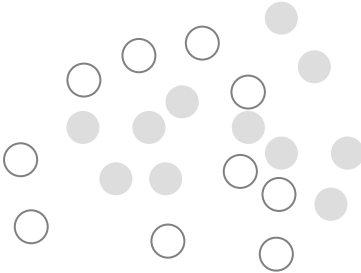
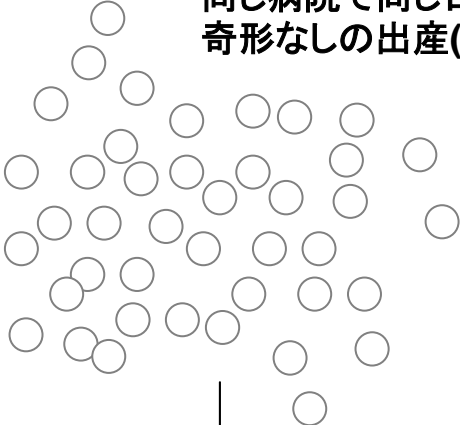
	奇形あり	全出生数	risk
北	6	387	153/10,000(1.53%)
南	7	1932	36/10,000(0.36%)

Outcome

奇形あり(case)



同じ病院で同じ日に生まれた
奇形なしの出産(control)



- 北
- 南

Exposure

図7 ケース・コントロール研究: 奇形のある児を生んだ妊婦と、奇形を生まなかった妊婦の妊娠中の住居を調査した。

- ある病院で奇形のある児と、同じ病院で同じ日に生まれた奇形の無かった児において、母親の住居を比較することにした(表3)。これは後ろ向きケース・コントロール研究である。
- Odds ratio (OR) = $(141 \times 4867) / (133 \times 1250) = 4.13$
- この場合のORは、汚染された井戸水の配水区域である北に住むことによって奇形児を生む危険性を示している。注意すべき点として、コントロールがstudy population からランダムに抽出されてきた場合はよいのであるが、この場合同じ出産日、同じ病院でマッチングしているため、study population の特徴を正しく代表していないことになる。その場合にはrelative risk (RR)を算出することができない。しかし、コントロールがstudy population 全体そのものであったり(表4)、ランダム抽出により全体を代表していたりすればRRを算出できる。そして、outcome 発生頻度が少なければ両者は近似する。何故ならodds = case/control であり、risk=case/total であり、total = case + control で、case が十分小さければcontrol とtotal は近似するからである。

表3 ケース・コントロール研究の結果

	case	control
北	141	133
南	1250	4867
合計	1391	5000

$$\text{Odds ratio} = (141 \times 4867) / (133 \times 1250) = 4.13$$

$$[\text{Risk ratio} = (141 / 141 + 133) / (1250 / 1250 + 4867) = 2.52^{***}X]$$

表4 study population からランダムにコントロールを抽出

	case	control	total	Risk
北	141	31750	31891	4.42/1000
南	1250	1161850	1163100	1.07/1000
合計	1391	1193600	1194991	RR = 4.13

$$\text{Odds ratio} = (141 \times 1161850) / (1250 \times 31750) = 4.13 = \text{RR}$$

Absolute risk vs. relative risk

- Absolute risk はリスク差 (risk difference) のことである。一方、relative risk はリスク比である。例えば川崎病に罹患した子供の10%が冠動脈瘤を合併するとする。そして、初期よりガンマグロブリンを使用することにより冠動脈瘤合併率を5%にまで減らすことができるとする。Relative risk は0.5 であり、risk difference は0.05である。Risk difference の逆数は number needed to treat (NNT) と呼ばれ、「何人に1人がその治療の恩恵にあずかるか？」を知ることができる。これが副作用であれば、number needed to harm (NNH) であり、「その治療により何人に1人が副作用をもつか？」に相当する。冠動脈瘤発生を半分に減らせるというインパクトが大きいですが、ガンマグロブリンを20人の川崎病患児に投与し、その恩恵にあずかるのは僅か1人しかいないとすると、血液製剤投与であることとのトレードオフが問題となる。

オッズ比 : Odds ratio vs. 相対比 : Relative risk

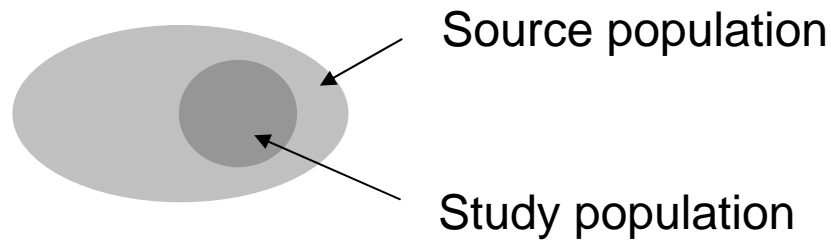
- 先に説明した通りで、大雑把に言うと、ケース・コントロール研究では odds ratio (OR), コホート研究, 臨床試験では relative risk (RR) を使うことが多い。主な理由は先に述べた通りである。更に、relative risk を用いると number needed to treat (NNT) などを計算できる利点もある。しかし、研究者はコホート研究や臨床試験で OR を用いることもあるかもしれない。対象に比して outcome が少なければ RR と OR は近似する (rare case assumption)。しかし、どの程度稀であるかのカットオフ値は決まっていない。そして、通常は同じ分子の数に比して、OR の分母の方が RR の分母より小さくなるのであるから、OR の方が RR よりインパクトのある数値になる。つまり、研究者によっては結果のインパクトをより大きく表現するために OR を好んで使うかもしれない。また、OR だと logistic regression analysis により多因子を同時に解析できる長所もある。

Source population vs. study population

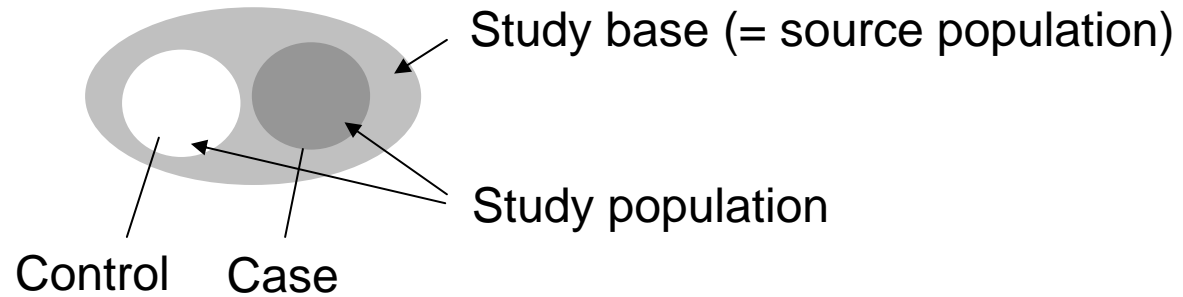
- しばしばどの集団を対象に臨床研究を行ったか不明確な場合がある。しかし、対象集団を明確にすることは臨床研究をデザインする上で最も重要なことである(図8)。Target population: 例えば20代から60代のナースを対象に調査を行なう場合、研究者は追跡調査しやすいという点でナースを選んだが、結論を全女性にまで拡張することを考えている。すなわち、結果をどこまでに適応できるか(あるいはしたいのか)、それがtarget populationということになる。閉経前の女性を対象に行なう経口避妊薬の冠動脈発生に関する調査結果を閉経後も含めた全年代層の女性をtarget populationとして適応することはできない。
- Study population: コホート研究においてperson time に組み込まれる人々のことを指す。すなわちある臨床研究において1000人の応募者中から診察の結果基礎疾患のない健康人500人を選んできて諸々の調査をしたとすると500人はstudy population であり、1000人はsource population ということになる。またstudy population = source population である場合も時にあるが、通常のcohort study においてはstudy population はsource population の一部を成す。
- Case control study においてはcase + control = study population ということになる。Case とcontrol はstudy base から抽出される。

- Source population: study population を抽出した源の人々を指す。Study population が source population よりランダムに抽出された場合、study population の結果をもって source population を推論できる。例えば港区でランダムに携帯電話に連絡をして1000人を抽出し港区の一般的傾向を推論する場合である。しかし、携帯電話でコンタクトをとれる人は、必ずしも港区住民の平均にあたるかどうかは判らない。よって一般的には study population を source population から完全にランダムに抽出することにこだわらないことが多いようである。このような状況を conditional occurrence relation と呼ぶ。また、ケース・コントロール研究において、ケースとコントロールの source population が異なることがある。これはバイアスを生じる可能性がある。ケース・コントロール研究の論文の方法を読む際、ケースとコントロールの base (or source) が異なっていないか注意深く検討するべきである。
- Base population: 所謂 person time になる人々を指す。cohort study においては base population = study population であり、case control study においては study base の中から person time に貢献する人々を指す。
- Study base: cohort study においては study base = study population = base population で、person time を構成する人々を指す。よってコホート研究においてはこんなに多くの名前と呼ぶ必要はなく、かえって混乱の元である。一方、ケース・コントロール研究において、ケースとコントロールは共通の study base から抽出する。抽出の際、病気の有無のみで抽出され、exposure と独立してはならない。

Cohort study



Case-control study



☒8 **Source population vs. study population**

Closed cohort vs. open cohort

- コホート研究はClosed (fixed) とopen (dynamic) の2種類に分けられる。
- コホート研究においては最初にstudy population を固定するが、open cohort においては一定以上の調査期間を設け、途中から加わることも可能であり、途中で追跡不能になってもかまわない。
- たとえば東海村に臨界事故のあった際、生活していた人々を追跡調査する場合closed cohort であり、それ以降移住してきた人々は含まない。また途中追跡不能になった場合は調査対象から除外しなくてはならない。臨床研究論文の方法にしばしばその旨が記載されている。1万人応募して5000人が適応であり、研究参加に同意し4000人で開始したところ最後まで経過を追えたのは3000人であった、といった具合である。
- 一方ある病院Aに1990年から2000年までの間にカルテのある人々を対象に研究する場合はopen cohort になる。例えばAさんは調査の開始から終了までA病院にかかっていたが、Bさんは1995年から最後まで、Cさんは1993年から1995年までかかったとする。open cohort では3人とも対象となる。そしてこのような場合person time として10年間追跡調査したAさんは2年間しか追跡されなかったCさんよりより大きく分母の数値に貢献することになる。また10年間の追跡調査で喫煙者が追跡開始より2年で禁煙してしまった場合でもperson time に組み込むことはできるが、cumulative incidence では対象からはずさなくてはならない。よってopen cohort に対しては通常person time が用いられる。

- Closed study ではincidence rate とcumulative incidence を測定できるが、incidence rateでは各々の観察期間が判っていないのではなく、cumulative incidenceでは追跡不能者が少なく、competing risk がなく、観察期間中のexposure risk が同じであり、対象は皆同じ期間観察されなくてはならない。例えば東海村放射能漏れ事故で多くの死亡が不幸にしてでてしまった場合、cumulative incidence で測定しようとするとう白血病の発生率は真の値より過小評価されてしまう。また他の場所へ引っ越す人々も多いかもしれない。長期追跡調査では加齢の問題もからんでくる。よってclosed study においてもincidence rate で算出する方が有利な場合がある。Cumulative incidence を使うとすれば短期で結果発生頻度が高い場合、例えば食中毒などの臨床研究に向いている。

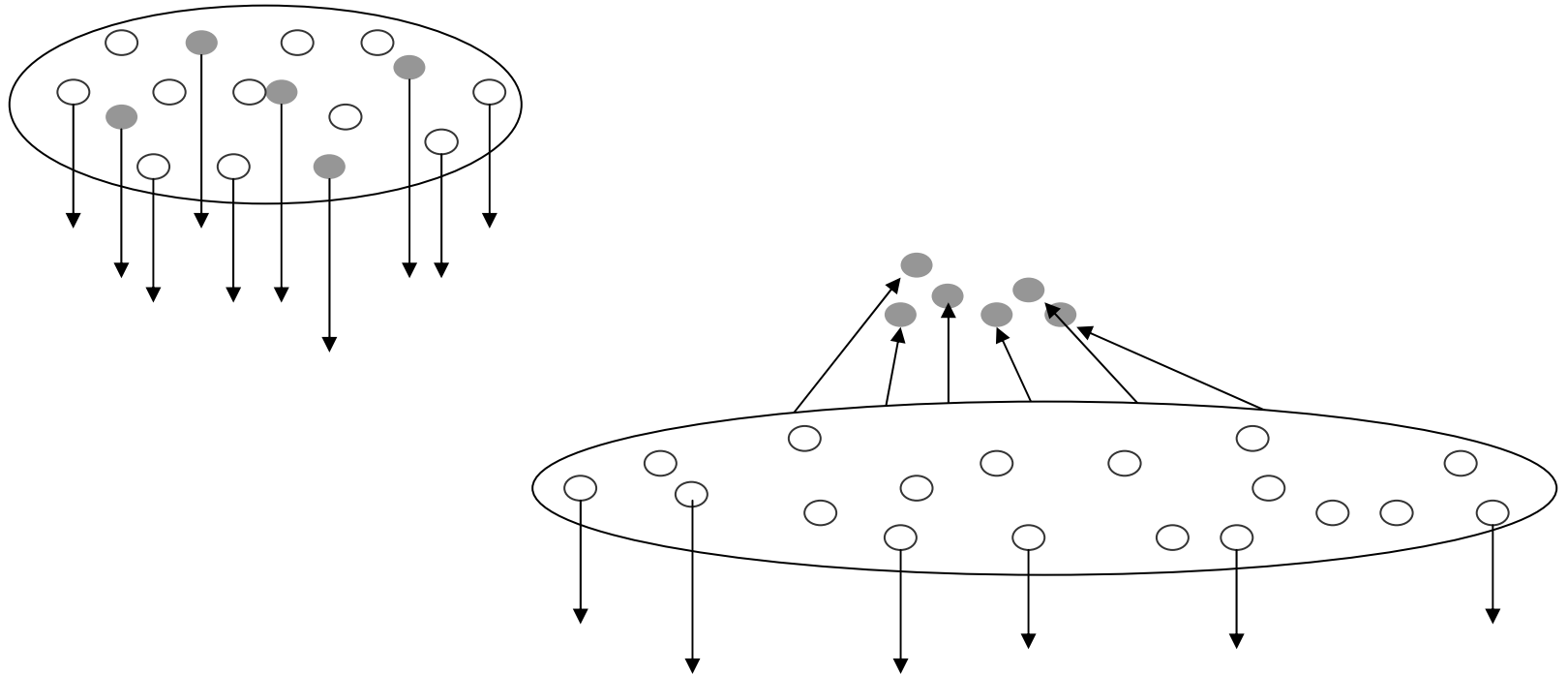
ケース vs. コントロール

- ケースもコントロールもexposure とは無関係に抽出されなくてはならない。コントロールはランダムに抽出すれば何の問題もない。しかし、ケース・コントロール研究の利点としてexposure 以外の因子を一致させて選んでいくことができる(マッチング)。その際exposure vs. no-exposure の比率がstudy base と異ならないようにしなくてはならない。例えば小児喘息とアトピー素因の関係を調べるのに歳の近い喘息をもたない兄弟をコントロールとすることを考える。しかし、アトピー素因は遺伝的要素が強いので、兄弟をコントロールとした場合、一般人口よりもコントロール群にアトピー素因が多く含まれる。よってアトピー素因と小児喘息の関係を過小評価してしまうことになる。

Primary vs. Secondary study base

- まずbase population を設定してその中からケースとコントロールを選定する場合をprimary study baseと呼ぶ(図9)。nested case control study (How to use クリニカル・エビデンスを参照)や、ある地域で発生する病気を調べるために、住民台帳からコントロールをランダムに抽出する場合が相当する。
- 一方caseが最初に存在し、それに基づいてstudy base を想定するものをsecondary study base (case-defined study base)と呼ぶ。病院ベースのケース・コントロール研究などはこれに当たる。例えば、入院中の肺がん患者を対象に、同じ病院に入院中の肺がん以外の疾病を持つ患者さんをコントロールとする。つまり、コントロールをケースが発生したのと同じpopulationから抽出する。しかし、紹介システムは複雑であり、ケースが発生したstudy base と、コントロールを抽出したstudy base が同じであるとは限らない。ケース・コントロール研究をデザインあるいは論文の批判的吟味をする際、ケースとコントロールが同じstudy base から発生しているか否かを検討することが重要である。

Primary study base method:最初にstudy base を明確にする。



Secondary study base method:最初のcaseを抽出し、
あとからそのcaseのstudy baseを考えcontrolを抽出する。

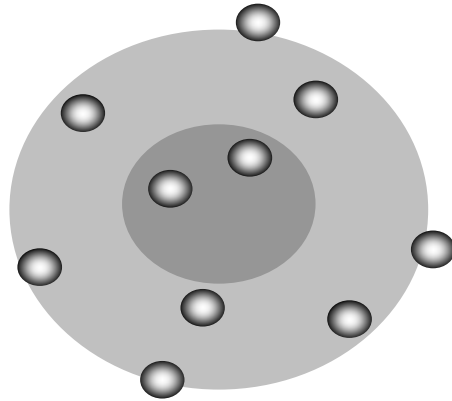
図9 Primary vs. Secondary study base

妥当性: Valid or 正確性: Accurate

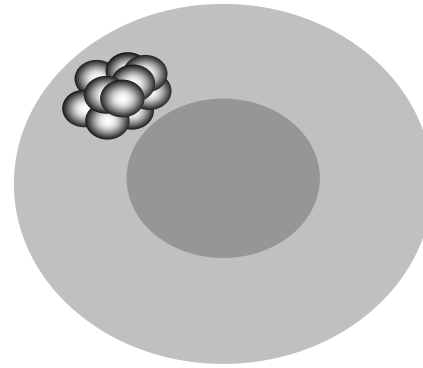
- バイアスのない状態をvalidと呼ぶ。これはefficiency (=accuracy, precise)とは性格を異にする(図10)。Validにするためには研究デザインを工夫し、バイアスの少ない質の高いデータを得る必要がある。データにバイアスが含まれてしまうと、どんなに高価なコンピュータを用いて解析しても、妥当性を上げることはできない。

内的妥当性Internal validity vs. 外的妥当性external validity (=generalizability)

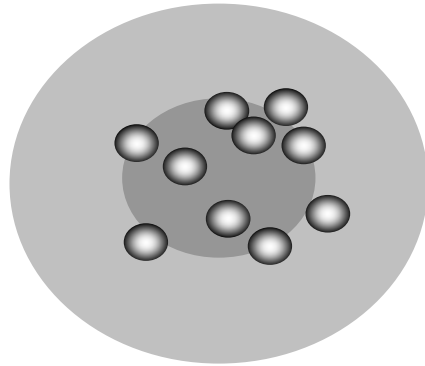
- 例えば、ある中国の二分脊椎の多い村で妊娠中より葉酸を内服させたところ、二分脊椎の発生を20%にまで減少させることができた。これはその中国の村ではvalidであるが(内的妥当性: internal validity)、東京で同じデータが再現できるとは限らない(外的妥当性: external validity)。何故なら、その中国の村の対象は東京と二分脊椎の発生しやすさという点で異なると予想されるからである。
- ランダム化二重盲検臨床試験では、内的妥当性を高めることができるが、参加者をプロトコル中の特殊な条件により絞り込んだ際、対象疾患の中でも偏りをもった集団となってしまうことがある。つまり、内的妥当性を上げるのに苦慮しすぎると、外的妥当性を落とすことになり得る。偏頭痛に対する新薬を試すために患者リクルートを行った。しかし、ほとんどの患者さんに断られ、説明した人の10%でしか同意を得られず、そのほとんどの患者さんは標準治療では頭痛の改善しない人たちであった。この条件でランダム化二重盲検臨床試験を行った場合、得られた結果は内的妥当性を有するが、この新薬の効果が全ての頭痛患者さんでもあてはまるとは限らない。つまり、外的妥当性をおとすことになる。
- 例えば、ダイオキシンの影響を出生児の奇形合併頻度で比較したところ有意な差を得なかったとする。これに対して論文査読者が、「奇形をもつ児は既に自然流産してしまうはずであり、この研究はその点を関知していないので妥当ではない」と指摘したとする。この研究結果は生きて生まれた児に対して適応できる結果であり(内的妥当性)、自然流産に関しては新たな研究をしなくてはならない(外的妥当性)。



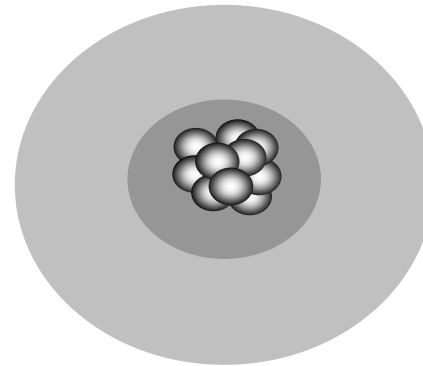
Not accurate not valid
的外れ&不正確



Accurate but not valid
的外れ&正確



Valid but not accurate
妥当&不正確



Valid and Accurate
妥当&正確

図10 妥当(valid)であることと正確(accurate)であることの相違