

疾患発生頻度の測定

臨床研究あるいは疫学においてしばしば用語の混同がみられます。本書ではハーバード大学の言葉の定義に従うようにします。これらの混同した用語を日本語に訳すとよけい混乱を来たしますので、なるべく英語のまま解説をしていきたいと思ひます。

Prevalence

ある時点（期間ではない）におけるある疾患の割合 (proportion)。

例

1. 1991 年の癌による死亡率は 23.8% だった。
2. 1960 年フレミングムにおける 52 歳から 85 歳の 2477 人中 310 人が白内障に罹患していた。

しばしば cross sectional study に際利用されます。また慢性疾患の場合有利ですが、インフルエンザのような急性疾患の場合その月の prevalence が低くてもその年のインフルエンザが少ないとは限りません。

Time referent: prevalence を算出した時間に関する数値のこと。

例

1985 年 40—45 歳の白内障の prevalence は 0.5% であり、80—85 歳では 43% であつた。Time referent は？

→1985 年と、年齢グループ両方が time referent にあたります。

Risk and Cumulative incidence

Risk

ある人がある一定期間にある病気になる確率。

Cumulative incidence

ある集団における平均のリスク

例

1. アメリカ女性の 8 人に 1 人、12.5% が生涯乳癌に罹患する。
2. 誰も死亡するリスクは 100% である。
3. 女性が生涯睾丸腫瘍になるリスクは 0% である。

4. ステージ I の悪性リンパ腫が 5 年間に再発するリスクは 35%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ます。すなわち、後者において、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ります。そしてどれくらい観察したかも大切な点です。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ますが、1 年間のリスクが 1%といえば高いと言えます。すなわち教室の中 100 人のうち 1 年後には 1 人病気を発症しているわけですから。また異なる研究データを比較するときも注意が必要です。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%といった場合と等価ではないのです。論文を比較する際注意が必要です。しかし観察期間が同じであれば、仮に小児期であろうと老年期であろうと限定する必要はありません。例えば生涯で乳癌になるリスクが 12.5%であるといっても、65 歳以前に 12.5%なのと 65 歳以前は 0%で 65 歳以降で 12.5%であるのとでは雲泥の差です。しかし単にリスクといったときこの 2 つを区別できません。

ある疾患になるリスク (or cumulative incidence: CI) は cohort study (コホート研究) により算出されます。

$t\text{-year } C = \frac{\text{ある一定期間に発生した病気の人数}}{\text{その期間病気になる可能性のある人数}}$

cohort study においてある一定期間の最初、病気を持っている人は対象から除外します。また病気になりえない人も除外します。

Framingham Heart Study*において 30—39 歳の 789 人を 12 年間に渡って追跡調査したところ 40 人が冠動脈疾患に罹患しました。12 年間の cumulative incidence は $40/789 = 5\%$ です。調査開始時既に冠動脈疾患を持っていたら対象から除外します。

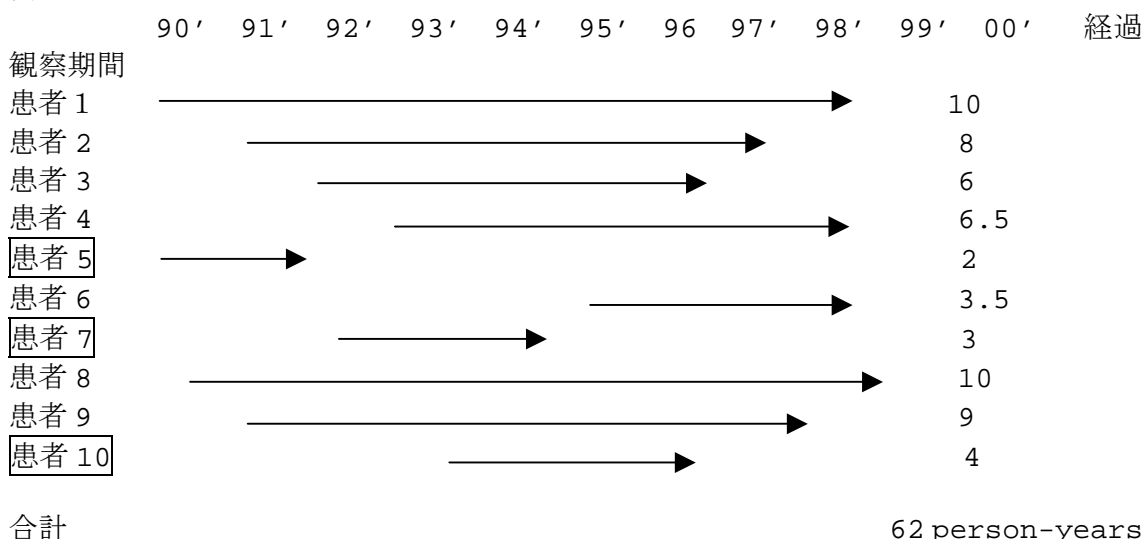
*Framingham はボストン郊外の街で、戦後この街に暮らす多くの人の健康状態を追跡調査し主に冠動脈疾患のリスクファクターが明らかにされました。大規模 cohort study の草分けとして知られます。

40 歳代 1000 人を対象に子宮癌に影響する環境因子を研究しようと思います。その際、何らかの理由で子宮摘出術を受けた女性は子宮癌にはなりえないので対象から除外します。

Incidence rate

リスクは最も臨床研究で用いられる測定法です。しかしながら、うまく測定できない場合があります。途中から研究に加えない場合や(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します。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ます

例 1



3 人の患者が腎不全を合併して死亡したとしますと、
incidence rate = 3 /62 person-years
になります。

分母はしばしば 10,100,,10,000 など示されます。その際一般的に分子は 1-100 の間に来るようにします。ですから上の例は
=48.4/1,000 person-years
と書く方が一般的です。

このような場合においてのみ rate を使用します。そして rate は person-time の単位をもちます。他の教科書で cumulative incidence rate, prevalence rate といった使い方も見うけられますが、本書では使用しないことにします。また rate 以外の prevalence, cumulative incidence などは割合ですから 0-1 (0-100%) で示され、単位を持ちません。Rate は割合 (proportion) と違って 0 から無限大までに及びます。

$$\text{Incidence rate} = (\text{ある期間に発症した患者数}) / (\text{person-time の合計})$$

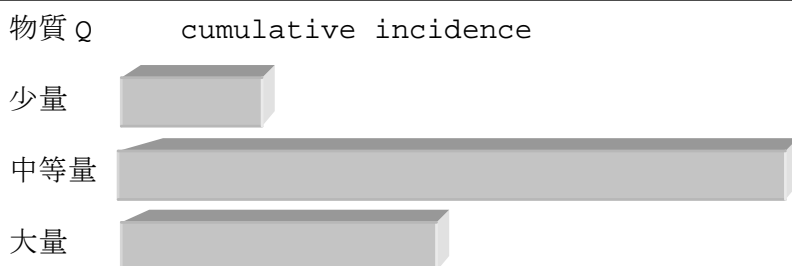
もちろん先の例のように子宮の無い人は子宮癌のリスクを持たないので対象から除外します。

1975 年、Framingham study において 35 歳女性 5000 人を追跡調査していましたが、General Motors が閉鎖したために 2500 人が他に職を求めて引っ越してしまいました。この研究では 1990 年までに 50 人が心筋梗塞を発症したので、15 年の

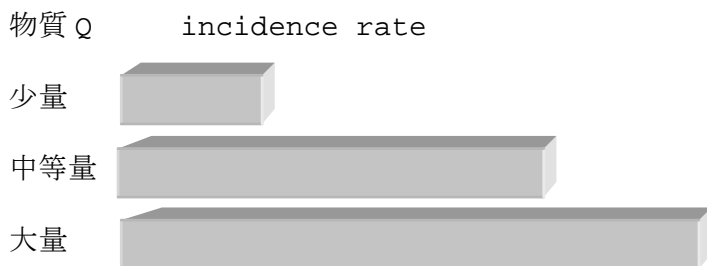
cumulative incidence は 1%ということになります。しかし 5 年後、Harvard School of Public Health の学生が、判らなくなっていた 2500 人の追跡調査をして更に 100 人の心筋梗塞例を追加しました。よって CI は 3%に増えました。Loss of follow up は大きな bias となります。しかし incidence rate を用いるとその問題は解決されます。

Competing risksも cumulative incidence や risk において問題になります。

例えば物質 Q はマウスに腫瘍を発生させます。そこでマウスに物質 Q を 2 年間与えて影響を比較しました。すると大量投与では肝毒性が強くて多くが半未満で死亡してしまったためほとんど腫瘍の発生をみませんでした。一方中東量与えたものでは 2 年後多くのマウスで肝臓腫瘍が観察されました。この結果をどう考えたらよいのでしょうか？



このような場合にも incidence rate が効果を発揮します。肝毒性で早期に死亡してしまったマウスは person-year (mouse year?) で考えるとそのウエイトは小さくなります。よって肝臓腫瘍を発生したマウスのウエイトがあがり、competing risks の問題を解決することができます。



非喫煙者の冠動脈疾患で死亡する割合 (cumulative incidence) が喫煙者のそれより高いとします。喫煙は冠動脈疾患を抑制すると結論してしまってもよいのでしょうか？

喫煙者の多くは冠動脈疾患を発症する前に肺癌や慢性呼吸器疾患で死亡することが多く、cumulative incidence でみると喫煙の冠動脈疾患発生を過小評価してしまう可能性があります。よって上記のように結論するのは危険です。このような場合 incidence rate で評価すべきです。

よって competing risks の存在下で cumulative incidence を直接計算すると incidence rate (IR)あるいは IR から計算しなおした cumulative incidence と

比べ、実際のデータを過小評価してしまうことになります。

以上のように 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ます。

Incidence rate と Cumulative incidence の関係

Incidence rate から Cumulative incidence に変換することはできないもの
 でしょうか？確かに IR は便利なのですが、CI は 20% のリスクなど直感的に理解しや
 すい利点があります。

以下の公式で変換できます。

$$T - \text{year } C = 1 - \exp.\{- \sum I_j \Delta T_j\}$$

例えば 50 歳の白人アメリカ人が 20 年間で乳癌になるリスクは何%ですか？

年齢	Incidence rate (1983-7)	観察期間 ΔT_j	$I_j \Delta T_j$
50-5 4	221.1/100,000 PY	5	1105.5/100,000 PY
55-5 9	272.1/100,000 PY	5	1360.5/100,000 PY
60-6 4	334.8/100,000 PY	5	1674.0/100,000 PY
65-6 9	392.3/100,000 PY	5	1961.5/100,000 PY

$$20 - \text{year } C = 1 - \exp.\{- \sum I_j \Delta T_j\} = 5.9\%$$

よって 50 歳のアメリカ白人女性が 70 歳までに乳癌になるリスクは 6% であるといえ
 ます。アメリカ人女性では乳癌が最も多い癌です。

上記表では年齢が上がる毎に incidence rate も上がりましたが、常に一定だとした
 ら単純に

$$t - \text{year } C = 1 - \exp.\{- IT\}$$

で表されます。

例えば上の例で 65 歳の女性が 5 年間に乳癌になるリスクはどれくらいでしょう？ 65
 から 69 歳までの incidence rate は一定であると仮定して計画してみてください。

$$20 - \text{year } C = 1 - \exp.\{- 0.003923 \times 5\} = 1.94\%$$

もし

$$\sum I_j \Delta T_j < 0.1 \text{ であれば}$$

$$t - \text{year } C = \sum I_j \Delta T_j$$

で近似できます。

更に I が一定で $IT < 0.1$ ならば

$$t - \text{year } C = IT$$

で近似できます。

Incidence rate と Prevalence の関係

Incidence rate (I), Prevalence (P), Duration (D)の間には以下の公式が成り立ちます。

$$P = ID / (1 + ID) \approx ID \quad (P < 0.1)$$

Or

$$P/(1 - P) = I \times D \quad (\text{odds of disease})$$

ここでもし病気の罹患率(prevalence)が0に近付けば(希な病気)、 $P/(1-P) \approx P$ と近似できます。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ます。 $P \approx ID \quad (P < 0.1)$

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ります。

	Incidence rate (I)	Duration (D)	Prevalence (P)
多発性硬化症	3/100,000 PY	25 years	75/100,000
肺癌	40/100,000 PY	1 years	40/100,000
咽頭炎	10,000/100,000 PY	3 days	82/100,000

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 は同じになります。肺癌の方が多発性硬化症より10倍もIが高いのに罹病期間が短いため prevalence はむしろ多発性硬化症の半分です。つまりサーベイを行なうと、咽頭炎、多発性硬化症、肺癌の順に多いという解答が得られるはずです。

関係の測定

ある因子に暴露された場合とそうでない場合、あるいは薬物投与群とプラシーボ群で比較する場合どのような方法があるでしょうか？

Difference measure

Risk difference (RD) = $C_e - C_u = \text{attributable risk}$ (同義語)
 C_e : cumulative incidence of exposed group, C_u : cumulative incidence of unexposed group

$$-1 < RD < +1$$

母親の血中のポリッサッカライドに対する抗体価と乳児期の下痢の発生を比較しました。RDを計算し解釈を加えてください。

	低力価	高力価	
下痢	12	7	19
合計	14	16	30
C	12/14	7/16	

$$RD = 12/14 - 7/16 = 42\%$$

母親のポリッサッカライドに対する抗体価が低い場合、42%の乳児が余計に下痢を発症する。

Rate difference (RD) = $I_e - I_u$
 I_e : incidence rate of exposed group, I_u : incidence rate of unexposed group

$$-\infty < RD < +\infty$$

何らかの治療目的で胸部に放射線照射を受けた人に乳癌が多く発生するかどうか調べました。RDを計算してください。

	放射線照射+	放射線照射-
乳癌	41	15
Person-years	28,010	19,017
IR	41/28,010	15/19,017

$$RD = 6.71 / 10,000 \text{ PY}$$

Ratio measure

Risk ratio (RR) = $C_e / C_u = \text{relative risk}$

先の例でRDでなくRRはどうですか？

	低力価	高力価	
下痢	12	7	19

合計	14	16	30
C	12/14	7/16	

$$RR = (12/14) / (7/16) = 1.96$$

RR が 1 – 2 の間の場合には
母親の抗体価が低いことによって乳児下痢のリスクが 96%増加する と表現します。

一方 RR が 2 以上の場合は何倍という表現を好んで用います。「喫煙により肺癌のリスクは 10 倍になる」など。

また母親の抗体価が高い場合、49%リスクが減ります とも言えます{RR=0.51}

incidence rate の場合も同様に
incidence rate ratio (RR) = I_e / I_u

relative risk とも呼ばれます。

	放射線照射+	放射線照射+
乳癌	41	15
Person-years	28,010	19,017
IR	41/28,010	15/19,017

$$RD = \{41 \cdot 28,010PY\} / \{15/19,017\} = 1.85$$

胸部放射線照射は乳癌の発生を 2 倍近くまで増やします。あるいは胸部放射線照射により 85%乳癌の発生が増えますとも報告できます。

もし暴露と結果の間に相関がなければ RR は 1 に近づきます。一方 RD は 0 に近づきます。RR は RD と異なり PY はつきません。

1970 年代、vinyl chloride に暴露されると肝血管肉腫になる確率が 1,000 倍に上がる事が発見されました。一般的な肝血管肉腫の発生頻度は年間 10—20、すなわち $3.33/10^8PY$ としますと、 $RD = 1000 \times 3.3/10^8PY - 3.3/10^8PY = 3.297 / 10^5 PY$ となります。RD は非常に小さく vinyl chloride を除去しても目に見えて肝血管肉腫が減るわけではありません。何故なら vinyl chloride に暴露されて発生する肝血管肉腫自体が少ないからです。

下の表は喫煙と冠動脈疾患、肺癌の関係を示したものです。

冠動脈疾患

年齢	Incidence rate / 10 ⁵ person-years		RD/ 10 ⁵ PY	RR
	非喫煙者	喫煙者 (21-39 cig/day)		
55-64	501	969	468	1.93
65-74	1,015	1,710	695	1.68

年齢によって RD と RR は異なります。しかし RD は大きな違いですが、RR は小さな違いでしかありません。このような RD の違いを effect modification と呼びます。

肺癌

年齢	Incidence rate / 10 ⁵ person-years		RD/ 10 ⁵ PY	RR
	非喫煙者	喫煙者 (21-39 cig/day)		
55-64	40	400	360	10
65-74	80	720	640	9

RD は冠動脈疾患も肺癌も近似しています。しかし、喫煙によって肺癌は 10 倍にまで押し上げられました。病因論的に喫煙は肺癌とより強固な関係にありそうです。

冠動脈疾患、肺癌ともに共通して年齢による RR はあまり変わらないのですが、RD は年齢によって大きくかわっています。このように effect modification は RD に認められれば RR には認められず、逆も真なりです。RD と RR 両方が年齢によって上昇するとす場合は理論上あり得ますが、臨床においては希です。

Attributable risk (rate) & population attributable risk (rate)

Attributable risk は risk difference (RD) と同義です。もしある人がその危険因子に暴露されなければ病気の発生がどれくらい変化するかを示します。

$$AR = RD = [C_{\text{exposed}} - C_{\text{unexposed}}]$$

もしも母親の抗体価が低い場合と高い場合で、前者における子供が下痢を起こす（母乳栄養で育った乳児）リスクが 0.86、後者のリスクが 0.44 であったとしますと、risk difference (attributable risk) は 0.42 となります。乳児の下痢の 42% は母親の抗体価が低かったことによるといえます。

Population attributable risk (PAR) は一般人口からその危険因子が除去された場合、どれくらい病気が減少するかを示しています。

$$PAR = Pr (E) \times [C_{\text{exposed}} - C_{\text{unexposed}}]$$

ある地域で母親の 25% は低い抗体価を示していました。もしこの地域の母親全員の抗体価が高ければ乳児の下痢を $0.25 \times 0.42 = 0.105$ (10.5%) 減らすことができます。

Attributable risk percent (AR%) = etiologic fraction, relative excess incidence, attributable proportion

$$AR\% = [C_{\text{exposed}} - C_{\text{unexposed}}] / C_{\text{exposed}} \times 100 = [RR - 1] / RR \times 100 = [OR - 1] / OR \times 100$$

もう一度喫煙と肺癌、冠動脈疾患の例に戻ります。

$$\text{肺癌(65-74 歳)} \quad AR\% = (9 - 1) / 9 \times 100 = 89\%$$

$$\text{冠動脈疾患(65-74 歳)} \quad AR\% = (1.68 - 1) / 1.68 \times 100 = 40\%$$

「喫煙者（ヘビースモーカー）に発生した肺癌の 90% は喫煙による。一方冠動脈疾患でのそれは 40% である」といえます。この言い方が一般人を説得するのには便利かもしれません。

Population attributable risk percent (PAR%) は exposure が無くなれば除去される病気の率です。PAR% は AR% より低い値をとります。

$$PAR\% = AR\% \times Pr (E=1/D=1)$$

すなわち AR% に病気の人の中で危険因子に暴露された人の割合をかけてやれば算出できます。ですから先の乳児下痢の例では、RR=1.96, 乳児下痢 19 人中 12 人の母親が低い抗体価しかしめしていなかったのもので、

$$APR\% = [(1.96 - 1) / 1.96] \times 12 / 19 = 0.31$$

もしも母親の抗体価が高ければ乳児の下痢は約 1/3 になると予想されます。

Q1：年齢分布が安定した状態で、病気の発生率 incidence が変わらないのに prevalence が減少しているとしたらどのようなことが考えられますか？

治療の改善などにより罹病期間(duration)が短くなったことが考えられます。逆に早期死亡が増えたとしても同じ結果になります。

Q2:ある疾患に対する予防効果を判定するのに prevalence と incidence rate とどちらの方が有用ですか？

Incidence rate は病気のない状態から病気が発生するまでの時間的経過をみており、予防はまさにこの incidence rate を下げることにあります。一方 prevalence の変化は予防と治療の両方に影響されます。

Q3:65 歳 1000 人を対象にマスキングを施行し、100 人に大腸ポリープがみつかりました。さらに 10 年間経過観察したところ新たに 200 人に大腸ポリープがみつかりました。疾患発生を測定してください？

スクリーニングの時点での prevalence は $100/1000 = 0.10$ であり、cumulative incidence は調査開始時点での病気を含まませんので 100 人は除外して考えます。すなわち $200 / 900 = 0.22$ です。

Q4：ある研究者は一般家庭では 100 に対して 5 の割合なのに親が慢性気管支炎であると 20 家庭に 3 の割合で子供が乳児期慢性気管支炎を発症することに気が付きました。疾患はどのように測定しますか？

Cumulative incidence で測定します。患者は最初の 1 年病気をもたず 1 年間に発症しています。期間の比較的短く発症率も極端に少ないわけではないので incidence rate よりも cumulative incidence の方が適当でしょう。

Q5:町の男性の間で最近 5 年間に 270 人が十二指腸潰瘍になりました。町の男性人口は 18500 人であり、調査終了時には 21500 人になっていました。どのように疾患を測定しますか？

町の人口の推移が判らないので正確な値をだすことはできませんが、同じペースで増加したとすると平均をとるべきでしょう。

$$\text{Person-years} = 5 \times [(18500 + 21500) / 2] = 100,000 \text{ person-years}$$
$$\text{Incidence rate} = 270 / 100,000 = 27 / 10,000$$

Q6：偏頭痛を持つ 10 人をいろいろな長さで頭痛の回数を経過観察しています。偏頭痛の回数を person-year で計算してください。

患者	経過観察期間 (月)	偏頭痛の頻度
1	3	0
2	6	2
3	2	1
4	4	2

5	4	1
6	8	2
7	4	1
8	5	3
9	5	2
10	3	0

14 / 44 person-months x 12 months = 3.82 /person-year

Q7:ある騒音の著しい工場で、作業員は耳栓をするように義務付けられていました。しかし500人中100人は耳栓をしていないことが判りました。そこで作業員全員に聴力検査を施行したところ耳栓をしている作業員16人に、耳栓をしていなかった作業員40人に聴力障害がみつかりました。騒音がおこる前全員の聴力は正常でした。

a. Relative risk (RR) はどれくらいでしょうか？

$$RR = (40/100) / (16/400) = 10$$

b. 耳栓をしていないことにより聴力低下を来たす確率は？

$$AR = RD = (40/100) - (16/400) = 0.36$$

c. 耳栓をしていなかった人達が耳栓をしていなかったことにより聴力低下を来たした確率は？

$$AR\% = (C_1 - C_0)/C_1 = (0.4 - 0.04) / 0.4 = 0.90$$

$$AR\% = (RR - 1)/RR = (10 - 1) / 10 = 0.90$$

要するに耳栓をせずに聴力低下を示した作業員の9割は耳栓をしていなかったことが原因であると結論付けられます。しかし1割は工場側にも責任があるのです。

d. もしも全員が耳栓をしていたら聴力低下はどれくらい減らすことができたでしょうか？

$$PAR = Pr(E) \times AR = (100/500) \times AR = 0.072$$

つまり作業員全員を100%としたとき7.2%は減っていたこととなります。つまり36人は聴力低下を起こさなかつたろうと思われます。

e. 作業員全員の中で、耳栓をしていなかったことによる聴力低下の割合は？

$$PAR\% = Pr(E) \times AR\% = (40/56) \times 0.90 = 0.64$$

聴力低下の約2/3は耳栓を全員がしていたら防げた問題といえます。