

Case control study

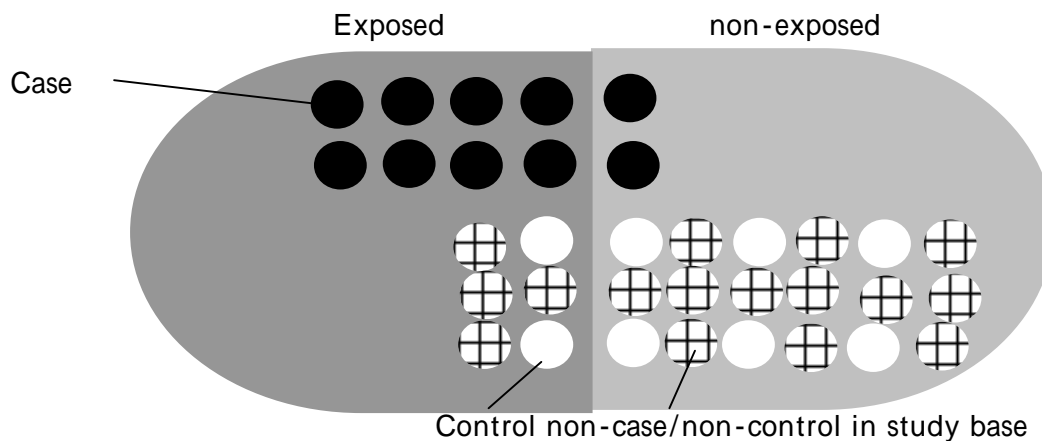
Person-time を用いた case control study

Cohort study においては person time を推奨してきました。Case control study においてもその cohort から case と control を抽出すれば person time で表記できます。まず cases を選定し暴露があったか無かったかで x_1, x_2 に分けます。Control は case が発生したのと同じ study base より person time の形で抽出し、exposure の有無で分配します。その際 exposure とは無関係であることが重要で、そうでないと bias を発生する結果となります (後述)。

	Exposure	No-exposure
Cases	x_1	x_0
Person-time	T_1	T_0

$$IRR = I_1/I_0 = (x_1/T_1) / (x_0/T_0) = (x_1/x_0) / (T_1/T_0)$$

Control は exposure と無関係に study base より抽出されますが、control は study base の一部になります。このことを sampling fraction と呼びます。Exposure 群より f_1 , no-exposure 群より f_0 の割合で control を抽出するとします。先に述べたように control は exposure と無関係に抽出されなくてはなりませんから $f_1 = f_0$ になるはずですが (no selection bias)。もしも study base の半数が検討したい因子に暴露されていたとすると、control も同率で暴露されていなくてはなりません。一般人口の喫煙率が 30% であれば、control の喫煙率も 30% でなくてはならないのです。そして case の喫煙率が 80% だと、exposure と outcome は何か関係がありそうだということになります。



Control の exposed と non-exposed の人数の比率はその study base の exposed と

non-exposed の比率と同じでなければなりません。

$$IRR = I_1/I_0 = (x_1/T_1) / (x_0/T_0) = (x_1/x_0) / [(T_1/T_0) \times (f_1 / f_0)] = (x_1/x_0) / (Y_1/Y_0) = OR$$

この場合、Y は control であり(分母に case は入っておらず) IRR = OR となります。Open cohort study でみた分母には case も含まれるため IRR = OR とするためには「疾患発生が希である」とする仮定が必要となってきます。逆に $1 < IRR < OR$, $1 > IRR > OR$ の関係が成り立ちます。つまり case-control study では IRR = OR だが cohort study では rare assumption が必要であると覚えてください。

Case と control の抽出

くどいようですが、Cases も control 同様 exposure とは無関係に抽出されなくてはなりません。Control は study base の中から直接ランダムに抽出すれば何の問題もありません。しかし case control study の利点として exposure 以外の因子を一致させて選んでくることができます。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りません。例えば喫煙と肺癌の関係を調べるのに同じ病院内で肺癌以外の病気で入院している患者の一部をコントロールにしようと思います。その際日常の飲酒状態でマッチングさせたとしますと、もし飲酒と喫煙が正の相関関係にあれば control 内に study base と比較してより多くの喫煙者が含まれてしまうかもしれませぬ。また control に含めようとしている疾患に喫煙がリスクファクターであるものが含まれていると bias を生じます。よってこのような場合喫煙と関連のある疾患は control から除外されるべきです。極端な話、肺癌患者さんに対する control を抽出するのに非喫煙者ばかり、あるいは喫煙者ばかり選んでしまうと正しく評価できません。喫煙およびそれと相関するファクターに関しては目をつぶって見ないようにしながら control を選んでこなくてはならないということです。また control は病気発生しえた人を選ばなければなりません。例えば子宮癌 control 選択の際子宮筋腫のため子宮を切除している人(at risk)を選んでは fair でないからです。

Open cohort に基く case control study において incidence case ではなく prevalence

case を用いた場合、OR が bias なく正しく測定されるためにはどのような条件が必要でしょうか？もちろん control は exposure に関係なく適切に抽出されたとします。

もし罹病期間が同じであれば prevalence を用いることもできます。 $D_1 = D_0$

$$OR = I_1/I_0 = I_1D / I_0D = [(P_1/1 - P_1)] / [(P_0/1 - P_0)]$$

Primary vs. Secondary study base

まず base population を設定してその中から case を同定する場合を primary study base とよびます。nested case control study やある地域で発生する病気を調べるような場合はこれにあたります。台帳などを基に population 自体を control とします。例えば random digit dialing という方法がありますが、これも bias が入らないわけではありません。例えば電話を持たない家は喫煙が多く、社会的地位が低いかもしれません。このような場合電話を持たない case を除外すればこの点に関する bias は解消されます(restriction)。一方 case が最初に存在しそれに基づいて study base を同定するものを secondary study base (case-defined study base)と呼びます。病院ベースの case control study はこれに当たります。しかし紹介システムは複雑であり study base 同定はしばしば困難です。