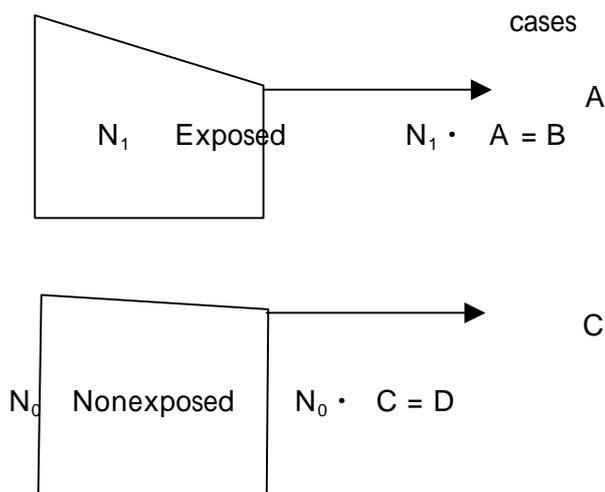


# Case-Control Studies Nested in Closed Cohorts

Nested case control study とは、非常によく調査された集団(closed cohort)から cases と control が発生した場合に行なうことができます。通常の case control study では study population が不明確であったり、case と control で異なっていることさえありますが、case-control study ではあたかも cohort study の如く study population が明確です。



Nested case control study とするには、ある集団で経過観察中に outcome を持った (case)全てと control 全てを把握している場合に行なうことができます。現実問題として、もともと cohort study を行なっていたのだけれども、可能性のある confounder をフォローできていなかったために、closed cohort の中から control を選んで case control study を行ないます。データはインタビューやカルテによって収集されます。また confounder でなく exposure を再検討する場合にも有効です。これは最初の目的とは異なった目的が新たに発生した場合などに便利です。例えば、もともと 4480 人の高血圧患者を治療法を変えて経過観察していたのですが、この中から癌患者が発生しました。幸い全員の血清を保存してあり、血清 selenium と発癌の関連を調査しました (Lancet 16:130-4,1983)。もちろんここで全員の血清 selenium を測定してもよいのですが、コストと血清を保存するため case control study を行ないます。今後多くの人の遺伝子を保存しておいて、何か outcome が発生してからその case とこれに対する control の遺伝子を保存タンク(nested cohort)の中から選んでできます。DNA chip などを使った検査はとても高価なので、nested case-control study が適しています。ただし遺伝子は基本的に生涯変化するものでもありませんので、必ずしもストックすることもないかもしれませんが。。

# Impact of breast feeding on admission for pneumonia during postneonatal period in Brazil: nested case-control study

Cesar JA, et al. BMJ 318:1316-1320,1999

## 要旨

目的：母乳栄養が乳児肺炎を予防するか？その予防効果は年齢によるか？について検討しました

デザイン：nested case control study

場所：Pelotas, southern Brazil

対象：152 人の生後 28 - 364 日で肺炎のため入院した 152 人の乳児

コントロールは population based case-control study により 2391 人選びました。

方法：オッズ比

結果：母乳栄養以外の乳児は母乳栄養乳児と比較して 17 倍肺炎に罹患しやすく、3 ヶ月未満に限ってみると 61 倍であり、それ以降の 10 倍と比較して大きな相違がありました。固形食では全ての乳児で 13.4 倍肺炎のリスクを上げ、3 ヶ月未満に限ってみれば 175 倍でした。

結論：母乳影響は乳児を肺炎から守り、特に乳児期初期重要であると思われました。

## Study questions

- 1 . Cases と control のもととなった source population は何ですか？また cohort の経過観察期間はどれくらいですか？
- 2 . Cases はどのように選ばれましたか？
- 3 . Control はどのように選ばれましたか？
- 4 . Table 2 から、社会的クラス(social class)が cases 側により多く分布しているようにみえませんか？
- 5 . どのような環境下において、社会クラスの違いが本研究のように confounding となりえるでしょうか？
- 6 . Table 3 の栄養のタイプと Table 4 のデータの関連について説明してください？ Table 4 を加えることによって何がいえませんか？

## 方法

### Study population

1993年、ブラジル南部のPelotasの市病院で出産した全ての女性に対してインタビューを行ないました。市の出産の99%は市の病院でした。655人を抽出し生後1,3ヶ月時に家庭訪問しています。他生後1ヶ月時5256人(市全体の99.1%)、3ヶ月時5214人(市全体の98.3%)に対してインタビューを行なっています。さらに生後6ヶ月時にもインタビューを行なっています。

### Cases の定義

1993年に生まれた小児で生後28日から364日までの間に市の病院に肺炎のため入院したものをcaseとしました。2人の独立した小児科医が審査官として可能な限りの情報を入院カルテより収集しました。肺炎の診断は臨床症状(呼吸促拍、呼吸音の異常)検査、レントゲン所見によってなされました。2人の審査官の意見が異なった場合には、もう1人に診断を下してもらいました。

### Control の定義

Controlはcohortの中から選びました。Caseの中で生後28-89日のものに対しcontrolは生後1ヶ月の家庭訪問から、生後90-179日のものに対しcontrolは生後3ヶ月の家庭訪問から、生後180-364日のものに対しcontrolは生後6ヶ月の家庭訪問から抽出しています。よって個々に年齢をマッチさせてはいませんが、年齢層は一致しています。一方この方法だと最初controlであり後にcaseになる場合もありました。

### サンプルサイズ

controlに25%のexposuresが存在しORを2.0を検出すると仮定し、 $\alpha$  error 0.05, power 80%と設定したところ、143 cases, 572 controlが必要であると算出されました。

### 統計学的解析

OR 95% CI: chi-square test

Adjustment: uncontrolled multiple logistic regression

## 結果

調査対象とした 5304 の乳児のうち 152 人 (2.9%) が乳児期肺炎のため病院に入院しました。250 種類の変数のうち、社会クラス、世帯収入、母親の学校教育年数、生活費に占める食料費の割合(parity)、妊娠中の体重増加が乳児肺炎による入院と関係していました(table 1 略)。

Table 2 はこれらの因子にしたがって cases と control の分布の違いを示したものです。6-11ヶ月が control において多くなっていますが、これは解析に影響しませんでした。

Table 2. Social class distribution in cases and controls (他は省略)

	cases	controls
Bourgeois	26 (17)	439 (21)
Atypical proletariat	39 (26)	1034 (43)
Typical proletariat	42 (27)	518 (22)
subproletariat	30 (20)	174 (7)

Table 3 は母乳栄養、混合、人工栄養の乳児の乳児期肺炎による入院率の相対比を示しています。混合あるいは人工栄養は母乳栄養に対してそれぞれ 3.8 倍、16.7 倍の relative risk をもっていました。

Table 3

	肺炎	コントロール	crude	adjusted
母乳栄養	9	779	1.0	1.0
混合栄養	23	563	4.2 (2.1 – 9.9)	3.8 (1.7-8.9)
人工栄養	120	1049	4.5 (1.4 – 14.5)	1.3 (0.3 – 4.9)
p 値			<0.001	<0.001

Table 4

年齢により層化したものです。年齢と栄養方法の違いにより肺炎発生頻度に差があるのが理解できます。

1 – 2.9 ヶ月	肺炎	コントロール	crude	adjusted
母乳栄養	5	392	1.0	1.0
混合栄養	7	169	3.2 (1.0 – 10.4)	2.9 (0.8 - 10.5)
人工栄養	35	88	31.2 (11.9 – 81.9)	61.1 (19.0 – 195.5)
p 値			<0.01	<0.01

3-5.9ヶ月	肺炎	コントロール	crude	adjusted
母乳栄養	3	212	1.0	1.0
混合栄養	11	164	4.7	3.4
			(1.3-17.2)	(0.9-13.5)
人工栄養	48	268	12.5	10.1
			(3.8-40.8)	(2.8-36.2)
p値			<0.01	<0.01
6-11.9ヶ月	肺炎	コントロール	crude	adjusted
母乳栄養	1	175	1.0	1.0
混合栄養	5	230	3.8	3.7
			(0.4-32.9)	(0.4-33.8)
人工栄養	37	693	9.3	9.2
			(1.3-68.6)	(1.2-69.7)
p値			<0.001	<0.01

## 考察

### デザイン上の問題点

case-control study はいくつかのバイアスにより影響されます。Reverse causality bias-呼吸器に問題があったが故に母乳栄養から他の栄養法に変更しなくてはならなかった-については、入院2ヶ月前に呼吸器感染症のために母乳を中止した場合は母乳栄養群に含めて考えました。Case に関しては過去にさかのぼって栄養方法を尋ねているため recall bias を生じ易いのですが、control は1ヶ月、3ヶ月、6ヶ月の時点での栄養方法を尋ねているので recall bias は生じにくいと考えられます。kappa index = 0.81 であり、比較的高い一致率、すなわち recall bias は小さいと考えられました。

\*\*\* Kappa Statistic \*\*\*

US ナース研究で、1回目サーベイと2回目サーベイにおける回答の一致率について検討しました。

第一回サーベイ	第二回サーベイ		合計
	週に1回以上牛肉を食べる	週に1回未満牛肉を食べる	
週に1回以上牛肉を食べる	136	92	228
週に1回未満牛肉を食べる	69	240	309
合計	205	332	537

上の表より、第一回サーベイにて週に一回以上牛肉を食べる人は  $a_1 = 228/537 = 0.425$ , 週に一回未満牛肉を食べる人は  $a_2 = 309/537 = 0.575$ , 第二回サーベイにて週に一回以上

牛肉を食べる人は $b_1 = 205/537 = 0.382$ , 週に二回未満牛肉を食べる人は $b_2 = 332/537 = 0.618$ 。第一回目と第二回目のサーベイにおいて回答一致期待率は  $Pe = (0.425 \times 0.382) + (0.575 \times 0.618) = 0.518$  と計算されます。一方第一回と第二回サーベイで実際に回答が一致したのは $P_o = (136 + 240)/537 = 0.700$  です。

$$K = [P_o - Pe]/[1 - Pe] = [0.700 - 0.518] / [1 - 0.518] = 0.182/0.482 = 0.378$$

更に standard error は以下の公式により求められます。

$$se(k) = \{1/[n(1-Pe)^2]\} \times \{Pe + Pe^2 - [a_i b_i (a_i + b_i)]\} = \{1/[537(1-0.518)^2]\} \times \{0.518 + 0.518^2 - 0.555\} = 0.0430$$

統計学的有意差は以下によって求められます。

$$Z = 0.378 / 0.0430 = 8.8, p < 0.001$$

よって統計学的に有意であるといえます。

Landis & Koch が kappa の評価について以下のように提案しています (Biometrics 33:2385-2391,1977)。

$K > 0.75$  : excellent reproducibility

0.4      $k$      0.75 : good reproducibility

0.0      $k < 0.4$  : marginal reproducibility

しかしながら、2つのデータのうち1つが golden standard であれば specificity と sensitivity を考えた方が kappa より適切です。

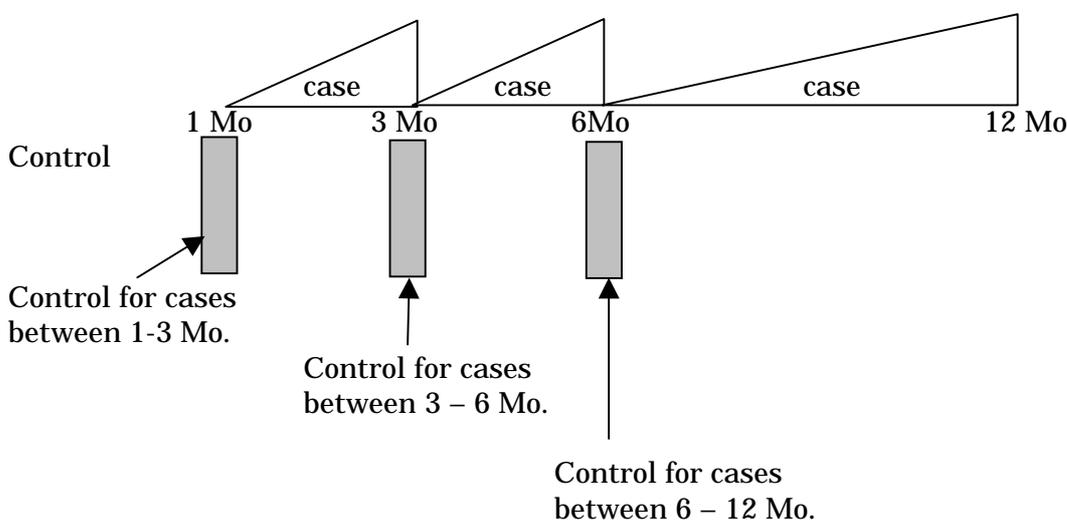
## Study questions に対する解答

完全に遂行された case-control study は cohort study の特別な形とみますこともできます。通常 case-control study においては recall bias が問題となりますが、cohort study ではその可能性は低くなります。例えば乳癌と経口避妊薬の使用を調査しているとします。Case に経口避妊薬の使用を尋ねれば真剣に思い出そうとしてくれるでしょうが、control は思い出そうとする努力を払おうとしないかもしれません。Exposure は case および control の記憶に頼る場合が多くなるため、プラスをマイナス、マイナスをプラスに分類してしまう、所謂 misclassification が発生しやすく bias を生じます。本研究では、case と control で異なった方法で情報収集しています。Case に対しては入院後、control に対しては家庭訪問です。しかしながら質問は比較的最近の出来事に対してであり、recall bias は生じにくいと考えられます。

1 . Cases と control のもととなった source population は何ですか？また cohort の経過観察期間はどれくらいですか？

- (1) Pelotas の市病院で 1993 年に生まれ、
- (2) 母親はその市の都市部に住み
- (3) 生後 28 日以上生存したもの

本研究は以下のような 3 つの連続した cohort study であると考えた方が理解しやすいと思います。



本研究では都市部に生まれたベビーだけに絞ったことにより、著者らはほぼ 100% のベビーをカバー（正確に把握する）することができました。もしも郊外の家族にまで広げると、source population の辺縁が曖昧になり、validity を下げることとなります。

2 . Cases はどのように選ばれましたか？

肺炎症例は日々の市病院サーベイランスに基いて行なわれました。肺炎の診断自体は臨床症状と可能であれば検査も加え判断されました。特に本研究では肺炎にて入院した患

者さんをケースとして選んでいます。人工栄養で育てている母子の方が病院を受診しやすいなどの傾向があると問題です。また本当は肺炎があるのに病院を受診していない場合もあります。

3 . Control はどのように選ばれましたか？

655 人の乳児が 1, 3, 6 ヶ月の時点でランダムに選ばれました。調査中ファンドが増えたので、結局 1 ヶ月で 655 回、3 ヶ月で 655 回、6 ヶ月で 1144 回、合計 2454 回の訪問を行なっています。そのうち 2391 回の訪問が成功しています。これらの訪問によって得られたデータの中から case と出産日の同じ乳児が control として選ばれています。

4 . Table 2 から、社会的クラス(social class)が cases 側により多く分布しているようにみえませんか？

Case は失業中や職不定に多く、また手に職のある仕事(atypical proletariat)や比較的高い社会地位(Bourgeois)は control において多くみられました。

5 . どのような環境下において、社会クラスの違いが本研究のように confounding となりえるでしょうか？

もしも社会クラスの相違が母乳栄養に関係し、かつ肺炎のリスクファクターと関連していれば、社会クラスは confounder となりえます。Table 1 から subproletariat では肺炎のリスクが 3.5 倍(95% CI 2.15 – 5.70), typical proletariat 1.78 (95% CI 1.14 – 2.76)と高くなっています。

社会クラスが confounder であるとするためには、さらに栄養方法との関連をみなくてはなりません。その際 case あるいは control どちらにおいて調べるべきでしょうか？

Control において調べるべきです。何故なら control がより正確に source population を反映していると考えられるからです。

6 . Table 3 の栄養のタイプと Table 4 のデータの関連について説明してください？ Table 4 を加えることによって何がいえませんか？

Table 3 では栄養方法によって母乳、混合、人工と分けていますが、Table 4 では、これを更に年齢によって層化しています。よって Table 4 の各年齢層の case, control の合計人数は Table 3 に当たります。Table 4 を加えることにより effect modification, すなわち幼弱乳児で栄養法と肺炎の関連が強く、年長乳児で関連が弱いことが示唆されました。

#### コメント

母乳が新生児、乳児の免疫を助けることは有名な話で誰もが知っています。ですから乳児肺炎を減らす可能性は十分あります。しかしその真偽について、あるいは人工栄養だと肺炎発症率がどれくらい上昇するのかなど誰も答えられません。この研究はそういった意味で我々の盲点をつく研究だったかもしれません。研究方法もよく考えられていますが、決して難しいものではなく我々も十分実行可能です。そう考えると我々の周り

には調査すべき題材がたくさんあるのではないのでしょうか？また母親に母乳栄養の重要性を説くときも数値で示せばとても説得力があるのではないのでしょうか？この研究では EBM のおもしろさが集約されているように思います。