

Special Contribution

Basic Epidemiology

—Methods and Their Application to Epidemiology on Cancer and Radiation (4)

Suminori Akiba

Department of Epidemiology and Preventive Medicine, Kagoshima University Graduate School of Medical and Dental Sciences
8-35-1, Sakuragaoka, Kagoshima 890-8544, Japan

Received 17 February 2015; revised 11 June 2015; accepted 18 June 2015

XI. Confounding

Confounding is distortion of the estimated effect of an exposure on an outcome, caused by the presence of an extraneous factor associated both with the exposure and the outcome (DE4)².

1. A simple example

Figure 4-1 shows the heights of boys and girls. Based on this figure, you may conclude that the heights of boys are the same as those of girls. Figure 4-2 shows the same data according to age. As can be seen in this figure, boys are taller than girls at the same age. In this example, age distorts the sex difference of height, acting as a confounder.

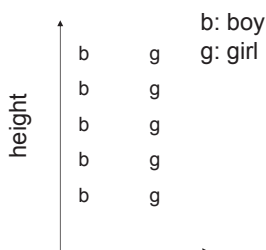


Fig. 4-1. Heights of boys and girls.

Suminori Akiba: Department of Epidemiology and Preventive Medicine, Kagoshima University Graduate School of Medical and Dental Sciences
8-35-1, Sakuragaoka, Kagoshima 890-8544, Japan
E-mail: akiba@m.kufm.kagoshima-u.ac.jp

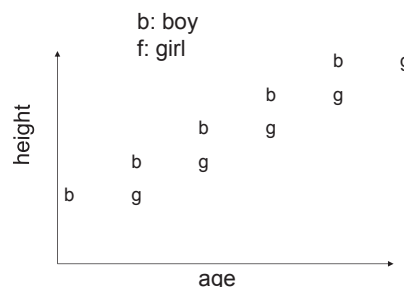


Fig. 4-2. Height of boys and girls according to age.

2. Numerical examples of confounding

Example 1

Table 12-1 summarizes the results of a fictitious cohort study. In this example, the RR of coronary heart disease (CHD) in relation to radiation exposure is 1 among smokers. Among non-smokers, the RR is also 1. In other words, neither among smokers nor among non-smokers, radiation is related to CHD risk. However, if smoking is ignored, the RR is 3.3. In this example, the confounding effect of smoking distorts the relationship between CHD risk and radiation exposure. Regarding this data set, the following two points should be noted: i) 90% of the exposed is smokers and only 10% of the unexposed is smokers (smoking is related to radiation); ii) the RR of CHD for smoking in this example is 5 as shown in Table 12-2. Figure 5 presents a schematic explanation for this example. Although radiation is not related to CHD risk (in this fictitious data), it appears to cause CHD because radiation happened to be related to smoking, which is a risk factor of CHD.

Table 12-1. Confounding by smoking (RR of CHD for smoking = 5)

Smoking	Radiation	N of subjects	CHD cases	RR of CHD for radiation exposure
Yes	Yes	900	450	$RR_{smk=yes} = 1$
	No	100	50	
No	Yes	100	10	$RR_{smk=no} = 1$
	No	900	90	
Total	Yes	1,000	460	$RR_{smk-total} = 3.3$
	No	1,000	140	

Table 12-2. RR of CHD for smoking

Smoking	N of subjects	CHD cases	RR of CHD for smoking
Yes	1,000	500	$RR = 5$
No	1,000	100	

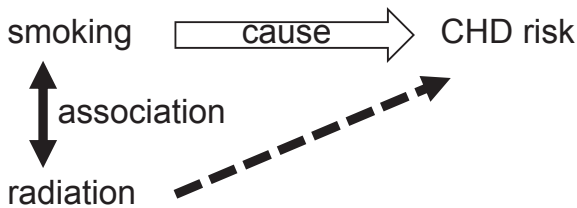


Fig. 5. An example of confounding (Tables 12 and 13).

Example 2

In Table 13-1, if smoking is ignored, the RR is 1.4 (see the total). Note that the RR of smoking for CHD in this example is 10 as shown in Table 13-2. In this example, smoking is a stronger risk factor of CHD than in the previous example (in Table 12-2). Note, however, smokers make up 60% of the exposed subjects and 40% of the unexposed; the association of smoking with radiation exposure is weaker in this example. As a consequence, there is a smaller deviation of $RR_{smk-total} (=1.4)$ from the true RR (=1) when compared to Table 12-1 ($RR=3.3$).

Table 13-1. RR of CHD for radiation exposure is confounded by smoking

Smoking	Radiation	N of subjects	CHD cases	RR of CHD for radiation exposure
Yes	Yes	600	60	$RR_{smk=yes} = 1$
	No	400	40	
No	Yes	400	4	$RR_{smk=no} = 1$
	No	600	6	
Total	Yes	1,000	64	$RR_{smk-total} = 1.4$
	No	1,000	46	

Table 13-2. RR of CHD for smoking

Smoking	N of subjects	CHD cases	RR of CHD for smoking
Yes	1,000	100	$RR = 10$
No	1,000	10	

Example 3.

In the examples of confounding described above, radiation exposure was not related to CHD risk when analysis was conducted separately for smokers and non-smokers. However, the absence of such an association is not the condition necessary for confounding. In the example shown in Table 14, radiation exposure is related to CHD risk and the RR is 2 when analysis was conducted separately for smokers and non-smokers. However, when smokers and non-smokers were combined, the RR is 4.1, which is approximately 2-fold larger than the RR obtained from analysis conducted separately for smokers and non-smokers.

In some situations, a confounded RR can be less than 1 whereas the true RR is larger than 1.

Table 14. Confounding by smoking (RR of CHD for smoking = 5)

Smoking	Radiation	N of subjects	CHD cases	RR of CHD for radiation exposure
Yes	Yes	900	450	$RR_{smk=yes} = 2$
	No	100	25	
No	Yes	100	20	$RR_{smk=no} = 2$
	No	900	90	
Total	Yes	1,000	460	$RR_{smk-total} = 4.1$
	No	1,000	140	

3. Criteria for confounding

Rothman and Greenland defines the criteria for confounding as follows (p123-125 in ME2[®]). Confounding is caused by a factor satisfying the following three conditions:

- 1) A confounding factor must be a risk factor for the disease. – smoking is a risk of CHD in Table 13-1.
- 2) A confounding factor must be associated with the exposure under study in the source population (the population at risk from which the cases are derived). – radiation is related to smoking in Table 13-1.
- 3) A confounding factor must not be affected by the exposure or disease. In particular, it cannot be an intermediate step in the causal path between the exposure and the disease.

In the example shown in Figure 6, the elevated blood pressure is a risk factor for cerebro-vascular attack (the disease of interest), and is also related to exposure (high

high salt intake – cause -> elevated blood pressure – causes -> CVA

Fig. 6. High salt intake causes cerebro-vascular attack (CVA) through elevated blood pressure.

salt intake). Nevertheless, it cannot be considered a purely confounding factor, since the effect of high salt intake is mediated through the effect of blood pressure. Note, however, the causal path may not be single. Therefore, even if a factor is an intermediate step in a causal path, it may not be so in another causal path. If that is the case, this factor can cause confounding.

4. Is this confounding?

In the example shown in Table 15, the RR is 1 among smokers and is 0.5 among non-smokers. In this example, the RRs are modified by smoking status. Therefore, you should present the RR of each smoking category. It is not advisable to present a RR combining different smoking categories in this case.

Table 15. The effects of smoking and radiation on CHD risk

Smoking	Radiation	N of subjects	CHD cases	RR of CHD in relation to radiation exposure
Yes	Yes	900	450	RR = 1
	No	100	50	
No	Yes	100	10	
	No	900	180	
				RR = 0.5
Total	Yes	1,000	460	RR = 2
	No	1,000	230	

5. Treatment of confounding

There are two methods to eliminate (or reduce the magnitude of) confounding. They are to use multivariate analysis and to conduct stratification.

5.1. Treatment by statistical modeling

In the fictitious data shown in Table 16, the difference between average blood cholesterol levels in Populations #1 has lower blood cholesterol levels than Population #2, and the difference is 10 mg/dl (see the total in the table). This difference can be referred to as the “crude difference”. Figure 7 illustrates the data presented in Table 16. In this figure, the data points at ages 30, 40, 50 and 60 in Population #1 are connected by a single line; and the data points at ages 60, 70, 80, and 90 in Population #2 are also connected by another single line. In Populations #1 and #2, the blood cholesterol level increases with age. Since these two lines are parallel, the magnitude of age-dependence

(of blood cholesterol levels) in the two populations are the same. Therefore, the difference of blood cholesterol levels between the two populations is not age dependent (is not modified by age), and the comparison can be made at any given age. For example, the difference at age 60 is 50 mg/dl (note that Population #1 has higher blood cholesterol levels than Population #2 in this comparison). In this example, the comparison of cholesterol levels in the two populations is distorted by age. Therefore, a proper comparison can only be made after taking age into account.

Table 16. Average blood cholesterol levels in Populations #1 and #2

Age (years)	Blood cholesterol concentration (mg/dl)	
	Population #1	Population #2
30	180	
40	200	
50	220	
60	240	190
70		210
80		230
90		250
Total	Mean = 210	Mean = 220

The data shown in Figure 7 can be expressed in the following multivariate regression model:

$$Y = 70 + 50X_1 + 2X_2$$

Y : blood cholesterol level

X₁ = 1 for Population #1, X₁ = 0 for Population #2

X₂ = age (in years)

The coefficient of X₁ (= 50 mg/dl) gives the difference of blood cholesterol level in Populations #1 and #2, taking age into account. This is called the age-adjusted value and is different from the crude value (=10 mg/dl) in this dataset.

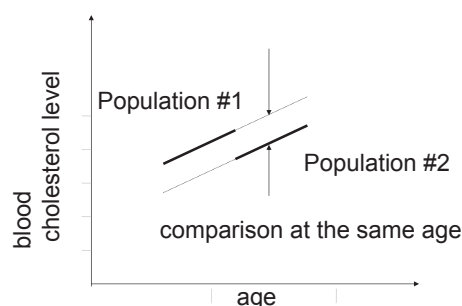


Fig. 7. Blood cholesterol and age in Populations #1 and #2.

5.2. Treatment of confounding by stratification

Suppose that radiation is not related to CHD risk. Therefore, OR=1 in Table 17 is the true value. How can we calculate a combined estimate of OR_{smk=yes} and OR_{smk=no}?

Table 17. Confounding effect of smoking on the relationship between CHD risk and radiation exposure

Smoking	Radiation	CHD cases	Controls	OR of CHD for radiation exposure
Yes	Yes	45	90	OR _{smk=yes} = 1
	No	5	10	
No	Yes	10	10	OR _{smk=no} = 1
	No	90	90	
Total	Yes	55	100	OR _{smk-total} = 0.6
	No	95	100	

Table 18. Confounding effect of smoking on the relationship between CHD risk and radiation exposure

Smoking	Radiation	CHD cases	Controls	OR of CHD for radiation exposure
Yes (= 1)	Yes	a ₁	b ₁	OR _{smk=yes} = 1
	No	c ₁	d ₁	
No (= 2)	Yes	a ₂	b ₂	OR _{smk=no} = 1
	No	c ₂	d ₂	

Table 19. Theoretical examples of confounding

RR of disease (cancer) by confounder is assumed to be 2.0 the proportion of the subjects with confounder (smoking) in the unexposed group=20% (odds of confounder among unexposed is assumed to be 2/8)	the proportion of the subjects with confounder in the exposed group				
	25%	27%	29%	31%	33%
crude RR / adjusted RR	1.042	1.058	1.075	1.092	1.108

For example, $1.042 = (2 \times 0.25 / (1 - 0.25) + 1) / (2 \times 2/8 + 1) / (0.25 / (1 - 0.25) + 1) \times (2/8 + 1)$

The answer is Mantel-Haenszel OR.

In Table 18, Mantel-Haenszel OR (OR_{MH}) can be calculated as follows:

$$OR_{MH} = \frac{\sum (a_i d_i / T_i)}{\sum (b_i c_i / T_i)}$$

where $T_i = a_i + b_i + c_i + d_i$.

In the data shown in Table 17, it can be calculated as follows:

$$OR_{MH} = \left\{ \frac{45 \times 10}{150} + \frac{10 \times 90}{200} \right\} / \left\{ \frac{90 \times 5}{150} + \frac{10 \times 90}{200} \right\} = 1.$$

6. Theoretical example of confounding

Many risk factors, such as smoking, are much more strongly associated with cancer risk than low-dose radiation exposure. As a result, even a weak association of radiation exposure with such a factor can confound the relationship between radiation exposure and cancer risk.

A numerical examples may be useful for understanding the magnitude of confounding. In the example shown in Table 19, the RR of cancer by a confounder is assumed to be 2, which is an approximate RR value observed for the relationship between smoking and solid cancer risk in many countries. In addition, the proportion of smokers among the unexposed group is assumed to be 20%. This table shows the ratios between crude RRs of cancer (the crude RR is the ratio between the proportions of cancer cases among the radiation-exposed and unexposed groups) and adjusted RRs of cancer for the different proportions of smokers among the exposed. When smokers account for 25% of the exposed group, the ratio between the crude RR and the adjusted RR is 1.042. This result indicates that even if the true RR is 1, a crude RR (without adjustment for confounding) can be 1.042. This RR corresponds to an ERR of 0.042, which corresponds to 120 mGy among men and 72 mGy among women if the ERR/Gy obtained from the LSS is applied to this population. As shown in this example, even a weak association between radiation and smoking cannot be ignored when examining the solid cancer risk associated with low-dose radiation.

References

- WHO. Health topics: Epidemiology [homepage on the internet.] Available from: <http://www.who.int/topics/epidemiology/en/>
- Last J (2001) Dictionary of Epidemiology, 4th edition. International Epidemiological Association.
- Porta M (2008) A Dictionary of Epidemiology, 5th edition. Oxford University Press.
- National Cancer Institute. NCI Dictionary of Cancer Terms: observational study [homepage on the internet.] Available from: <http://www.cancer.gov/dictionary>
- Brown JR and Thornton JL (1957) Percivall Pott (1714-1788) and Chimney Sweepers' Cancer of the Scrotum. Br J Ind Med. 14(1):68-70.
- Lind J (1753) A Treatise of the Scurvy in Three Parts. Containing an inquiry into the Nature, Causes and Cure of that Disease, together with a Critical and Chronological View of what has been published on the subject. Millar, London.
- Hempel S (2013) John Snow. The Lancet [serial online]. 381(9874):1269-1270. Available from: [http://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(13\)60830-2/fulltext](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(13)60830-2/fulltext)
- Rothman KJ and Greenland S (1998) Modern Epidemiology second edition. Lippincot, Wiliamd and Wilkins.
- Gardner MJ and Winter PD (1984) Mortality in Cumberland

- during 1959-78 with reference to cancer in young people around Windscale, *Lancet*, i, 217.
10. Gardner MJ, Hall AJ, Downes S, Terrell JD (1987) Follow up study of children born to mothers resident in Seascale, West Cumbria (birth cohort), *Br Med J*. 295:822-827.
 11. Gardner MJ, Hall AJ, Downes S, Terrell JD (1987) Follow up study of children born elsewhere but attending schools in Seascale, West Cumbria (schools cohort), *Br Med J*. 295:819-822.
 12. Lagakos SW, Mosteller F (1986) Assigned shares in compensation for radiation-related cancers. *Risk Anal*. 6(3):345-357.
 13. Kreyszig E (1970) Introductory mathematical statistics. Principles and methods. John Wiley & Sons. Inc. New York.
 14. Miettinen OS (1985) Theoretical Epidemiology. Wiley, New York.
 15. Rothman KJ and Greenland S (2008) Modern Epidemiology third edition. Lippincot, Williams and Wilkins.
 16. Beebe G and Usagawa M (1968) The major ABCC samples. TR12-68. Atomic Bomb Casualty Commission.
 17. Active Research Protocols [Internet]. Radiation homepage on the Effects Research Foundation. Available from: http://www.rerf.or.jp/programs/rp_e/index.html
 18. Nair MK, et al. (1999) Population study in the high natural background radiation area in Kerala, India. *Radiat Res*. 152:S145-148.
 19. Azizova TV, et al. (2010) Cerebrovascular diseases in the cohort of workers first employed at Mayak PA in 1948-1958. *Radiat Res* 174: 851-864.
 20. Krestinina LY, et al. (2007) Solid cancer incidence and low-dose-rate radiation exposures in the Techa River cohort: 1956-2002. *Int J Epidemiol*. 36:1038-46.
 21. Bauer S, et al. (2005) Radiation exposure due to local fallout from Soviet atmospheric nuclear weapons testing in Kazakhstan: solid cancer mortality in the Semipalatinsk historical cohort, 1960-1999. *Radiat Res*. 164:409-419.
 22. Tung CJ, et al. (1998) Dose reconstruction for residents living in ⁶⁰Co-contaminated rebar buildings. *Health Phys*. 74:707-713.
 23. Chen W.L. (2002) Radiation surveys and dose equivalent assessments for ⁶⁰Co-contaminated rebar buildings. *Appl Radiat Isot*. 56:901-906.
 24. Hwang SL, et al. (2006) Cancer risks in a population with prolonged low dose-rate gamma-radiation exposure in radiocontaminated buildings, 1983-2002. *Int J Radiat Biol*. 82:849-58.
 25. Azizova TV, et al. (2010) Cardiovascular diseases in the cohort of workers first employed at Mayak PA in 1948-1958. *Radiat Res* 174:155-168.
 26. Ivanov VK, et al. (2000) Radiation-epidemiological analysis of incidence of non-cancer diseases among the Chernobyl liquidators. *Health Phys* 78:495-501.
 27. Russian national medical and domestic registry [homepage on the Internet]. National radiation and epidemiological registry. Available from: http://www.nrnr.ru/rgmndr_eng.html
 28. Hosoda Y, Kuba M, Miyake T, et al. (1997) First Analysis of Mortality of Nuclear Industry Workers in Japan, 1986-1992. *Journal of Health Physics* 32:173-184.
 29. UNSCEAR (2006) Report to the General Assembly. Effects of Ionizing Radiation. Annex A: Epidemiological studies of radiation carcinogenesis. United Nations Scientific Committee on the Effects of Atomic Radiation, United Nations, New York.
 30. Clark GN, Briggs A, Cooke AM (2005) A History of the Royal College of Physicians of London, Oxford University Press, ISBN 0-19-925334-X.
 31. Gardner MJ, Snee MP, Hall AJ, et al. (1990) Results of case-control study of leukemia and lymphoma among young people near Sellafield nuclear plant in West Cumbria, *BMJ*, 300, 423-429.
 32. Doll R, Evans J, Darby SC. (1994) Paternal exposure not to blame, *Commentary, Nature* 367:678-680.
 33. HSE investigation of leukemia and other cancers in the children of male workers at Sellafield. HSE BOOKS 1994.
 34. Pobel D, Viel JF. (1997) Case-control study of leukaemia among young people near La Hague nuclear reprocessing plant: the environmental hypothesis revisited. *BMJ*. 314:101-106.
 35. Miettinen O. (1976) Estimability and estimation in case-referent studies. *Am J Epidemiol*. 103:226-235.
 36. Kendall GM, Little MP, Wakeford R, et al. (2013). A record-based case-control study of natural background radiation and the incidence of childhood leukaemia and other cancers in Great Britain during 1980-2006. *Leukemia*. 27(1):3-9.
 37. Sato Y, Akiba S, Kubo O, et al. (2011) A case-case study of mobile phone use and acoustic neuroma risk in Japan. *Bioelectromagnetics*. 32(2):85-93.
 38. Walker AM (1982) Anamorphic analysis: sampling and estimation for covariate effects when both exposure and disease are known. *Biometrics* 38:1025-1032.
 39. White JE (1982) A two stage design for the study of the relationship between a rare exposure and a rare disease. *Am J Epidemiol* 115:119-128.
 40. Langholz B & Clayton D (1994) Sampling strategies in nested case-control studies, *Environ Health Perspect* 102 (Suppl. 8); 47-51.
 41. Ebi KL, Zaffanella LE, Greenland S (1999) Application of the case-specular method to two studies of wire codes and childhood cancers. *Epidemiology* 10(4):398-404.
 42. Cohen B L (1995) Test of the linear no-threshold theory of radiation carcinogenesis for inhaled radon decay products. *Health Phys*. 68:157-174.
 43. Darby S and Doll R (2000) Reply to 'Explaining the lung cancer versus radon exposure data for USA counties' *J Radiol Prot* 20:221-222.
 44. Ahmad OB, et al. (2001) Age standardization of rates: a new WHO standard. GPE Discussion Paper Series: No.31 EIP/GPE/EBD World Health Organization.
 45. Brenner AV, et al. (2011) I-131 dose response for incident thyroid cancers in Ukraine related to the Chernobyl accident. *Environ Health Perspect* 119:933-939.
 46. Tao Z, et al. (2012) Cancer and non-cancer mortality among Inhabitants in the high background radiation area of Yangjiang, China (1979-1998). *Health Phys*. 102:173-181.
 47. Iwasaki T, et al. (2003) Second analysis of mortality of nuclear industry workers in Japan, 1986-1997. *Radiat Res*. 159:228-238.
 48. Boice JD Jr, et al. (2006) Mortality among Rocketdyne workers who tested rocket engines, 1948-1999. *J Occup Environ Med*. 48:1070-1092.
 49. Mei-Shu Lai (2010) Cancer monitoring and control planning in Taiwan. JACR Monograph No 16. Japanese Association of Cancer Registries. Tokyo.
 50. Cornfield J (1951). A method for estimating comparative rates from clinical data. Applications to cancer of the lung, breast and cervix. *J. Nat. Cancer Inst*. 11 1269-1275.
 51. Prentice R L (1986). A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika* 73:1-11.
 52. Blot WJ and Day NE (1979). Synergism and interaction: are they equivalent? [letter] *American Journal of Epidemiology* 110:99-100.

53. Rothman KJ, Greenland S, Walker AM (1980) Concepts of interaction. *Am J Epidemiol.* 112:467-470.
54. Preston DL, et al. Solid cancer incidence in atomic bomb survivors: 1958-1998. *Radiat Res.* 2007; 168:1-64.

Appendix

Radiation (= X)	Exposed (= x1)	Exposed (= x1)	Unexposed (= x0)	Unexposed (= x0)
Confounder (= C)	Present (= c1)	Absent (= c0)	Present (= c1)	Absent (= c0)
The number of disease	Dx1c1	Dx1c0	Dx0c1	Dx0c0
The number of subjects	Px1c1	Px1c0	Px0c1	Px0c0

The crude relative risk (RR) of radiation for the disease is expressed as RR_{ofDbyX_crude} , which is calculated by the following formula:

$$RR_{ofDbyX_crude} = ((Dx1c1 + Dx1c0) / (Px1c1 + Px1c0)) / ((Dx0c1 + Dx0c0) / (Px0c1 + Px0c0))$$

The RR of radiation for the disease among the subjects with the confounder is expressed as RR_{ofDbyX_c1} . The RR of radiation for the disease among without the confounder is expressed as RR_{ofDbyX_c0} . We assume that the presence of confounder C does not modify the RR of radiation for the disease; therefore, $RR_{ofDbyX_c0} = RR_{ofDbyX_c1}$. This RR will be expressed as $RR_{ofDbyX_adjusted}$.

Odds of the confounder among subjects with radiation exposure is expressed as $ODDS_{ofC_x1}$. This odds can be calculated as $Px1c1 / Px1c0$.

Similarly, $ODDS_{ofC_x0}$ can be calculated as $Px0c1 / Px0c0$.

The relative risk of the confounder for the disease among subjects without radiation exposure can be calculated as $(Dx0c1 / Px0c1) / (Dx0c0 / Px0c0)$, which can be rewritten $(Dx0c1 / Dx0c0) / ODDS_{ofC_x0}$, and is referred to as RR_{ofDbyC_x0} . Therefore, $(Dx0c1 / Dx0c0) / ODDS_{ofC_x0} = RR_{ofDbyC_x0}$.

This equation can be transformed as follows:
 $Dx0c1 = RR_{ofDbyC_x0} \times ODDS_{ofC_x0} \times Dx0c0$.

By addition of $Dx0c0$ to both sides of this equation, the following equation is obtained:

$$Dx0c1 + Dx0c0 = (RR_{ofDbyC_x0} \times ODDS_{ofC_x0} + 1) \times Dx0c0. \quad - (A)$$

Similarly, the following equation is obtained:

$$Dx1c1 + Dx1c0 = (RR_{ofDbyC_x1} \times ODDS_{ofC_x1} + 1) \times Dx1c0. \quad - (B)$$

$Px1c1 + Px1c0$ can be calculated as $(ODDS_{ofC_x1} + 1) \times Px1c0. \quad - (C)$

$Px0c1 + Px0c0$ can be calculated as $(ODDS_{ofC_x0} + 1) \times Px0c0. \quad - (D)$

RR_{ofDbyX_crude} can be rewritten, using (A), (B), (C) and (D), as follows:

$$\begin{aligned} & ((RR_{ofDbyC_x1} \times ODDS_{ofC_x1} + 1) \times Dx1c0) / ((ODDS_{ofC_x1} + 1) \times Px1c0) \\ & / ((RR_{ofDbyC_x0} \times ODDS_{ofC_x0} + 1) \times Dx0c0) \times \\ & ((ODDS_{ofC_x0} + 1) \times Px0c0). \end{aligned}$$

This formula can be rewritten as follows:

$$\begin{aligned} & (RR_{ofDbyC_x1} \times ODDS_{ofC_x1} + 1) / (RR_{ofDbyC_x0} \times \\ & ODDS_{ofC_x0} + 1) \\ & / ((ODDS_{ofC_x1} + 1) \times (ODDS_{ofC_x0} + 1)) \\ & \times (Dx1c0 / Px1c0) / (Dx0c0 / Px0c0). \end{aligned}$$

The last line of the formula described above is $(Dx1c0 / Px1c0) / (Dx0c0 / Px0c0) = RR_{ofDbyX_c0} = RR_{ofDbyX_adjusted}$.

Since there is no modification of RR_{ofDbyC} by X, $RR_{ofDbyC_x1} = RR_{ofDbyC_x0}$. Hereinafter, this value is expressed as RR_{ofDbyC} .

$ODDS_{ofC_x1} = Px1c1 / Px1c0$, and

$ODDS_{ofC_x0} = Px0c1 / Px0c0$.

When those expressions are used,

$$\begin{aligned} & RR_{ofDbyX_crude} / RR_{ofDbyX_adjusted} \\ & = (RR_{ofDbyC} \times Px1c1 / Px1c0 + 1) / (RR_{ofDbyC} \times \\ & ODDS_{ofC_x0} + 1) / (Px1c1 / Px1c0 + 1) \times (ODDS_{ofC_x0} + 1). \end{aligned}$$

When $Px1c1 / (Px1c1 + Px1c0)$ is expressed as Cpr (the prevalence of the confounder among the exposed), $Px1c1 / Px1c0 = Cpr / (1 - Cpr)$.

$$\begin{aligned} & \text{Therefore, } RR_{ofDbyX_crude} / RR_{ofDbyX_adjusted} \\ & = (RR_{ofDbyC} \times Cpr / (1 - Cpr) + 1) / (RR_{ofDbyC} \times ODDS_{ofC_x0} + 1) / (Cpr / (1 - Cpr) + 1) \times (ODDS_{ofC_x0} + 1). \end{aligned}$$