Today's Presentation

NORTHERN NEW ENGLAND CLINICAL & TRANSLATIONAL RESEARCH NETWORK

Case-Control vs Cohort Studies: Design & Analysis

Peter W. Callas, PhD
Associate Professor
Director, Medical Biostatistics
Larner College Of Medicine
University Of Vermont

Project ECHO®

Case-Control vs Cohort Studies: Design & Analysis

ECHO 25 January 2023

Peter W. Callas, PhD
Associate Professor
Director, Medical Biostatistics
Larner College of Medicine
University of Vermont

Study Designs

Descriptive

Case report / case series

Ecologic

Cross-sectional

Analytic

Case-control

Cohort

Clinical Trial

Observational

Experimental

Independent Variable (X)

Predictor

Risk factor

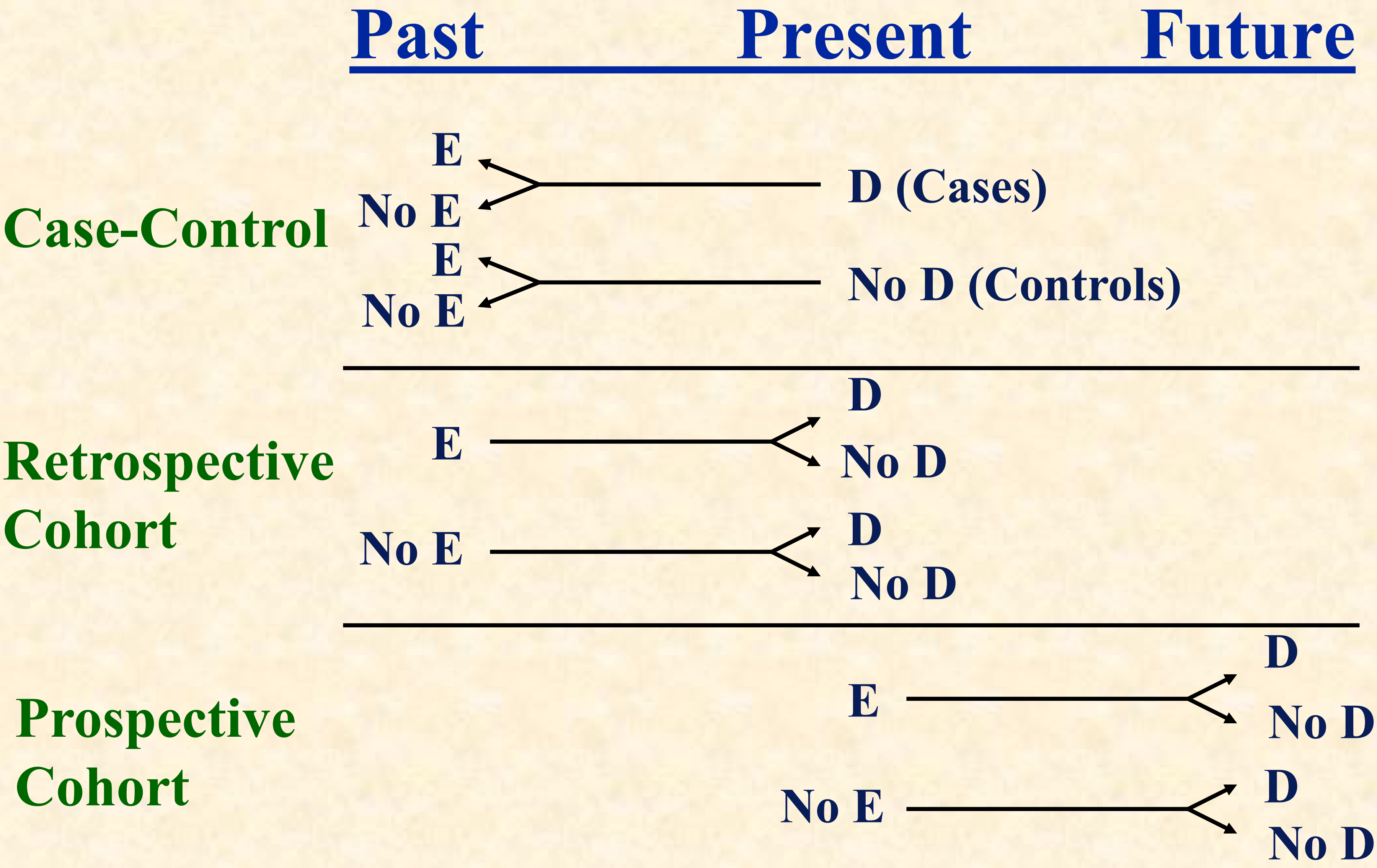
Preventive factor

Exposure – E

Dependent Variable (Y)

Incidence, prevalence

Outcome – D


Past
Present
Future
Case-Control
E
No E
D (Cases)
E
No E
No D (Controls)
Retrospective
Cohort
E
D
No D
No E
D
No D
Prospective
Cohort
E
D
No D
No E
D
No D

Case-Control

⊕

- Relatively quick and inexpensive
- Good for rare outcomes

⊖

- Selection bias – e.g., selection of control group
- Information bias – e.g., recall bias
- Not good for rare exposures
- Can only study one outcome
- Cannot directly compute incidence rates

Cohort

⊕

- Estimate incidence so can compute relative risk, etc
- Good for rare exposures
- Temporality
- Minimize information, selection bias
- Examine multiple outcomes

⊖

- Expensive, time-consuming - prospective
- Not good for rare outcomes
- Retrospective – need good records for exposure and potential confounders

	D	No D	
E	a	b	a+b
No E	c	d	c+d

$$RR = \frac{\text{probability of D in E}}{\text{probability of D in no E}} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}}$$

RR

- Risk ratio
- Rate ratio
- Relative risk

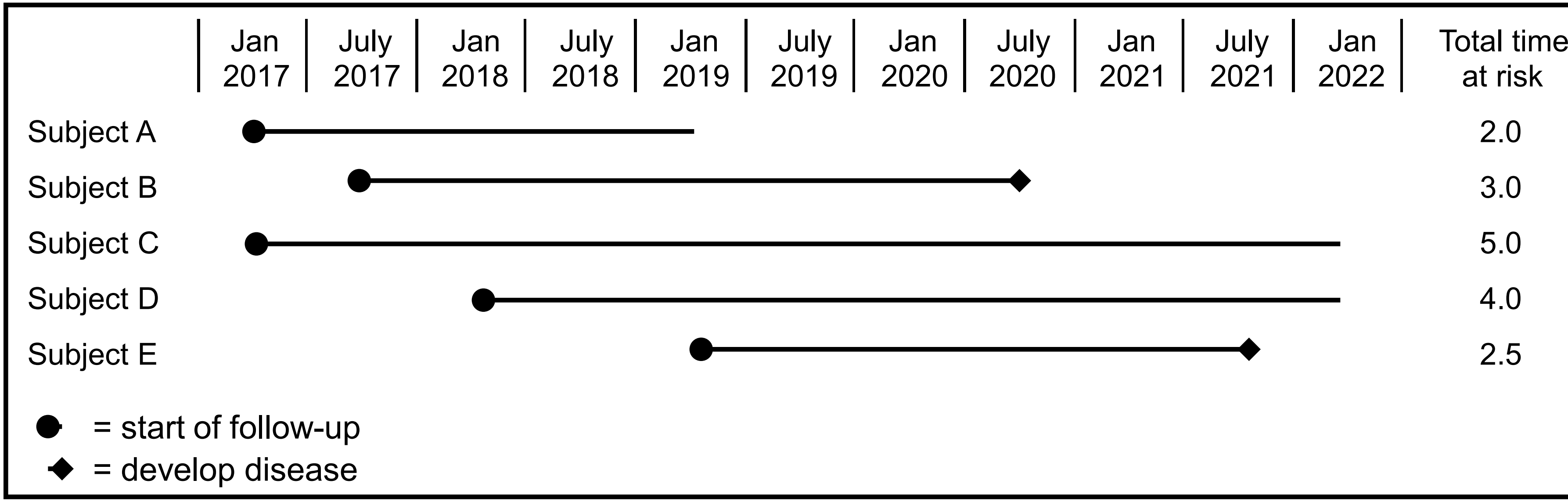
HR

- Hazard ratio

Silica Exposure

	Silicosis	No	
High	51	5269	5320
Low	4	1728	1732

$$RR = \frac{\text{probability of D in E}}{\text{probability of D in no E}} = \frac{\frac{51}{5320}}{\frac{4}{1732}} = 4.15$$



$$\textbf{Incidence Proportion} = \frac{\textbf{2 new cases}}{\textbf{5 at risk}} = \textbf{0.4}$$

$$= \textbf{40 per 100 per 5 years} = \textbf{8 per 100 per 1 year}$$

$$\textbf{Incidence Rate} = \frac{\textbf{2 new cases}}{\textbf{16.5 py at risk}} = \textbf{0.121 yr}^{-1}$$

$$= \textbf{12.1 per 100 person-years}$$

	D	No D	
E	a	b	a+b
No E	c	d	c+d

OR

- Odds ratio

$$OR = \frac{\text{odds of D in E}}{\text{odds of D in no E}} = \frac{\frac{a/(a+b)}{b/(a+b)}}{\frac{c/(c+d)}{d/(c+d)}} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}$$

Silica Exposure

	Silicosis	No	
High	51	5269	5320
Low	4	1728	1732

$$OR = \frac{\text{odds of D in E}}{\text{odds of D in no E}} = \frac{\frac{51}{5269}}{\frac{4}{1728}} = \frac{51 \text{ x } 1728}{5269 \text{ x } 4} = 4.18$$

	D	No D	
E	a	b	a+b
No E	c	d	c+d

$$RR = \frac{\text{probability of D in E}}{\text{probability of D in no E}} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}}$$

$$OR = \frac{\text{odds of D in E}}{\text{odds of D in no E}} = \frac{\frac{a/(a+b)}{b/(a+b)}}{\frac{c/(c+d)}{d/(c+d)}} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}$$

If rare D: $\dfrac{\frac{a}{b}}{\frac{c}{d}} \approx \dfrac{\frac{a}{\cancel{a}+b}}{\frac{c}{\cancel{c}+d}}$

so OR ≈ RR

RR = 4.15
OR = 4.18

	Cases	Controls	
E	a	b	$\cancel{a+b}$
No E	c	d	$\cancel{c+d}$
	a+c	b+d	

$$\mathrm{OR} = \frac{\text{odds of D in E}}{\text{odds of D in no E}} = \frac{\;\frac{a}{b}\;}{\;\frac{c}{d}\;} = \frac{ad}{bc}$$

$$\mathrm{OR} = \frac{\text{odds of E in D}}{\text{odds of E in no D}} = \frac{\;\frac{a/(a+c)}{c/(a+c)}\;}{\;\frac{b/(b+d)}{d/(b+d)}\;} = \frac{\;\frac{a}{c}\;}{\;\frac{b}{d}\;} = \frac{ad}{bc}$$

Case Control Study: OR =

Odds of E in cases relative to odds of E in controls

Odds in D in exposed relative to odds of D in unexposed

≈ Risk of D in exposed relative to risk of D in unexposed

Odds Ratio in a Case-Control Study:

Example: OR = 4.18

- **Odds of exposure in cases 4.18 times higher than odds of exposure in controls**

- **Odds of disease in exposed 4.18 times higher than odds of disease in unexposed**

- **Risk of disease in exposed 4.18 times higher than risk of disease in unexposed**

Silica Exposure		Silicosis	No	
	High	51	5269	5320
	Low	4	1728	1732

$$OR = \frac{\text{odds of D in E}}{\text{odds of D in no E}} = \frac{51 \text{ x } 1728}{5269 \text{ x } 4} = 4.18$$

Silica Exposure		Silicosis	No
	High	51	122
	Low	4	40
		55	162

$$OR = \frac{\text{odds of E in D}}{\text{odds of E in no D}} = \frac{51 \text{ x } 40}{122 \text{ x } 4} = 4.18$$

For each case and control in nested case-control study, lifetime work history determined using:

- **Surveillance data from VT Department of Health Division of Industrial Hygiene**
- **Self-reported work histories from a pulmonary function study**
- **Pension records**
- **Autopsy reports**
- **Death certificates**
- **Obituaries**

Table 1 Estimated exposure concentrations of respirable free silica by time period

Job class	Location	<1940 N	<1940 mg/m^3	1940–1949* mg/m^3	≥1950 N	≥1950 mg/m^3
Bit grinder†	Quarry	1	0.17			
Blacksmith†	Quarry	4	0.03			
Boxer	Shed	14	0.08	0.06	103	0.04
Carver	Shed	19	0.37	0.22	149	0.07
Channel bar	Quarry	3	0.15	0.08		0.01‡
Crane	Shed	9	0.16	0.11	32	0.05
Cutter	Shed	331	0.39	0.23	1569	0.07
Draftsman	Shed	12	0.01	0.01		0.01
Driller	Quarry	120	1.07	0.54	7	0.01
Foreman	Shed		0.12	0.09	9	0.05
Grinder	Shed	31	0.19	0.13	5	0.07
Jackhammer	Quarry	10	1.05	0.56	7	0.06
Labourer	Shed		0.24	0.17	8	0.10
Lumper	Shed	5	0.30	0.18	138	0.06
Maintenance	Shed	12	0.24	0.16	28	0.07
Quarry (general)	Quarry	22	0.13	0.07		0.01‡
Office worker	Shed	29	0.04	0.04		0.04
Polisher	Shed	35	0.12	0.10	570	0.07
Sandblaster	Shed	43	0.24	0.16	337	0.07
Sawyer	Shed	13	0.13	0.10	634	0.06
Shed (general)	Shed	153	0.12	0.09	491	0.05
Surfacer	Shed	150	0.28	0.18	101	0.08

*Estimates are averages of those for the earlier and later periods because few measurements were available from 1940 to 1949.
†Job not performed after 1939.
‡Trend applied using jackhammer and driller data.

		Silicosis	No	
Silica Exposure	High	51	5269	5320
	Low	4	1728	1732

$$\text{OR} = \frac{\text{odds of D in E}}{\text{odds of D in no E}} = \frac{51 \times 1728}{5269 \times 4} = 4.18 \quad p = 0.003$$

95% CI 1.51-11.59

		Silicosis	No
Silica Exposure	High	51	122
	Low	4	40
		55	162

$$\text{OR} = \frac{\text{odds of E in D}}{\text{odds of E in no D}} = \frac{51 \times 40}{122 \times 4} = 4.18 \quad p = 0.006$$

95% CI 1.42-12.29

⊕ Nested Case-Control

- Good if assessing E or other variable that is expensive, complicated, invasive, etc.
- E measured before D, so
  - no differential bias
  - temporality
- Controls from same population as cases so↓ selection bias
- Efficient

Mortality in Vermont granite workers and its association with silica exposure

Pamela M Vacek,[1] Dave K Verma,[2] William G Graham,[3] Peter W Callas,[1] Graham W Gibbs[4,5]

ABSTRACT

Objectives To assess mortality in Vermont granite workers and examine relationships between silica exposure and mortality from lung cancer, kidney cancer, non-malignant kidney disease, silicosis and other non-malignant respiratory disease.

Methods Workers employed between 1947 and 1998 were identified. Exposures were estimated using a job—exposure matrix. Mortality was assessed through 2004 and standardised mortality ratios (SMRs) were computed. Associations between mortality and exposure to silica were assessed by nested case—control analyses using conditional logistic regression.

Occup Environ Med 2011;68:312-318