SOC 3811/5811:
BASIC SOCIAL STATISTICS

Causal Inference and Observational Data

Example #1


Abo tion Increases
Brea st Cancer Risk
Toll Free 1-8 NO CANCER


Abortion LINKED t
BREAST CANCER!
28 studies reveal increased risk
www.abortionbreastcancer.com
www.togethertolife.ca

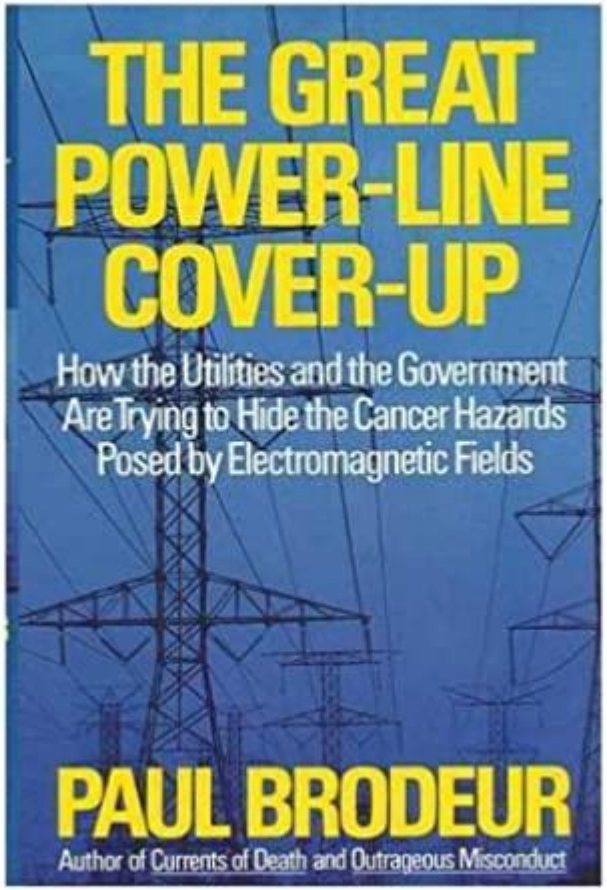
ABORTION & BREAST CANCER

- 44% HIGHER RISK
- RISK RISES WITH NUMBER OF ABORTIONS

NATIONAL REVIEW ONLINE

Does having an abortion increase a woman's risk of breast cancer?

Example #2


THE GREAT POWER-LINE COVER-UP
How the Utilities and the Government Are Trying to Hide the Cancer Hazards Posed by Electromagnetic Fields
PAUL BRODEUR
Author of Currents of Death and Outrageous Misconduct

Does living near high-voltage power lines increase children leukemia risk?

Causality and Counterfactuals

At the heart of all case and effect statements is a *counterfactual—the situation that would have existed had the explanatory variable not changed*

For example, by making the causal claim:

> "The COVID pandemic caused higher unemployment rates"

we are simultaneously making the counterfactual claim:

> "If we had not experienced the COVID pandemic, unemployment rates would not have increased"

A fundamental problem in explanatory research is that we never actually observe the counterfactual

Criteria for Establishing Causality

Three conditions that must be met in order to establish that X causes Y:

1. X and Y must be empirically associated (criteria of association)
2. Change in X must precede change in Y in time (criteria of temporal ordering)
3. There must be no third variable, Z, which acts as a "confounder"—or which induces "spuriousness"—in the association between X and Y (criteria of nonspuriousness)

Research Design & Causal Inference

We know how to measure associations between X and Y

Establishing temporal ordering is a matter of designing research appropriately (and especially of measuring X and Y thoughtfully)

Confounding is entirely avoided in experimental designs (because the control group serves as a counterfactual)

Confounding can be at least partly overcome by statistical control in observational designs

Example #1

How could we study this?

Example #1

"Abortion causes breast cancer"

1. What is the counterfactual?

2. How would you conduct an experiment that would establish the causal effect of abortion on breast cancer? Why might it be impossible to conduct that experiment?

3. Using observational data (which is really all we will ever have): Are there variables or factors that might confound the association between abortion and breast cancer risk?

Springer Link

Original paper | Published: 24 November 2013

A meta-analysis of the association between induced abortion and breast cancer risk among Chinese females

Yubei Huang, Xiaoliang Zhang, Weiqin Li, Fengju Song, Hongji Dai, Jing Wang, Ying Gao, Xueou Liu, Chuan Chen, Ye Yan, Yaogang Wang & Kexin Chen

Cancer Causes & Control **25**, 227–236(2014) | Cite this article

2292 Accesses | **25** Citations | **198** Altmetric | Metrics

Abstract

Objective

To evaluate the association between induced abortion (IA) and breast cancer risk among Chinese females.

Methods

We searched three English databases (PubMed, ScienceDirect, and Wiley) and three Chinese databases (*CNKI*, *WanFang*, and *VIP*) for studies up to December 2012, supplemented by

Our results might be confounded by additional factors. First, some abortions performed before marriage might be included. However, these abortions were very few, and probably would not be reported in China [9], as they are less socially acceptable and are associated with more stigmas. Second, though inadequate choices of the reference group might be the main reason why there was no association in the strongest studies, i.e., cohorts, NOS of 8–9, and those conducted in Shanghai, the positive result of association between IA and breast cancer risk still might be overstated. Third, the pooled ORs might be confounded by other factors, including age, parity, and age at first birth. Although meta-analysis based on adjusted ORs could theoretically get a clearer conclusion, crude ORs from univariate logistic regression were used in the primary analysis based on the following three reasons: (1) some of the included studies did not report the adjusted ORs, including those not focusing on IA and those concluding negative ORs after multiple adjusting due to small sample size or inadequate choices of the reference group. In fact, only 13 of the 36 studies had reported the adjusted ORs, and summary based on these 13 adjusted ORs was similar to the primary result, suggesting that the primary result was not substantially confounded by the un-adjusted factors. (2) The adjustment terms varied greatly in the included studies. Summarizing these results from different calculation methods would inevitably incur more confounding rather than get a clearer result. (3) ORs from cross-table were also crude ORs equal to ORs calculated from univariate logistic regression. In order to get a more comparable result with cross-table, crude ORs from univariate logistic regression rather than adjusted ORs from multivariate logistic regression should be used. Therefore, these results should be interpreted with caution, and future prospective cohort studies with more adequate reference group were needed to investigate the association further.

Induced Abortion and Breast Cancer Risk

Committee Opinion | Number 434 | June 2009

By reading this page you agree to ACOG's Terms and Conditions. Read terms

Share

Search page

Jump to:

Abstract

References

Download PDF

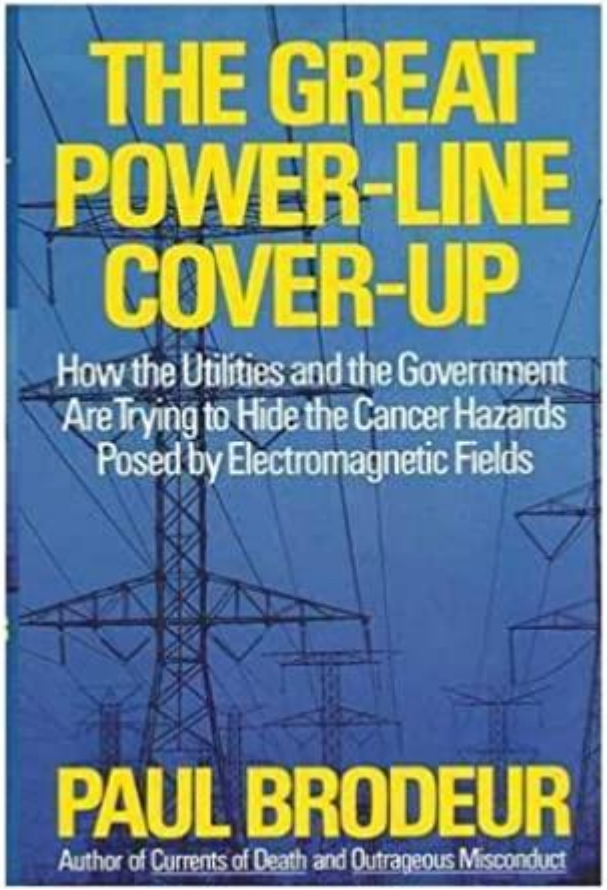
Number 434 *(Replaces No. 285, August 2003. Reaffirmed 2019)*

Committee on Gynecologic Practice

This document reflects emerging clinical and scientific advances as of the date issued and is subject to change. The information should not be construed as dictating an exclusive course of treatment or procedure to be followed.

ABSTRACT: The relationship between induced abortion and the subsequent development of breast cancer has been the subject of a substantial amount of epidemiologic study. Early studies of the relationship between prior induced abortion and breast cancer risk were methodologically flawed. More rigorous recent studies demonstrate no causal relationship between induced abortion and a subsequent increase in breast cancer risk.

Example #2



How could we study this?

Example #2

"Living near high-voltage lines causes leukemia"

1. What is the counterfactual?

2. How would you conduct an experiment that would establish the causal effect of living near power lines on risk of leukemia? Why might it be impossible to conduct it?

3. Using observational data: Are there variables or factors that might confound the association between whether you live near power lines and leukemia risk?

Article | Open Access | Published: 29 May 2018

Epidemiology

Proximity to overhead power lines and childhood leukaemia: an international pooled analysis

Aryana T Amoon, Catherine M Crespi, [...] Leeka Kheifets

British Journal of Cancer **119**, 364–373(2018) | Cite this article

20k Accesses | **14** Citations | **67** Altmetric | Metrics

Abstract

Background

Although studies have consistently found an association between childhood leukaemia risk and magnetic fields, the associations between childhood leukaemia and distance to overhead power lines have been inconsistent. We pooled data from multiple studies to assess the association with distance and evaluate whether it is due to magnetic fields or other factors associated with distance from lines.

Methods

We present a pooled analysis combining individual-level data (29,049 cases and 68,231 controls) from 11 record-based studies.

Results

There was no material association between childhood leukaemia and distance to nearest overhead power line of any voltage. Among children living < 50 m from 200 + kV power lines, the adjusted odds ratio for childhood leukaemia was 1.33 (95% CI: 0.92–1.93). The odds ratio was higher among children diagnosed before age 5 years. There was no association with calculated magnetic fields. Odds ratios remained unchanged with adjustment for potential confounders.