

Ch 13: V1 2019 USCOTS Workshop 1

Statistics Literacy For Decision Makers

13: Confounding & Cornfield

by
Milo Schield

*Half-Day Workshop
USCOTS May 16, 2019*

www.StatLit.org/pdf/2019-Schield-USCOTS-Slides13.pdf

Ch 13: V1 2019 USCOTS Workshop 2

Workshop Schedule

1:00 Ch 1 Statistical Literacy – Introduction
 1:30 Ch 2 Statistical Literacy – Details

2:15 Ch 3 Measurements
 2:45 Ch 4 Ratios

3:30 Ch 13 Standardizing
 4:00 Feedback

Ch 13: V1 2019 USCOTS Workshop 3

Confounding: Chapter 13 Outline

Cornfield-Fisher debate

Cornfield conditions

Standardizing percentages, rates and averages

Standardizing percentage & number attributable

Statistical significance and confounding

Ch 13: V1 2019 USCOTS Workshop 4

Stat Literacy: Study Statistics as Evidence in Arguments

The Point or the Target

The more disputable the point,
the stronger the evidence must be.

Statistic As Evidence

“All Statistics are Socially Constructed”
So, “Take CARE”!!

Statistics may be influenced by:

C	A	R	E
Context	Assembly	Randomness	Error

Ch 13: V1 2019 USCOTS Workshop 5

Cornfield-Fisher Debate

Doctors had noticed the strong association between smoking and lung cancer. Statisticians argued that this evidence strongly supported the claim that smoking was a cause of lung cancer.

Fisher, a smoker, noted that *association is not causation in observational studies*.

Fisher produced data. Identical twins were more likely to share a smoking preference than were fraternal twins. This statistic supported genetics as an alternate explanation for the association.

Ch 13: V1 2019 USCOTS Workshop 6

Cornfield-Fisher Debate

Now when the world’s leading statistician says something that every statistician agrees is true, most reasonably-minded statisticians would back off.

And when the world’s leading statistician produces data indicating a plausible confounder, it seems incredible that anyone would reply.

Jerome Cornfield did!

Ch 13: V1 2019 USCOTS Workshop 7

Cornfield Conditions

Cornfield **proved** that the relative risk of lung cancer had to be greater for a confounder (e.g., genetics) than for the predictor (e.g., smoking) in order to nullify or reverse the observed association.

Cornfield pointed out that smokers were about 10 times as likely to get lung cancer as non-smokers.

Fisher's data involved a factor of two.

Fisher never replied.

Ch 13: V1 2019 USCOTS Workshop 8

Contributions to Human Knowledge

“Cornfield's minimum effect size is as important to observational studies as is the use of randomized assignment to experimental studies.

No longer could one refute an ostensive causal association by simply asserting that some new factor (such as a genetic factor) might be the true cause.

Now one had to argue that the relative prevalence of this potentially confounding factor was greater than the relative risk for the ostensive cause.”

Schield (1999). [This was written 20 years ago!]

Ch 13: V1 2019 USCOTS Workshop 9

Confounder Distribution

Since confounders may be unknown, there is no way to derive or infer their distribution.

Schield (2018) argued that we needed a standard for confounder: a standard confounder distribution.

He proposed an exponential (one factor determined) with a mean relative risk of 2.

This applied if predictor and confounder are binary.

Ch 13: V1 2019 USCOTS Workshop 10

Confounder Distribution Unknown & Unknowable

Standard Confounder Distribution: Mean = 2

Minimum Relative Risk:
4: Significant 3: Strong
2: Moderate 1.5 Weak
1.0 Ignorable.
Schield 2018

Ch 13: V1 2019 USCOTS Workshop 11

Controlling for a Confounder: Graphical Technique

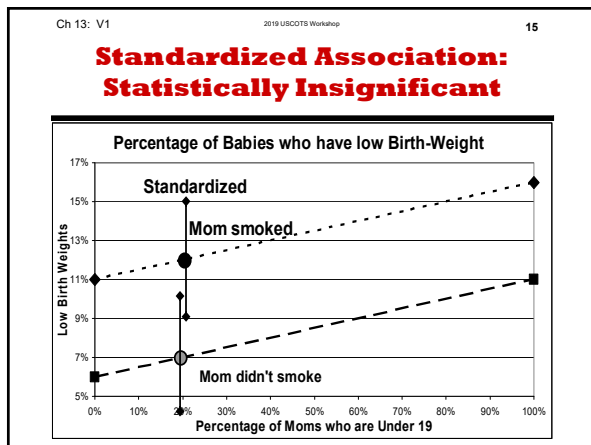
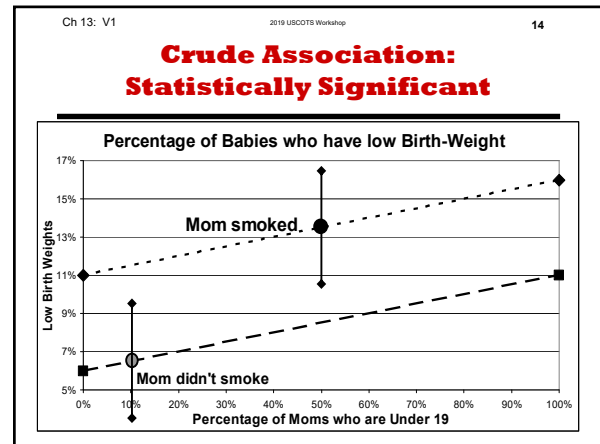
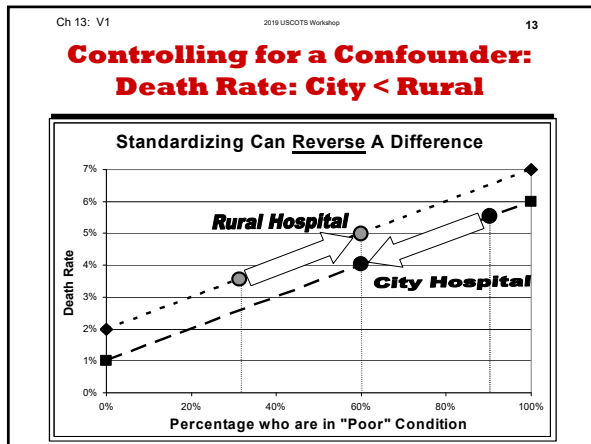
Wainer introduced a simple graphical technique that made the control of a binary confounder a relatively simple matter.

Schield (2006). Presenting Confounding Graphically Using Standardization, *STATS* magazine.
www.statlit.org/pdf/2006SchieldSTATS.pdf

Ch 13: V1 2019 USCOTS Workshop 12

Crude Association: Death Rate: City > Rural

A Confounder can Influence a Difference



Ch 13: V1 2019 USCOTS Workshop 16

Confounder Effect on Statistical Significance

Controlling for a confounder can transform a statistically-significant association into an association that is statistically insignificant.

Although statistical educators are clearly aware of this, there is nothing in any introductory textbook that alerts students to this possibility.

The failure to show a significance reversal is *statistical negligence*.

Statistics Literacy For Decision Makers

13: Confounding & Cornfield

by
Milo Schield

*Half-Day Workshop
USCOTS May 16, 2019*

www.StatLit.org/pdf/2019-Schild-USCOTS-Slides13.pdf

Workshop Schedule

1:00 Ch 1 Statistical Literacy – Introduction

1:30 Ch 2 Statistical Literacy – Details

2:15 Ch 3 Measurements

2:45 Ch 4 Ratios

3:30 Ch 13 Standardizing

4:00 Feedback

.

Confounding: Chapter 13 Outline

Cornfield-Fisher debate

Cornfield conditions

Standardizing percentages, rates and averages

Standardizing percentage & number attributable

Statistical significance and confounding

Stat Literacy: Study Statistics as Evidence in Arguments

The Point or the Target

The more disputable the point,
the stronger the evidence must be.

Statistic As Evidence

“All Statistics are Socially Constructed”

So, “Take CARE”!!

Statistics may be influenced by:

C	A	R	E
Context	Assembly	Randomness	Error

Cornfield-Fisher Debate

Doctors had noticed the strong association between smoking and lung cancer. Statisticians argued that this evidence strongly supported the claim that smoking was a cause of lung cancer.

Fisher, a smoker, noted that *association is not causation in observational studies*.

Fisher produced data. Identical twins were more likely to share a smoking preference than were fraternal twins. This statistic supported genetics as an alternate explanation for the association.

Cornfield-Fisher Debate

Now when the world's leading statistician says something that every statistician agrees is true, most reasonably-minded statisticians would back off.

And when the world's leading statistician produces data indicating a plausible confounder, it seems incredible that anyone would reply.

Jerome Cornfield did!

Cornfield Conditions

Cornfield **proved** that the relative risk of lung cancer had to be greater for a confounder (e.g., genetics) than for the predictor (e.g., smoking) in order to nullify or reverse the observed association.

Cornfield pointed out that smokers were about 10 times as likely to get lung cancer as non-smokers.

Fisher's data involved a factor of two.

Fisher never replied.

Contributions to Human Knowledge

“Cornfield's minimum effect size is as important to observational studies as is the use of randomized assignment to experimental studies.

No longer could one refute an ostensive causal association by simply asserting that some new factor (such as a genetic factor) might be the true cause.

Now one had to argue that the relative prevalence of this potentially confounding factor was greater than the relative risk for the ostensive cause.”

Schild (1999). [This was written 20 years ago!]

Confounder Distribution

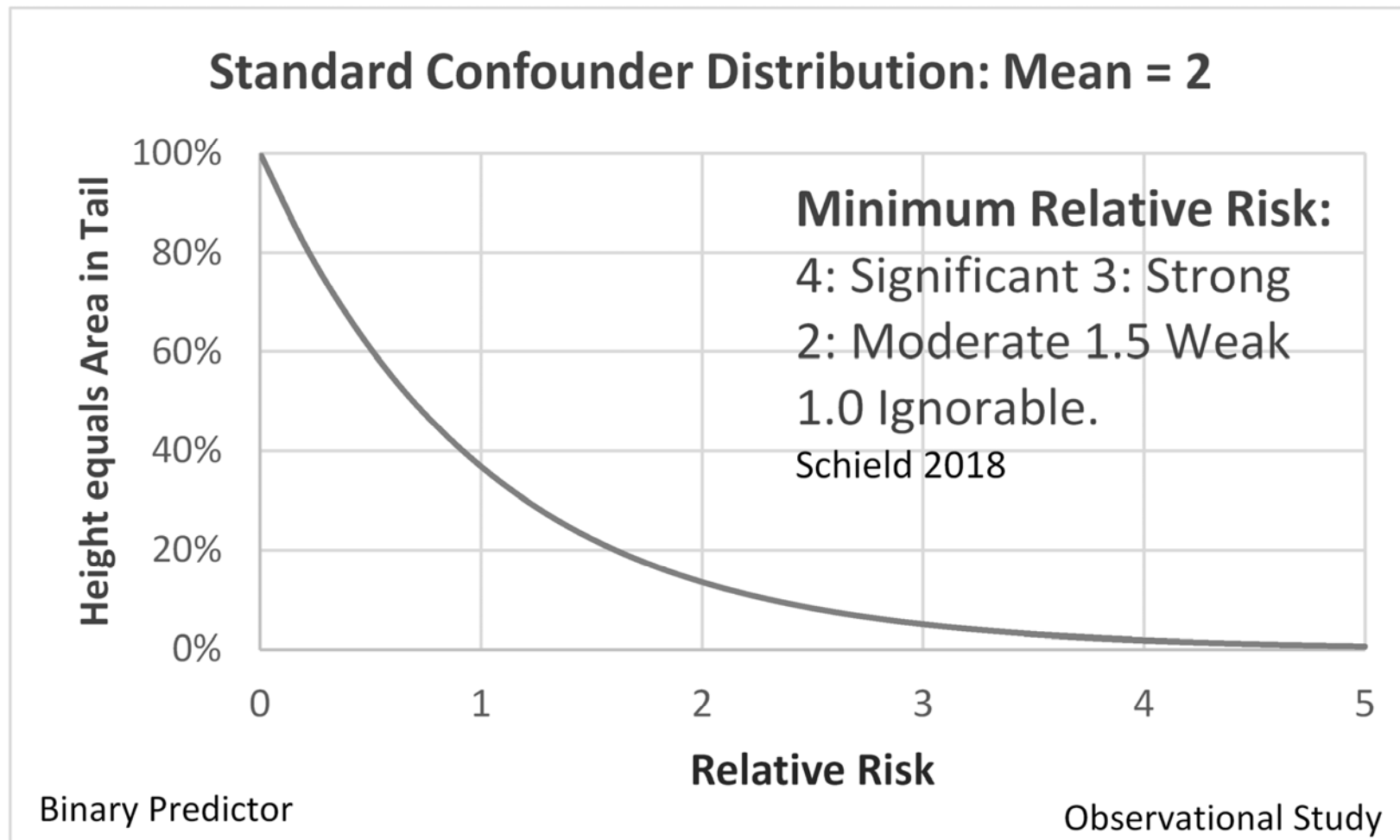
Since confounders may be unknown, there is no way to derive or infer their distribution.

Schild (2018) argued that we needed a standard for confounder: a standard confounder distribution.

He proposed an exponential (one factor determined) with a mean relative risk of 2.

This applied if predictor and confounder are binary.

Confounder Distribution Unknown & Unknowable



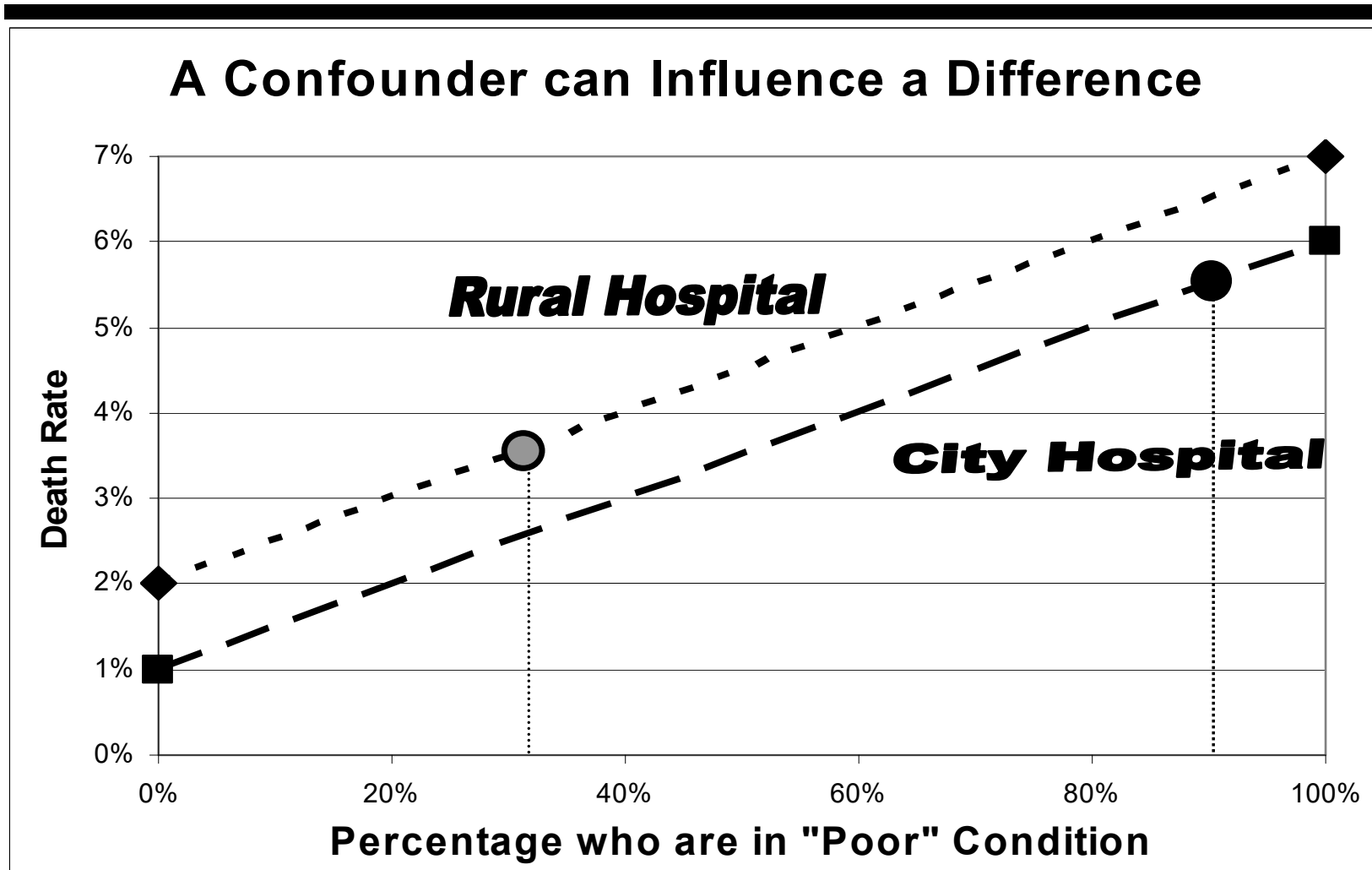
Controlling for a Confounder: Graphical Technique

Wainer introduced a simple graphical technique that made the control of a binary confounder a relatively simple matter.

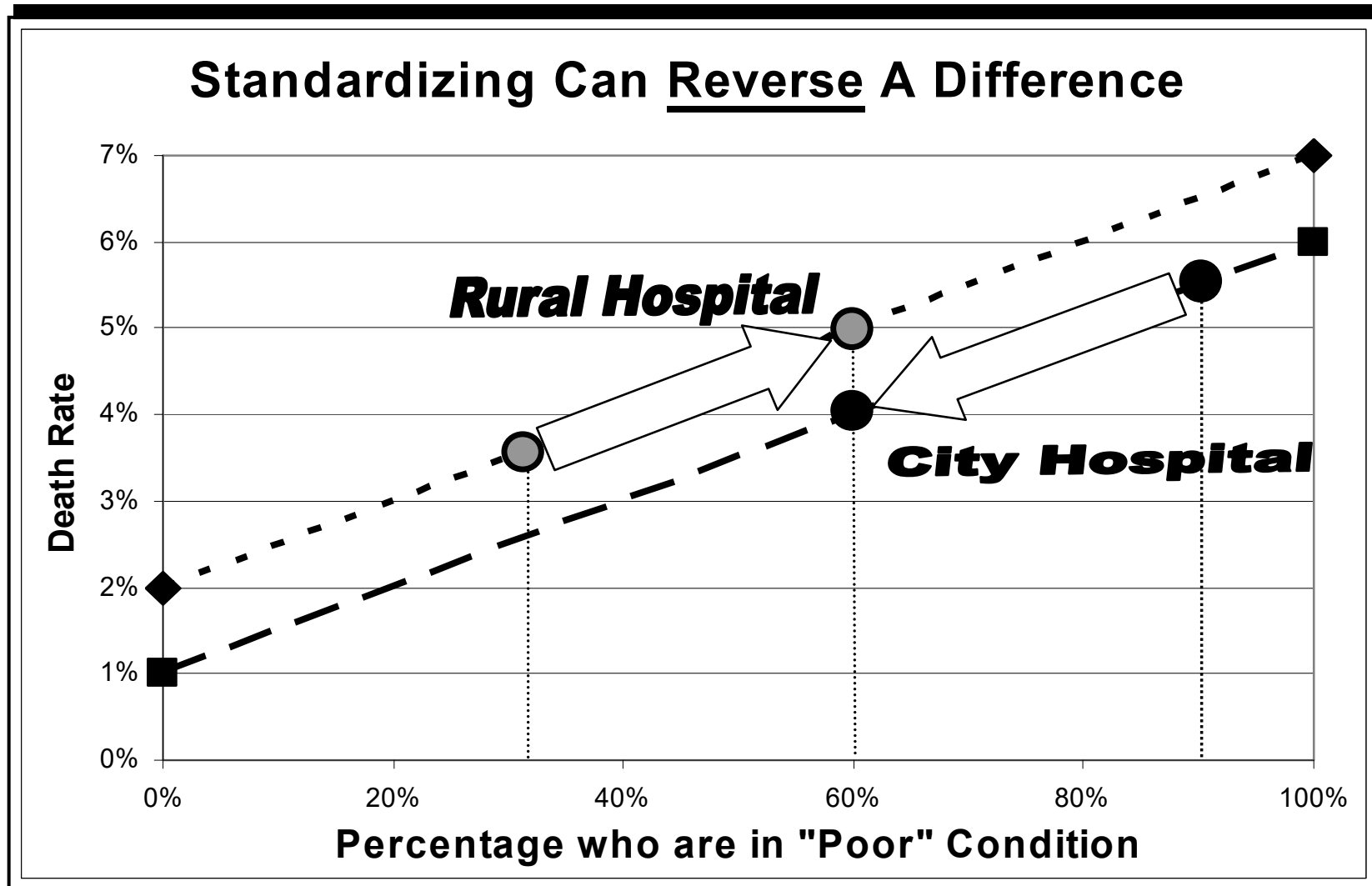
Schild (2006). Presenting Confounding Graphically Using Standardization, *STATS* magazine.

www.statlit.org/pdf/2006SchildSTATS.pdf

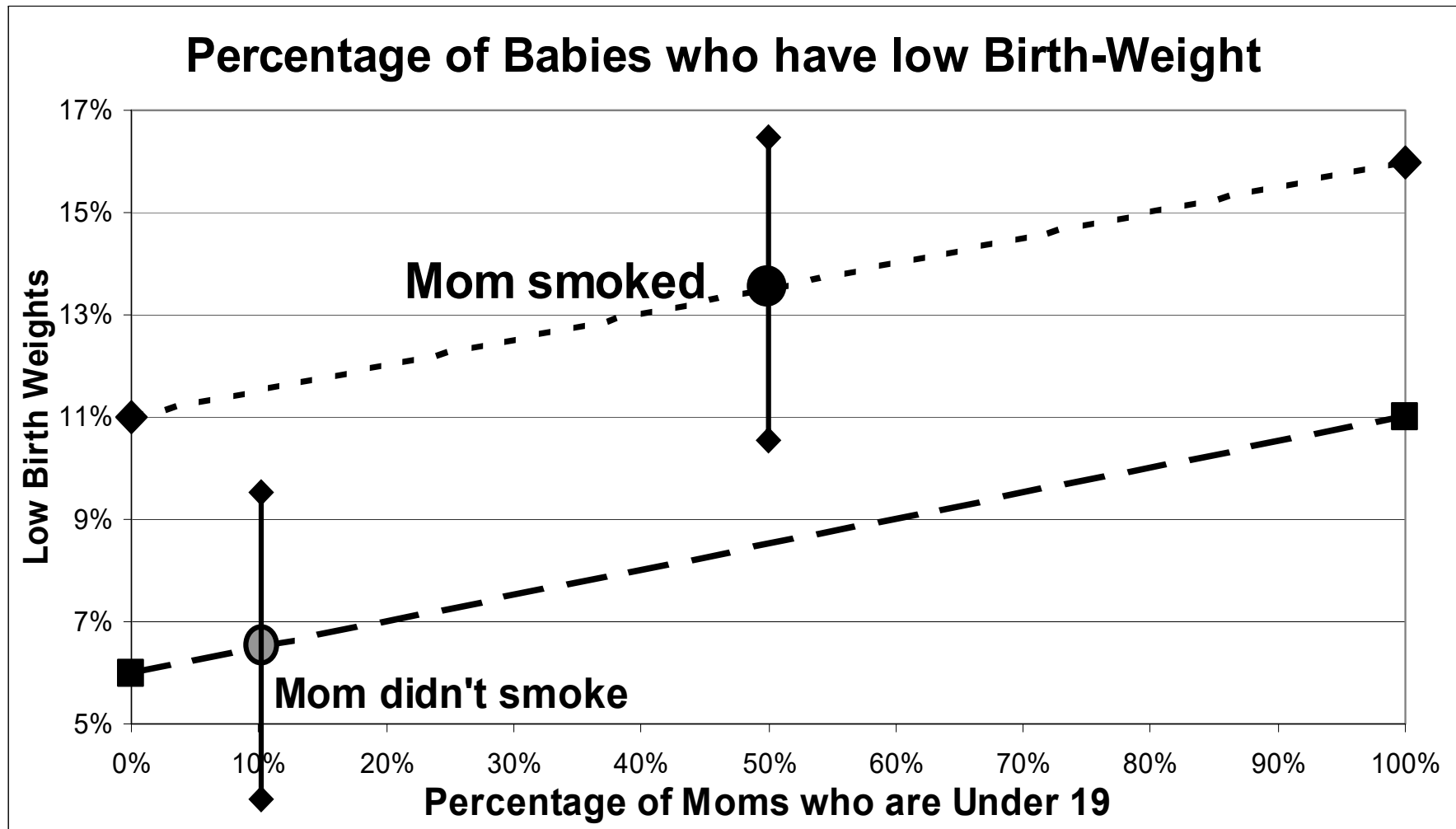
Crude Association: Death Rate: City > Rural



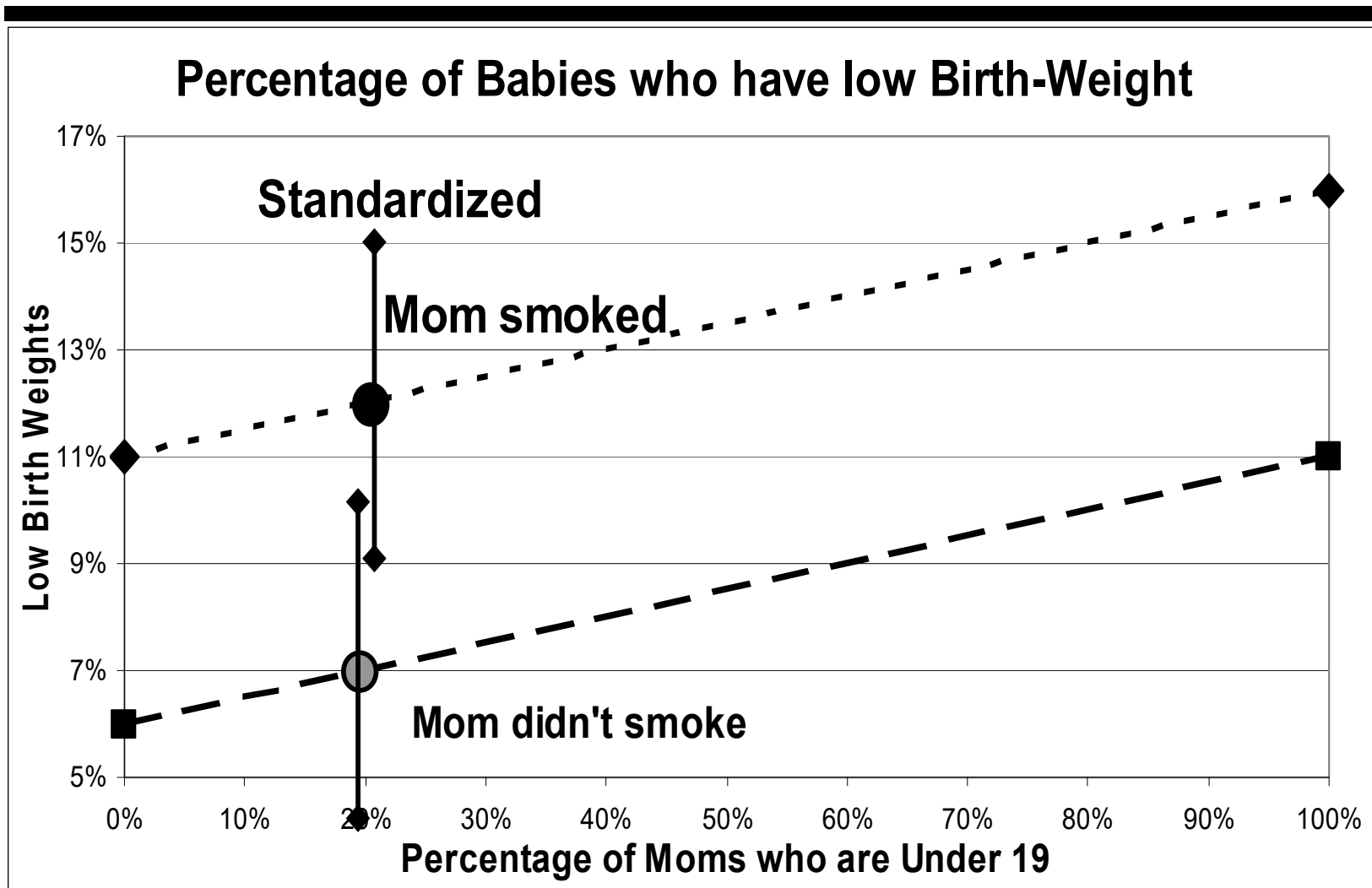
Controlling for a Confounder: Death Rate: City < Rural



Crude Association: Statistically Significant



Standardized Association: Statistically Insignificant



Confounder Effect on Statistical Significance

Controlling for a confounder can transform a statistically-significant association into an association that is statistically insignificant.

Although statistical educators are clearly aware of this, there is nothing in any introductory textbook that alerts students to this possibility.

The failure to show a significance reversal is *statistical negligence*.