Epidemiological Reasoning

MILO SCHIELD

Augsburg College

Director, W. M. Keck Statistical Literacy Project

www.Augsburg.edu/StatLit **www.StatLit.org** **Schield@augsburg.edu**

The 2004 Surgeon General's Report

The Objectivist Center (TOC) Summer Seminar

Union College

12 July 2005

www.StatLit.org/pdf/2005SchieldUnion6Up.pdf

Association versus Causation

"*Association is not causation*." This ambiguous claim includes at least 4 different statements.

Association (without manipulation) is …

1. **not ever** [sufficient to infer] causation.
2. **not always** [sufficient to infer] causation
3. **sometimes** [sufficient to infer] causation.
4. **sometimes** [evidence for] causation.
5. **always** evidence of causation somewhere.

#1 is false (statistically); #2 - #5 are true.

Using Association as evidence of Causation

Using association as evidence for causation is necessary when manipulation is:

1. impossible (cosmology).
2. possible but unethical (smoking).
3. possible & ethical but expensive.
4. possible & ethical but takes time.
5. possible & ethical but requires choice.

Epidemiology

Epidemiology typically involves

- observational studies.
- ordinal, nominal or binary predictors.
- binary outcomes (live or die)
- relative risk, odds ratio, attributable fraction

Epidemiology seldom involves

- random assignment
- Pearson correlation coefficient

Epidemiology and Public Health

Epidemiology, epidemics & communicable diseases:

- 1796: Small Pox & William Jenner
- 1854: Cholera & John Snow
- 1944: Tuberculosis (TB)

Epidemiology studies personal health:

- 1964: Smoking and Cancer
- 1981: “Causes of Cancer” by Doll & Peto
- 1992: Second-hand smoke (ETS) EPA Assessment

Epidemiology studies lifestyle:

- 1990s onward

Epidemiology: Study of Health Risk Factors

Health - Reuters

Smoking Ups Impotence Risk in Younger Men

Health - HealthDay

Mom's Poor Diet Can Up Diabetes Risk in Child

Epidemiology: Conflicting Claims

5/96: British Medical Journal (BMJ) article cites “**overwhelming evidence**" that **excessive consumption [of salt] causes high blood pressure, heart disease and strokes**.

5/96: Journal of American Medical Assoc. (JAMA) article concludes that "**dietary salt intake has little effect on blood pressure in the population at large**."

The New Epidemiology: Study of Social Risk Factors

TV violence is

- a contributing factor to increases in violent crime and antisocial behavior. .
- a "risk factor" that contributes to increasing a person's aggressiveness.

just as every cigarette increases the chance that someday the smoker will get lung cancer, so every exposure to violence increases the chances that someday a child will behave more violently than they otherwise would.

Epidemiological Reasoning: Probabilistic Causation

Suppose:

- 80% of lung cancer deaths involve smokers
- 20% of smokers die of lung cancer

Thus, in relation to death from lung cancer, smoking may be (speaking loosely):

- very ‘necessary’ but far from ‘sufficient’

Q. Is smoking a cause of lung cancer death?

We can’t run an experiment.

Epidemiological Reasoning: Criteria

- Statistically significant (statistical criteria)

Hill (1965) criteria:

- Strength of Association (Relative Risk)
- Coherence, Plausibility & Analogy (Fits with known facts & theories)
- Dose-response relationship
- Consistency (repeatability)
- Temporality (cause precedes effect)

2004 Surgeon General's Report

Systematized reporting of epidemiological results as a basis for public health decisions.

1 Gave current statistics on effects of smoking

2 Reviewed language used in previous reports.

3 Classified strength of evidence for causation into four categories…

4 Was silent on minimum strength of association needed to control for confounding.

See www.cdc.gov/tobacco/sgr/sgr_2004/

#1: Quantitative Effects of (Active) Smoke

Tobacco use remains the leading preventable cause of disease and death in the US,

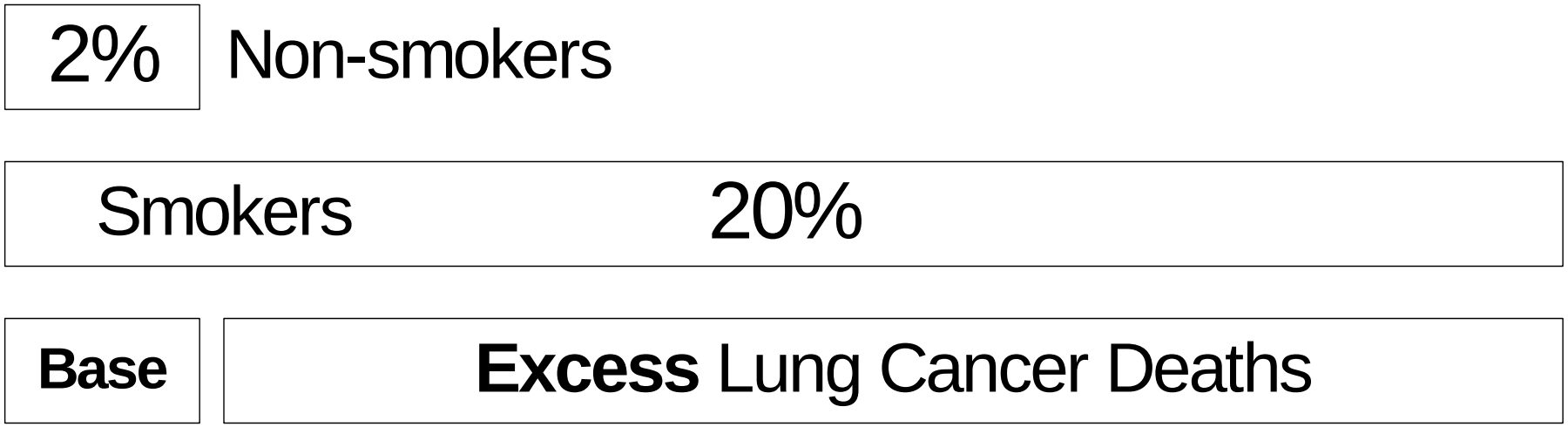
- causing approximately 440,000 deaths/year
- costing approximately $157 billion in annual health-related economic losses.

Nationally, smoking results in more than 5.6 million years of potential life lost each year.

Ch 1. 2004 US Surgeon General's Report

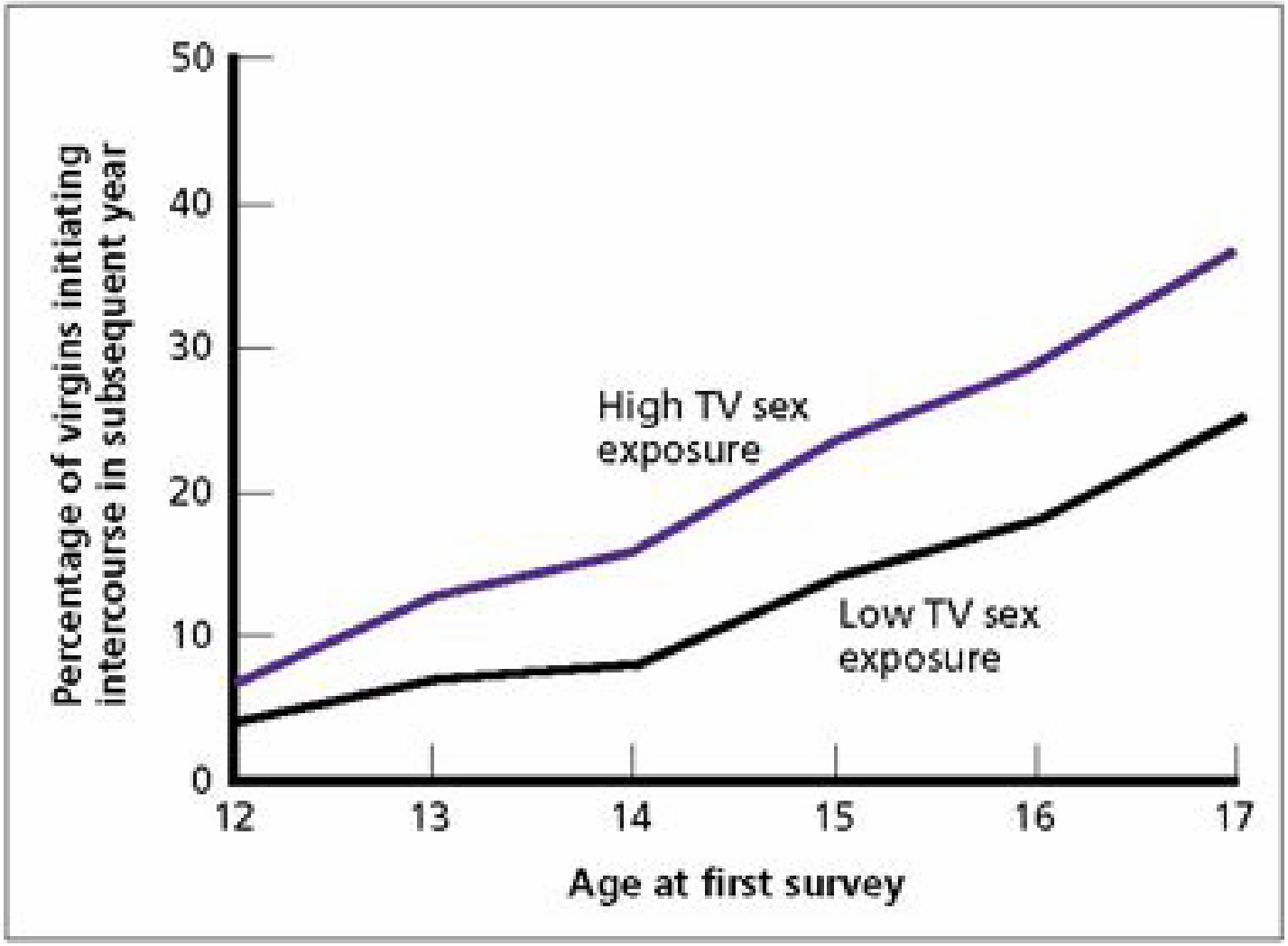
Epi Reasoning: "Deaths Attributed"

Hypothetical Deaths due to lung cancer



90% of these smoker deaths due to lung cancer are *attributable to* smoking

Twice as likely: 50% attributable to TV Sex



#2: Grammar of Association vs. Causation

Reviewed 250 statements in reports: 1964 – 2002

1. Association/relation; associated/related
2. Comparison: more, greater, stronger, bigger, etc.
3. Link, connection, factor: risk factor
4. Changes (active verb): Increases, cuts, contributes
5. Logical connection: ‘due to’ (*Risk due to* …)
6. *Causal* factor, time relation (leads to, results in)
7. *Causally* associated/related; a *causal* relation
8. *A cause*: *judged to be causal*, *causes*, *effect of*.

Change/Compare Grammars Observational Studies

Change and compare grammars imply causation

1. Electrical fields ***increase*** miscarriages
2. Night shift work ***Ups*** Breast Cancer Risk (40%)
3. HRT ***raises*** Cancer (24%) and Stroke (41%) Risks
4. Alcohol ***Ups*** Breast Cancer Risk (6% /glass-wk)
5. Eggs & Hot Dogs ***Cut*** Breast Cancer Risk (18%)
6. Broccoli ***Reduces*** Breast Cancer Risk (24%)
7. Parental tobacco ***leads to*** brain tumors (22%)
8. Non-shavers ***more likely*** to have stroke (70%)

Epidemiological Studies: Cheaper; Always Possible

NON-EPI ṠTUDIES:	**OUTCOME**	
PREDICTOR	Repeatable (*Migraine*)	Non-Repeatable (*Death*)
Changeable (*Medicine, $, Education*)	*Before/After*: Physical Exp. or Stat Exp.	Statistical Exp. / Clinical Trial: *Random Assign*
Unchangeable Physical: *Race* Moral: *Smoker*		

#3: Strength of Evidence for Causation (4 levels)

A. Evidence is **inadequate** to infer the presence or absence of a causal relationship (which encompasses evidence that is sparse, of poor quality, or conflicting).

B. Evidence is **suggestive of no causal relationship**.

C. Evidence is **suggestive but not sufficient** to infer a causal relationship.

D. Evidence is **sufficient** to infer a causal relationship

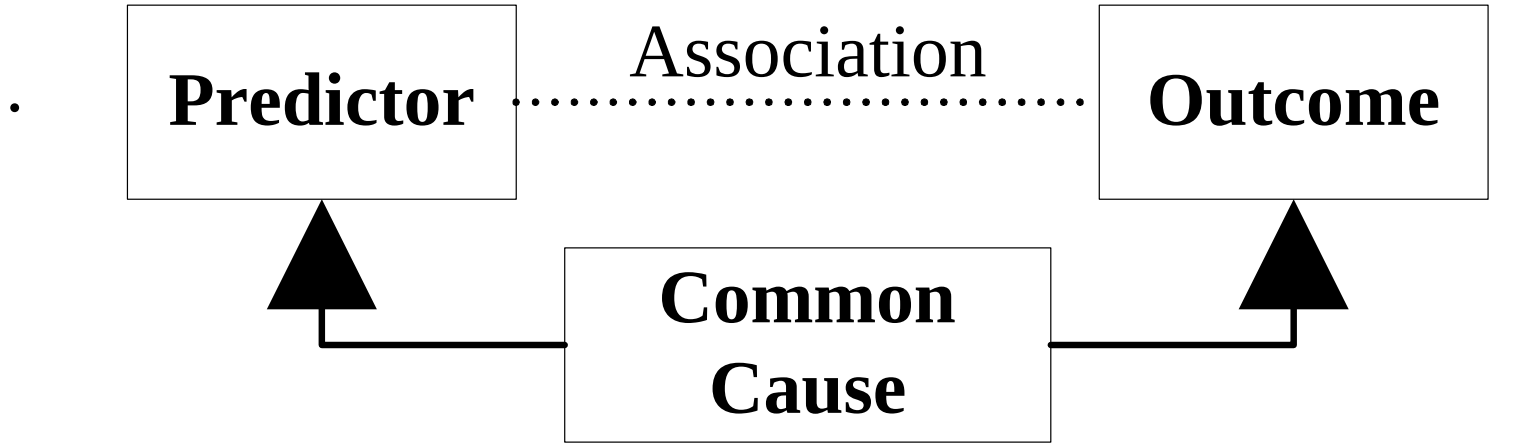
Suggestive but not sufficient to Infer Causation

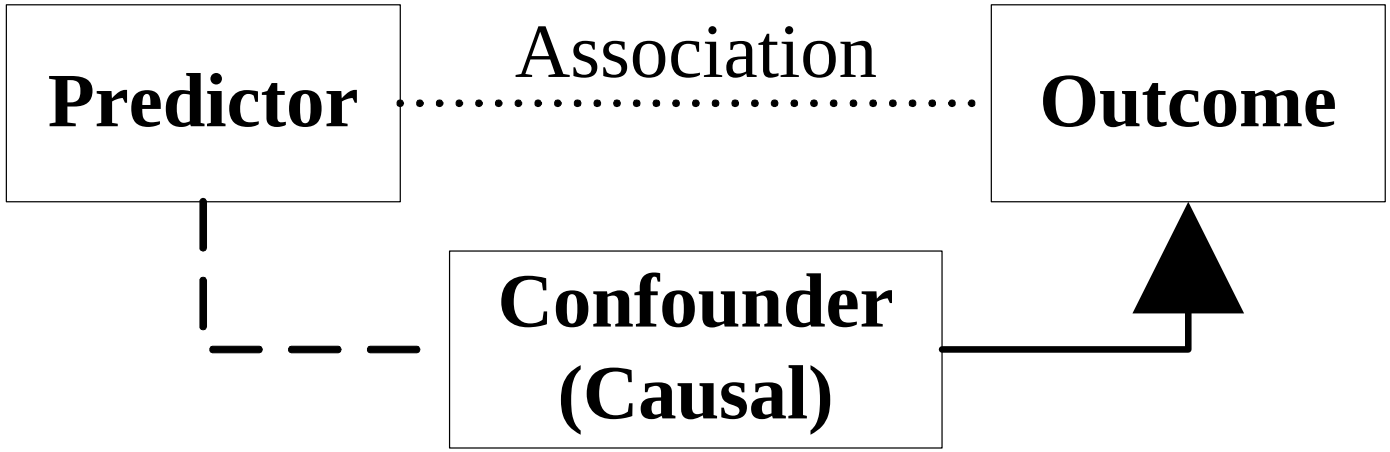
Male cigarette smokers have higher death rate from coronary artery disease than non-smoking males.

1964: "…more prudent from the public health viewpoint to **assume** that the established **association has causative meaning**, than to suspend judgment until no uncertainty remains."

2004: "placing it in the "suggestive" category … although it falls ***short of proving causation***, this evidence still makes causation ***more likely than not***."

Epidemiology & Context: Untangling Confounders





Relative Risk = 1.19
16% Attributable

1993: EPA & ETS (second-hand smoke):

- "3,000 American nonsmokers die each year from lung cancer caused by ETS.
- 150,000 to 300,000 children under 18 months of age get pneumonia or bronchitis.
- Second-hand smoke worsens condition of up to one million asthmatic children.

#4: Criteria for Causation Strength of Association

In Epidemiology, strength of association is typically measured using Relative Risk.

RR > 3 is a rule of thumb to avoid spurious results due to **confounding**. (Taubes)

RR > 10: lung cancer among smokers

RR < 2: problems from 2nd-hand smoke

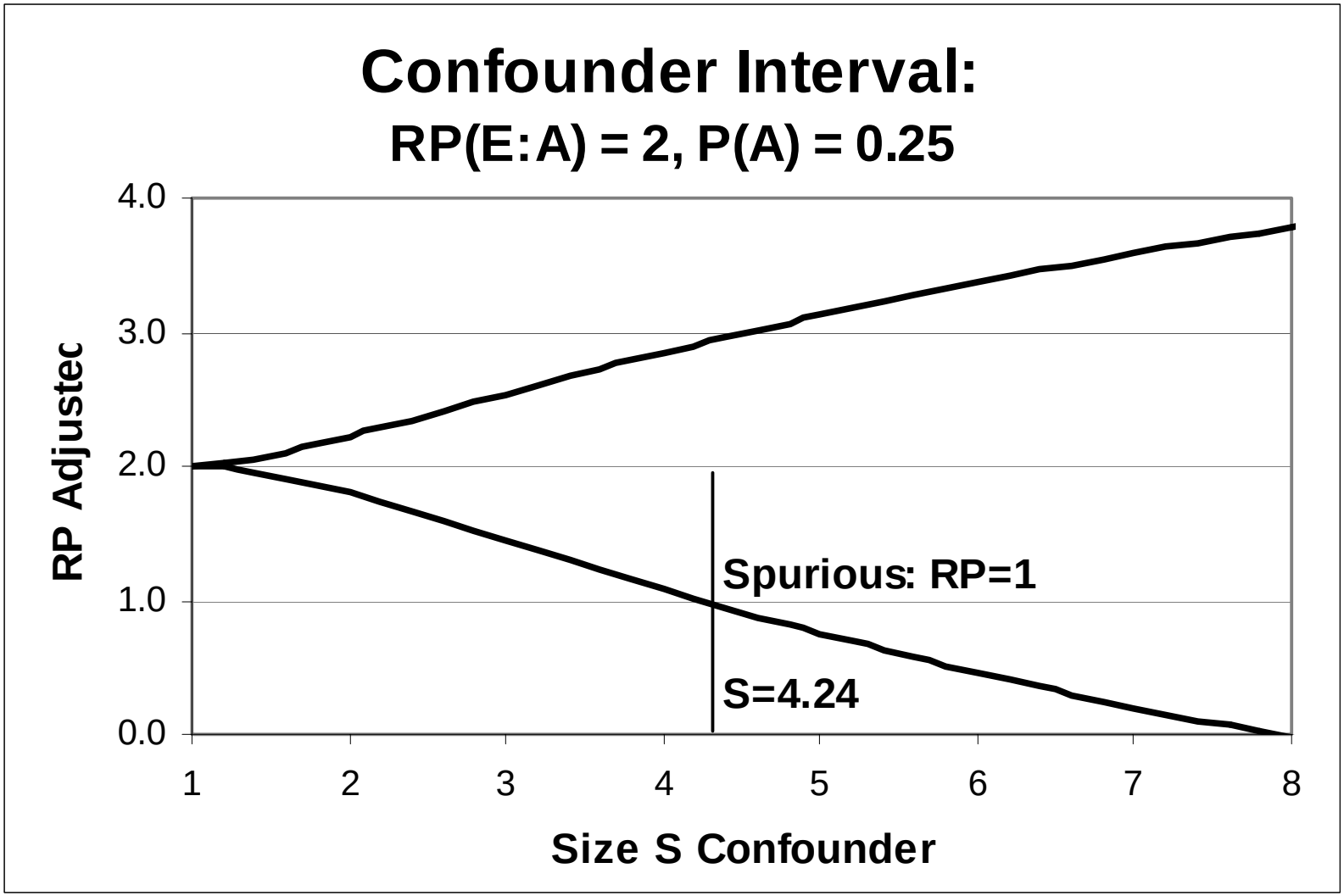
Problem Selecting Epidemiological Minimum

Sir Richard Doll: No single study is persuasive unless the lower limit of its 95% confidence level falls above a **threefold** increased risk.

As a rule of thumb, says Angell of the New England Journal, "we are looking for a relative risk of **3 or more"** before accepting a paper.

John Bailar, epidemiologist: "*There is no reliable way of identifying the dividing line*."

Model of a Confounder: Schield & Burnham (2004)



Goal: Help People make better Decisions using Data

Decisions based on observational data:

- Public health: quarantine, medical testing
- Education policy: best practices, league ratings
- Public policy (people): E.g., The Bell Curve, "More Guns; Less Crime"
- Public policy ($): Economics and finance
- Business: data mining and data modelling
- Personal Health: Vitamins, Supplements
- Personal Medical: HRT, chemo, radiation

Epidemiological Reasoning

Allows "inductive inference" without

- examining weight of evidence
- identifying plausible mechanisms

Enables government intervention (public policy) based on public health/safety grounds:

- bypasses individual rights & freedoms
- extends ideas of 'threat' & 'coercion'

Journalistic Assessment

Note as "Soft Science" (observational study)

- Not a manipulative experiment (no before/after)
- Not a statistical experiment (no random assign)

Evaluate strength of association:

- "Very weak": less than a factor of 2 (100% more)
- "Modest": factor of 2 to 3 (100% to 200% more)
- "Moderate: factor of 3 to 5 (200% to 400% more)
- "Strong": more than a factor of 5 (400% more)

Epi Reasoning in Education

"Pro and Con" Should be

- a chapter in intro logic/thinking texts
- a chapter in intro statistics texts.
- required in college general education.
- introduced at the secondary level
- used as a wedge for contextual thinking
- used as a wedge for "conceptual literacy" (inductive reasoning about core values)

Recommendation: Statistical Literacy

Study statistics used in **everyday** arguments

Students in introductory statistics need to

- focus on observational studies & decisions
- focus on context and confounding
- see statistics used to argue for causation
- see statistical prevarication & opportunism
- understand epidemiological reasoning.

Epidemiology References

2004 Report: www.cdc.gov/tobacco/sgr/sgr_2004/

Abrahamson, J. H (1994). *Making Sense of [Epidemiological] Data*. 2nd Edition, Springer Press.

McKim and Turner (1997) Editors. . Causality in Crisis: Statistical Methods and the Search for Causal Knowledge in the Social Sciences. Notre Dame Press

Brignell, John (2004). *The Epidemiologists: Have They Got Scares For You*. Amazon UK only; not Amazon US

Schield, Milo and Tom Burnham (2004). *Confounder Resistance and Confounder Intervals for a Binary Confounder*. Proceedings of ASA Section on Education