Weighing & Expressing Scientific Evidence

Barnett S. Kramer, M.D., M.P.H.

Office of Disease Prevention

National Institutes of Health

Disclosure/Disclaimer

- **No financial conflicts (no consulting or speaker honoraria, stocks, patents, or patent applications)**

- **Outside activity: editor-in-chief, JNCI**
  - **No connection with the NCI or Federal government**

- **Opinions are mine, not official positions of the Federal Government or National Institutes of Health**
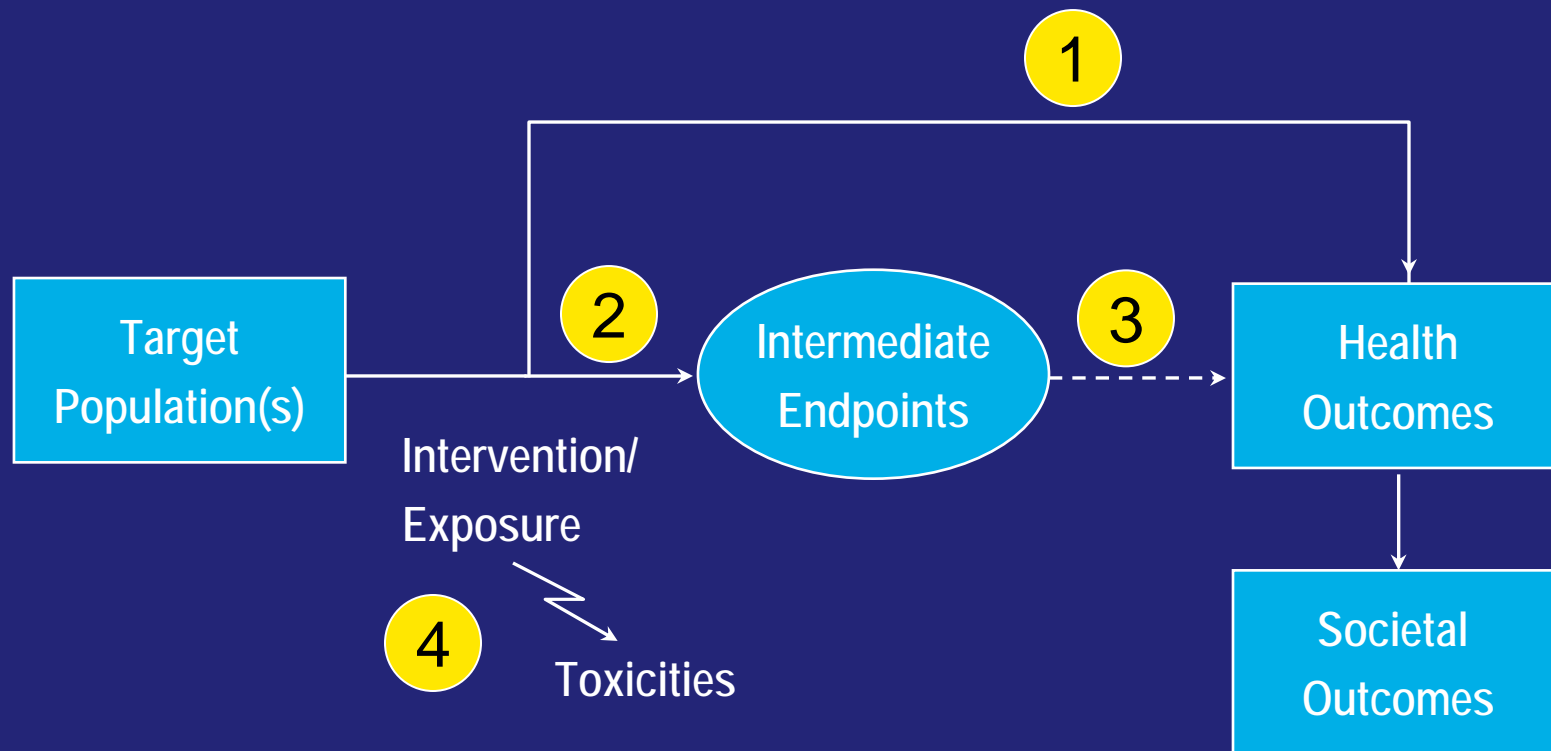
Levels of Decision Making

Level I: *"Would you have this done for yourself or for someone else in your immediate family?"*

Influenced by one's personal experience with the disease and capacity to deal with risk.

Affects few people.

Level II: *"What would I recommend to my patient/client?"*

Physician making a recommendation for his/her patients. Influenced by prior experience, but the scientific evidence may play a greater role.

Affects possibly hundreds of people.

Level III: *"What would I recommend to the nation, the world?"*

Across-the-board recommendations for a population.

Must be based on rigorous assessment of the scientific evidence.
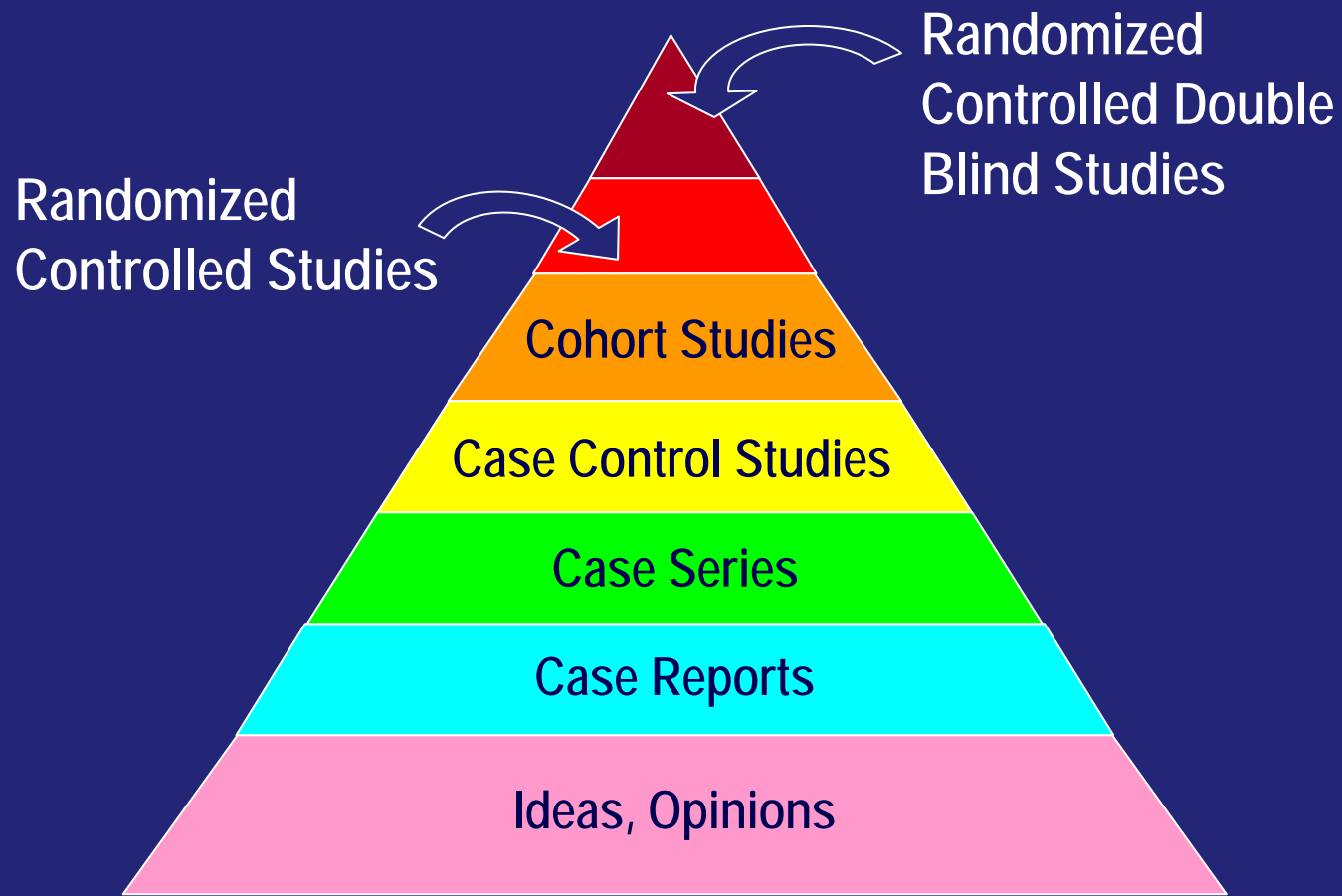
Affects hundreds of thousands, even millions of people.

I. Weighing Evidence

Questions to Ask About Medical Research

1. What is the exposure and what is the outcome?

2. How certain is it that the exposure causes the outcome?

3. How important is the outcome?

4. How big is the effect?

5. To whom does it apply?

Analytic Framework

Randomized Controlled Double Blind Studies

Randomized Controlled Studies

Cohort Studies

Case Control Studies

Case Series

Case Reports

Ideas, Opinions

Meta-analysis

SUNY Downstate's Evidence Based Medicine Course

Main Study Types

- **Observational**
  - Case reports
  - Case series (with or without historical controls)
  - Ecologic studies

- **Experimental**
  - Randomized, controlled studies
  - Randomized, controlled, double-blinded studies

Case Report

- Typically reports striking observations from a single person

  "A report of an aggressive brain tumor in a woman who frequently used a cellular phone suggested to the neurosurgeon that microwaves caused her cancer."

Uncontrolled Case Series

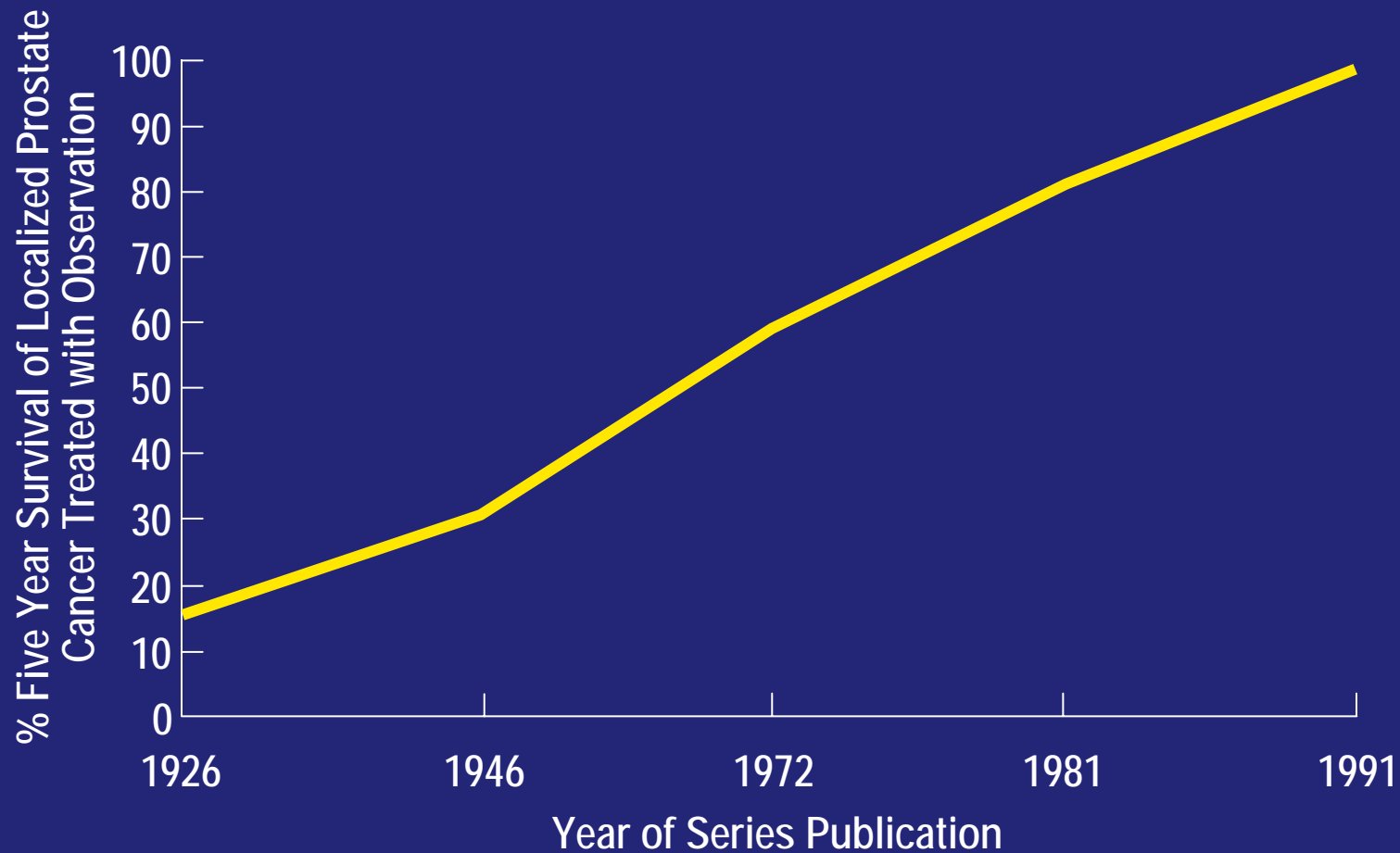■ Group of patients had the same exposure, but no comparison group

"An oncologist reported that a surprisingly high proportion of her kidney cancer patients reported living near electrical transformers…"

Case Series with Historical Controls

- Before-after study using past experience as a comparison group

"A study published today reports that rates of childhood brain tumors began to rise shortly after diet drinks came onto the market. Parents reported high intakes of diet cola in their children."

Improving 5-Year Survival of Patients with Localized Prostate Cancer Managed by "Watchful Waiting"

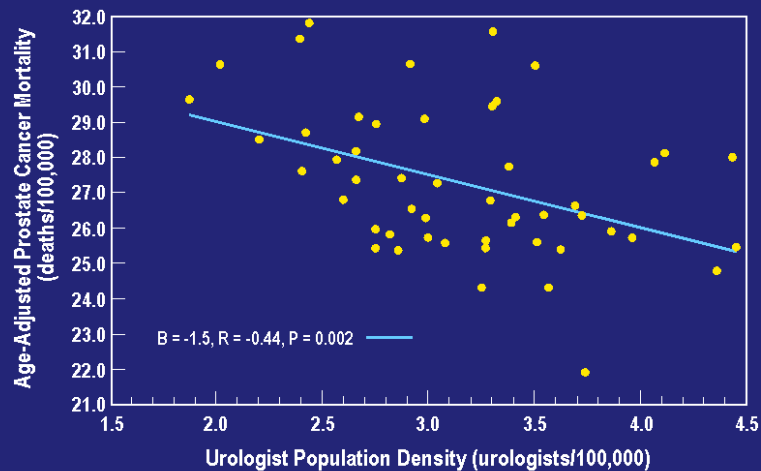(L. Thompson, AUA CME Book, 1998)

Ecologic Studies

- Studies in which the unit of analysis is populations and not individuals

   "People living in countries that eat Mediterranean diets are half as likely to suffer strokes than residents of countries that eat American-style diets. 'Something in the diet is protective,' said Dr. Ben Jones…"
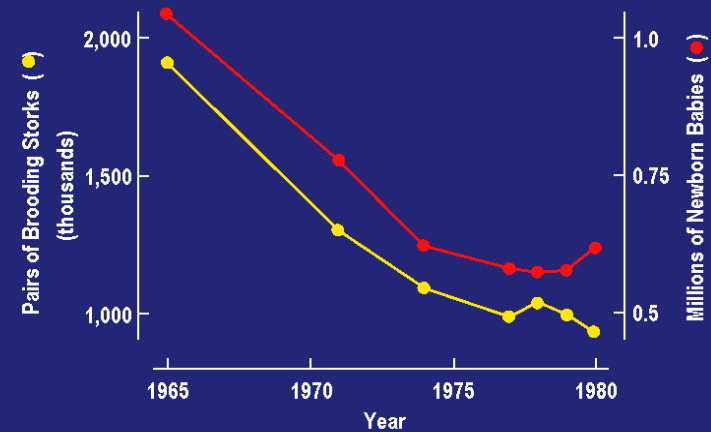
Ecologic Studies:
Support for Prostate Screening?

Prostate Cancer Mortality versus Urologist Population Density in the U.S.

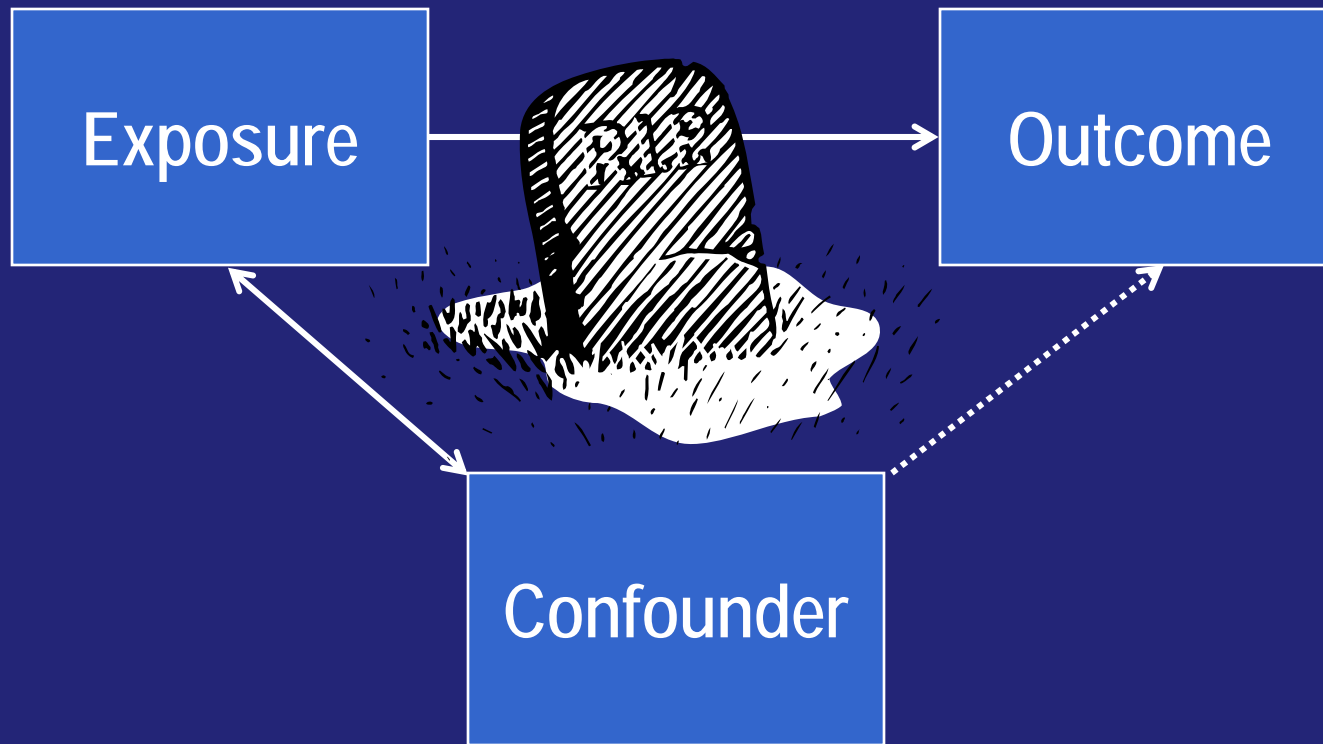B = -1.5, R = -0.44, P = 0.002

JL Colli and CL Amling, Prostate Cancer and Prostatic Disease, 2008

Brooding Stork Population Versus New Births in West Germany

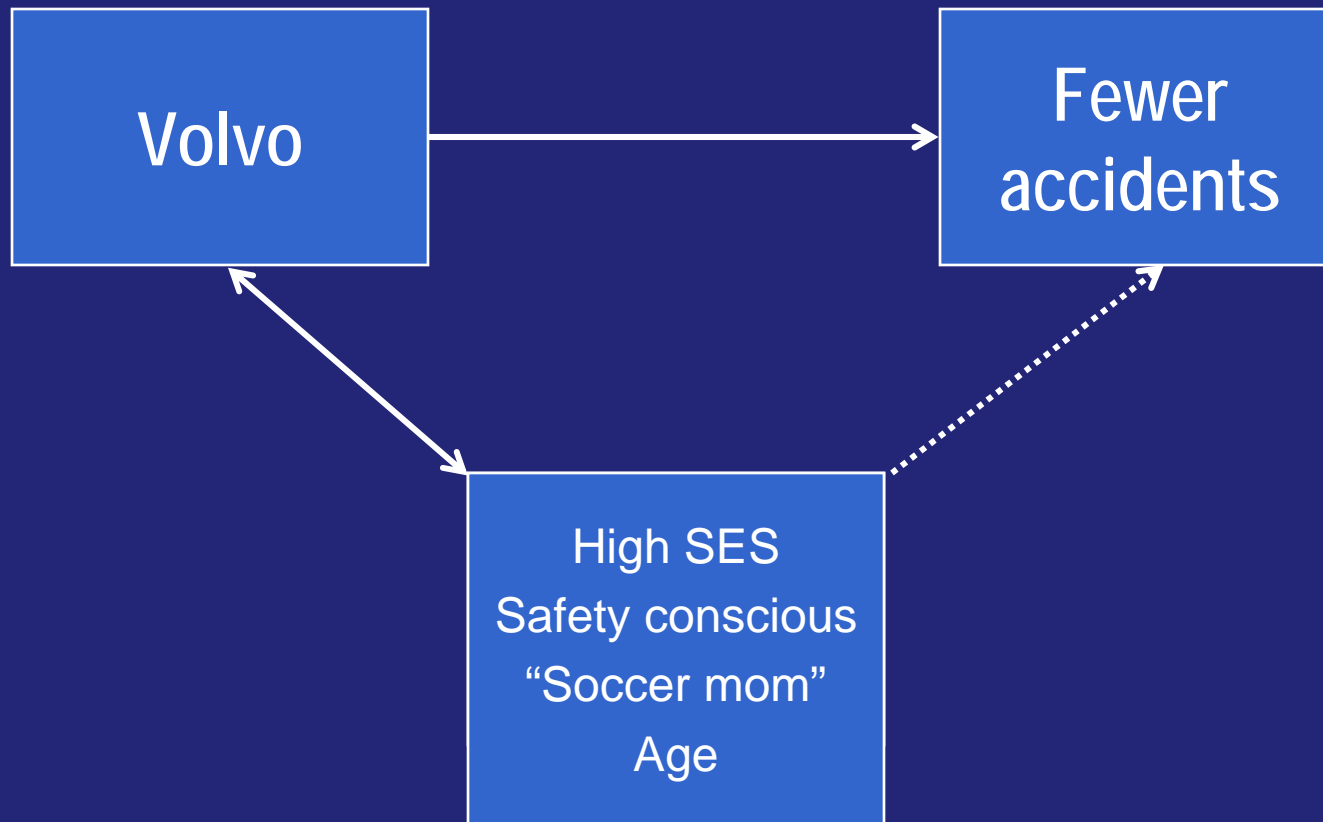Source: Helmut Sies, *Nature*, April 1988

Confounding Variables
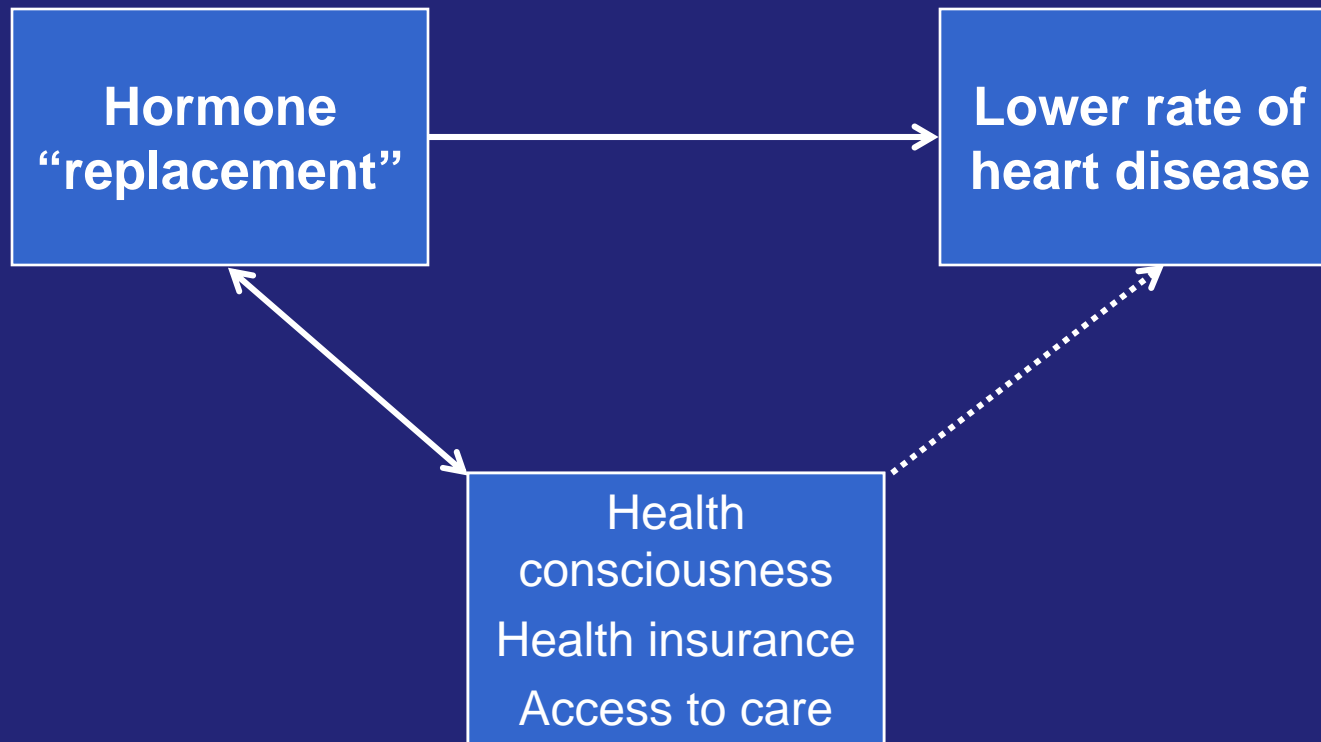

Exposure → Outcome

Confounder

Confounding is the death of any study!

Confounding Variables

Confounding Variables

Hormone "replacement" → **Lower rate of heart disease**

Health consciousness
Health insurance
Access to care

Confounding is a concern in any observational study!

"[They] are guilty until proven innocent."

David Ransohoff, JNCI, 2006

A Closer Look at Outcomes

Figure adapted from HG Welch

Surrogate Endpoint:

A non-health outcome or biomarker used to draw conclusions about the effect of an intervention or exposure on a health outcome observed later in time.

Validated Surrogate Endpoint:

A surrogate endpoint that would yield the same inference about the effect of treatment or exposure as the health outcome.

Advantages of a Validated Surrogate Endpoint in Prevention

Could guide further research

Could permit shorter, more efficient trials (in theory)

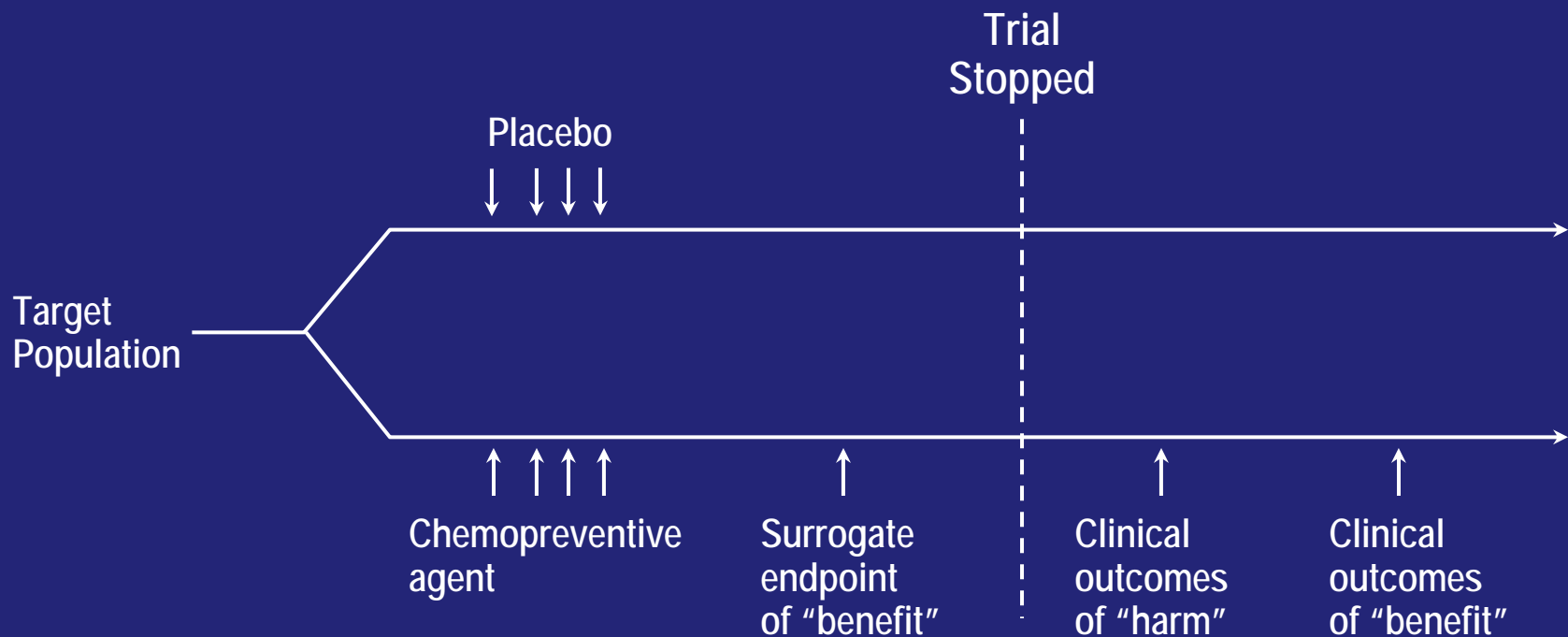But could introduce "noise" if the surrogate infrequently progresses to a health outcome

PSA

polyps

mammographic breast density

Challenges to Surrogate Endpoint Use

- Even statistically "validated" surrogate endpoints only predict for one health outcome

- Chemotherapeutic agents may have mixed effects on health outcomes

Pitfalls of Surrogate Endpoints for Cancer Prevention

B. Levin, JNCI, 2003

Discordance Between Surrogate Endpoint Biomarkers and Health Outcomes

Intervention	Effect on Surrogate Endpoint		Effect on Health Outcome	
Encainide, flecainide	↓	Cardiac arrhythmias (PVCs)	↑	Sudden Death
Postmenopausal Estrogen + Progestin	↓ ↓ ↑	Cholesterol LDL HDL	↑	Coronary heart disease

Discordance Between Surrogate Endpoint Biomarkers and Health Outcomes (2)

Intervention	Effect on Surrogate Endpoint	Effect on Health Outcome
Low fat, high vegetable diet (Women's Health Initiative)	↓ **Polyps**	⟷ **Colon Cancer**
Torcetrapib	↑ **HDL cholesterol**	↑ **Deaths, heart failure**

Summary of Surrogate Endpoint Use

- Biomarker "surrogate" endpoints have important uses

    - Drug development: screen for promising agents

    - Provide justification for large scale trials

- But challenges for use as endpoints in definitive clinical trials remain

    - Few (none??) can replace health outcomes at present

    - May give a qualitatively wrong result

    - Do not capture both health benefits and harms

II. Expressing Evidence

"Level of Evidence" for Cancer Prevention & Screening Physician Data Query (PDQ)

■ **Definition:** certainty of the editorial board's estimate of the health effects of implementing an intervention

■ **Steps**

    I. Description of the evidence (5 Domains)

    II. Summary assessment for both benefits and harms

Description of Evidence in PDQ: Five Domains

1. Study design: ranked by design strength

2. Internal validity: "quality" of execution within study design (good, fair, poor)

3. Consistency (coherence)/volume of evidence
   - One vs. multiple studies
   - Small vs. large studies
   - Consistent direction of outcomes

4. Magnitude of effects: prefer absolute vs. relative effects
   - Change from 1% to 0.5%, or from 4/1000 to 2/1000 [Not: 50% decrease]

5. External validity (good, fair, poor)
   - Applicability in usual practice with same effect?

PDQ Summary on Hormone Replacement Therapy and Breast Cancer

Based on solid evidence, combination hormone replacement therapy is associated with an increased risk of developing breast cancer. The evidence concerning the association between estrogen-only therapy and breast cancer incidence is mixed.

Study Design: Randomized controlled trials

Internal Validity: Good

Consistency: Good

External Validity: Good

Summary

- Beware of surrogate endpoints
    - They are harder to validate than most think

- Look to the strongest level of evidence in generating public policy (Level III decisions)

- Express evidence as explicitly as possible for the public

"A wise man proportions his belief to the evidence."

David Hume, Scottish philosopher, 1711-1776