

22 August 2011 ISI 1

## Epidemiological Models Generate Spotty Statistics

---

**MILO SCHIELD,**  
*Augsburg College*

*US Rep: International Statistical Literacy Project*  
*VP: National Numeracy Network*  
*Webmaster: www. StatLit.org*  
**August 22, 2011**  
[www.StatLit.org/2011SchieldISI6up.pdf](http://www.StatLit.org/2011SchieldISI6up.pdf)

22 August 2011 ISI 2

### 1. Linear-model statistics are found in the media


---

- “the rate of [Alzheimer’s] decline unfolded 4 percent more quickly *for each* additional year of education.” Reuters 10/22/2007.
- “*For every* can or glass of sugar-sweetened beverage a child drank [a day] ..., a child’s ... chance of becoming obese increased 60%.” The Lancet, 2001; 357:505-508.
- “*each* hour of television watched per day at ages 1-3 increases the risk of attention problems, such as ADHD, by almost 10 percent at age 7.” Science Daily 4/6/2004.

22 August 2011 ISI 3

### 2. Epidemiological-model statistics are also found in the media

---



22 August 2011 ISI 4

### 2. How Epidemiological Models Generate Statistics

---

Percentage *Attributed* to Exposure = 100% \*  
 $(\text{RateExposed} - \text{RateControl}) / \text{RateExposed}$

Consider these death rates:

- 10% for a city hospital (4% for a rural hospital).

Based on this, one can say

- “60% of the deaths at city hospital are *attributed to that hospital.*”
- Excess / Higher =  $(10\% - 4\%) / 10\% = 60\%$

22 August 2011 ISI 5

### 2. How Epidemiological Models Generate Statistics #2

---

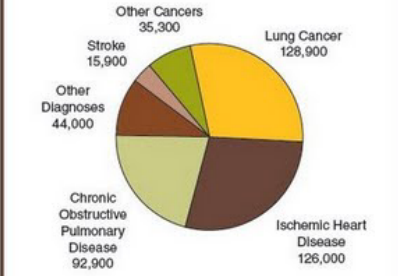
If lung-cancer deaths are 10 times as prevalent for smokers as for non-smokers, then **90%** of smoker lung-cancer deaths are *attributable* to smoking.

Suppose the *coroner-certified* lung-cancer deaths among smokers number **100,000**.

The lung-cancer deaths among smokers *attributable* to smoking number 90,000: **90% of 100,000**.

22 August 2011 ISI 6

### About 443,000 U.S. Deaths Attributable Each Year to Cigarette Smoking\*



Category	Number of Deaths
Lung Cancer	128,900
Ischemic Heart Disease	126,000
Chronic Obstructive Pulmonary Disease	92,900
Other Diagnoses	44,000
Stroke	15,900
Other Cancers	35,300

\* Average annual number of deaths, 2000-2004.  
Source: MMWR 2008;57(45):1226-1228.

22 August 2011 ISI

**Responsible for ...**

---

**Study blames diesel for deaths**  
 By Jon Brodtkin / Daily News Staff  
 Wednesday, February 23, 2005

**D**iesel pollution is responsible for more deaths than drunk drivers and homicides, according to a new study that estimates how many premature deaths, asthma attacks and heart attacks are caused by diesel pollution in every U.S. county.

Nationwide, diesel pollution causes 21,000 premature deaths each year, including 475 in Massachusetts and 81 in Middlesex County, robbing those who die of an average of 14 years of their lives.



22 August 2011 ISI

8

**2. Epidemiological statistics are common!**  
**US Annual Deaths Attributable To:**

---

Smoking: 467,000	Blood pressure: 395,000
Overweight: 216,000	Inactivity: 191,000
Blood sugar: 190,000	LDL cholesterol: 113,000
Dietary salt: 102,000	Low omega-3 : 84,000

High dietary trans fatty acids: 82,000  
 Alcohol use: 64,000 (90,000 less 26,000 averted)  
 Low intake of fruits and vegetables: 58,000  
 Low poly-unsaturated fatty acids: 15,000

www.emaxhealth.com/2/24/30740/  
 smoking-high-blood-pressure-obesity-top-preventable-death-causes.html

22 August 2011 ISI

9

**2. Web Articles Involving Deaths**  
**[Google matches]**

---

### WEB SEARCH	### WEB SEARCH
7 M deaths <b>associated with</b>	591 K deaths attributed
87 K deaths connected with	246 K deaths attributable
6 M deaths linked to	65 K attributable deaths
5 M deaths because of	384 K unnecessary deaths
<b>39 M deaths due to</b>	363 K excess deaths
28 M deaths <b>caused by</b>	1.2 M premature deaths

Google Search (4/2011) Exact Phrase

22 August 2011 ISI

10

**2. Epidemiological statistics**  
**are easily generated**

---

Two rates: complementary wholes and a common part:

**Percentage attributable:**

- Graduation: % attributable to having educated parents
- Dropouts: % attributable to having a single-parent
- Incarceration: % attributable to having low IQ

**Cases attributable:**

- Premature babies attributable to mom being a smoker
- Auto accidents attributable to having DWI conviction
- Deaths attributable to living in US vs. Mexico

22 August 2011 ISI

11

**Causes; Account for**

**YAHOO! NEWS**

HOME U.S. BUSINESS WORLD ENTERTAINMENT SPORTS TECH POLITICS ELECTIONS

U.S. Video Local News Education Religion Politics Crimes and Trials

**Junk food causes a third of heart attacks** 

Mon Oct 20, 4:08 pm ET

 WASHINGTON (Reuters) – Diets heavy in fried foods, salty snacks and meat account for about 35 percent of heart attacks globally, researchers reported on Monday.

Their study of 52 countries showed that people who ate a "Western" diet based on meat, eggs and junk food were more likely to have heart attacks, while those who ate more fruits and vegetables had a lower risk.

**USA TODAY Your Life Health**

Your Life Health Fitness & Food Family & Parenting Sex & Relationships

Alternative medicine Alzheimer's Breast cancer Cold & flu Depression Diabetes

**Coffee, sex, smog can all trigger heart attack, study finds**

By Steven Reinberg HealthDay Updated Feb 26, 2011 11:00 AM | 106 | 2 |  

A major analysis of data on potential triggers for heart attacks finds that many of the substances and activities Americans indulge in every day — coffee, alcohol, sex, even breathing — can all help spur an attack.

 Because so many people are exposed to dirty air, air pollution while stuck in traffic topped the list of potential heart attack triggers, with the researchers pegging 7.4% of heart attacks to roadway smog.

But coffee was also linked to 5% of attacks, booze to another 5%, and pot smoking to just under 1%, the European researchers found.

22 August 2011 ISI 13

## 2. Epidemiological statistics encourage seductive grammar

---

Consider these titles of news stories:

- 45,000 deaths *attributable to* uninsurance
- 45,000 deaths *associated with* lack of insurance
- Lack of insurance *linked to* 45,000 deaths
- 45,000 die ... *because of* lack of health insurance
- Lack of Health Insurance *Kills* 45,000 a Year
- Lack of Health Insurance *cause* 44789 deaths
- Lack of insurance *to blame* for almost 45,000 deaths

Source: [www.StatLit.org/pdf/2010SchieldICOTS.pdf](http://www.StatLit.org/pdf/2010SchieldICOTS.pdf)

22 August 2011 ISI 14

## 3. Epidemiological models generate “spotty” statistics

---

Model-based numbers are sensitive to the assumptions. Epidemiological models are extremely sensitive to what is – and is not – taken into account.

“**Spotty**” statistics are statistics generated by epidemiological models: they conceal more than they reveal.

“**Spotty statistics**” look like real statistics, but they are model-based. They look like they identify real causes, but they merely indicate the size of an association.

22 August 2011 ISI 15

## 3. “Spotty” statistics are very sensitive to assumptions

---

**2004: 400,000 deaths attributable to obesity.**  
*“Obesity might soon pass smoking as the country's leading cause of preventable death.”*

**2005: 27,000 deaths attributable to obesity.**  
 This finding **transformed obesity from fearsome killer to pitiable also-ran**, ranked in 7th place.

What changed? Perhaps they took something into account that is closely associated with death rates. Something such as age ... ☺

22 August 2011 ISI 16

## 3. “Spotty” statistics are hard to detect

---

Spotty statistics – epidemiologically-based statistics – are common – but hidden – in the news.

- They have no unique grammar or keywords.
- They look plausible -- coroners might count them.
- We treat counts as facts.
- Journalists and politicians don't question them.
- We don't question them.

22 August 2011 ISI 17

## 3. Statistical education should focus more on “spotty statistics”

---

By remaining silent on these “spotty” statistics, **the most common type of model-based statistics in the everyday media**,

statistical education may be seen as an archaic subject stuck in the 20<sup>th</sup> century fields of Fisher's Rothamsted agricultural station.

22 August 2011 ISI 18

## 4. Statistical Literacy focuses on numbers in the everyday media

---

**Statistical literacy** is the ability to read and interpret summary statistics in the everyday media: in graphs, tables, statements, surveys and studies.

Statistical literacy studies the use and misuse of “spotty statistics” in the everyday media.

#### 4. Statistical Educators Should Promote Statistical Literacy

---

Almost 20% of US four-year colleges offer a course in statistical literacy.

Statistical educators may say  
“*This is not our job.*”

**But if not you, then who?**

**If not now, then when?**

#### Bibliography

---

Schild, Milo (2009). Confound Those Speculative Statistics. *2009 American Statistical Association Proceedings of the Section on Statistical Education. [CD-ROM] P. 4255-4266.* See [www.StatLit.org/pdf/2009SchildASA.pdf](http://www.StatLit.org/pdf/2009SchildASA.pdf)

Schild, Milo (2010). Assessing Statistical Literacy: TAKE CARE in *Assessment Methods in Statistical Education: An International Perspective*. Edited by P. Bidgood, N. Hunt and F. Joliffe. Wiley Publishers, Ch. 11, p. 133-152. Excerpts: [www.statlit.org/pdf/2010SchildExcerptsAssessingStatisticalLiteracy.pdf](http://www.statlit.org/pdf/2010SchildExcerptsAssessingStatisticalLiteracy.pdf)