

July 2010 ICOTS8 1

Association-Causation Problems in News Stories

Milo Schield
W. M. Keck Statistical Literacy Project
Augsburg College

Intl. Conference on Teaching Statistics
Ljubljana Slovenia ICOTS-8 12 July 2010
www.StatLit.org/pdf/2010SchieldICOTS6up.pdf

July 2010 ICOTS8 2

Topic

This paper investigates related problems – inaccuracies, omissions and ambiguities – in number-based news stories.

It investigates causation-related problems involving causation words and action verbs.

July 2010 ICOTS8 3

1: Confusing Million and Billion

In a three year period there were 23 mix-ups of the words *million* and *billion* in the Los Angeles Times and 38 mix-ups in the New York Times.

Analysis: This is a big mix-up! Data is needed on how readers understand these big numbers.

In a convenience survey, 200 Augsburg College students were asked “*How big is a billion?*” They said: 1,000 million (59%), 100 million (18%), a million-million (10%), 10 million (7%) and “Don’t know” (6%).

July 2010 ICOTS8 4

2: Confusion of the Inverse

AP: 9/30/09. **Too much candy could lead to prison**

LONDON, England — Willy Wonka would be horrified. Children who eat too much candy may be more likely to be arrested for violent behavior as adults...

*Of children who ate candies daily at age ten, **69% were arrested** for violent offenses by age 34.*

The real statistic:

69% of those arrested for violent offenses by age 34 ate candies daily at age ten.

July 2010 ICOTS8 5

3: Missing Context

Soft Drinks Could Boost Pancreatic Cancer Risk.
*People who down two or more soft drinks a week may have **double the risk** of developing deadly pancreatic cancer, compared to non-soda drinkers.*

But the overall number of people developing the malignancy remains low...”

Estimated risk: 14 per 100,000
Doubling this risk means ONE more person in a group of 7,200 soda drinkers may develop pancreatic cancer.

July 2010 ICOTS8 6

A: Times Less Comparisons

*Circumcised men were **two to three times less likely** to contract HIV. AFP 5/28/2009.*

*Patients who had end-of-life talks were **three times less likely** to spend their final week in intensive care ...*

Times Less Than:
 Math/stat: Three times less than 6 is minus 12
 Non-math: Three times less than 6 is 2
Statistical educators and journalists need to talk !!!

July 2010 ICOTS8 7

B: Times More Comparisons

Teenagers who spent more than 100 hours swimming in chlorinated pools were up to six times more at risk of having asthma [12%] than other teens [2%].

Retired [NFL] players are 19 times more likely to struggle with memory problems [1.9%] than similarly aged men who never played professional football [0.1%].

Times More/Less Than:
 Math/stat: Six times more than 2% is 14%.
 Non-math: Six times more than 2% is 12%.
Statistical educators and journalists need to talk !!!

July 2010 ICOTS8 8

Incomplete Comparisons

LA Times: Jan 31, 2010. Opinion: Doug Smith:
The NY Times was more likely to overstate the case
 a. “than to understate the case”
 b. “than the LA Times was.”

More doctors like Crest ...
 “than [they like] any other toothpaste”
 “than nurses like Crest”

Most doctors like Crest ... [Hidden comparison]

July 2010 ICOTS8 9

C: Incomplete Comparisons with multiple groups

1. *Taller People Earn More Money. Reuters 8/25/2006.*
 “Earn more” than what group: shorter people.
 Single groups are no problem.

Multiple groups are a problem:

2. *Obese women less likely to be screened for cancer.*
 “Less likely” than obese men or non-obese women?

3. *Study: Young girls more likely to be fat. AP 2007*
 “More likely” than older girls or young boys?

July 2010 ICOTS8 10

D: Confusing “Frequently” with “Likely”

1. *1995 Honda Civic: Most Frequently Stolen Car. State Farm Insurance. 7/9/2008.*

2. *New car study lists most likely to be stolen – ‘96 Honda Civic. Mountain Times 8/27/2009*

3. *Study: Cadillac Escalade most likely stolen. AP*

How can two cars both be “most likely” to be stolen?
 Confusion between “frequently” and “likely”
Frequently compares counts; Likely compares ratios.
 #1 and #3 are correct; #2 is wrong.

July 2010 ICOTS8 11

E: Slope-based Comparisons

Slope compares: As X increases, Y increases/decreases.
Each Daily Soda Increases Obesity Risk 60%.
 For every can or glass of sugar-sweetened beverage a child drank [a day] ..., a child’s ... chance of becoming obese [as an adult] increased 60%.

60% sounds big – but 60% of what?
 If the chance of obesity is 5% for non-drinkers, then a 60% increase means an 8% chance.
 60% gets more attention than 3 percentage points!

July 2010 ICOTS8 12

Distinguish Causation from Association

Causation (8%): *cause, effects, results, prevents*
 Association (2%): *associate, relate, correlate,*

Between (67%):
Action verbs: ups, cuts, raises, boosts, increases ...
Other: due to, because of, attributed to

Inappropriate use of “causes”:

- Obesity **causes** later onset of puberty in boys
- *Junk food causes a third of heart attacks.*
- *Obesity growing to be top cancer cause.*

July 2010 ICOTS8 13

“Between” Words in Headlines

- *Study: Estratest doubles breast cancer risk*
- *Gene increases depression risk: study.*

In these cases, before-after studies are impossible.

- *Weddings boost mood: study.*
- *Expanding waist worsens kids' sleep apnea.*

In these cases, before-after studies are possible

Journalists should distinguish these two cases.

July 2010 ICOTS8 14

Distinguish ‘Due to’ From ‘Result of’ or ‘Caused by’

**US Healthcare:
Third Leading Cause of Death**

225,000 Americans die each year die as a result of their medical treatments:

- 7,000 deaths/year due to hospital medication errors
- 12,000 deaths/year due to unnecessary surgery
- 20,000 deaths/year due to other errors in hospitals
- 80,000 deaths/year due to infections in hospitals
- 106,000 deaths/year due to negative effects of drugs

Reference: Starfield, B. (2000, July 26). Is US health really the best in the world? *Journal of the American Medical Association*, 284(4), 483-485.

July 2010 ICOTS8 15

Confusing Association with Causation

1. Study: 45,000 Uninsured *Die* a Year (CBS News)
2. 45,000 deaths *attributable to* uninsurance
3. 45,000 US deaths *associated with* lack of insurance
4. No health coverage *tied to* 45,000 deaths a year
5. Lack of insurance *linked to* 45,000 deaths
6. Study: 45,000 U.S. Deaths *From* Lack of Insurance
7. One death every 12 minutes *due to* no health insurance
8. 45,000 ... die *because of* lack of health insurance
9. Lack of Health Insurance *Kills* 45,000 a Year
10. Lack of Health Insurance *cause* 44,789 deaths
11. Lack of insurance *to blame for* almost 45,000 deaths

August 2009 ASA 16

Conclusion

More research is needed to see how journalists and educated adults understand the difference between association and causation, and the various devices used to indicate this difference.

Statistical educators, journalism faculty and quantitative journalists should join together to analyze statistical illiteracy in the everyday media **so they can give better guidance to journalists.**

August 2009 ASA 17

Background

These slides were presented at ICOTS-8 in Ljubljana, Slovenia. They accompanied the paper: *Association-Causation Problems in News Stories*. This paper is available at:

www.StatLit.org/pdf/2010SchildICOTS.pdf