Introduction: Not everything you read is true

David Schlesinger, Ph.D.

Lars Leksell Gamma Knife Center
University of Virginia

University of Virginia Health System

Department of Radiation Oncology

Conflicts of Interest

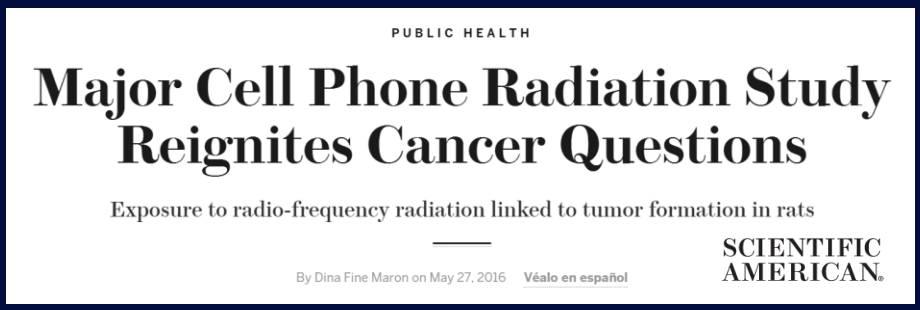
Research support: Elekta Instrument, AB

Session Educational objectives

1. Learn about the presence of statistical problems in published studies
2. Identify common signs and symptoms of potential problems in various types of statistical tests
3. Learn methods for correctly implementing statistical analyses of the type commonly found in clinical publications

The truth is hard to come by

Cellphone use causes cancer



PUBLIC HEALTH

Major Cell Phone Radiation Study Reignites Cancer Questions

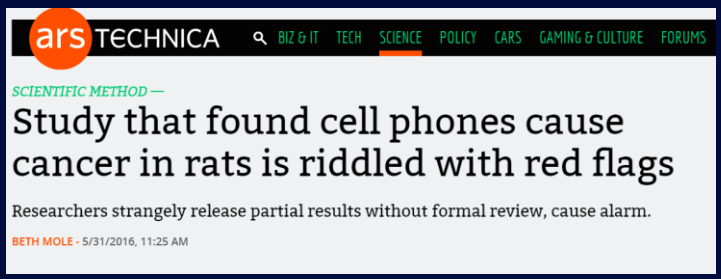
Exposure to radio-frequency radiation linked to tumor formation in rats

By Dina Fine Maron on May 27, 2016 Véalo en español

SCIENTIFIC AMERICAN.

http://www.cancer.gov/about-cancer/causes-prevention/risk/radiation/cell-phones-fact-sheet

Cellphone use causes cancer – maybe?



ars TECHNICA BIZ & IT TECH SCIENCE POLICY CARS GAMING & CULTURE FORUMS

SCIENTIFIC METHOD —

Study that found cell phones cause cancer in rats is riddled with red flags

Researchers strangely release partial results without formal review, cause alarm.

BETH MOLE - 5/31/2016, 11:25 AM

- Study released before peer-review
- Control rats showed less than expected natural rate of tumor incidence and died early
- Incidence of tumor development correlates with age
- Early control death magnified the statistical findings

http://arstechnica.com/science/2016/05/study-that-found-cell-phones-cause-cancer-in-rats-is-riddled-with-red-flags/

Human studies are mostly one-sided

Publication Year	Study	Type	# participants	Outcome
2010	Interphone Study Group	Case-control study	~5000 cases; ~5000 matched controls; 13 countries	No overall risk*
2001 (updated 2007, 2011)	Danish cohort study	Cohort study	358,000	No association
2013 (updated 2014)	Million Women Study	Prospetive cohort study	791,710	Yes (acoustic neuroma), then no association
2014	CERENAT	Multicenter case control	447 cases, 892 matched controls	No association with regular use; yes association with heaviest use
2011	Swedish pooled analysis	Pooled analysis of 2 case control studies	1251 cases, 2438 controls	Increased risk of glioma

The result of studies of thousands of animals and hundreds of thousands of people is that we have no definitive answer to the question of cellphone use and cancer.

So….

How confident can we be about studies like this:

SRS for lung cancer: Does morning or afternoon make a difference?

Abstract

BACKGROUND: Circadian cell-cycle progression causes fluctuating radiosensitivity in many tissues, which could affect clinical outcomes. The purpose of this study was to determine whether outcomes of single-session gamma knife radiosurgery (GKRS) for metastatic nonsmall cell lung cancer (NSCLC) differ based on treatment time.

METHODS: Fifty-eight patients received GKRS between 10:00 am and 12:30 pm and 39 patients received GKRS between 12:30 pm and 3:00 pm. The mean peripheral dose was 18.6 Gy. The mean tumor size was 7.3 cm^3. Magnetic resonance imaging was used to score local control at 3 months. Cause of death (COD) was categorized as central nervous system (CNS)-related or systemic.

RESULTS: Demographic and disease characteristics of the 2 groups were similar. Local control at 3 months was achieved in 97% (35/36) of patients who underwent GKRS early in the day versus 67% (8/12) of patients who underwent GKRS later in the day (chi-square, P = .014). Early GKRS was associated with better survival (median 9.5 months) than late GKRS (median 5 months) (Kaplan-Meier log-rank test, P = .025). Factors contributing to better survival in a Cox regression model included early treatment time (P = .004) and recursive partition analysis class (P < .001). Cause of death in the early treatment group was CNS-related in 6% (3/47) of patients versus 24% (8/34) of patients in the late treatment group (chi-square test, P = .026).

CONCLUSIONS: GKRS for metastatic NSCLC had better local control, better survival, and a lower rate of CNS-related cause of death when given earlier in the day versus later in the day. These retrospective data should encourage future study in brain radiosurgery and non-CNS stereotactic body radiotherapy series.

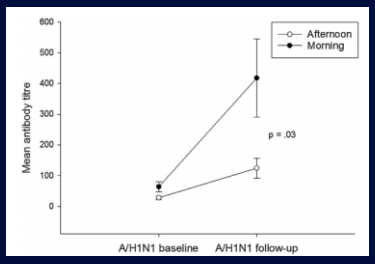
D. Rahn, et al., Cancer 177(2), 2011.

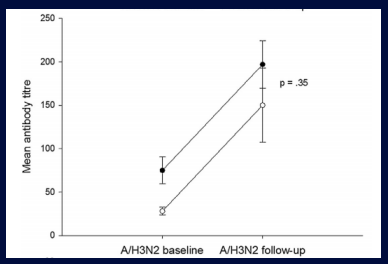
But…maybe we were onto something in this case…..

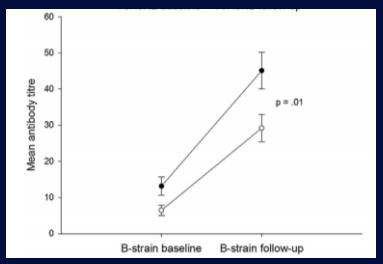
Vaccine. 2016 May 23;34(24):2679-85. doi: 10.1016/j.vaccine.2016.04.032. Epub 2016 Apr 26.

Morning vaccination enhances antibody response over afternoon vaccination: A cluster-randomised trial.

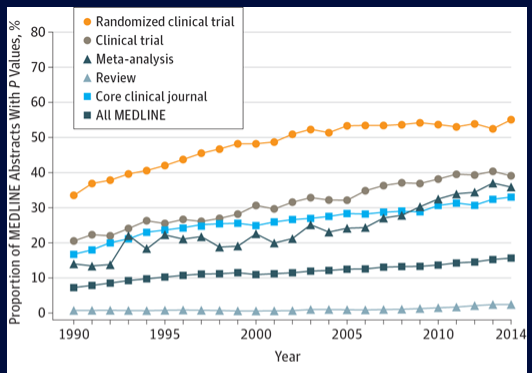
Long JE[1], Drayson MT[2], Taylor AE[3], Toellner KM[2], Lord JM[4], Phillips AC[5].







Medicine increasingly relies on p-values

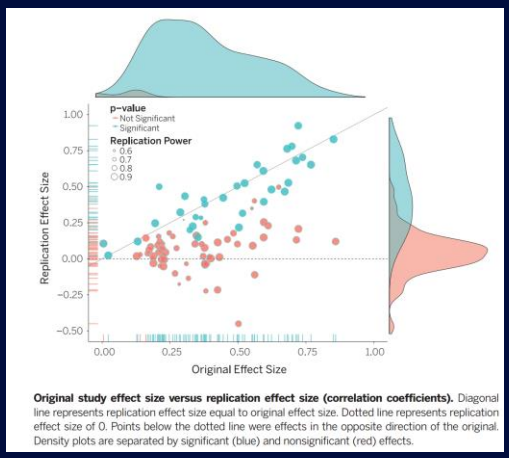


Article Category	Abstracts, No.	Abstracts Reporting *P* Values, No. (range per year)
Randomized clinical trial	332 262	165 968 (2056-12 101)
Clinical trial	197 921	60 317 (736-3700)
Meta-analysis	52 639	16 018 (30-2960)
Review	1 396 965	17 125 (176-2245)
Core clinical trial	796 103	210 045 (2907-10 090)
All MEDLINE	12 821 790	1 608 736 (20 709-138 218)

http://jama.jamanetwork.com/article.aspx?articleid=2503172

What we think is the truth often can't be replicated

One study's result is not necessarily the truth



Original study effect size versus replication effect size (correlation coefficients). Diagonal line represents replication effect size equal to original effect size. Dotted line represents replication effect size of 0. Points below the dotted line were effects in the opposite direction of the original. Density plots are separated by significant (blue) and nonsignificant (red) effects.

Open Science Collaboration, Science, 49, 2015.

RESEARCH ARTICLE

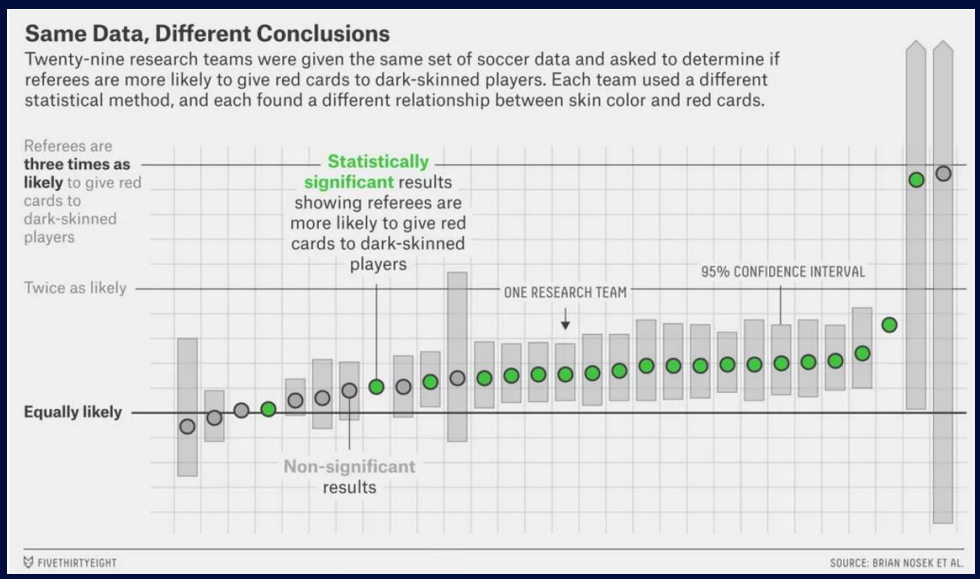
PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration*†

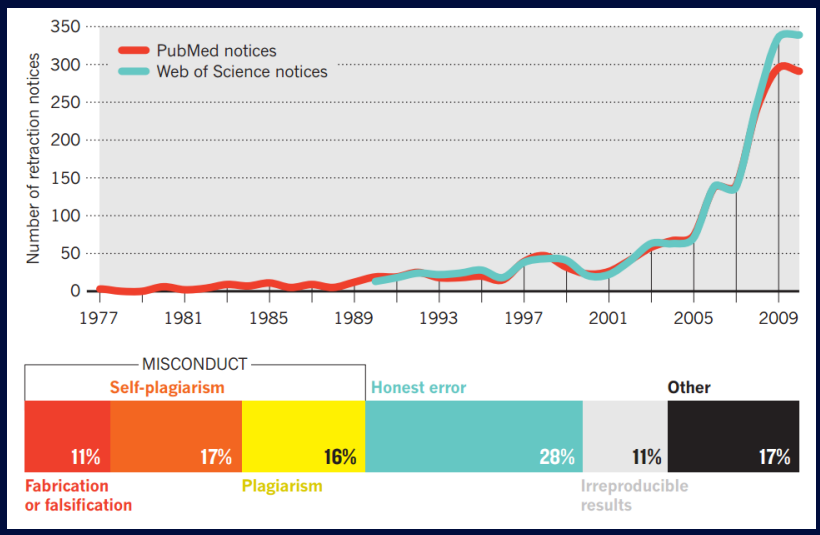
http://news.harvard.edu/gazette/story/2016/03/study-that-undercut-psych-research-got-it-wrong/)

http://projects.iq.harvard.edu/psychology-replications/



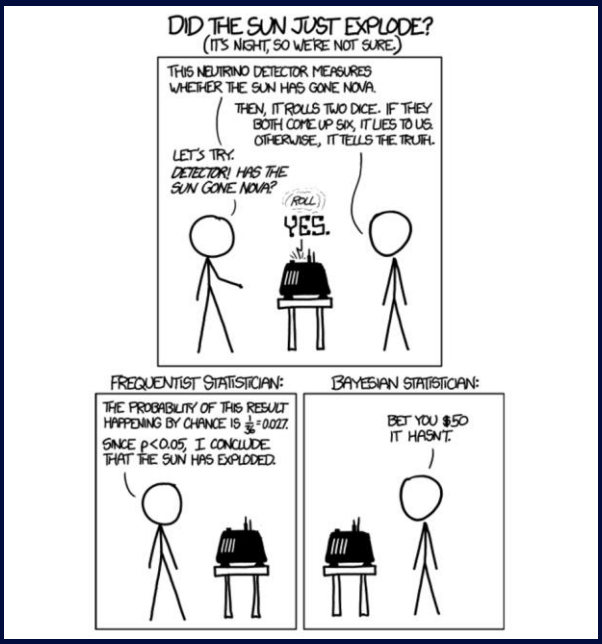
http://fivethirtyeight.com/features/science-isnt-broken/#part2

The number of retractions is sharply rising



R. Van Noorden, Nature, 478, 2011.

A lack of statistical fluency may be part of the problem



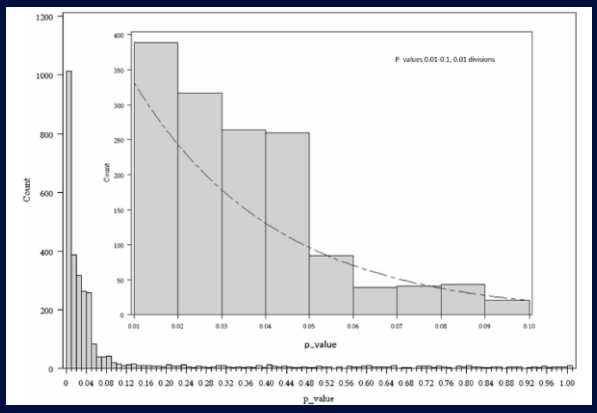
HTTP://XKCD.COM/1132

Many medical physicists receive little training in practical statistics as applied to clinical outcomes studies

However…these studies are at the heart of our profession.

How to recognize when the statistics don't quite add up?

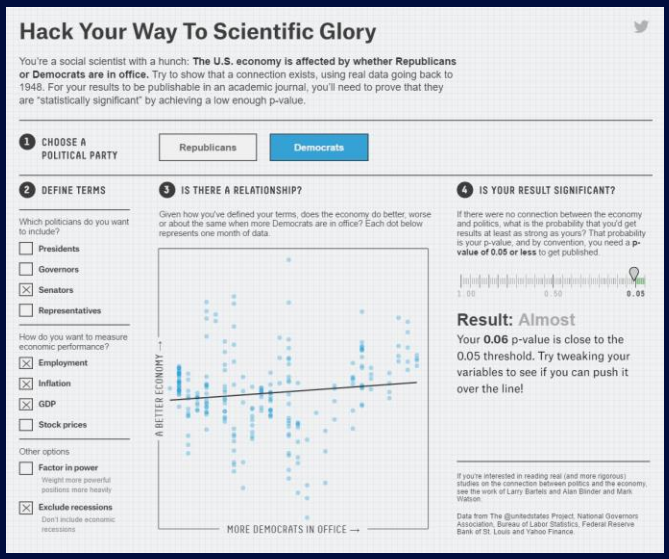
p-values just below 0.05 are over-represented



The distribution of probability values in medical abstracts: an observational study

Bastiaan Ginsel[1*†], Abhinav Aggarwal[1†], Wei Xuan[2] and Ian Harris[1†]

There are many ways to achieve a desired story



http://fivethirtyeight.com/features/science-isnt-broken/#part2

But…we can learn to be better