The misuse of statistics is unethical

Tony Panzarella, MSc, PStat
Assistant Professor and MSc Program Director,
Division of Biostatistics, Dalla Lana School of Public Health,
University of Toronto

UNIVERSITY OF TORONTO
DALLA LANA SCHOOL OF PUBLIC HEALTH

Objective

- To argue that the misuse of statistics is unethical

BRITISH MEDICAL JOURNAL

Medicine and Mathematics

Statistics and ethics in medical research

Misuse of statistics is unethical

DOUGLAS G ALTMAN

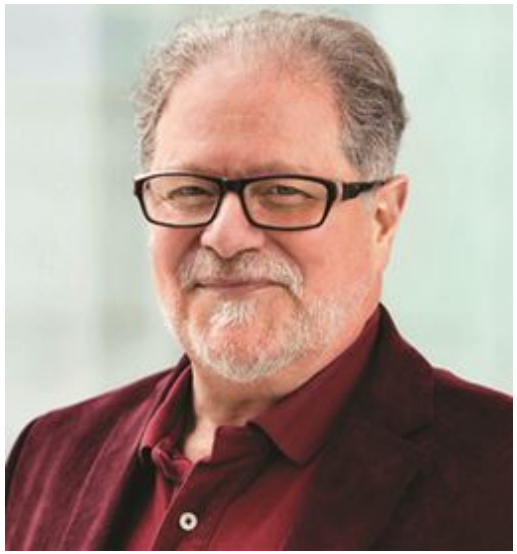
A brief introduction

- Biostatistician who worked over 32 years in a clinical cancer research environment
- Former member of UHN Oncology REB (> 15 years)
- Former member of OCREB (6 years)
- Professionally accredited statistician by the Statistical Society of Canada since 2008

REB Chairs that I have served under

Ron Heslegrave

Jack Holland



Richard Sugarman

Content Areas of Ethical Review

1. Is the research project asking a reasonable or important question?
2. **Will the research project, as designed, answer the question being asked?**
3. Are the procedures which research participants must undergo acceptable?
4. Will compensation be made to research participants?
5. Is adequate care being taken over the research participants' confidentiality?
6. What procedures will be followed to seek the consent of the research participants?

Foster and Holley, 1998

Content Areas of Ethical Review

- The second question demonstrates that the assessment of scientific validity is also required as part of the process of ethical review
- In a survey of IRBs in the US 'poor study design' was the next most frequent reason after 'improperly designed consent form' for protocols requiring revision (Jones et al, 1996)
- For some applications, the REB review may be the only opportunity for assessing the scientific validity of the proposal.

Statistics and medical ethics

- It is well appreciated that ethical considerations may affect the design of an experiment
  - In clinical trials (CT), for example, we cannot carry out controlled trials of cigarette smoking
  - As another example, we could not study the effect of seatbelt use on fatality in vehicle collisions using a CT design.

Altman DG *BMJ* 1980

The usual situation

Ethics

affects

Statistics

Ethical study

Appropriate design

Statistics and medical ethics

- Far less frequently discussed is how statistical aspects of a study affect ethics

This talk

Ethics

Statistics

Unethical

Bad Statistics

Statistics and medical ethics

- It is unethical to carry out bad scientific experiments
- Statistical methods are one aspect of this. If the statistical methods are substandard than the research will be unethical.

- There are 2 principal reasons for this:
  - Misuse of patients (or animals) and other resources
  - It is incorrect to publish results that are incorrect or misleading

Statistics and medical ethics

- Misuse of patients (or animals) and other resources

- "…one of the most serious ethical problems in clinical research is that of placing subjects at risk of injury, discomfort, or inconvenience in experiments where there are too few subjects for valid results, too many subjects for the point to be established, or an improperly designed random or double-blind procedure."

WW May *J Med Ethics* 1975

Statistics and medical ethics

- It is incorrect to publish results that are incorrect or misleading

- Errors in the use of statistics may occur at all stages of an investigation, and one error can be sufficient to render the whole exercise useless.

Quality of statistics in papers in medical journals

- Schor and Karten, JAMA (1966)
- 149 papers reporting analytic studies in several journals
  - 28% were judged acceptable
  - 67% were deemed deficient
  - 5% totally unsalvageable

The editor of JAMA wrote therefore

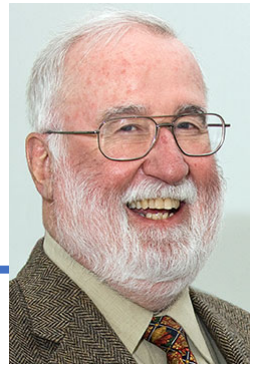
- "The study is an indirect argument for greater knowledge and appreciation of statistics by the medical author, for a reiteration on his part that the biostatistician is not a worrisome censor, but a valuable ally, and that biostatistics, far from being an unrelated mathematical science, is a discipline essential to modern medicine – a pillar in its edifice."

Statistics and medical ethics

- There are many ways in which the statistical content of research may be deficient

Biases in analytic research

J. Chronic Dis 1979



David Sackett

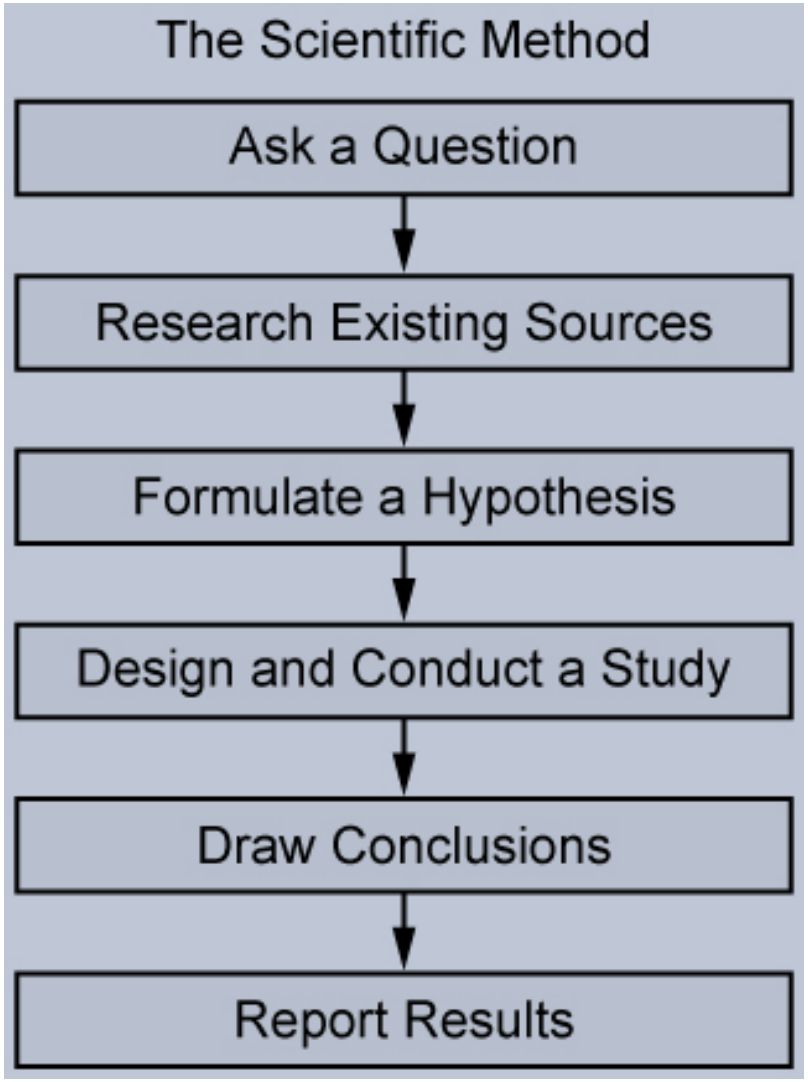
- Planning (5)
- Design (22)
- Execution / Data collection (18)

66%

- Data Analysis / Presentation (5)
- Interpretation / Publication (6)

"This distribution reflects … the relative seriousness of statistical errors at each stage, and indicates where there is greatest need for statistical expertise"

Altman DG *BMJ* 1980


The Scientific Method
Ask a Question
Research Existing Sources
Formulate a Hypothesis
Design and Conduct a Study
Draw Conclusions
Report Results

The role of statistics in the scientific method

- Statistics is essential to good experimental design
  - e.g. randomized clinical trials

- Statistics is important in analyzing and interpreting data

Experimental design

- To obtain a rigorous test of a hypothesis it is important to obtain data that can provide evidence for or against a hypothesis

- If its to be concluded that drug A is more effective than drug B the results must be statistically significant

Experimental design

- If the sample size is too small the results are unlikely to be statistically significant
  - Suppose subjects are subjected to health risks
  - If the experiment is not well designed such that no meaningful conclusions can be drawn, the potential benefits will not outweigh the potential harms
  - In providing informed consent subjects usually believe the research is valuable and may advance science

Analysis and interpretation

- It is relatively easy to misuse statistical methods
  - Applying a method correctly involves
    - a good understanding of the variables used,
    - the sampling process, and
    - a sound understanding of the theory and assumptions underlying that method
- Statistics applied correctly can distinguish between random noise in the data and the real signal

Analysis and interpretation

- Misuses of statistics can also occur in the absence of erroneous or distorted results

- Misuse can arise from a failure to provide the research community with important information about the methods used
  - Handling of outliers
  - Handling of missing data

Misuse

- Misuse, for our purposes, is an “incorrect use”

- Not all misuses have equivalent ethical implications
  - Misconduct
  - Negligence
  - Incompetence

Misconduct

- Intentional deception
  - Fabrication (making up data or results)
  - Falsification (manipulating, changing, or omitting data or results)
  - Honest error is never misconduct

Negligence

- Excessive errors due to haste, ignorance, or sloppiness

Incompetence

- Lacking the proper statistical knowledge

Results of careless or deceptive use of statistics

- Poor science
- Wasting time and energy of other researchers
- Funds could have been allocated to more deserving projects

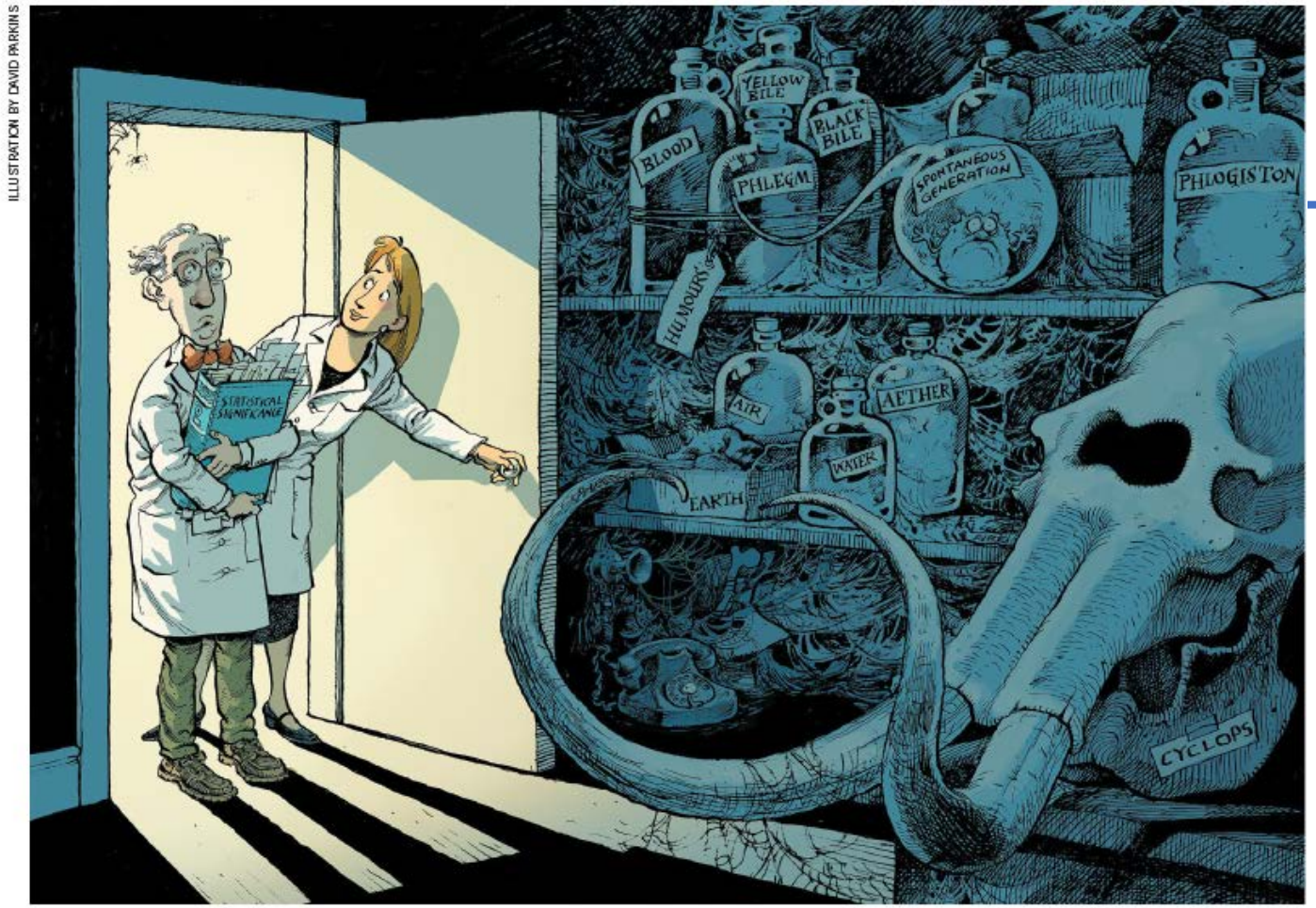
Contributing factors to misuse

- Pressures to publish, obtain grants
- Career ambitions or aspirations
- Conflicts of interest and economic motives
- Inadequate supervision, education, or training
- User-friendly statistical software packages
- Journals tend to publish only "significant" results (i.e., $P < 0.05$)

Question #1

- Which one of the following do you feel contributes most to the misuse of statistics by researchers:

A. Pressures to publish, obtain grants
B. Career ambitions or aspirations
C. Conflicts of interest and economic motives
D. Inadequate supervision, education, or training
E. User-friendly statistical software packages
F. Journals tend to publish only "significant" results (i.e., $P < 0.05$)



Retire statistical significance

Valentin Amrhein, **Sander Greenland**, **Blake McShane** and more than 800 signatories call for an end to hyped claims and the dismissal of possibly crucial effects.

(*Nature*, March 2019)

You have probably heard this before

- "…there was 'no difference' between two groups because the difference was 'statistically non-significant'?

Statistics Notes

Absence of evidence is not evidence of absence

Douglas G Altman, J Martin Bland

What the authors stated

- For example, consider a series of analyses of unintended effects of anti-inflammatory drugs.

- Because their results were statistically non-significant, one set of researchers concluded that exposure to the drugs was "not associated" with new-onset atrial fibrillation (the most common disturbance to heart rhythm) and,

- that the results stood in contrast to those from an earlier study with a statistically significant outcome.

What the data showed

- The statistically non-significant results found a risk ratio of 1.2 (that is, a 20% greater risk in exposed patients relative to unexposed ones).

- They also found a 95% confidence interval that spanned everything from a trifling risk decrease of 3% to a considerable risk increase of 48% ($P = 0.091$).

- The researchers from the earlier, statistically significant, study found the exact same risk ratio of 1.2. That study was simply more precise, with an interval spanning from 9% to 33% greater risk ($P = 0.0003$; our calculation).

- “It is ludicrous to conclude that the statistically non-significant results showed “no association”, when the interval estimate included serious risk increases; it is equally absurd to claim these results were in contrast with the earlier results showing an identical observed effect.”

Question #2

- A controlled trial of a new treatment led to the conclusion that it is significantly better than placebo: $p < 0.05$. Which of the following statements do you prefer

A. It has been proved that the treatment is better than placebo
B. If the treatment is not effective, there is a less than a 5% chance of obtaining such results
C. The observed effect of the treatment is so large that there is a less than 5% chance that the treatment is no better than placebo
D. None of the above
E. I do not really know what a *p*-value is and do not want to guess

Question #3

- Given its rampant misinterpretation how do you feel about banning the use of the term "P < 0.05" in research publications?

A. Strongly Agree
B. Agree
C. Somewhat Agree
D. Somewhat Disagree
E. Disagree
F. Strongly Disagree

What's the problem?

- Reliance on thresholds of statistical significance can mislead us

Ethical Guidelines for Statistical Practice, ASA 1994

- Responsibility of individuals
- Responsibility of those employing statisticians
- Guidelines become tools for research integrity when they are integral to actual practice

CODE OF ETHICAL STATISTICAL PRACTICE

Introduction

The mission of the Statistical Society of Canada (SSC) encourages the development and use of statistics and probability. To promote high standards in statistical practice in Canada, practitioners must demonstrate ability and professional integrity. Hence, the SSC has adopted this Code of Ethical Statistical Practice to guide its accredited members in their everyday professional life. All accredited members must agree to abide by this Code of Ethical Statistical Practice as a necessary condition of their accreditation.

Authority

The authority for the SSC Code of Ethical Statistical Practice derives from its formal adoption by the SSC on March 20, 2004. The Society is thus committed to uphold these principles with respect to the conduct of its accredited members.

Elements of Ethical Statistical Practice

The elements have been grouped into four major areas of responsibility that define ethical statistical practice.

A. Responsibility to Society

1. Conform to procedures that protect human rights and dignity. In particular, ensure that the collection and storage of information and the publication of results adhere to relevant privacy laws or privacy standards set out by the SSC or other relevant bodies.
2. Strive to advance public knowledge and understanding of information by the application of appropriate statistical methods and interpretation of results, and by providing assistance in discrediting false or misleading information.
3. Maintain objectivity and strive to avoid procedural or personal bias. The creation of valid data-based information is vital to informed public opinion and policy.
4. Acquire appropriate knowledge and understanding of relevant legislation, regulations and standards in the practitioner's field of application and comply with these requirements.

Statistical Society of Canada

B. Responsibility to Employers and Clients

1. Carry out and document work with due care and diligence in accordance with the requirements of the employer or client.
2. Avoid disclosure or authorization to disclose, for personal gain or benefit to a third party, confidential information acquired in the course of professional practice without the prior written permission of the employer or client, or as directed by a court of law.
3. Declare any interest, financial or otherwise, that could be perceived as influencing the outcome of work undertaken for a client or employer.
4. Advise clients or employers of any potential or actual conflict between the ethical standards of statistical practice and the interests of the client or employer.
5. Exercise care to prevent the use of any misleading summary of the data. Strive to ensure that all assumptions and limitations relevant to the data, the analysis and the results are fully disclosed.

C. Responsibility to Other Statistical Practitioners

1. Uphold the reputation of statistical practice and seek to improve professional standards by participating in their development, use and enforcement. Avoid any action that will adversely affect the good standing of Statistics and Statisticians.
2. Refrain from speaking in the name of the Society without the authorization by the Executive of the Society.
3. Encourage and support fellow statisticians in their professional development and, wherever possible, encourage recruitment and provide opportunities for new entrants to the profession.
4. Act with integrity toward fellow statisticians and other professionals, avoiding any activity that is incompatible with professional standards. Ensure that due credit is ascribed to fellow professionals. While question and debate are encouraged, criticism should be directed toward procedures rather than persons. Avoid publicly casting doubt on the professional competence of others.

D. Professionalism

1. Adhere to the Guidelines for Professional Development, seeking to upgrade professional knowledge and skills and to be informed of technological developments, procedures and standards relevant to the field of application.
2. Seek to exercise recognized good practice, upholding quality standards and encouraging fellow-practitioners to do likewise.
3. Only undertake work and provide services that are within the limits of professional competence; and do not lay claim to any level of competence not possessed.
4. Accept responsibility for work and give objective and reliable information on procedures in any professional review or assessment.
5. Refuse to enter into, or comply with, any arrangement where financial or other rewards are contingent upon the outcome of a proposed statistical inquiry.

Summary

This code establishes basic principles to help practitioners maintain the highest standards of professional conduct and describes the behaviour that peers may expect from their statistical colleagues. An accredited member's behaviour is expected to conform always with the expectations of informed, respected and experienced peers. In response to a complaint, the procedures set out in the documentation of the SSC Accreditation Committee will apply.

For more information, please contact the SSC:

info@ssc.ca
Tel.: (613) 733-2662
Fax: (613) 733-1386
www.ssc.ca

Statistical Society of Canada

TCPS REB membership – minimum requirements

Tri-Council Policy Statement
Ethical Conduct for Research Involving Humans
TCPS2 2018
Canadian Institutes of Health Research
Natural Sciences and Engineering Research Council of Canada
Social Sciences and Humanities Research Council

- The REB shall consist of at least five members, including both men and women, of whom:
- **at least two members have expertise in relevant research disciplines, fields and methodologies covered by the REB**;
- at least one member is knowledgeable in ethics;
- at least one member is knowledgeable in the relevant law (but that member should not be the institution's legal counsel or risk manager). This is mandatory for biomedical research and is advisable, but not mandatory, for other areas of research; and
- at least one community member who has no affiliation with the institution.

Statistician membership in REBs in Canada

- Cross-sectional survey, March 2003
- 55% response rate (77/140)
- **78%** (59/76; 95% CI 0.69-0.86) of the REBs reported having no statistician in their membership.
- All 59 who answered "no" reported they had never had a statistician

L. Thabane et al. 2005

Table 1.
Distribution of REBs Who Feel That They Do Not Need a Statistician

	Have a statistician	*Do not have a statistician*
Need a statistician	8	8
Do not need a statistician	4	50

Table 2.
Frequency Distributions of How REBs without a Statistician Deal with Statistical Issues

Response (n = 59)	*Frequency: Count (%)*
Within committee	30 (51)
Approach a statistician if necessary	26 (44)
Prior review by research and development committee	12 (20)
Other	8 (14)

Table 3.
Reasons Respondents Give for Not Having a Statistician on the REB

"There would be no disadvantage in having one."
"One is preferred but not always available."
"Occasionally it would be helpful."
"I think it is necessary to have at least a researcher who is an expert within statistical issues."
"By not having a statistician, we acknowledge that our methodological competence is not complete."
"We have tried (and failed) to recruit statisticians."
"Ethical decisions do not depend on such detailed scrutiny."
"We have extensive statistical training."
"Certainly such expertise is important in evaluating methodology in an assessment of cost-benefit analysis…With limited resources it is a struggle to simply get people to serve."
"…now that you mention it!"

Table 4.
Distribution of Statistical Issues in REB Checklists

Statistical Issue: (n = 21)	*Frequency: Count*
Study design and methodology	13
Study population/sample size	6
Randomization	5
Data analysis plan	3
Statistical justification for study design	2
Methods of data collection	2

Example 1 – Correlation does not confer causation

- Serum beta carotene and the incidence of lung cancer

- Observational studies showed a negative correlation i.e. individuals with higher levels of beta carotene had lower incidence of lung cancer, suggesting that individuals at risk for lung cancer would benefit from increased serum beta carotene levels

DeMets D 1999

Example 1 – Correlation does not confer causation

- This hypothesis was tested in three large RCTs
- ATBC trial
- CARET trial
- Physicians Health Study
- Baseline serum values of beta carotene were in fact predictive of lung cancer incidence.
- However, those who were randomized to receive beta carotene developed both a *higher* incidence of lung cancer and mortality compared to those who did not receive beta carotene.

Example 1 – Correlation does not confer causation

- Claiming a treatment benefit on this basis may inflict useless or harmful interventions on patient populations

Example 2 – Power of a study

- Freiman JA et al. 1978 conducted a survey of clinical trials published in *NEJM* and found 71 studies which claimed to find no treatment effect.
- Yet 50 of the 71 could have missed a 50% treatment effect due to inadequate power and small sample size
- These trials were inadequate tests of their respective treatments

Example 2 – Power of a study

- Patients participated in these poorly designed studies, thus wasting their contribution.

- The situation has likely improved, at least for trials that address major health problems, and that are funded by government or private industry, where there is peer review or regulatory review

Example 3 – CAST study

- Finding new effective treatments as quickly as possible is a goal that researchers and patients both share, especially for diseases which have high morbidity or mortality
- Studies may require large sample sizes and take years to complete.
- To speed up the research process and reduce costs, researchers have turned to surrogates for the clinical outcome

Example 3 – CAST study

- In cardiology patients with cardiac arrhythmias are at a higher risk for sudden death.

- Drugs were developed to suppress the arrhythmia. While not initially approved for broad use the practice of cardiology accepted the effects of these drugs on the surrogate, arrhythmia, to infer clinical benefit in preventing or reducing the risk of sudden death

Example 3 – CAST study

- The Cardiac Arrhythmia Suppression Trial (CAST was a randomized placebo-controlled trial designed to test the effect of three of the arrhythmia suppressing drugs in a large population of patients with cardiac arrhythmia).

- The trial was terminated abruptly when a safety monitoring committee observed a dramatic increase in all-cause mortality, and in sudden death for those patients on the arrhythmia suppressing drugs

What went wrong?

- Perhaps the causal pathway is different than initially believed
- Perhaps the drugs were acting through more than one mechanism

Conditions for an adequate surrogate

1. The surrogate outcome must correlate well with the clinical outcome
2. The effect of the treatment on the surrogate must predict the effect on the clinical outcome

This latter condition is often not met because either the wrong causal pathway has been identified or the treatment is acting on more than one pathway

The impact

- “Unfortunately, these drugs had been widely used for care of patients with arrhythmias beyond the initial approval indication without appropriate further testing and undoubtedly many patients were harmed. Without CAST, this practice may have been continued for a considerable period of time.”

DeMets D 1999

Final Thought

- "Medical research must be constantly vigilant about preventing such misuse. However, misuse is not so easy to detect so that ethical and honest use of the methods is essential. Just as no one may be able to check if you obey a stop sign late at night in a residential neighbourhood, rarely will anyone be able to double check in detail our statistical analyses and design. Ultimately, the proper use of statistical methods must rest with the research integrity of each individual."

Dave DeMets

References

- Altman DG. Statistics and Ethics in medical research. *BMJ* 281: 1182-1184, 1981.
- Cooper JA, McNair L. What is the Role of a Statistician on the IRB? *Journal of Empirical Research on Human Research Ethics* 9: 89-90, 2014.
- DeMets DL. Statistics and Ethics in Medical Research. *Science and Engineering Ethics* 5:97-117, 1999.
- Gardenier JS, Resnik DB. The misuse of Statistics: Concepts, Tools, and a Research Agenda. *Accountability in Research* 9:65-74, 2002.
- Thabane L, Childs A, Lafontaine A. Determining the level of Statistician Participation on Canadian-Based Research Ethics Boards. *IRB: Ethics and Human Research* 27: 11-14, 2005.