Role of a statistician in regulatory decision-making; views of a clinician

EFSPI conference, Basle

John Johnston, Clinical Assessor

September 2022

Using 'simple' language:

<mark>**a. Present the facts:**</mark>

There is a **point estimate 27% reduction** in the incidence rate ratio of falls (exercise group compared to the control group)

Incidence rate ratios of falls that are compatible with the data are **from 45% reduction to 17% increase** (95% CI)

<mark>**b. Wait / invite the clinician to form an opinion of the facts**</mark>

Using 'simple' language:

c. Now give your opinion on the facts:

The 'true' incidence rate ratio may be higher or lower than 1; the **data are inconclusive**

The 95% compatibility interval appears moderately wide; there is **moderate uncertainty**

d. invite clinician opinion

Using 'simple' language:

e. Now add to your opinion:

Overall, data are considered inconclusive and there is moderate uncertainty.

Further evidence is needed to evaluate the effect of a programme of exercise on reduction in falls because the association now shown may be affected by uncontrolled biases and / or other error.

Using 'simple' language:

<mark>**f. Now agree on the conclusion:**</mark>

It is possible that an exercise program in adult subjects with QRST disease is associated with reduction in falls; further study is warranted.

<mark>**g. extra:**</mark>

You may remind the clinician that he / she will need to address uncertainties not captured by the statistics as well as the potential costs and benefits of such a programme of exercise on the incidence of falls.

Medicines & Healthcare products Regulatory Agency

Thank you.

About us

We are the Medicines and Healthcare products Regulatory Agency (MHRA). We are the regulator of medicines, medical devices and blood components for transfusion in the UK. We are responsible for making sure these products meet set standards for safety, quality and efficacy (effectiveness).

We are an Executive Agency of the Department of Health and Social Care. We are the UK National Control Laboratory and the World Health Organization Collaborating Centre and International Laboratory for Biological Standards.

gov.uk/mhra

Copyright information

Role of a statistician in regulatory decision-making; views of a clinician

By John Johnston, Clinical Assessor, MHRA (UK)

John is a clinical assessor in the Innovative Medicines section of the Biological Unit at the UK's Medicines and Healthcare products Regulatory Agency.

Prior to Brexit, John served as the UK delegate at the Blood Products Working Party and the Committee for Advanced Therapies at the European Medicines Agency.

John was consultant editor and contributed to three chapters in 'The Law and Regulation of Medicines and Medical Devices', 2nd edition, 2021, edited by P Feldschreiber, published by Oxford University Press. The chapter on 'Clinical efficacy and safety: the concept of benefit risk' was co-authored with Khadija Rantell of MHRA (UK) and EFSPI.

The presentation to EFSPI 2022 describes:

- The input of a statistician to decision making
- The clinician as the 'target audience' for the statistician
- The need for a structured approach coupled to a simple form of language

Professor **George Box** promoted the use of **structure** (in the form of the scientific method) for the purpose of discovery

Professor **Sander Greenland** promotes the use of a simplified **language** to avoid confusion between (in this case) statistician and clinician. Thus:

- Recognise that we do not work in reality; we work in models that are simplified versions of reality; models may deceive.
- Avoid dichotomies
- Avoid the words 'significance' and 'confidence'
- Avoid 'nullism'
- Make more use of the word 'compatible' and its counterpart 'incompatible'

And to the above I add:

- come to an agreement with your clinical colleague on what each party means by measurement because the quality of measurement is an important yet under-recognised factor that contributes to the p-value; worthwhile measurement is much preferred.

Also:

Daniel Kahneman is a Nobel prize winner in economics for his work on human judgment and decision-making. He has concluded that humans essentially make decisions in two ways:

- a fast, intuitive method, and
- a slower, reasoning and evidence-based method

Kahneman's research insists that the first method, while more 'comfortable' and often faster, is more prone to error. The second method, which can take more time, is generally more reliable because it depends on clearer reasoning and makes space for evidence, including statistics.

Kahneman ultimately concludes both processes have value but that over-reliance on intuition — or decisions based on gut feelings — invites a higher risk of error.

A structured approach is preferred.

Suggested reading:

<u>Structure / The Scientific Method</u>

- Statistics for Experimenters, chapter 1. Box G, Hunter J & Hunter W. Wiley Interscience USA, 2005 Box.pdf.
- 1.- Quality and the Art of Discovery by George Box Part I.
- 2.- Quality and the Art of Discovery by George Box Part II.

George Box appears in other YouTube videos that can be found easily with a search engine of your choice

George Box was amongst the first (if not the first) statisticians to use the word 'compatible'.

<u>On a preferred style of language</u>

- Greenland S et al. To curb research misreporting, replace significance and confidence by compatibility. Prev Med. 2022 Jul 1;107127. Available at: ScienceDirect
- Amrhein V & Greenland S. Rewriting results in the language of compatibility. Trends Ecol Evol. 2022;37(7):567-568. https://pubmed.ncbi.nlm.nih.gov/35227533/
- Berner D, Amrhein V. Why and how we should join the shift from significance testing to estimation. J Evol Biol. 2022;35(6):777-787. Shift from significance testing to estimation.
- Rafi Z & Greenland S. Semantic and cognitive tools to aid statistical science: replace confidence and significance by compatibility and surprise. BMC Med Res Methodol. 2020;20(1):244. https://pubmed.ncbi.nlm.nih.gov/32998683/
- Greenland S. Invited Commentary: The Need for Cognitive Science in Methodology. Am J Epidemiol. 2017;186(6):639-645. doi: 10.1093/aje/kwx259. https://academic.oup.com/aje/article/186/6/639/3886035

See also:

The value of a p-value with Charles Poole.

Moving to a World Beyond "p < 0.05".

<u>Case example</u>

Canning CG et al. Exercise for falls prevention in Parkinson disease. Neurology. 2015;84(3):304-312 NEUROLOGY 2014.pdf

Canning et al (2015) as adapted by: Infanger D & Schmidt-Trucksäss A. P value functions: An underused method to present research results and to promote quantitative reasoning. Stat Med. 2019;38(21):4189-4197. https://onlinelibrary.wiley.com/doi/epdf/10.1002/sim.8293

Allen NE et al. Interventions for preventing falls in Parkinson's disease (Review). Cochrane Database of Systematic Reviews 2022, Issue 6. Art. No.: CD011574. Cochrane Library

Also:

Tukey JW. The Future of Data Analysis. The Annals of Mathematical Statistics. 1962;33(1):1-67 http://www.jstor.org/stable/2237638

Tukey JW. The Philosophy of Multiple Comparisons. Statistical Science. 1991;6(1):100-116. https://www.jstor.org/stable/2245714

Birnbaum A. Confidence Curves: An Omnibus Technique for Estimation and Testing Statistical Hypotheses. Journal of the American Statistical Association. 1961;56:246-249 http://www.jstor.com/stable/2282249

Extras

"Estimation and compatibility" in: Box, G.E., 1980. Sampling and Bayes' inference in scientific modelling and robustness (with discussion). Journal of the Royal Statistical Society: Series A (General) 143:383-404.
https://rss.onlinelibrary.wiley.com/doi/epdf/10.2307/2982063

"Consistency" in: Cox DR (1977) The role of significance tests. Scandinavian Journal of Statistics 4: 49–70. https://cox-1977-with-discussion-and-reply.pdf

"Estimation" in: Yates, F. (1951). The influence of 'Statistical Methods for Research Workers' on the development of the science of statistics. J Amer Statist Assoc 46, 19-34
https://www.jstor.org/stable/pdf/2280090.pdf?refreqid=excelsior%3A244c589a2eac0728812152bfe3a38d5a&ab_segments=&origin=&acceptTC=1

go to pp 32-34

> On the other hand the emphasis given to formal tests of significance throughout *Statistical Methods*, and to a great extent also in *The Design of Experiments*, has had two consequences which are not wholly satisfactory. In the first place it has resulted in what seems to me to be an undue concentration of effort by mathematical statisticians on investigations of tests of significance applicable to problems which are of little or no practical importance. Second, and more important, it has caused **scientific research workers** *to* **pay undue attention to the results of the tests of significance they perform on their data, particularly data derived from experiments, and too little to the estimates of the magnitude of the effects they are investigating.**

> The emphasis on tests of significance, and the consideration of the results of each experiment in isolation, have had the unfortunate consequence that scientific workers have often regarded the execution of a test of significance on an experiment as the ultimate objective. Results are significant or not significant and that is the end of it.

Frank Yates, Rothamsted Experimental Station Frank Yates

Some points to think about on measurement

- Measurement in Science (Stanford Encyclopaedia of Philosophy)
- Measurement Essentials 2nd Ed.
- MESA Memo 44: Ordinal and Interval.   MESA Memo 63: Probabilistic Models.