MRC Clinical Trials Unit

Smarter Studies
Global Impact
Better Health

UCL

Strengthening simulation studies in statistical research

2021 RSS conference, Manchester

Tim Morris @tmorris_mrc

(for the panel on simulation studies)

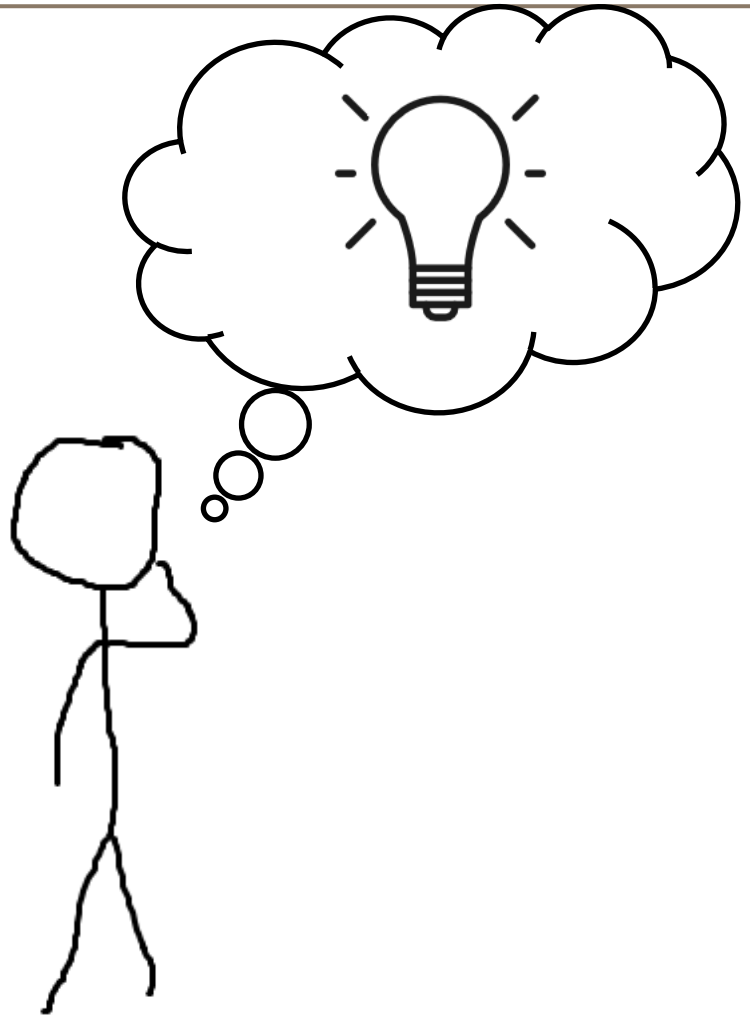
STRATOS INITIATIVE

You know what statisticians are really good at?

Not simulation studies, that's for damn sure.

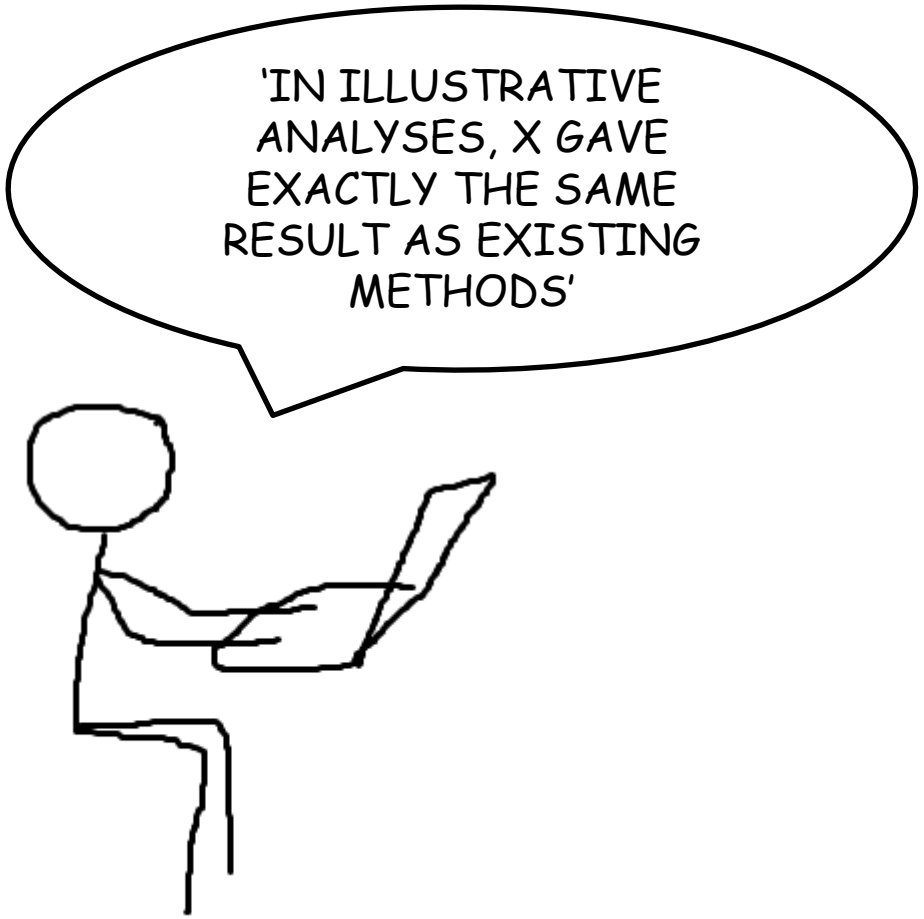
A statistician

A statistician



A statistician

A statistician



A statistician

‘I would like to draw an analogy between clinical research that is focused on treatments and methodological research that is focused on methods. Imagine that medical journals required authors to present new prototype treatments in all articles but rejected comparative clinical trials because the treatment’s principle is not new.
‘This is more or less what happens in methodological research. Neutral comparison studies are not really accepted as valuable research.’

—Anne-Laure Boulesteix, 2021

‘Towards neutral comparison studies in methodological research’

Special issue of Biometrical Journal.
Submissions until 31 Mar 2022.

Late phase clinical trials

In late-phase clinical trials, investigators have a burden to ensure fairness when comparing randomised groups. This is a key role of the statistician.

Allocation concealment, randomisation, blinding, placebos, shams, pre-specified statistical analysis plans, etc.

Importantly, the people who initially developed the intervention are rarely the investigators.

Reading a simulation study

‘4 | MONTE CARLO SIMULATIONS

‘We generate n independent vectors $(Y_i, X_i), i = 1, \dots, n$ from the model […]
‘We repeated the procedure 1,000 times […]
‘Results are shown in table 2 and in the appendix […]
‘The methods performs well when correctly specified […]’

So what?

Why should people take simulation studies seriously?

Why does this matter?

It is useful to remember that simulation studies are frequently used to justify what people actually do!

We want to get things right, whether or not we are incentivised to do so.

Received: 15 August 2017 | Revised: 20 October 2017 | Accepted: 22 October 2017

DOI: 10.1002/bimj.201700129

LETTER TO THE EDITOR

Biometrical Journal

On the necessity and design of studies comparing statistical methods

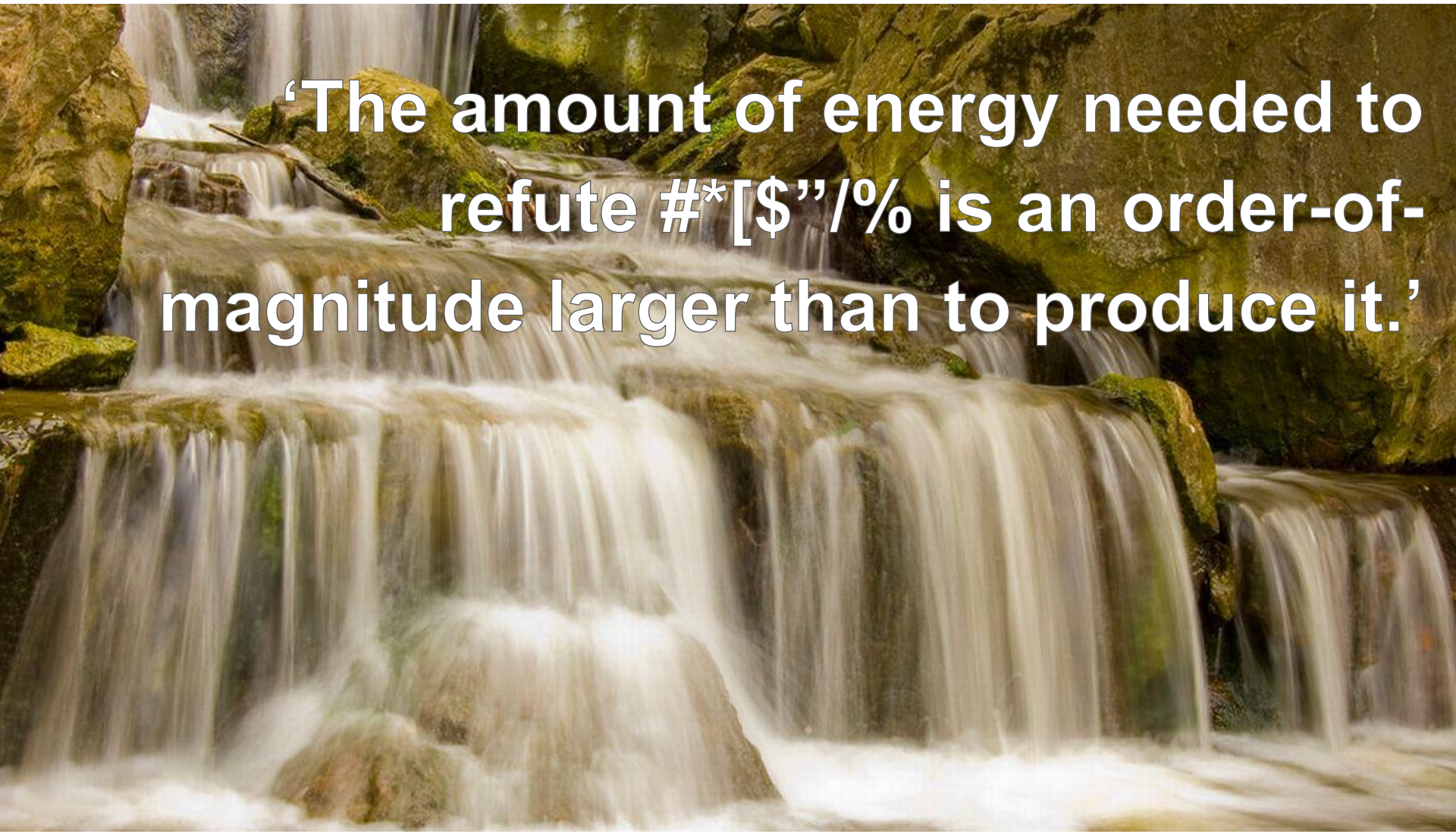
Anne-Laure Boulesteix[1]
Harald Binder[2]
Michal Abrahamowicz[3]
Willi Sauerbrei[2]
for the Simulation Panel of the STRATOS Initiative

What steps can we take?

A collection of opinions on some steps that would make a huge difference

You will think of more; write them down and use them to give your simulation studies a competitive edge!

Some are a particular challenge for multivariable model folks


‘The amount of energy needed to refute #*[$”/% is an order-of-magnitude larger than to produce it.’

Simulation studies need to be done well first time!

Received: 29 November 2017 | Revised: 23 August 2018 | Accepted: 2 November 2018

DOI: 10.1002/sim.8086

Check for updates

TUTORIAL IN BIOSTATISTICS

WILEY Statistics in Medicine

Using simulation studies to evaluate statistical methods

Tim P. Morris[1] | Ian R. White[1] | Michael J. Crowther[2]

1. Structure for your readers

Reporting guidelines help researchers to report various types of study.

Things that investigators regard as obvious will not be obvious to outsiders.

1. Structure for your readers

It makes it much easier to read any study if you have a framework.

When randomised trials are reported using Consort, I know roughly what to expect to read. (The fact that no trial is ‘vanilla’ does not detract from its usefulness.)

1. Structure for your readers

A – Aims

D – Data-generating mechanisms

E – Estimands

M – Methods of analysis

P – Performance measures

2. Justify your choices

A D E M P

In explaining what you did, you should also include justifications for your choices: **why did I do this?**

‘We did 1,000 repetitions [everyone else does]’

2. Justify your choices

Think of different phases: proof-of-concept (like pre-clinical work), trying to hone a method (like dose-finding), comparison of competing methods in non-ideal situations (phase III), understanding when a method breaks (phase IV)…

Justify your choices as a log for future-you.

3. MV challenge: Estimands

Estimands are not just for treatments in randomised trials

It is not enough to show that your modelling procedure was able to recover some aspect of the data-generating mechanism.

How do you anticipate this method being used in practice? Until you have specified this, it is impossible to evaluate.

4. MV challenge: the data-generating mechanisms you imagined do not represent all possible DGMs

5. MV challenge: Know/learn your performance measures

For me, a classic line is ‘the new method is unbiased and demonstrates good coverage’… I KNOW the authors are hiding its inefficiency

Remember that MSE and coverage depend heavily on sample size

5. MV challenge: Know/learn your performance measures

Some commonly-used performance measures:

Bias Empirical SE MSE	Properties of estimator $\hat{\theta}$
Average model-based SE	Property of $\widehat{SE}(\hat{\theta})$
Coverage Power	Properties of confidence interval
Convergence Computational speed	Computational/planning

6. MV challenge: If you used code, make it available!

The **second worst** thing is ‘code not available’

First worst is ‘code available from the authors on request’

Don’t do this to your readers

6. MV challenge: If you used code, make it available!

Replication is a far more concrete concept for simulation studies than for other empirical studies

If you are interested, see the **replisims.org/** initiative by Anna Lohmann, Rolf Groenwold and Kim Luijken

6. Quantify Monte Carlo error

Simulation studies involve drawing [pseudo-]random numbers. Results will depend (to some extent) on the particular numbers that were drawn.

We need to quantify uncertainty due to using a finite number of repetitions (Monte Carlo error).

7. Neutral schmeutral

Have you come across the ‘methodological attribution problem’?

From Gelman’s ‘Bayesian statistics then and now’ *Statistical Science* 2010 (doi:10.1214/09-STS308)

My second meta-principle of statistics is *the methodological attribution problem,* which is that the many useful contributions of a good statistical consultant, or collaborator, will often be attributed to the statistician's methods or philosophy rather than to the artful efforts of the statistician himself or herself. Don Rubin has told me that scientists are fundamentally Bayesian (even if they do not realize it), in that they interpret uncertainty intervals Bayesianly. Brad Efron has talked vividly about how his scientific collaborators find permutation tests and p-values to be the most convincing form of evidence. Judea Pearl assures me that graphical models describe how people really think about causality. And so on. I am sure that all these accomplished researchers, and many more, are describing their experiences accurately. Rubin wielding a posterior distribution is a powerful thing, as is Efron with a permutation test or Pearl with a graphical model, and I believe that

A final thought…

Some of us find it helpful to think about the analogy of simulation studies to drug development

I described one reason. The second is that it forces us to consider how we would do them if they were regulated and the burden was on us to verify to a cautious neutral party that a method does in fact work.

Acknowledgements

The simulation panel – especially Anne-Laure Boulesteix
Ian White
Georg Heinze
Anna Lohmann
Tra My Pham