



MRC  
Clinical  
Trials Unit



# Structured planning and reporting of simulation studies

Tim Morris, MRC Clinical Trials Unit at UCL

*Stratos @ Lorentz, Leiden*

16–20 Sep 2024

Smarter Studies  
Global Impact  
Better Health <sup>1</sup>

# Simulation studies

Simulation studies are **empirical** studies: they generate data, which is analysed, returning **noisy results**.

We do simulation studies for all sorts of reasons and on many topics.

Common in early-phase statistical research, often executed by people with mathematical inclinations, with less involvement in empirical research.

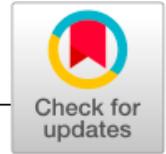
# Why I'm here

Received: 29 November 2017

Revised: 23 August 2018

Accepted: 2 November 2018

DOI: 10.1002/sim.8086



**TUTORIAL IN BIostatISTICS**

WILEY **Statistics**  
in Medicine

## Using simulation studies to evaluate statistical methods

Tim P. Morris<sup>1</sup>  | Ian R. White<sup>1</sup>  | Michael J. Crowther<sup>2</sup> 



# What I didn't know

Around the time we released a pre-print of the paper, Stratos had recently formed its Simulation Panel...

This early letter is a great perspective and gives us an idea of progress since.

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**

# Background

I started out as a trial statistician and was impressed by ideas about research integrity, reporting guidelines, etc.

The first simulation study reports I read were surprising. More like amateur murder mysteries than descriptions of scientific experiments.

No planning, just thinking-while-coding.

‘Come on, they all contain the same elements! But reports mix ideas up, fail to report lots, never explain decisions... we ultimately take conclusions on trust!’

# Main contribution: ADEMP

**A** – Aims

**D** – Data-generating mechanisms

**E** – Estimands / targets

**M** – Methods of analysis

**P** – Performance measures

# Aims

Consider the difference between:

1. 'How does my method perform in scenario X?'
2. 'For a particular applied problem, what are the merits and drawbacks of the available methods?'

# Data-generating mechanisms

Motivated by method or by applied problems?

1. Easy-to-pass test
2. Range of 'realistic' mechanisms (think outside your model!)
3. Very severe test

'It **should** return odd results if you do this and anyone smart enough to think of doing it should be smart enough to know not to!'

– Bill Gould, 13 Sep 2024

# Data-generating mechanisms

Many options! Parametric models, resampling methods, so-called ‘plasmode’ simulation.



TUTORIAL IN BIostatISTICS |  **Open Access** |   

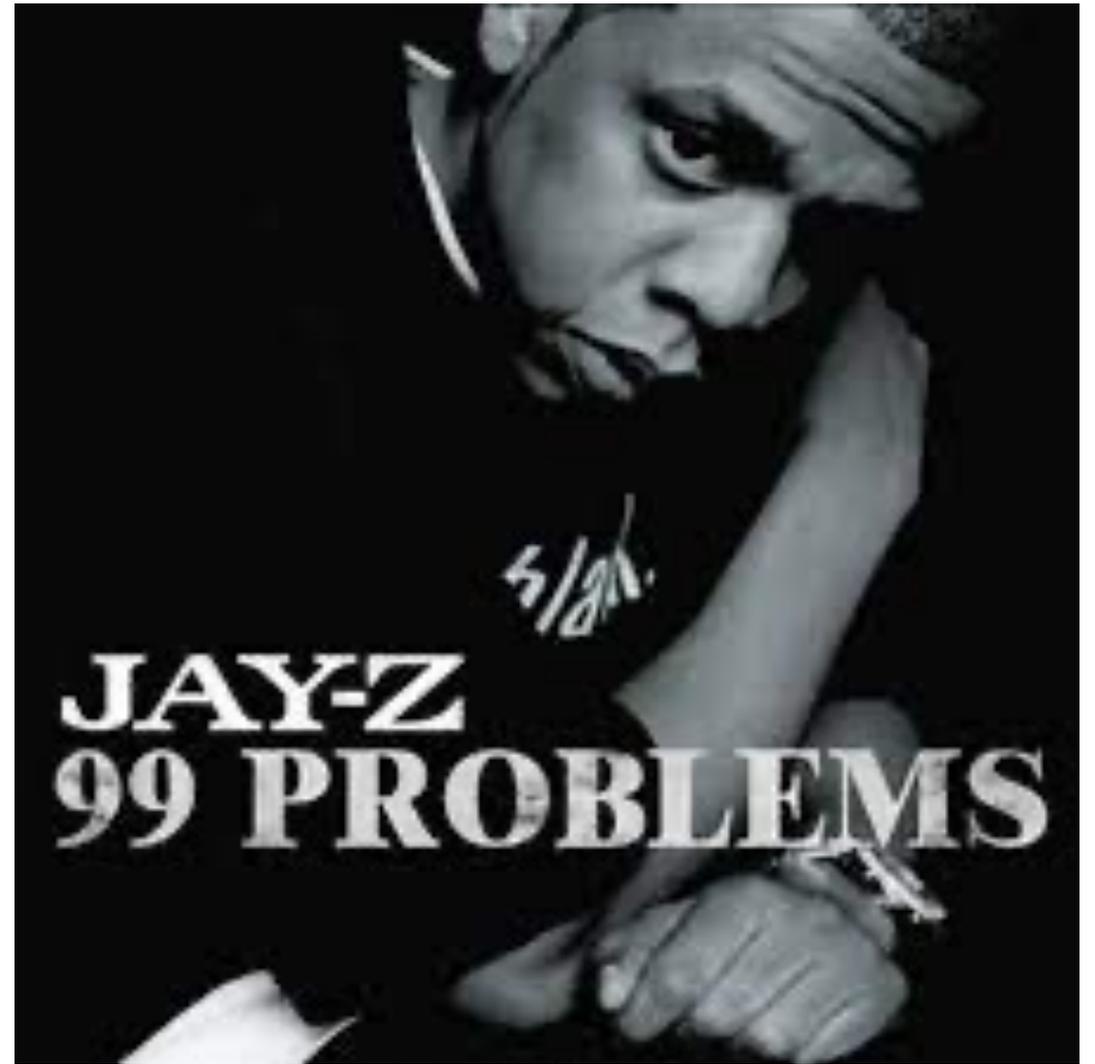
## **Statistical plasmode simulations–Potentials, challenges and recommendations**

Nicholas Schreck , Alla Slynko, Maral Saadati, Axel Benner

First published: 14 February 2024 | <https://doi.org/10.1002/sim.10012>

Nicholas Schreck and Alla Slynko are equally contributing authors.

**‘I got 99 problems but I’m  
ignoring all-but-one’  
– Methodologist**



# Estimands / targets

Estimand = what we wish to estimate (& the true value)

**Qualitative** 'Expected difference by treatment-initiation decision on quality of life at 12 weeks among UK men aged 50+ with non-metastatic prostate cancer...'

**Quantitative** 1.5

# Estimands / targets

Easy approach is to focus on ability to accurately recover 1+ parameter/s of a data-generating model. Avoids wasting energy on thinking!

Two classic examples of how this can be problematic:

1. Risk prediction: Evaluating ‘bias’ in individual parameter estimates may miss better out-of-sample performance.
2. Covariate adjustment in RCTs: classic ‘Estimand is the conditional log-odds ratio I happened to use to simulate data – oh look which method happens to perform well!’\*

(True story: as I wrote (2) my train passed the NHS ‘Logistics Support Unit’!)

Barrier: What if we don’t know the true value?

# Methods of analysis

- Have you chosen **relevant** methods and variations?
- Are included and omitted methods well **justified**?
- Are methods for **variance** and **interval** estimation clear (not just point estimation)?
- Have you considered **operator expertise**?
- Will you experience **method failure**? How will you deal with it?

# Comments welcomed!

<https://arxiv.org/abs/2408.11594>

On the handling of method failure in comparison studies

Milena Wunsch \* <sup>1,2</sup>, Moritz Herrmann <sup>1,2</sup>, Elisa Noltenius<sup>3</sup>, Mattia Mohr<sup>3</sup>,  
Tim P. Morris <sup>4</sup>, and Anne-Laure Boulesteix <sup>1,2</sup>

<sup>1</sup>Institute for Medical Information Processing, Biometry, and Epidemiology, Faculty of Medicine, LMU  
Munich (Germany)

<sup>2</sup>Munich Center for Machine Learning (MCML), Munich (Germany)

<sup>3</sup>Department of Statistics, LMU Munich, Munich (Germany)

<sup>4</sup>MRC Clinical Trials Unit, UCL, London (UK)

August 22, 2024

# Performance measures

How we measure the performance of methods. Things like proportion of non-failed reps, run time, bias, empirical SE, model SE, coverage, power, etc.

Important: learn how these relate to one other.

Consider how 'method failure' impacts definition of performance, and comparison of methods with differential failure.

# ADEMPI with Implementation?

A kind of miscellaneous category about ‘what we actually did’:

- Number of repetitions & justification
- Software used & version
- Packages & versions
- Operating system
- ...

# What are we comparing? Analyses

- The comparison of in
- Methods of analysis
  - Data generating mechanisms
  - Even estimands!
  - Others

Article

**A comparison of different population-level summary measures for randomised trials with time-to-event outcomes, with a focus on non-inferiority trials**

Matteo Quartagno , Tim P Morris , Duncan C Gilbert, Ruth E Langley, Matthew G Nankivell, Mahesh KB Parmar and Ian R White 

CLINICAL  
TRIALS

Clinical Trials  
1–9

© The Author(s) 2023



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/17407745231181907

[journals.sagepub.com/home/ctj](https://journals.sagepub.com/home/ctj)

 Sage

# ADEMP a few years on

I often hear people refer to ADEMP as a 'reporting guideline'. It certainly helps with reporting but is not a guideline!

Pre-registration? See ADEMP-PreReg  
<https://github.com/bsiepe/ADEMP-PreReg>

# Progress

EDITORIAL |  **Free Access**

## **Editorial for the special collection “Towards neutral comparison studies in methodological research”**

Anne-Laure Boulesteix , Mark Baillie, Dominic Edelmann, Leonhard Held, Tim P. Morris, Willi Sauerbrei

First published: 17 February 2024 | <https://doi.org/10.1002/bimj.202400031>

# You do not *have* to use ADEMP

‘By all means break the rules, and break them beautifully, deliberately and well.’

– Robert Bringhurst, *The Elements of Typographic Style*

ADEMP has become reasonably standard. See above if you don't want to. Nan and I discussed yesterday AEDMP for some of her work.