Phases of development for statistical methods

Tim Morris, MRC Clinical Trials Unit at UCL
Joint work with Georg Heinze, Anne-Laure Boulesteix, Michael Kammer and Ian White

Key idea

13 years ago, Ian White floated the core idea I am presenting today:
a methods development pipeline

I am absolutely not proposing regulation of statistical methods, but wonder…
How would we approach things if statistical methods were regulated?

ELSEVIER

Journal of Clinical Epidemiology 152 (2022) 23–29

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

A method was developed for correcting the bias in the usual study weights in meta-analyses

Stephen D. Walter[a,*], Narayanaswamy Balakrishnan[b]

[a]*Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada*
[b]*Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada*

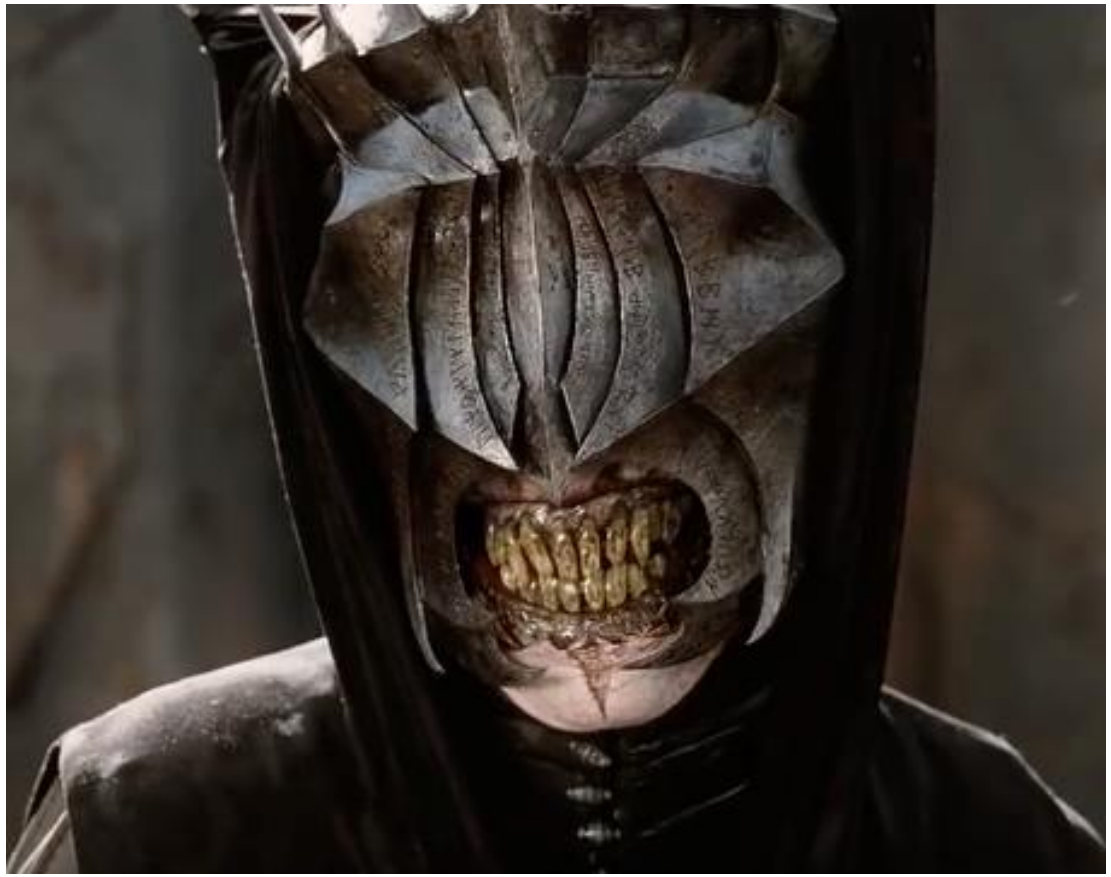
Accepted 29 August 2022; Published online 21 September 2022

Walter & Balakrishnan in brief

- The estimated (unscaled) weights in meta-analysis are biased upwards
- Find an expression for the bias and use this to derive a bias-correction
- Try it in two example meta-analyses with few, small studies; the weights are slightly different
- '*We recommend that our bias correction should be routinely adopted*' and '*Elimination of this bias will enhance the validity of conclusions from a meta-analysis, compared with the situation when the standard weights are used.*' ☺

Why do people not use new methods?

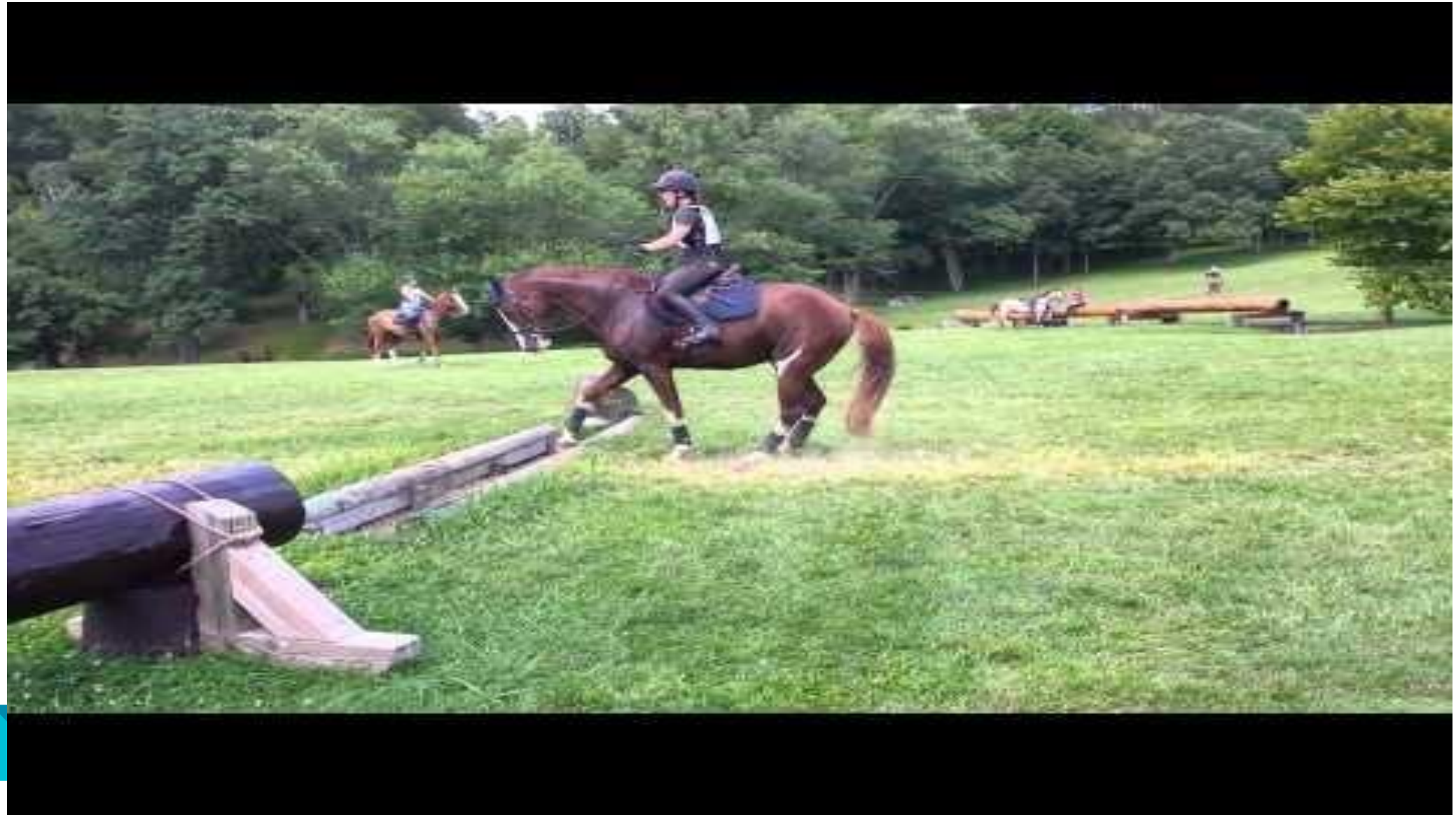
Why do people not use new methods?



A malevolent opponent?

It’s sometimes true that someone fundamentally disagrees with you and are actively working against your agenda (consider whether they have a good point)

Why do people not use new methods?



Why do people not use new methods?

‘They said it’s good in there and I should come… only they didn’t give me the key’

Let’s give applied researchers credit

Let’s start at the end. Applied researchers need to use a method that is suitable for their study, e.g. a trial I am designing, or this EHR dataset in front of me.

Suppose they are motivated to do something as well as they can, rather than just to get something past journal editors / regulators (which always amounts to precedence).

They might want to know **when a method is and is not preferred method; what diagnostics are available; pitfalls that may occur at application**.

Phases of drug development

PHASE I TRIALS: Initial studies to determine the metabolism and pharmacologic actions of drugs in humans, the side effects associated with increasing doses, and to gain early evidence of effectiveness; may include healthy participants and/or patients.
PHASE II TRIALS: Controlled clinical studies conducted to evaluate the effectiveness of the drug for a particular indication or indications in patients with the disease or condition under study and to determine the common short-term side effects and risks.
PHASE III TRIALS: Expanded controlled and uncontrolled trials after preliminary evidence suggesting effectiveness of the drug has been obtained, and are intended to gather additional information to evaluate the overall benefit-risk relationship of the drug and provide and adequate basis for physician labelling.
PHASE IV TRIALS: Post-marketing studies to delineate additional information including the drug's risks, benefits, and optimal use.

Received: 12 August 2022 | Revised: 9 December 2022 | Accepted: 22 January 2023

DOI: 10.1002/bimj.202200222

Biometrical Journal

RESEARCH ARTICLE

Phases of methodological research in biostatistics—Building the evidence base for new methods

Georg Heinze[1] | Anne-Laure Boulesteix[2] | Michael Kammer[1,3] | Tim P. Morris[4] | Ian R. White[4] | on behalf of the Simulation Panel of the STRATOS initiative




Phase …

Scope: A study in this phase will typically…	Elements: Typically, a study in this phase will consist of…	Outcome: after this phase, we know…

Phase I

Scope: A study in this phase will typically…	Elements: Typically, a study in this phase will consist of…	Outcome: after this phase, we know…
…introduce a new idea, demonstrate validity by investigation of (asymptotic or finite-sample) properties, show potential to improve on existing methods or to be the only solution.	…mathematical derivations and proofs; possibly some example data analyses.	…whether a method is valid – or not – from a theoretical point of view

Phase II

Scope: A study in this phase will typically…	Elements: Typically, a study in this phase will consist of…	Outcome: after this phase, we know…
…possibly refine and extend the method; demonstrate the use of the method with real data; it will consider a limited range of possible applications.	…simulation studies including limited comparisons with other methods, simple example data analyses.	…whether a method can be used with caution or should not be used in certain applied settings.

MRC Clinical Trials Unit

UCL

Phase III

Scope: A study in this phase will typically…	Elements: Typically, a study in this phase will consist of…	Outcome: after this phase, we know…
…compare a relatively new method with competitors and demonstrate its use in practice; considers a wide range of applications	…simulations with a wide range of scenarios, e.g. different outcome types (ideally set up to be ‘neutral’ rather than ‘showy-offy’), realistic comparative example data analyses.	…in which settings (among many) a method tends to outperform competing methods.

Phase IV

Scope: A study in this phase will typically…	Elements: Typically, a study in this phase will consist of…	Outcome: after this phase, we know…
…summarise the evidence about a method, also in comparison with competing methods; uncover previously unknown behaviour of the method; consider an extended range of possible and actual applications.	…a review of the existing evidence about a method, simulation studies with extended/unusual scenarios, complex comparative example data analyses.	…when a method is and is not preferred; what diagnostics are available and which pitfalls may occur during its application

Example: Firth's bias-correction for maximum likelihood estimators (problem description)

Biometrika (1967), **54**, 1 *and* 2, *p.* 181
Printed in Great Britain

On the bias of various estimators of the logit and its variance with application to quantal bioassay

BY JOHN J. GART AND JAMES R. ZWEIFEL
National Cancer Institute, Bethesda, Md

SUMMARY

The bias of several logit estimators and their corresponding variance estimators is investigated in small samples. Their use in quantal bioassay is similarly explored.

Example: Firth's bias-correction for maximum likelihood estimators

Take separation in logistic regression, when log-odds ratios go to $\pm\infty$. So in finite samples, it's infinitely biased. Firth's correction corrects this bias.

A neat side effect is that it gives finite estimates of regression coefficients in GLMs even with data constellations where maximum likelihood estimates do not exist.

Phase I: derived correction and gave some simple examples

Biometrika (1993), 80, 1, *pp.* 27–38
Printed in Great Britain

Bias reduction of maximum likelihood estimates

BY DAVID FIRTH
Department of Mathematics, University of Southampton, SO9 5NH, U.K.

SUMMARY

It is shown how, in regular parametric problems, the first-order term is removed from the asymptotic bias of maximum likelihood estimates by a suitable modification of the score function. In exponential families with canonical parameterization the effect is to

Phase II: simulation demonstrated that correction improved on existing methods

STATISTICS IN MEDICINE
Statist. Med. 2002; **21**:2409–2419 (DOI: 10.1002/sim.1047)

A solution to the problem of separation in logistic regression

Georg Heinze*,† and Michael Schemper

Section of Clinical Biometrics, Department of Medical Computer Sciences, University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria

Phase III: comprehensive 'neutral' simulation study on logistic regression

van Smeden *et al. BMC Medical Research Methodology* (2016) 16:163
DOI 10.1186/s12874-016-0267-3

BMC Medical Research Methodology

RESEARCH ARTICLE Open Access

No rationale for 1 variable per 10 events criterion for binary logistic regression analysis

Maarten van Smeden[1*], Joris A. H. de Groot[1], Karel G. M. Moons[1], Gary S. Collins[2], Douglas G. Altman[2], Marinus J. C. Eijkemans[1] and Johannes B. Reitsma[1]

Phase IV: explained why different methods lead to different results and gave advice on how to detect and to deal with separation in practice

American Journal of Epidemiology

Vol. 187, No. 4
DOI: 10.1093/aje/kwx299
Advance Access publication:
August 17, 2017

Practice of Epidemiology

Separation in Logistic Regression: Causes, Consequences, and Control

Mohammad Ali Mansournia, Angelika Geroldinger*, Sander Greenland, and Georg Heinze

* Correspondence to Dr. Angelika Geroldinger, Section for Clinical Biometrics, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Austria Spitalgasse 23, 1090 Vienna, Austria (e-mail: angelika.geroldinger@gmx.at).

Link to Stratos ‘levels’?

Arguably a method may be used with the dataset in front of you when it has not gone through all these phases.

There is perhaps a link to draw to Stratos ‘levels’, where we may be comfortable with level-3 researchers using methods that have gone through less rigorous testing than level-1 researchers.

‘What is the point of this talk?’

I hope that thinking about phases:

- Gives frustrated methods researchers doing ‘early-phase’ work a framework to understand why their method has not been universally adopted;
- Recognises ‘late-phase’ statistical methods research as valuable;
- Gives applied researchers a way to articulate their hesitation about using new methods.

Some encouragement

Examples?

The Cox m

Propensity

Multiple &

Jointly published by Akadémiai Kiadó, Budapest and Kluwer Academic Publishers, Dordrecht

Scientometrics, Vol. 59, No. 3 (2004) 467–472

Short communication

Sleeping Beauties in science

ANTHONY F. J. VAN RAAN

Centre for Science and Technology Studies, Leiden University, Leiden (The Netherlands)

MRC Clinical Trials Unit

UCL

Points to ponder

- Can you think of a great/useful method that is still ‘sleeping’?
- What work would be needed to ‘awaken’ it?
- Do you regard late-phase work as ‘marketing’ for attention-seekers?!

Just out!

Georg Heinze, Anne-Laure Boulesteix, Michael Kammer, Tim P Morris, Ian R White.
Phases of methodological research in biostatistics—building the evidence base for new methods.
Biometrical Journal. 2023 doi:10.1002/bimj.202200222