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

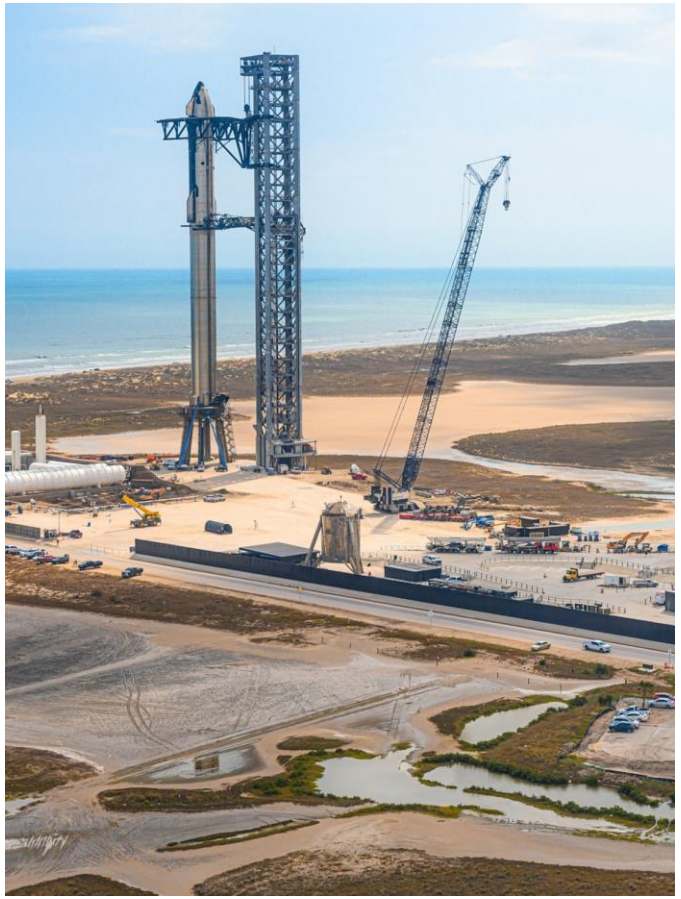
Georg Heinze,

Anne-Laure Boulesteix, Michael Kammer, Tim Morris, Ian White

for the Simulation Panel of the STRATOS Initiative

Questions of an investor

- Question I: can that starship theoretically take off from earth?
- Question II: does the starship actually take off from earth?
- Question III: can it take me to Mars?
- Question IV: will it take me back safely?



Novelty and innovation: drivers of scientific advancement?

- ‚Develop new methods!‘ they say:
  - Your funding agency
  - Your PhD evaluation committee
  - Your tenure track agreement
- We comply!
  - New methods fill our journals, our seminars, our journal clubs, CRAN, …
- BUT: Which of those methods actually enter the toolbox of a data analyst?

Applied researchers and data analysts:

- What are you looking for before using a method?
- Evidence that:
  - Method does what it is intended to do,
  - Method works in real analysis and some evidence that it is of advantage,
  - Method is widely a good choice,
  - Method is preferred over others in your application,
    diagnostics are available and pitfalls are well understood

Methodological researchers,

Are you supplying this evidence?

Honestly, mostly not!

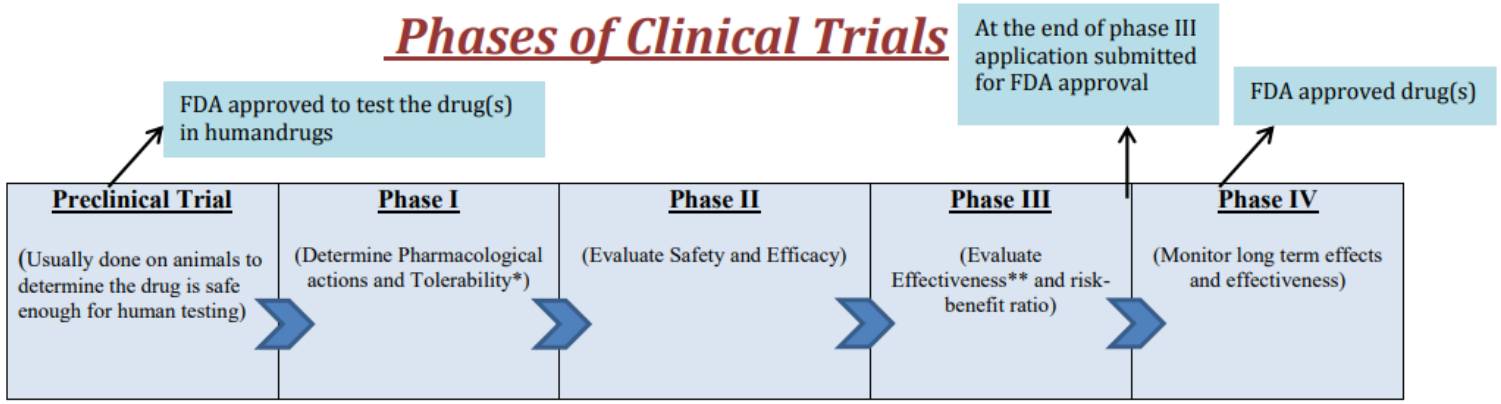
Can you supply this evidence (in a single paper)?

Honestly, no!

Do we have to supply this evidence in a single paper?

No you don‘t have to!

Just like with drugs, method development needs time



Duration	3-6 years since the drug discovery	Months	Months-Years	Years	Ongoing following FDA approval
Sample size	Not specific	Small	Large	Larger	Impacting larger beyond
Population	In vitro and In vivo animals	Healthy population / may be with targeted disease e.g. cancer, T.B. etc.	Population with target disease	Diverse population with target disease	Diverse population with target disease & new age groups, gender
Types of studies	Not specific	Unblinded & Uncontrolled	May be Placebo (inactive substance) & Active, Controlled	Randomized & controlled	Expended safety comparison
Factors to be identified	Mechanism of action, Efficacy (ability of drug to act against pathogen/disease), Safety (concerns the medical risk to the patient)	Pharmacodynamics (side effect / desire effect, mechanism of action of drug) Pharmacokinetics (absorption, distribution, metabolism & elimination of drug), Tolerated dose	Drug-Drug & Drug-Disease interaction, Efficacy at various doses, Patient safety	Dosage intervals, Risk-benefit information, Efficacy and safety for subgroups	Epidemiological data, Efficacy and safety within large diverse populations, Pharmacoeconomics (comparison of value of one pharmaceutical drug or drug therapy to another)

*Tolerability: Represents the degree to which overt adverse effects can be tolerated by the patient.
**Effectiveness: The extent to which a drug achieves its intended effects.

References:

1. http://www.nlm.nih.gov/services/ctphases.html
2. http://www.fda.gov/drugs/resourcesforyou/consumers/ucm143534.htm
3. http://www.innovation.org/drug_discovery/objects/pdf/RD_Brochure.pdf

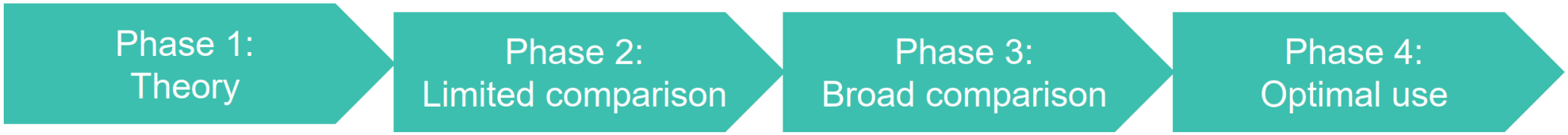
https://www.aku.edu/ctu/Documents/phases.pdf

Learning from drug development

Phases of research as a framework for building evidence

Drug development

Biostatistical methods



Methodological research: Phase I

Example:
Firth (1993): Bias reduction
of maximum likelihood estimates

- Aim:
  - Introduce new idea to solve a problem
  - Demonstrate its validity by investigating ist properties,
  - Show potential to improve on existing solutions
    or to be the only solution
- Elements of a study:
  - Mathematical derivations and proofs
  - Simple example data analyses
- After that phase we know:
  - Whether method is valid or invalid from a theoretical point of view

Methodological research: Phase II

Example:
Heinze and Schemper (2002):
A solution to the problem of separation in logistic regression

- Aim:
  - Demonstrate use of method with real data
  - Introduce refinements and extensions
  - Considering a limited range of possible applications
- Elements of a study:
  - Simulations including limited comparisons with other methods
  - Simple example data analyses
- After that phase we know:
  - Whether method can be used with caution
    or should not be used in certain applied settings

Methodological research: Phase III

- Aim:
  - Comparing a relatively new method with competitors
  - Demonstrating its use in practice
  - Considering a wide range of applications
- Elements of a study:
  - Refinements of method to broaden applicability
  - Simulations with a wide range of scenarios and different outcome types, set up as neutral comparison studies
  - Realistic comparative example data analyses
- After that phase we know:
  - In which settings (among many) a method can be safely used
  - In which settings it outperforms other methods

Examples:

- van Smeden et al (2016): No rationale for 1 variable per 10 events criterion for binary logistic regression analysis
- Puhr et al (2017): Firth‘s logistic regression with rare events: accurate effect estimates and predictions?

Methodological research: Phase IV

Example:
Mansournia et al (2018):
Separation in logistic regression:
Causes, consequences and control

- Aim:
  - Summarizing the evidence about a method, also in comparison with competing methods
  - Uncovering previously unknown behaviour with complex data
  - Considering an extended range of possible and actual applications
- Elements of a study:
  - Review of existing evidence about a method
  - Simulations with extended range of scenarios
  - Complex comparative example data analyses
- After that phase we know:
  - When a method is preferred and when it is not
  - What diagnostics are available
  - Which pitfalls may occur with its application

Experience with phases concept (1/3)

- Concept of phases is about summarizing the evidence about a *method*
- *Studies* (papers) may deal with several *methods*:
  - Later phase about established method: identifying a bug
  - Earlier phase about new method to solve the bug
- Hence, *papers* cannot be easily ‚categorized' into a single phase:
  - Phase X for Method A
  - Phase Y for Method B
- → Phases apply to specific *claims about methods in papers*

Experience with phases concept (2/3)

- With later phases, we found that
  single papers rarely *complete* a phase,
  but papers usually rather *contribute to* a phase,
- and knowledge about a method accumulates
  with several papers contributing to a phase.

- Contribution to a phase X should be preceded by contributions to phase X-1:
  - Before studying performance in simulations,
    describe the theoretical properties of a method!
  - Before applying the method in new target settings,
    try to understand how the method behaves in standard problems!

Experience with phases concept (3/3)

- In a paper, authors often claim to provide a ‚later' phase contribution (*inventor bias!*)
  while an independent assessor might attribute earlier phase
- This stresses the need for neutral comparison studies

- → Based on our experience, we encourage researchers to
  - first identify, in which phase the empirical evidence about a method currently is,
  - then to conduct research in this or the next phase.
  - Be transparent about the empirical evidence BEFORE AND AFTER your study!
  - Do we know of possible pitfalls?
    - Example: NRI which *post-marketing* was shown to have fundamental flaws (Pepe et al, 2015, Stat Biosci.)

Distinguish earlier from later phases

- While our concept is about methods (not software),
  it cannot be completely separated from the availability of software.
  - User-friendly package available? YES→III/IV NO→I/II

- Another aspect of a later phase is
  that the method has already been established before:
  - Comparison after extensive experience with method? YES→III/IV NO→I/II

- A crucial property of a later-phase contribution is neutrality:
  - Neutral comparison intended? YES→III/IV NO→I/II

Earlier phases

- If a new methodology is described,
  a paper most likely contributes to Phases I or II
- To distinguish Phase II (from Phase I):
  - Application to a realistic data example? YES → II
  - Comparison with other methods? YES → II
  - Code to apply to a similar data set available? YES → II

Later phases

- The border line between phases III and IV is a bit blurry
- A few hints towards identifying a phase IV contribution:
  - Broad phase III comparison study has been done before? YES→IV
  - Exploring new target settings, breakdown scenarios (in comparison to other methods)? YES→IV
  - Investigating new diagnostics for a method? YES→IV
  - Focus is on differential behaviour in specific settings rather than on ‚overall advantage' of a method? YES→IV

Outlook

- While early phase development studies are abundant,
  good Phase III and IV studies are still rare
- They are often not appreciated as ‚original research'
- They are difficult to design and conduct (not just ‚bigger simulation studies')

→ BUT THEY ARE NEEDED!!

- Funding agencies:
  - don't accept proposals that claim to cover all phases
    (from invention to roll-out into routine)!
  - But do accept good proposals that aim to evaluate existing methods!
- PhD evaluators, tenure track evaluators:
  - consider neutral comparison studies as valuable scientific contributions!

The phases of methodological research

TABLE 1 A brief description of the proposed scheme of phases of methodological research

Phase	Scope: A study in that phase will typically aim at …	Elements: Typically, a study in that phase will consist of…	Outcome: after that phase, we know…
I	… introducing a new idea, demonstrating its validity by investigation of (asymptotic or finite-sample) properties, showing potential to improve on existing methods or to be the only solution.	… mathematical derivations and proofs, very simple example data analyses.	… whether a method is valid or invalid from a theoretical point of view.
II	… demonstrating the use of the method with real data, probably introducing refinements and extensions; it will consider only a limited range of possible applications.	… simulations 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.
III	… comparing a relatively new method with competitors and demonstrating its use in practice; it will consider a wide range of applications.	… simulations with wide range of scenarios and different outcome types (ideally set up as neutral comparison studies), realistic comparative example data analyses.	… in which settings (among many) a method can be safely used and in which it outperforms competing methods.
IV	… summarizing the evidence about a method, also in comparison with competing methods; uncovering previously unknown behavior of the method in complex data analyses; considering an extended range of possible and actual applications.	… a review of the existing evidence about a method, simulations with extended range of scenarios, complex comparative example data analyses.	… when a method is and when it is not the preferred method; what diagnostics are available and which pitfalls may occur with its application.

Heinze et al,
Biometrical Journal 2023

Reference

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

Biometrical Journal, Special issue ‚Neutral comparison studies'