REPLICATION STUDIES
IMPROVING REPRODUCIBILITY IN THE EMPIRICAL SCIENCES

Jean Philippe de Jong, PhD
WCRI, 30 May 2017, Amsterdam
Background

- Royal Netherlands Academy of Arts and Sciences (KNAW) advises on proper conduct of science
- Concerns about lack of reproducibility of study results
- KNAW installed a committee ‘Replication studies’ to advise on replication studies
- Advice is based on scientific literature, reports by other advisory bodies, interviews with experts, an invitational workshop and its own deliberations
- This presentation contains preliminary findings.
- Report: October 2017
KNAW Committee replication studies

- Prof. dr. J.P. (Johan) Mackenbach (Erasmus MC), chair
- Prof. dr. C.M. (Cock) van Duijn (Erasmus MC)
- Prof. dr. H.R. (Harry) Büller (Academisch Medisch Centrum)
- Prof. dr. A.W. (Aad) van der Vaart (Universiteit Leiden)
- Prof. dr. E.J. (Eric-Jan) Wagenmakers (Universiteit van Amsterdam)
- Dr. P.Y.W. (Patricia) Dankers (Technische Universiteit Eindhoven)
- Prof. dr. L.M. (Lex) Bouter (Vrije Universiteit Amsterdam)
- Dr. J.Ph. (Jean Philippe) de Jong (KNAW), secretary
Scientific progress requires that results are reproducible

- No data across science ...
- Open Science Collaboration (psychology): 36%
- Bayer HealthCare: 25%
- Amgen: 11%
- High quality randomized clinical trials: 85%
The scientific community is concerned about the current degree of non-reproducibility of important research findings.
Definitions

• A *replication study* is designed to test the reproducibility of the results of a previous study.
• The *methods* should be similar, otherwise a meaningful assessment of reproducibility is impossible.
• Results are *reproduced* if they are similar enough not to raise concerns about the methodologies and assumptions.
• A replication study is *not* replication *within* a study, the results of which would not be published separately.
Truth and reproducibility

• So, similar studies should lead to similar results

However:

• Reproducible results are not necessarily true
  • but can increase confidence in findings
• Non-reproducible results are not necessarily untrue
  • but can decrease confidence and require a good explanation (potentially leading to important insights)
Impact of non-reproducibility

- Delaying scientific progress
- Polluting drug pipeline
- Unethical use of test subjects
- Wasting resources
- Diminishing public trust
- Harming individuals and environment
Non-reproducibility has many causes

- Setting up/conducting a study: random error/noise, human error, biases, changes in conditions
- Analysis of results: data dredging/p-hacking, and outcome-driven data collection
- Reporting results: failing to publish or only partially publishing results
- Underlying factors:
  - lack of proper training
  - incentive structure (funding and publications) disproportionately rewarding novel, positive results over robust approaches
**Unavoidable and avoidable causes of non-reproducibility**

- Some factors are inherent to the scientific endeavor and even rigorously conducted studies will yield a proportion of published results that cannot be reproduced.
- Others factors are avoidable and can be considered ‘questionable research practices’.
Approaches to improving reproducibility: prevention

- Improve study methods
  - Researchers should strengthen quality control mechanisms through automation, guidelines, checklists, validation studies and internal replications.
  - Institutions should improve researchers’ skills in rigorous study design, analysis and interpretation of results.
  - Institutions should provide independent methodological support and oversight on studies.

- Improving study reporting
  - Institutions and funders should require pre-registration of hypothesis-testing studies.
  - Journals should issue detailed guidelines and checklist for how to report study methods.
  - Institutions and journals should require storage of study data and methods in a repository.

- Improve the organization and culture of research
  - Journals should publish more studies with ‘negative’ results.
  - Funders should provide more long-term funding for researchers.
  - Institutions should reward researchers’ peer review activities and efforts to improve rigorous study execution.
Approaches to improving reproducibility: replication

• A certain degree of non-reproducibility cannot be avoided upfront

• Replication studies will therefore are thus necessary to:
  - check the results of an individual study and to get nearer to the truth
  - gain insight into and improve the functioning of science
Kinds of replication studies, three choices:

- **Who** will execute the study: the original investigator team, an independent team, a collaboration?
- **What** aspects of a study to replicate: sample collection, data collection, measurements, analysis, interpretation?
- **How** precisely to follow the original study: a considerable degree of similarity is needed, but a perfect copy might neither be desirable nor feasible?
What percentage of research efforts should be replication studies?

• Limited data on occurrence of replication studies
• At most a few percent in various disciplines
• Is this enough/too much?
The desired rate of replication studies depends on:

- the actual degree of non-reproducible results/likelihood that a result is non-reproducible
- to what extent conducting replication studies will contribute to societal goals
- whether it is an efficient use of research funds
- A comparison to alternatives:
  - doing innovative studies
  - taking ‘preventive’ measures to improve reproducibility.
Good replication practices require:

1. Information sharing
2. Know-how
3. Incentivizes
researchers need to adequately share information about original and replication studies

However...

- Publication bias distorts current evidence
- Lack of detail in reporting of methods and data of original study
researchers need to know when and how to perform a replication study

However...

• Difficulties in assessing the need for replications in terms of benefits and costs
• Insufficient skills in choosing study designs and interpreting results in terms of reproducibility
researchers need to experience the proper incentivizes

However...

• Researchers prefer ‘creative’ and ‘original’ studies
• Researchers view replications as an attack on colleagues
• Funding agencies focus on ‘innovative’ research
• Journals have a criterion of ‘originality’
• Career evaluations are based on (high impact) publications, funding and new findings
Recommendations to researchers

• Conduct replication studies when appropriate.
• Researchers should generate data on reproducibility and replication studies.
Recommendations to funders

• Funders should assess the need for replication studies within fields based on benefits, costs and alternatives.
• Funders should create better funding opportunities for replication studies.
Recommendations to journals

- Journals should encourage the submission of replication studies.
- Journals should issue detailed guidelines and checklist for how to report study methods
- Journals (and institutions) should require storage of study data and methods in a repository
**Recommendations to institutions**

- Institutions (and funders) should require pre-registration of hypothesis-testing studies.
- Institutions should educate researchers on how to design replication studies and assess reproducibility.
- Institutions should credit replication studies in career evaluations and awards.
THANK YOU

Email me at:
jean.philippe.de.jong@knaw.nl