

Why Most Published Research Findings Are False

Daniel Matheson

ACTSC 991 - Spring 2019

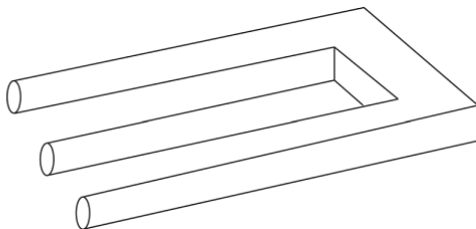
June 19th 2019



Why Most Published Research Findings Are False

John P. A. Ioannidis

Reprinted courtesy of the Public Library of Science



1. Measuring the 'Truth'
2. Detecting/Measuring False Positives
3. Corollaries
4. Fixing the Crisis



- ▶ In modern times, there is increased concern that research findings are not reproducible/replicable.
- ▶ Ioannidis (2005) does further than this and claims that *it can be proven that most claimed research findings are false*.



Definition (Positive Predictive Value)

The PPV is the post-study probability that the research finding is true. It is defined as

$$PPV = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false positives}}.$$

This is the metric Ioannidis uses to claim that most published research findings are false.



Corollary 1

The smaller the studies conducted in a scientific field, the less likely the research findings are to be true.

⇒ smaller studies have smaller power, and hence lower PPV.



Corollary 1

The smaller the studies conducted in a scientific field, the less likely the research findings are to be true.

⇒ smaller studies have smaller power, and hence lower PPV.

Corollary 2

The smaller the effect sizes in a scientific field, the less likely the research findings are to be true.

⇒ smaller effect size also lowers power.



Corollary 3

The greater the number and the lesser the selection of tested relationships in a scientific field, the less likely the research findings are to be true.
⇒ PPV is an increasing function of pre-study odds. Thus research findings are less likely to be true in 'hypothesis-generating' experiments.



Corollary 3

The greater the number and the lesser the selection of tested relationships in a scientific field, the less likely the research findings are to be true.

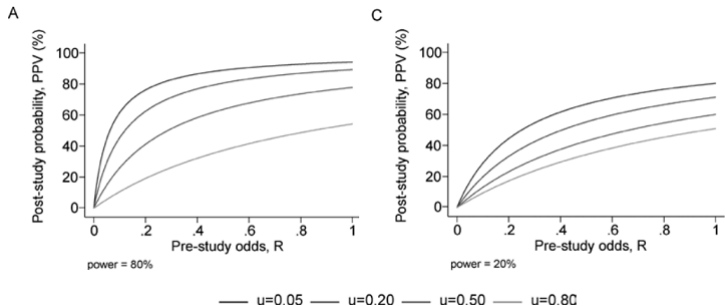
⇒ PPV is an increasing function of pre-study odds. Thus research findings are less likely to be true in 'hypothesis-generating' experiments.

Corollary 4

The greater the flexibility in designs, definitions, outcomes, and analytical modes in a scientific field, the less likely the research findings are to be true.

⇒ flexibility allows for bias to enter and the potential for transforming 'negative' results into 'positive' results.





Corollary 5

The greater the financial and other interests and prejudices in a scientific field, the less likely the research findings are to be true.



Corollary 5

The greater the financial and other interests and prejudices in a scientific field, the less likely the research findings are to be true.

Corollary 6

The hotter a scientific field (with more scientific teams involved), the less likely the research findings are to be true.



Corollary 5

The greater the financial and other interests and prejudices in a scientific field, the less likely the research findings are to be true.

Corollary 6

The hotter a scientific field (with more scientific teams involved), the less likely the research findings are to be true.

- ▶ This last corollary may seem paradoxical, but Ioannidis argues that since many teams are working on the same research questions, timing becomes very important to beat the competition. Therefore teams may prioritize releasing their most impressive 'positive' results as soon as possible, leading to bias or incorrect results.



Corollary 5

The greater the financial and other interests and prejudices in a scientific field, the less likely the research findings are to be true.

Corollary 6

The hotter a scientific field (with more scientific teams involved), the less likely the research findings are to be true.

- ▶ This last corollary may seem paradoxical, but Ioannidis argues that since many teams are working on the same research questions, timing becomes very important to beat the competition. Therefore teams may prioritize releasing their most impressive 'positive' results as soon as possible, leading to bias or incorrect results.
- ▶ The term "Proteus phenomenon" has been coined to describe rapidly alternating extreme research claims and equally extreme opposite refutations.



This all sounds very bleak, so what do we do?

- ▶ Reduce bias as much as possible. Easier said than done!



This all sounds very bleak, so what do we do?

- ▶ Reduce bias as much as possible. Easier said than done!
- ▶ Larger studies and meta-analyses.



This all sounds very bleak, so what do we do?

- ▶ Reduce bias as much as possible. Easier said than done!
- ▶ Larger studies and meta-analyses.
- ▶ Avoid emphasizing the statistically significant findings of any single research team when many others are also working on the same topic.



This all sounds very bleak, so what do we do?

- ▶ Reduce bias as much as possible. Easier said than done!
- ▶ Larger studies and meta-analyses.
- ▶ Avoid emphasizing the statistically significant findings of any single research team when many others are also working on the same topic.
- ▶ Stop chasing p-values! We should improve our understanding of the *pre-study odds*. That is, the researchers should consider what they believe a priori what the chances are that they are testing a true relationship. The PPV is an increasing function of these pre-study odds.



This all sounds very bleak, so what do we do?

- ▶ Reduce bias as much as possible. Easier said than done!
 - ▶ Larger studies and meta-analyses.
 - ▶ Avoid emphasizing the statistically significant findings of any single research team when many others are also working on the same topic.
 - ▶ Stop chasing p-values! We should improve our understanding of the *pre-study odds*. That is, the researchers should consider what they believe a priori what the chances are that they are testing a true relationship. The PPV is an increasing function of these pre-study odds.
-
- ▶ Do you have any ideas?



- ▶ Ioannidis, J. P. (2005). Why most published research findings are false. PLoS medicine 2 (8), e124.



Summary/Discussion points

Whether or not a research claim is true depends on:

- (1) study size,
- (2) effect size (smaller \Rightarrow less likely to be true),
- (3) number of tested relationships (higher \Rightarrow less likely to be true),
- (4) flexibility in designs, definitions, outcomes and analysis (more flexible \Rightarrow less likely to be true),
- (5) financial/other interests and prejudices in the field.
- (6) the number of research teams working on a the research problem in question.

