Understanding the Replication Crisis as a Base-Rate Fallacy

PhilInBioMed, Université de Bordeaux

6 June 2018

Alexander.Bird@kcl.ac.uk

Sowerby Foundation · Philosophy & Medicine · King's College London

introduction — the replication crisis

52% of 1,576 scientists taking a survey conducted by the journal Nature agreed that there was a significant crisis of reproducibility

Amgen successfully replicated only 6 out of 53 studies in oncology

And then there is social psychology . . .

introduction — the base rate fallacy

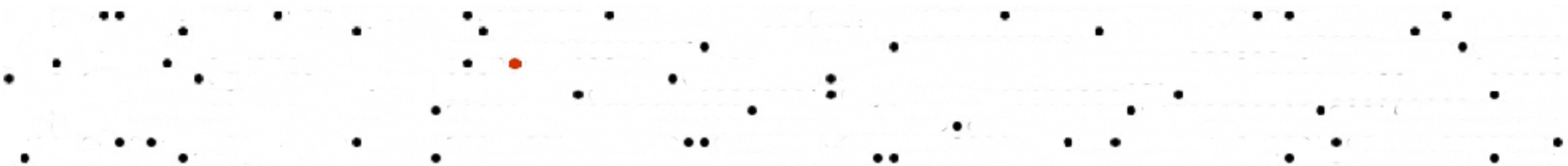
screening for a disease, which affects 1 in every 1,000 individuals, with a 95% accurate test

an individual *S* tests positive, no other risk factors; what is the probability that *S* has the disease?

Harvard medical students, 1978

11 out of 60 got the correct answer

introduction — the base rate fallacy



introduction — the base rate fallacy

base rate of disease = 1 in 1,000 = 0.1% (call this π)

false positive rate = 5% (call this α)

false positives among the 999 disease-free greatly outnumber the 1 true positive

from the base rate fallacy to the replication crisis

two types of error and accuracy

type of error	error rate	accuracy	type of accuracy
Type-I (false +ve)	α	$1-\alpha$	confidence level
Type-II (false −ve)	β	$1-\beta$	power

from the base rate fallacy to the replication crisis

do not conflate

False Positive Report Probability (FPRP)

Pr (S does not have the disease, given that S tests positive for the disease)

with

False positive error rate (α)

Pr (S tests positive for the disease, given that S does not have the disease)

from the base rate fallacy to the replication crisis

do not conflate

Pr (the temperature will be below 0°C, given that it will snow)

with

Pr (it will snow, given that the temperature will be below 0°C)

from the base rate fallacy to the replication crisis

do not conflate

Pr (this person has spots, given that they have measles)

with

Pr (this person has measles, given that they have spots)

from the base rate fallacy to the replication crisis

do not conflate

False Positive Report Probability (FPRP)

Pr (S does not have the disease, given that S tests positive for the disease)

98%

with

False positive error rate (α)

Pr (S tests positive for the disease, given that S does not have the disease)

5%

from the base rate fallacy to the replication crisis

mad scientist Dr M generates crazy hypotheses but tests then stringently

π = 0.1% (base rate of truth in Dr M's hypotheses)

accuracy = 95% (of the tests of his hypotheses)

Pr (one of Dr M's hypotheses is true, given that it passes a stringent test)

= 2%

from the base rate fallacy to the replication crisis

sane scientist Prof S generates hypotheses in a new field and tests then stringently

π = 10% (base rate of truth in Prof S's hypotheses)

accuracy = 95% (of the tests of her hypotheses)

Pr (one of Prof S's hypotheses is true, given that it passes a stringent test)

= 32%

from the base rate fallacy to the replication crisis

do not conflate

False Positive Report Probability (FPRP)

Pr (hypothesis is false, given that it passes the experimental test)

with

False positive error rate (α)

Pr (hypothesis passes the test, given that it is false)

from the base rate fallacy to the replication crisis

α	5%	
FPRP	~~5%~~	

(assuming $\pi = 0.1$ and $\beta = 0.05$)

from the base rate fallacy to the replication crisis

α	5%	
FPRP	32%	

(assuming $\pi = 0.1$ and $\beta = 0.05$)

from the base rate fallacy to the replication crisis

α	5%	?
FPRP	32%	5%

(assuming $\pi = 0.1$ and $\beta = 0.05$)

from the base rate fallacy to the replication crisis

α	5%	0.56%
FPRP	32%	5%

(assuming $\pi = 0.1$ and $\beta = 0.05$)

explaining the replication crisis

1. low π (background rate of truth)
2. non-negligible α

explaining the replication crisis

1. low π

hypotheses derived from underlying theory

" 1. If the standard model is correct, then the hypothesis that **the Higgs particle exists** should be true.

2. There is very strong evidence that the standard model is correct."

explaining the replication crisis

1. low π

hypotheses derived from underlying theory

weak underlying theory

- weak connection between theory and derived testable hypotheses
- weak evidence for the theory

  (example: Bapineuzumab)

explaining the replication crisis

1. low π

hypotheses derived from underlying theory

"1. If the amyloid cascade hypothesis is correct, it is possible that reduction in plaques will reverse Alzheimer's, and so conceivable that **Bapi will help Alzheimer's patients**.

2. The amyloid cascade hypothesis may be correct but the evidence is far from conclusive"

explaining the replication crisis

1. low π

external pressure to develop new clinical interventions

explaining the replication crisis

2. non-negligible α

convention: accept target hypothesis (reject null hypothesis) if *p*-value is less than pre-defined value for α (usually 0.05)

type-I errors will occur in 5% of false hypotheses tested

explaining the replication crisis

2. non-negligible α

compare physics

convention is 5σ

type-I errors will occur in 1 in 3 million false hypotheses tested

other explanations

1. low statistical power
2. publication bias
3. bias, questionable research practices, and fraud

what is to be done?

1. live with it

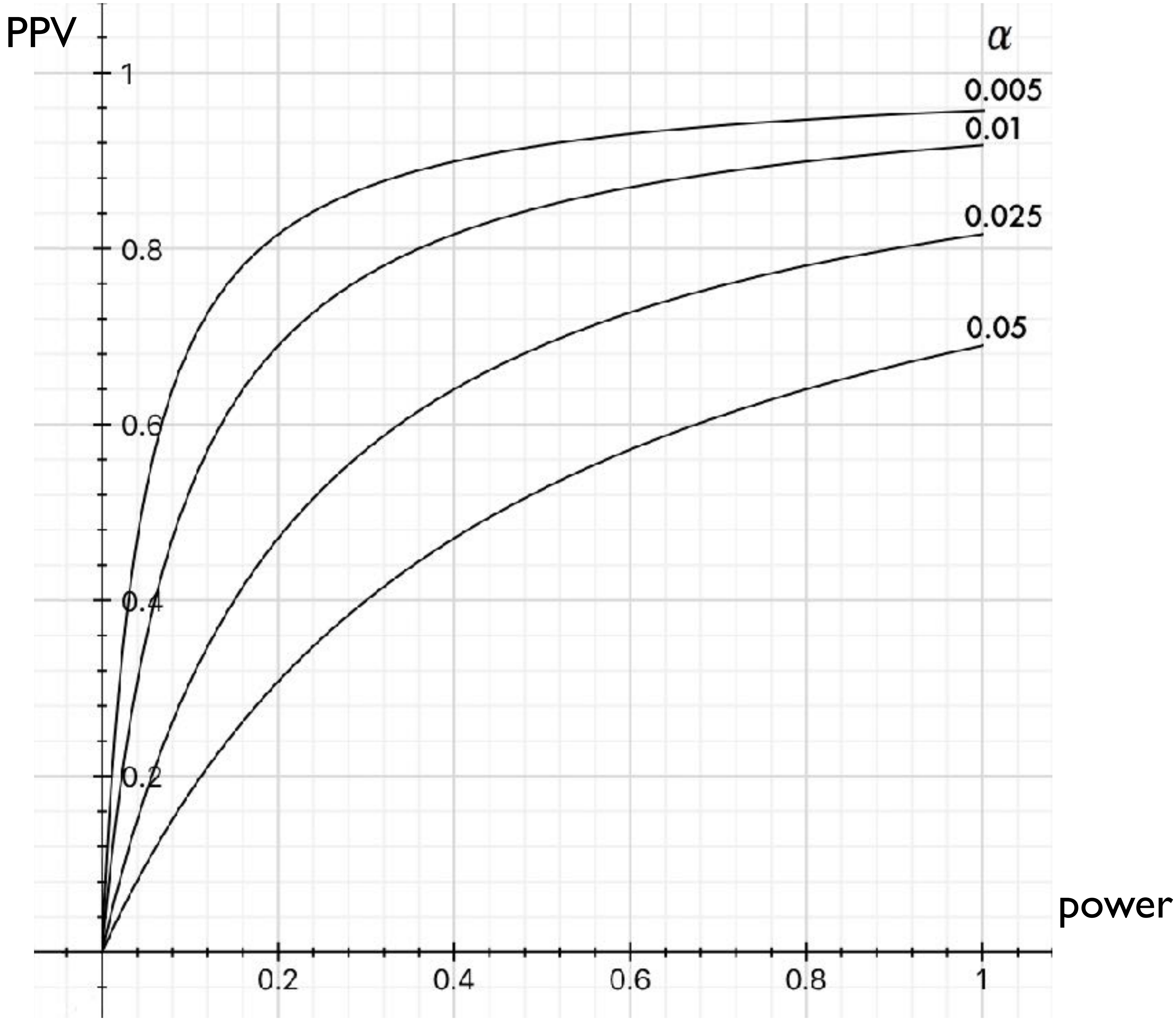
   but support replication better

2. increase π

   more basic science

3. decrease α

   trade-off between low α and effect size


PPV
α
1
0.8
0.6
0.4
0.2
0.005
0.01
0.025
0.05
power
0.2
0.4
0.6
0.8
1

from the base rate fallacy to the replication crisis

two types of error and accuracy

type of error	error rate	accuracy	type of accuracy
Type-I (false +ve)	α	$1-\alpha$	confidence level
Type-II (false −ve)	β	$1-\beta$	power

from the base rate fallacy to the replication crisis

	hypothesis is true	hypothesis is false
passes test	$\pi(1-\beta)$	$(1-\pi)\,\alpha$
fails test	$\pi\,\beta$	$(1-\pi)(1-\alpha)$

$$\text{FPRP} = \frac{(1-\pi)\alpha}{(1-\pi)\alpha + \pi(1-\beta)}$$

from the base rate fallacy to the replication crisis

False Positive Report Probability (FPRP) =

Pr (hypothesis is false, given that it passes the experimental test)

often the fallacy is made of confusing FPRP with:

Pr (hypothesis passes the test, given that it is false)

the latter is α

if $\pi = 0.1$ and $\beta = 0.05$

for FPRP < 0.05

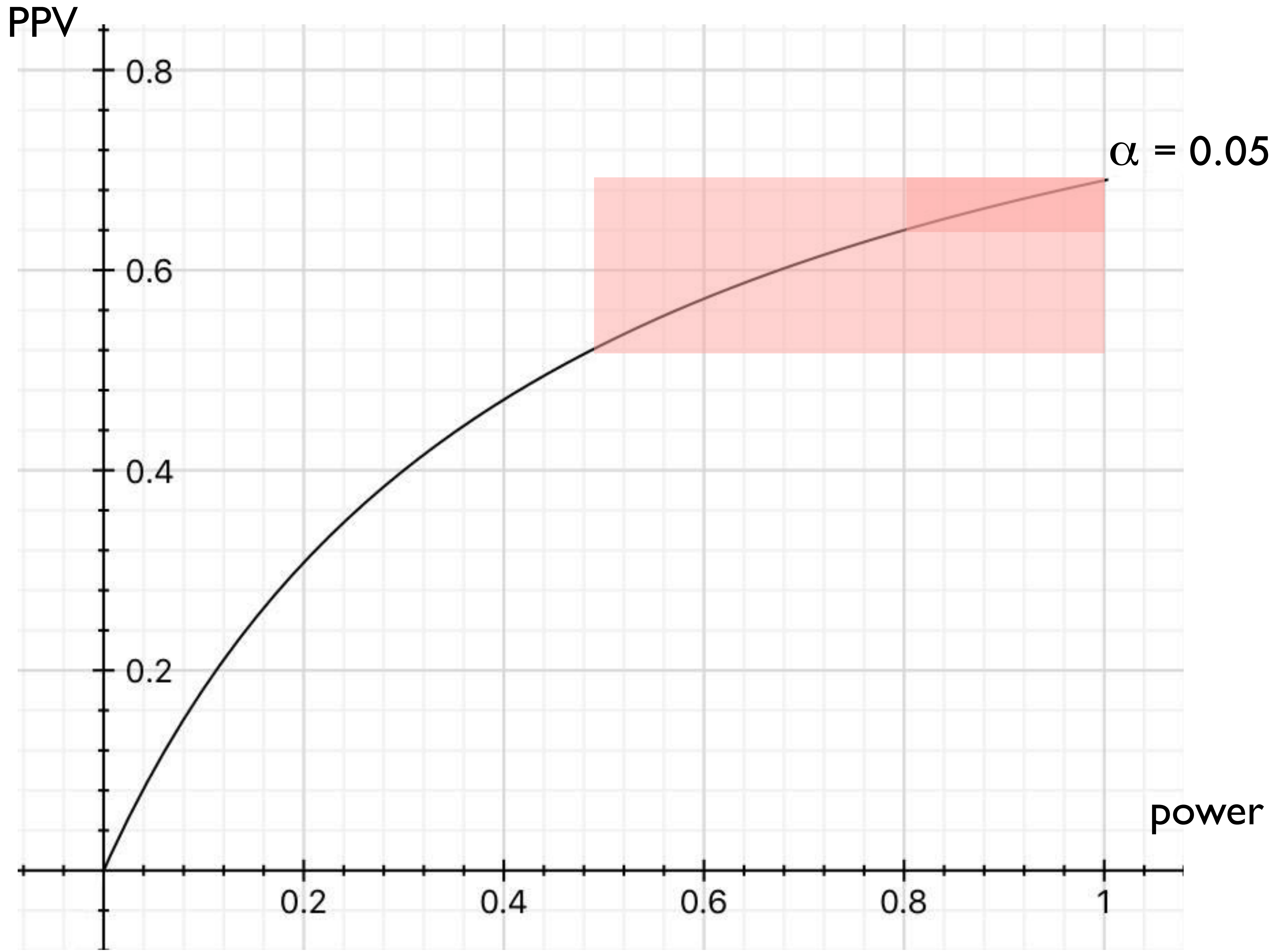
then $\alpha < 0.0056$

other explanations

1. low statistical power

$$\text{PPV} = \frac{\pi(1-\beta)}{(1-\pi)\alpha + \pi(1-\beta)}$$

decrease β → increase PPV



PPV
0.8
0.6
0.4
0.2
α = 0.05
power
0.2
0.4
0.6
0.8
1

other explanations

2. publication bias

other explanations

3. bias, questionable research practices, and fraud

explaining the replication crisis

1. low priors

psychology: hypotheses suggested by observational studies, unsystematic observation or intuition

explaining the replication crisis

1. low priors

hypotheses modelled on other hypotheses

suggestive but not a source of strong hypotheses

falsity feedback effect: hypotheses modelled on false hypotheses

header

1. history of science
2. use of psychology
3. socializing science

note that it is the *use* made of these that is original in *Structure*

theoretical history of science