Are most published research findings false?!

Vincent Detours

vdetours@ulb.ac.be

- IRIBHM, UNIVERSITE LIBRE DE BRUXELLES

A lot of biomedical research rests on purely statistical evidence

- Gene-disease association studies
- Epidemiology
- Search for molecular markers of disease status, progression, etc.
- Clinical trials for treatment effectivness
- Etc.

Researchers face a number of potential pitfalls

- Multiple testing issues
- Poorly defined outcome, a posteriori interpretation
- Data drenging
- Confirmation bias
- Difficulty to publish negative results. Pressure to publish strong claims
- Conflict of interest
- Etc.

The research community *as a whole* faces a number of potential pitfalls

- Multiple testing issues for ‘hot topics’
- Herd behaviors
- Co-optation
- Conflict of interest
- Prominence of initial claims. Invisibility of negative results
- Etc.

The research community *as a whole* may be modeled mathematically

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Ioannidis J.P., *PLoS Medicine*, 2005

Bayesian approach to scientific research

Basic question:

- What is the probability that a scientific claim is actually true *given* that a published study reports it is true?

Or, stated in mathematical terms:

- What is the *positive predictive value* (PPV) of formal statistical significance?

Bayasian approach to scientific research

Example

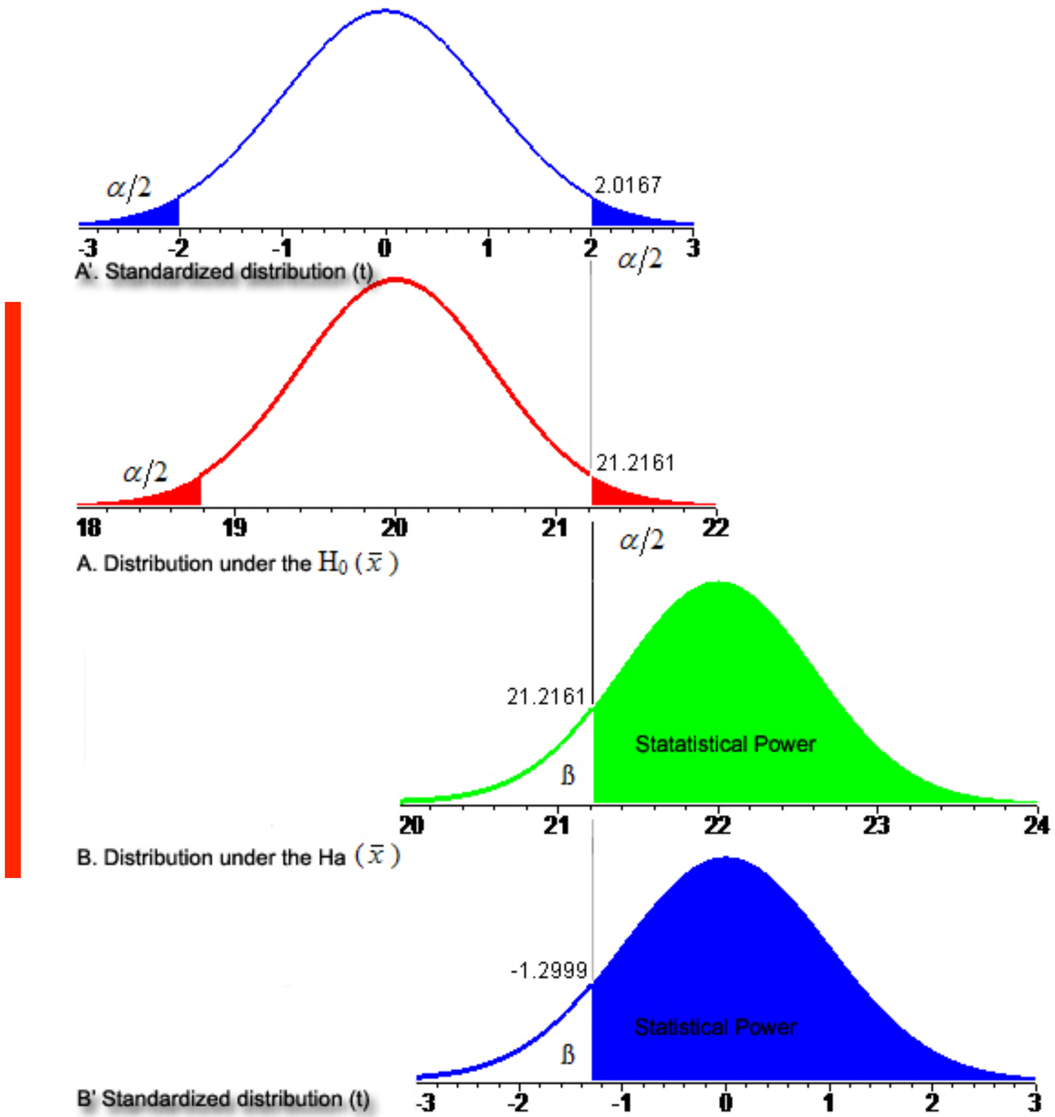
- 10 polymorphisms in 100.000 are likely to be associated with schizophrenia

> pre-study odds are 10^{-4}

Knowing that a statistical association with mutation in Gene Z was reported in *Nature*, what is the probability that gene Z mutation actually *does* predispose to schizophrenia?

Parameters of the model

- R = (true association)/(no association), it is research field-specific
- α = type I error rate (typically 0.05)
- 1-β = power of the study
- c = number of tests per study
- u = bias, i.e. proportion of studies reporting a significant result when there is none
- n = number of independent studies



From http://www.indiana.edu/~statmath/stat/all/power/statistical_power.jpg

Model 1: no bias, no multiple study

		Ground truth		
		Association	No Association	Total
Research	Significant			
	Not Significant			
	Total			

R defined as : $R = \dfrac{N_T}{N_F}$ P(true association) $= \dfrac{N_T}{N_T + N_F} = \dfrac{N_T/N_F}{N_T/N_F + 1} = \dfrac{R}{R+1}$

p(a AND B) = p(A) x P(b), if A and B are independent

P(A OR B) = p(a) + p(B), if A and B are independent

Model 1: no bias, no multiple study

		Ground truth		
		Association	No Association	Total
Research	Significant	???		
	Not Significant			
	Total			

R defined as : $R = \dfrac{N_T}{N_F}$ P(true association) $= \dfrac{N_T}{N_T + N_F} = \dfrac{N_T/N_F}{N_T/N_F + 1} = \dfrac{R}{R+1}$

p(there is an association AND scientist detects it as significant) x c

$$= \frac{R}{R+1} \cdot (1-\beta) \cdot c = c \cdot \frac{R(1-\beta)}{R+1}$$

Model 1: no bias, no multiple study

		Ground truth		
		Association	No Association	Total
Research	Significant	$c \cdot \frac{R(1-\beta)}{R+1}$	???	
	Not Significant			
	Total			

R defined as : $R = \frac{N_T}{N_F}$ P(true association) $= \frac{N_T}{N_T + N_F} = \frac{N_T / N_F}{N_T / N_F + 1} = \frac{R}{R+1}$

p(there is no association AND scientist detects it as significant) x c

$$= c \cdot \left(1 - \frac{R}{R+1}\right) . \alpha = \frac{c \cdot \alpha}{R+1}$$

Model 1: no bias, no multiple study

		Ground truth		
		Association	No Association	Total
Research	Significant	$c \cdot \frac{R(1-\beta)}{R+1}$	$\frac{c \cdot \alpha}{R+1}$	$c \cdot \frac{R+\alpha-\beta R}{R+1}$
	Not Significant	???		
	Total			

R defined as : $R = \frac{N_T}{N_F}$ P(true association) $= \frac{N_T}{N_T + N_F} = \frac{N_T/N_F}{N_T/N_F + 1} = \frac{R}{R+1}$

p(there is association AND scientist fails to detect it as significant) x c

$$= c \cdot \frac{R}{R+1} \cdot \beta$$

Model 1: no bias, no multiple study

		Ground truth		
		Association	No Association	Total
Research	Significant	$c \cdot \frac{R(1-\beta)}{R+1}$	$\frac{c \cdot \alpha}{R+1}$	$c \cdot \frac{R+\alpha-\beta R}{R+1}$
	Not Significant	$= c \cdot \frac{R}{R+1} \cdot \beta$	???	
	Total			

R defined as : $R = \frac{N_T}{N_F}$ P(true association) $= \frac{N_T}{N_T + N_F} = \frac{N_T/N_F}{N_T/N_F + 1} = \frac{R}{R+1}$

p(there is no association AND scientist does not detect it as significant) x c

$$= c \cdot \left(1 - \frac{R}{R+1}\right) \cdot (1-\alpha) = c \cdot \frac{1-\alpha}{R+1}$$

Model 1: no bias, no multiple study

		Ground truth		
		Association	No Association	Total
Research	Significant	$c \cdot \frac{R(1-\beta)}{R+1}$	$\frac{c \cdot \alpha}{R+1}$	$c \cdot \frac{R+\alpha-\beta R}{R+1}$
	Not Significant	$= c \cdot \frac{R}{R+1} \cdot \beta$	$c \cdot \frac{1-\alpha}{R+1}$	$c \cdot \frac{1-\alpha+\beta R}{R+1}$
	Total	$\frac{cR}{R+1}$	$\frac{c}{R+1}$	c

$$PPV = \frac{TP}{TP+FP} = \frac{\frac{(1-\beta)Rc}{R+1}}{\frac{(1-\beta)Rc}{R+1} + \frac{c\alpha}{R+1}} = \frac{(1-\beta)R}{(1-\beta)R+\alpha}$$

Model 1: no bias, no multiple study

Table 1. Research Findings and True Relationships

Research Finding	True Relationship Yes	No	Total
Yes	$c(1-\beta)R/(R+1)$	$c\alpha/(R+1)$	$c(R+\alpha-\beta R)/(R+1)$
No	$c\beta R/(R+1)$	$c(1-\alpha)/(R+1)$	$c(1-\alpha+\beta R)/(R+1)$
Total	$cR/(R+1)$	$c/(R+1)$	c

Model 1: no bias, no multiple study

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

Schizophrenia example

- $R = 10^{-4}$
- c = 100.000
- $\alpha = 0.05$
- $1-\beta$ = 60% at effect size 1.3

Post-study probability = $12 * 10^{-4}$!

Model 2: effect of bias

Table 1. Research Findings and True Relationships

Research Finding	True Relationship Yes	No	Total
Yes	$c(1-\beta)R/(R+1)$	$c\alpha/(R+1)$	$c(R+\alpha-\beta R)/(R+1)$
No	$c\beta R/(R+1)$	$c(1-\alpha)/(R+1)$	$c(1-\alpha+\beta R)/(R+1)$
Total	$cR/(R+1)$	$c/(R+1)$	c

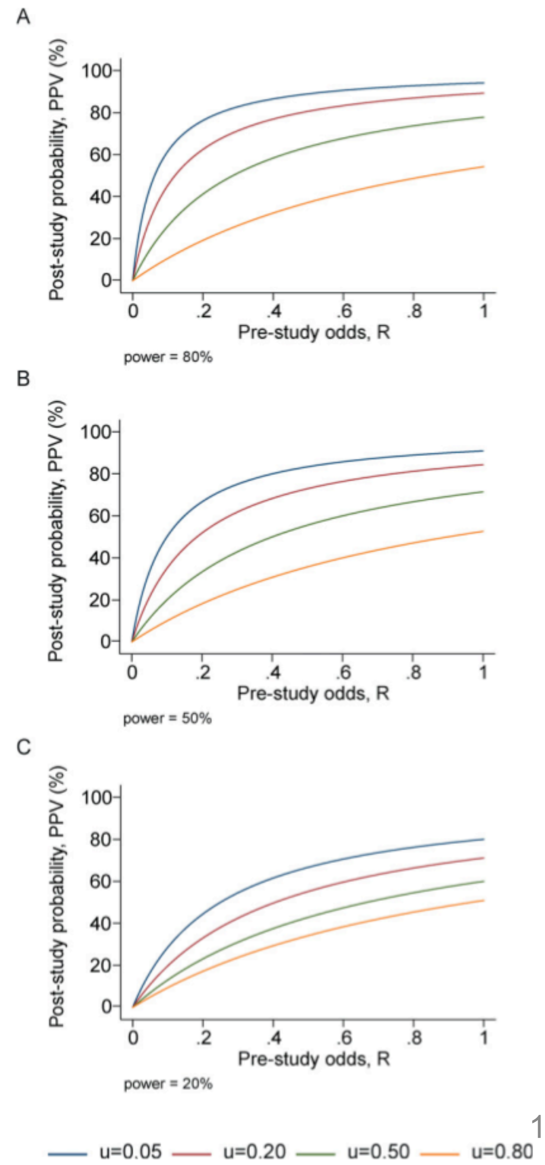
Table 2. Research Findings and True Relationships in the Presence of Bias

Research Finding	True Relationship Yes	No	Total
Yes	$(c[1-\beta]R+uc\beta R)/(R+1)$	$c\alpha+uc(1-\alpha)/(R+1)$	$c(R+\alpha-\beta R+u-u\alpha+u\beta R)/(R+1)$
No	$(1-u)c\beta R/(R+1)$	$(1-u)c(1-\alpha)/(R+1)$	$c(1-u)(1-\alpha+\beta R)/(R+1)$
Total	$cR/(R+1)$	$c/(R+1)$	c

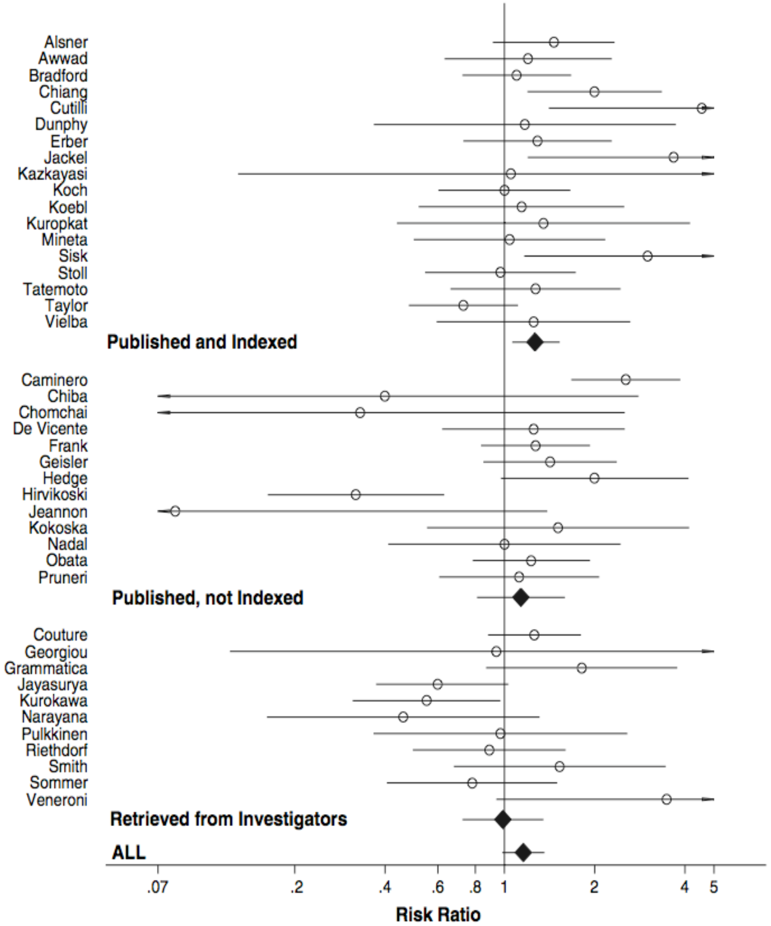
Model 2: effect of bias

$$PPV = \frac{(1-\beta)R + u\beta R}{R - \beta R + \alpha + u - u\alpha + u\beta R}$$

Bias strongly reduces PPV



Biases *do* exist in actual published studies



- Meta-analysis of studies of p53 status association of head and neck cancer outcome

- Results are stratified by publication status

(from Kyzas et al. JNCI, 2005)

Model 3: effect of multiple studies

Table 1. Research Findings and True Relationships

Research Finding	True Relationship Yes	No	Total
Yes	$c(1-\beta)R/(R+1)$	$c\alpha/(R+1)$	$c(R+\alpha-\beta R)/(R+1)$
No	$c\beta R/(R+1)$	$c(1-\alpha)/(R+1)$	$c(1-\alpha+\beta R)/(R+1)$
Total	$cR/(R+1)$	$c/(R+1)$	c

Table 3. Research Findings and True Relationships in the Presence of Multiple Studies

Research Finding	True Relationship Yes	No	Total
Yes	$cR(1-\beta^n)/(R+1)$	$c(1-[1-\alpha]^n)/(R+1)$	$c(R+1-[1-\alpha]^n-R\beta^n)/(R+1)$
No	$cR\beta^n/(R+1)$	$c(1-\alpha)^n/(R+1)$	$c([1-\alpha]^n+R\beta^n)/(R+1)$
Total	$cR/(R+1)$	$c/(R+1)$	c

Model 3: effect of multiple studies

$$PPV = \frac{(1-\beta^n)R}{R + 1 - (1-\alpha)^n - \beta R^n}$$

Studies in ‘hot’ fields have reduced PPV

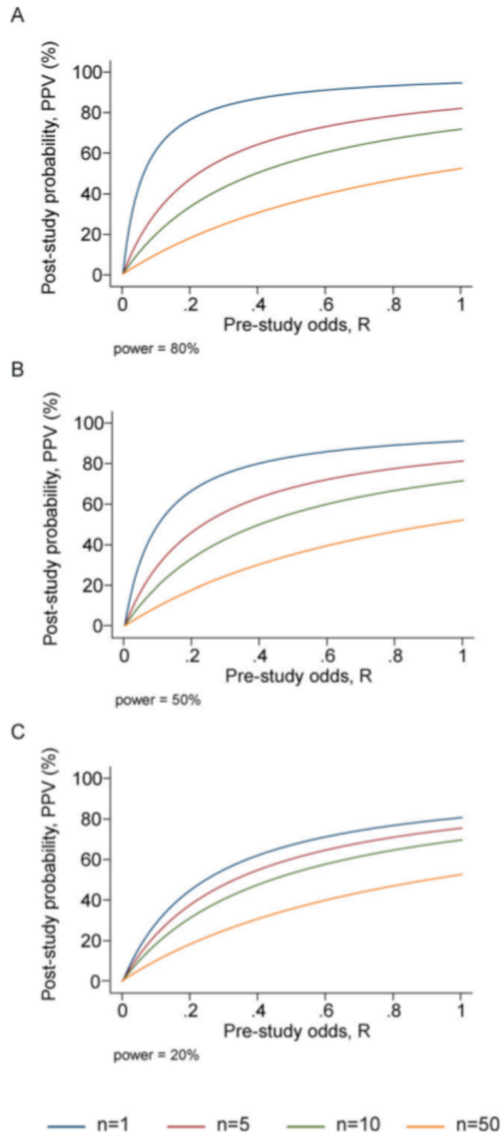


Table 4. PPV of Research Findings for Various Combinations of Power ($1 - \beta$), Ratio of True to Not-True Relationships *(R)*, and Bias *(u)*

$1 - \beta$	*R*	*u*	Practical Example	PPV
0.80	1:1	0.10	Adequately powered RCT with little bias and 1:1 pre-study odds	0.85
0.95	2:1	0.30	Confirmatory meta-analysis of good-quality RCTs	0.85
0.80	1:3	0.40	Meta-analysis of small inconclusive studies	0.41
0.20	1:5	0.20	Underpowered, but well-performed phase I/II RCT	0.23
0.20	1:5	0.80	Underpowered, poorly performed phase I/II RCT	0.17
0.80	1:10	0.30	Adequately powered exploratory epidemiological study	0.20
0.20	1:10	0.30	Underpowered exploratory epidemiological study	0.12
0.20	1:1,000	0.80	Discovery-oriented exploratory research with massive testing	0.0010
0.20	1:1,000	0.20	As in previous example, but with more limited bias (more standardized)	0.0015