# EVIDENCE BASED MEDICINE: HEILIGE GRAAL OF GEVAARLIJKE ILLUSIE?

Yvo Smulders

Academie Nieuwezorg

Epidemiologie:
Waarneming in groepen

EBM: geneeskunde die leunt
op epidemiologisch verkregen bewijs

Academie
Nieuwezorg

**EBP…heeft de systematische reflectie op de gevolgen van het medisch handelen sterk bevorderd, maar…**

EBM

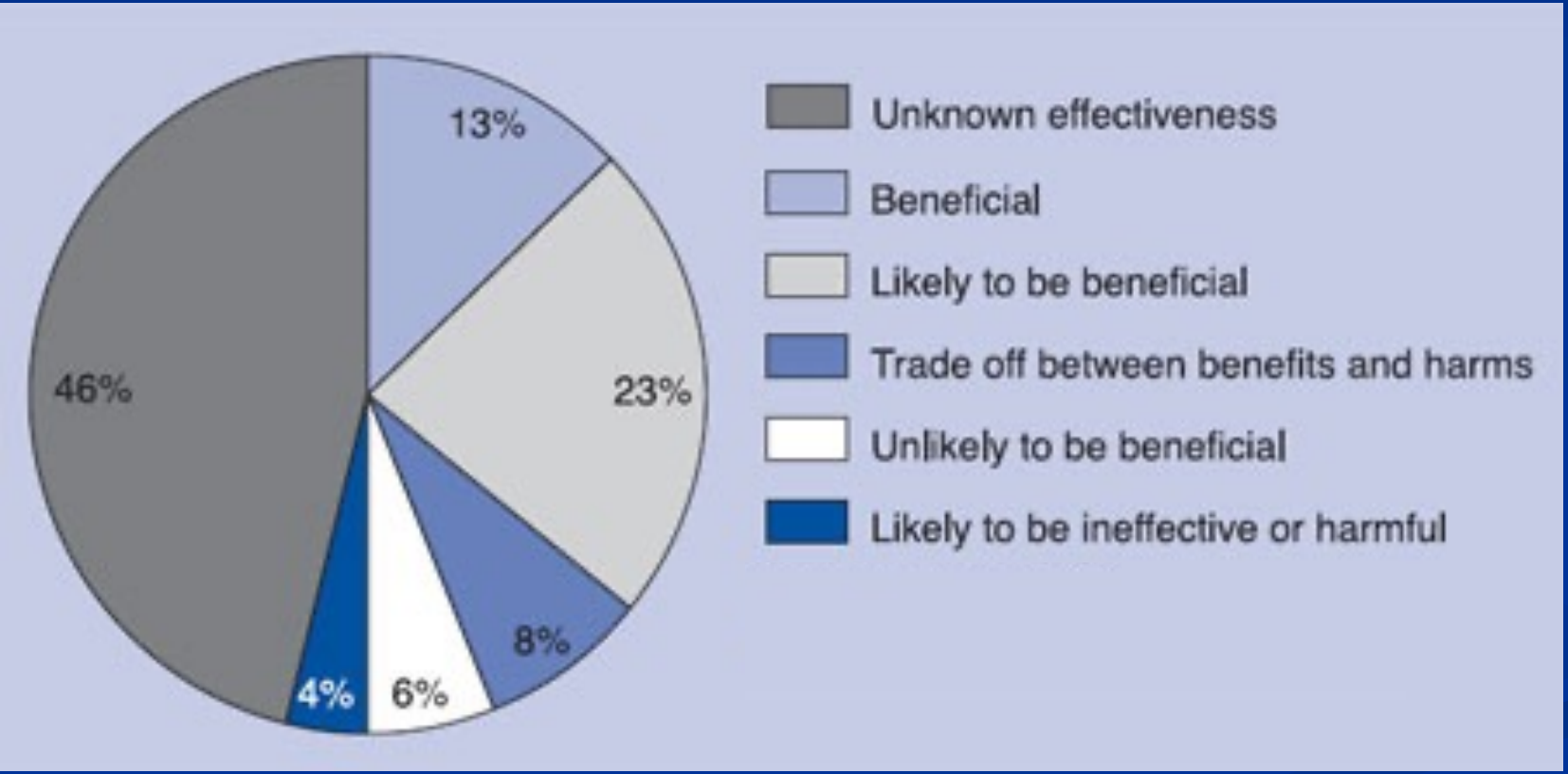
- … is een reductie van de werkelijkheid….
- … houdt onvoldoende rekening met verschillen tussen patiënten …
- … stuurt de zorgpraktijk in de richting van datgene
  - wat kan worden onderzocht
  - waar geld voor is/commercieel interessant is

# Problemen met epidemiologisch bewijs

- ...is er vaak niet
- ...is er wel, maar niet voor jouw patiënt
- ...onbetrouwbaar

Academie Nieuwezorg

# Epidemiologisch bewijs voor veel-voorkomende therapieën




http://clinicalevidence.com

Academie Nieuwezorg

JAMA®

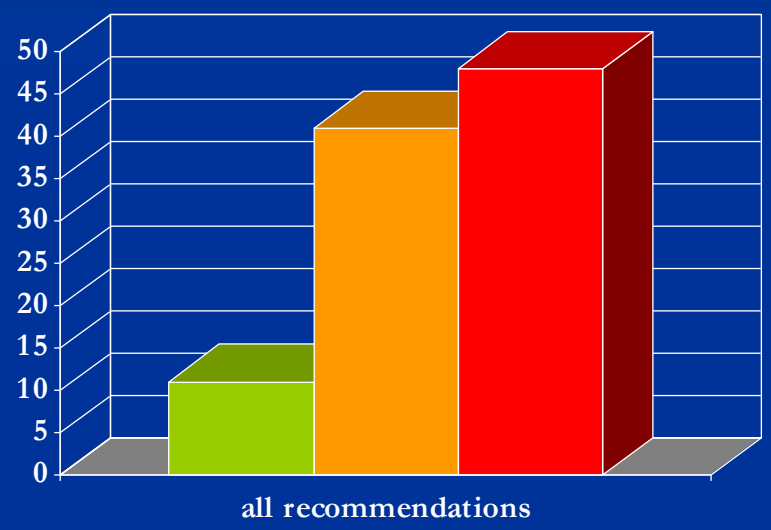
Online article and related content current as of March 2, 2009.

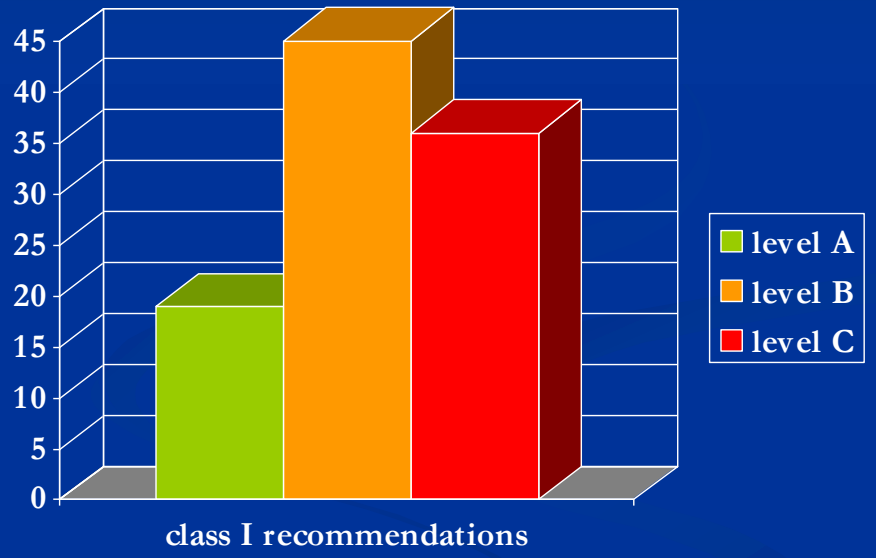
## Scientific Evidence Underlying the ACC/AHA Clinical Practice Guidelines

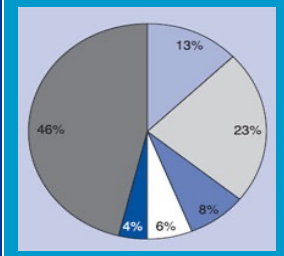
Pierluigi Tricoci; Joseph M. Allen; Judith M. Kramer; et al.

*JAMA*. 2009;301(8):831-841 (doi:10.1001/jama.2009.205)

http://jama.ama-assn.org/cgi/content/full/301/8/831




**Meeste is niet onderzocht**
**Dat halen we nooit in!**

Academie Nieuwezorg

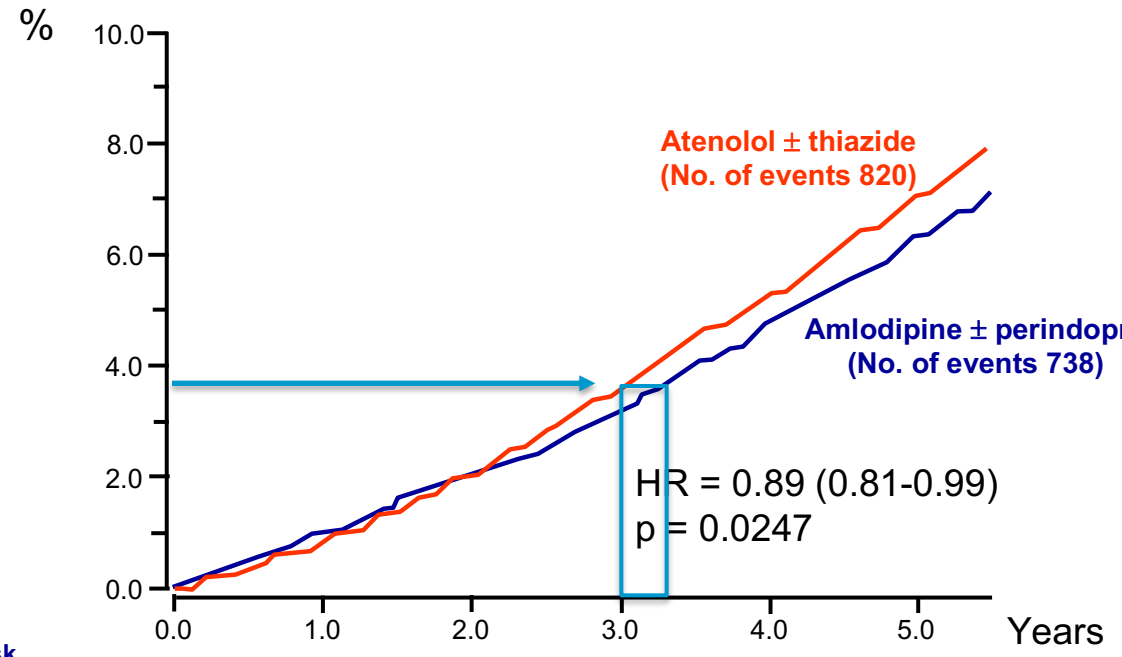
# Beschikbaar bewijs: *Focus*

- Ja: beste modaliteit
- Nee: beste *wijze* van behandeling

Academie Nieuwezorg

# Behandel-modaliteit

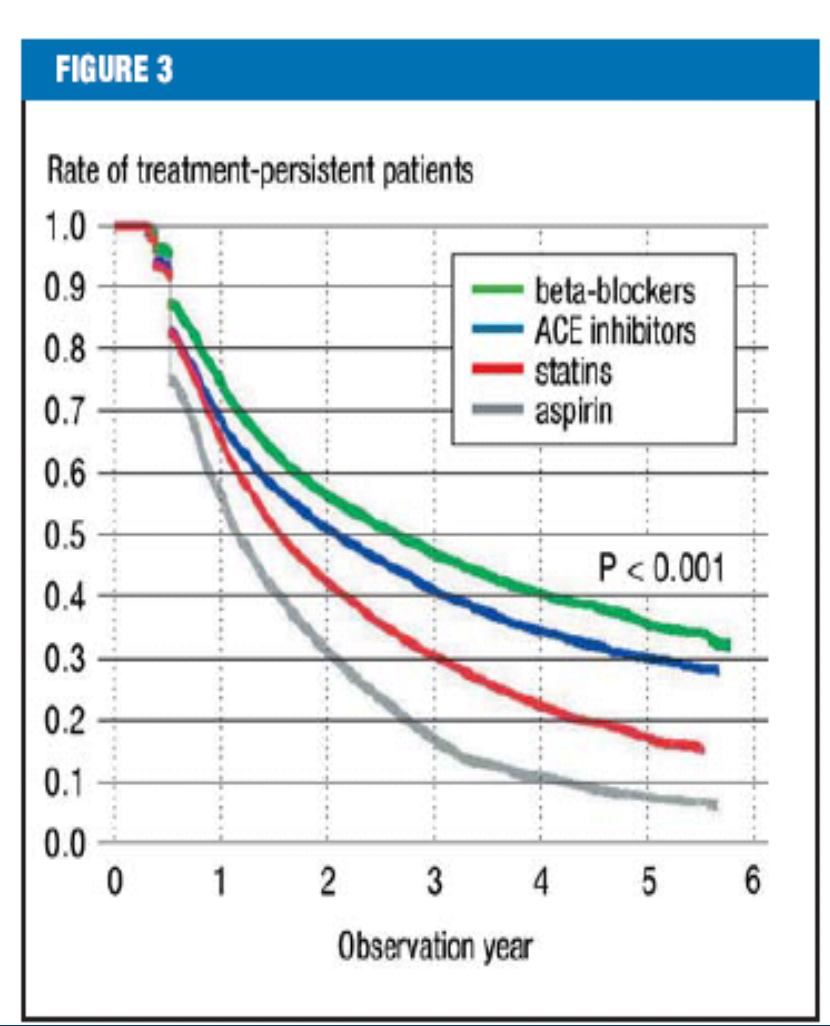
## All-cause mortality



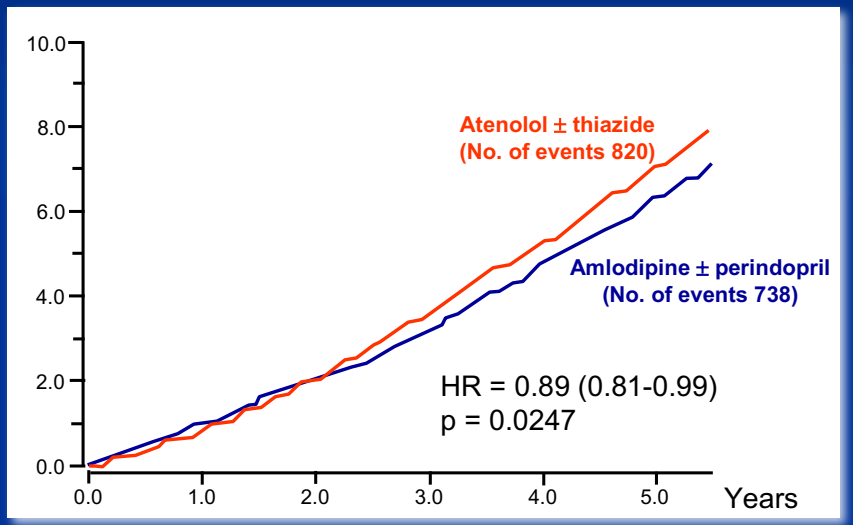

| Number at risk | | | | | | |
|---|---|---|---|---|---|---|
| Amlodipine ± perindopril | 9639 | 9544 | 9441 | 9332 | 9167 | 8078 |
| Atenolol ± thiazide | 9618 | 9532 | 9415 | 9261 | 9085 | 7975 |

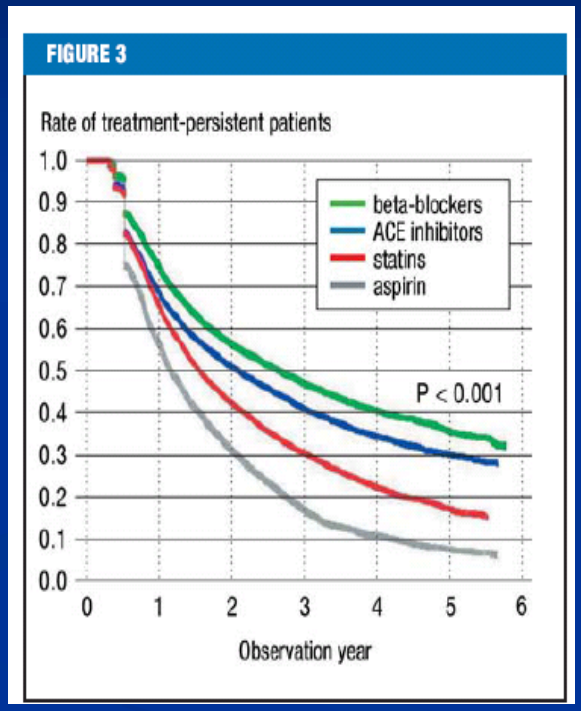
ascot

# Medicatie-trouw na hartinfarct




Academie Nieuwezorg




Academie Nieuwezorg

REVIEW ARTICLE

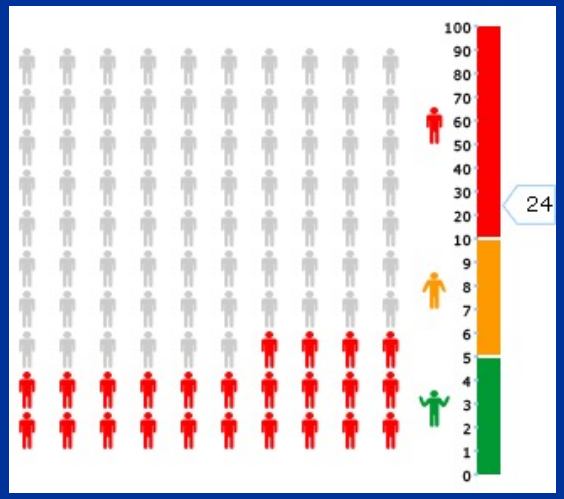
# Interventions to Enhance Medication Adherence in Chronic Medical Conditions

*A Systematic Review*

*Sunil Kripalani, MD, MSc; Xiaomei Yao, MD; R. Brian Haynes, MD, PhD*

*Arch Int Med 2007;167;540*

informatie




x3

Pt zélf terugzien

x6

Electronic pill box

x3

# Beschikbaar bewijs: *Focus*

- Ja: beste modaliteit
- Nee: beste *wijze* van behandeling
- Al helemáál niet: diagnostiek

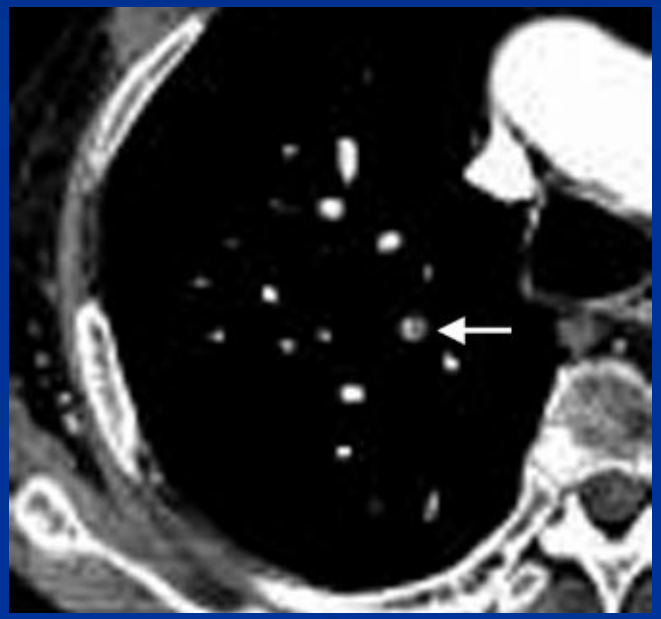
Academie Nieuwezorg

LESS IS MORE

# Time Trends in Pulmonary Embolism in the United States

*Evidence of Overdiagnosis*

*Arch Intern Med. 2011;171(9):831-837*



Veel meer diagnoses longembolie
Gelijk aantal doden door longembolie
Sterke toename in doden door bloedverdunners

Academie Nieuwezorg

# Problems of Epidemiological Literature

- ...is er niet wanneer er echt behoefte is
- ...is er wel, maar niet voor jouw patiënt
- ...onbetrouwbaar

Academie Nieuwezorg

# Het bewijs is er…maar voor wie??

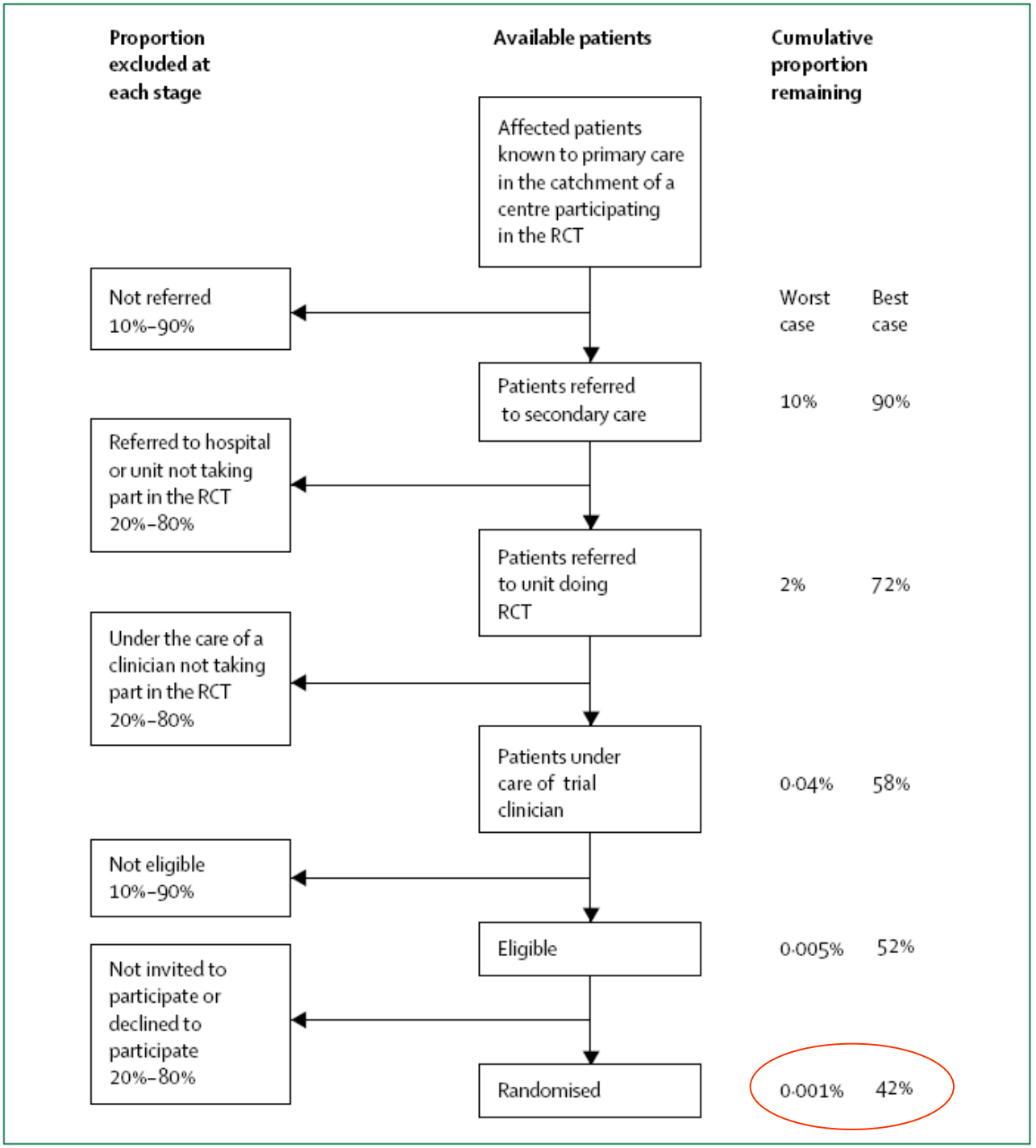
- …voor de gemiddelde trial-deelnemer

## Treating Individuals 1

**External validity of randomised controlled trials: "To whom do the results of this trial apply?"**

*Lancet* 2005; 365: 82–93 *Peter M Rothwell*

Academie Nieuwezorg


Proportion excluded at each stage
Available patients
Cumulative proportion remaining
Affected patients known to primary care in the catchment of a centre participating in the RCT
Not referred 10%–90%
Worst case
Best case
Patients referred to secondary care
10%
90%
Referred to hospital or unit not taking part in the RCT 20%–80%
Patients referred to unit doing RCT
2%
72%
Under the care of a clinician not taking part in the RCT 20%–80%
Patients under care of trial clinician
0·04%
58%
Not eligible 10%–90%
Eligible
0·005%
52%
Not invited to participate or declined to participate 20%–80%
Randomised
0·001%
42%

# Het bewijs is er…maar voor wie??

- …de gemiddelde trial-deelnemer

## Treating Individuals 1

**External validity of randomised controlled trials: "To whom do the results of this trial apply?"**

*Lancet* 2005; 365: 82–93 *Peter M Rothwell*

→ '10% or less…'

JAMA®

Online article and related content current as of April 20, 2009.

**Eligibility Criteria of Randomized Controlled Trials Published in High-Impact General Medical Journals: A Systematic Sampling Review**

Harriette G. C. Van Spall; Andrew Toren; Alex Kiss; et al.

*JAMA*. 2007;297(11):1233-1240 (doi:10.1001/jama.297.11.1233)

http://jama.ama-assn.org/cgi/content/full/297/11/1233

'…exclusie-criteria zijn veel-voorkomend én relevant…': leeftijd, bijkomende ziekten, contra-indicaties

Academie Nieuwezorg

+

=

Gezond verstand
Kennis van mechanismen
Ervaring
Intuïtie
~~'geloof'~~
~~'magie'~~

Academie Nieuwezorg

# Het bewijs is er…maar voor wie??

- Je moet ALTIJD extrapoleren
- Die extrapolatie is niet evidence-based, per definitie
- Echte Evidence-based Medicine kan eigenlijk alléén op groepsniveau…

Academie Nieuwezorg

AGGREGATIENIVEAU
Academie
Nieuwezorg

# EBM … TUSSENSTAND

- <1/3 is ooit bestudeerd
- Toepasbaar op <10% van je patiënten

→ 1/30 of 'frequently applied treatments' is truly evidence-based

Academie Nieuwezorg

# Problemen met epidemiologisch bewijs

- ...is er niet wanneer er echt behoefte is
- ...is er wel, maar niet voor jou
- ...onbetrouwbaar

Academie Nieuwezorg

# wantrouwen?

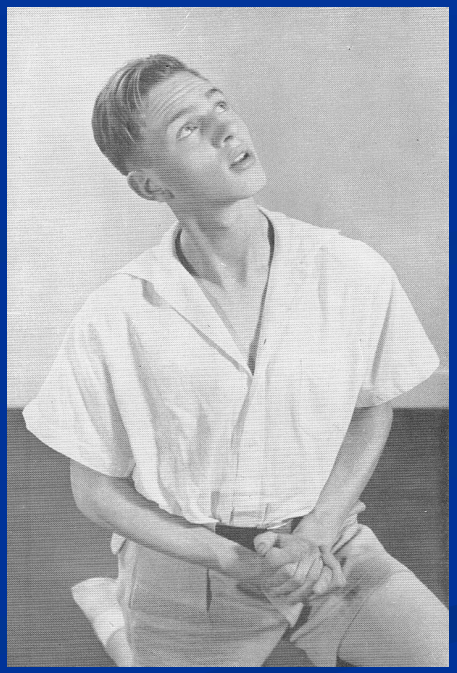
- Bias en Berekening
- Belangen
- Boerenbedrog

Academie Nieuwezorg

# wantrouwen

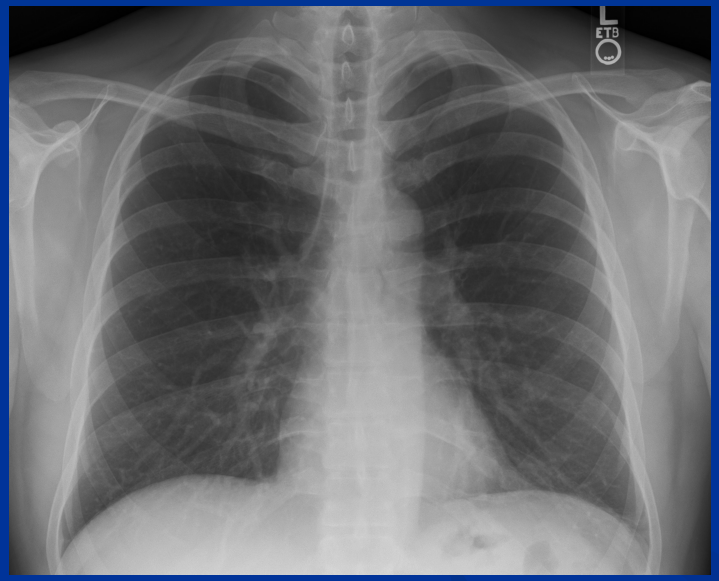
- Stats and Bias: een ‘ongeladen’ start…
- Conflicts of interest
- Outright fraud

*P-value*



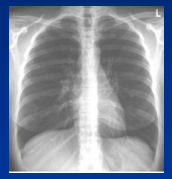
Academie Nieuwezorg

# Thomas Bayes



Academie Nieuwezorg

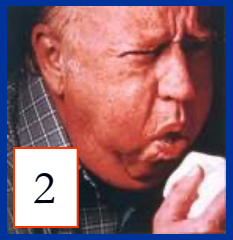
# Sensitiviteit en specificiteit



To diagnose pneumonia: sensitivity: 0.95 / specificity 0.80

Bayes Criterion:
Odds prior to test x test characteristics pos./neg. test = Odds for disease áfter test

→ x (0.95/0.2) = 0.5 → P-pneumonia=33%



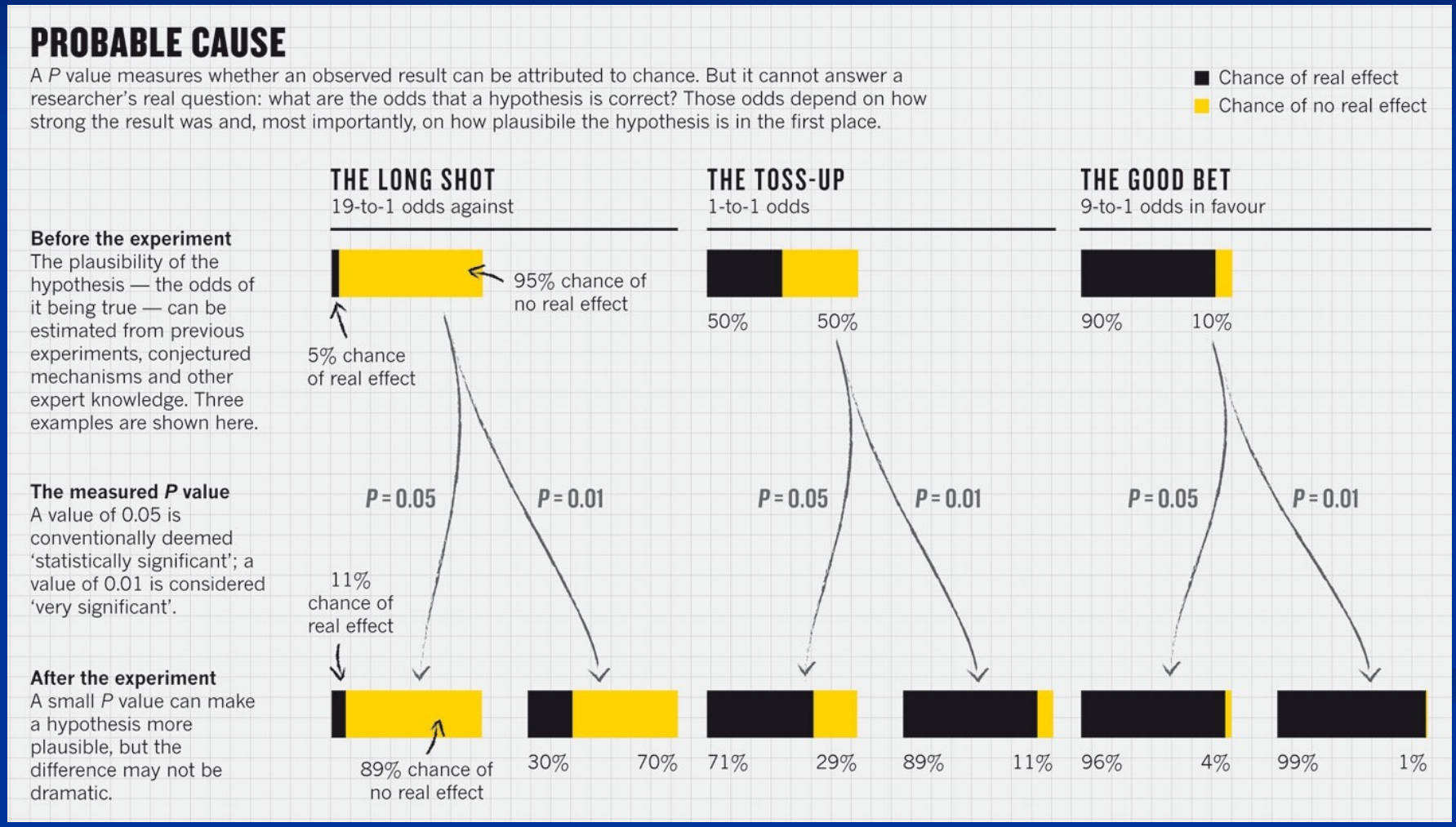

→ x (0.95/0.2) = 9.5 → P-pneumonia=90%

Academie Nieuwezorg

# Een onderzoek is een test

Academie Nieuwezorg

# Het belang van plausibiliteit



Essay

# Why Most Published Research Findings Are False

John P. A. Ioannidis

Betrouwbaarheid van statistisch significante resultaten hangt af van:

- Power
- Grootte van gevonden effect
- (selectie van) aantal getestte associaties
- 'flexibiliteit' in design, keuze uitkomstmaten, analyses
- Vooroordelen en belangen
- 'hot'ness van onderwerp (tijdsdruk, napraten)
- (Biologische) plausibiliteit

Academie Nieuwe zorg

Ioannidis, PLoS Medicine 2005; 2: 696-701

Essay

# Why Most Published Research Findings Are False

John P. A. Ioannidis

Alles statistisch significant

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

| 1 – β | *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 |

# wantrouwen?

- Bias en Berekening: *meeste onwaar*
- Belangen
- Boerenbedrog

Academie Nieuwezorg

SPECIAL ARTICLE

# A National Survey of Physician–Industry Relationships

Eric G. Campbell, Ph.D., Russell L. Gruen, M.D., Ph.D., James Mountford, M.D., Lawrence G. Miller, M.D., Paul D. Cleary, Ph.D., and David Blumenthal, M.D., M.P.P.

- N=3200 US physicians
- 94% financial COIs
  - 83% food/lunches, etc
  - 78% drug samples
  - 35% conference or CME support
  - 28% consulting-lecture fees etc

Wie?

NEJM 356:1742

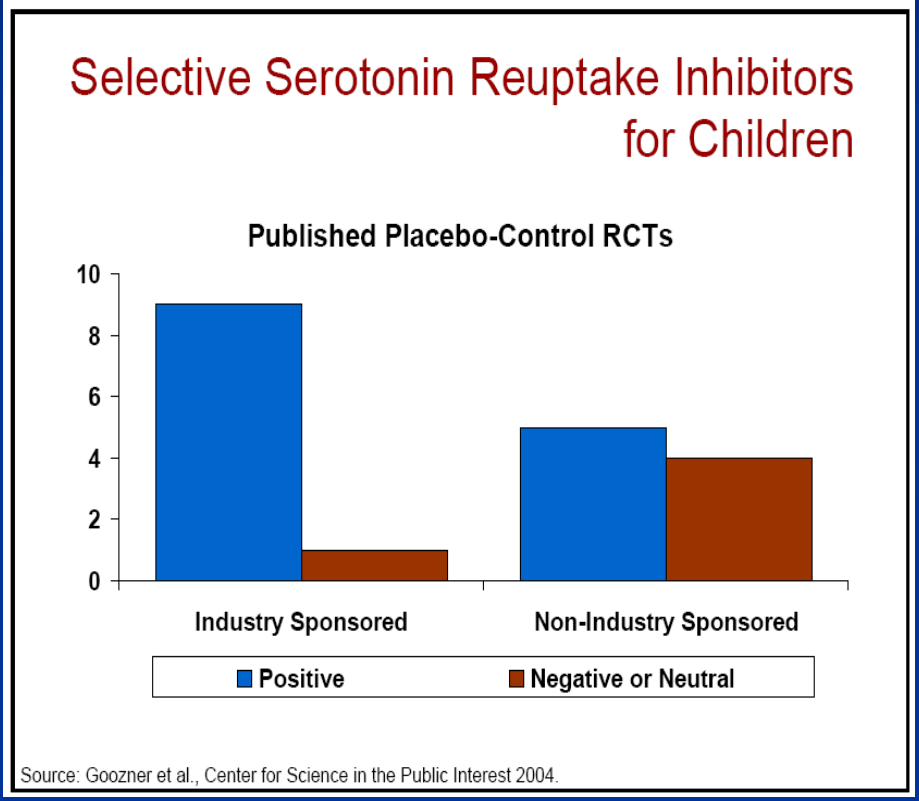
Academie Nieuwezorg

**Table 3.** Multivariate Predictors of Physician–Industry Relationships.*

| Characteristic | Type of Physician–Industry Relationship | | | |
|---|---|---|---|---|
| | Samples | Gifts | Reimbursements | Payments |
| | *odds ratio (95% CI)* | | | |
| **Personal** | | | | |
| Sex | | | | |
| Male | 1.00 | 1.00 | 1.00 | 1.00 |
| Female | 0.93 (0.61–1.41) | 1.18 (0.75–1.84) | 1.01 (0.73–1.42) | 0.37 (0.24–0.58)† |
| Race or ethnic group | | | | |
| Non-underrepresented | 1.00 | 1.00 | 1.00 | 1.00 |
| Underrepresented | 1.80 (0.97–3.34) | 1.69 (0.83–3.43) | 1.77 (1.14–2.74)† | 1.46 (0.89–2.40) |
| **Professional** | | | | |
| No. of years in practice | | | | |
| <20 yr | 1.00 | 1.00 | 1.00 | 1.00 |
| ≥20 yr | 0.92 (0.67–1.27) | 0.58 (0.40–0.84)† | 1.18 (0.89–1.56) | 1.18 (0.86–1.61) |
| Specialty | | | | |
| Family practice | 1.00‡ | 1.00 | 1.00‡ | 1.00‡§¶ |
| Internal medicine | 0.72 (0.42–1.25)‡ | 0.54 (0.32–0.90)† | 1.26 (0.87–1.83)‡‖ | 1.35 (0.89–2.04)‡¶‖ |
| Pediatrics | 0.56 (0.33–0.94)†‡ | 0.67 (0.40–1.12) | 0.59 (0.41–0.86)†** | 0.51 (0.33–0.78)†§** |
| Anesthesiology | 0.05 (0.03–0.09)†§¶‖**†† | 0.89 (0.49–1.64) | 0.31 (0.20–0.48)†§¶**†† | 0.21 (0.12–0.36)†§**†† |
| Cardiology | 1.64 (0.79–3.41)‡¶ | 1.14 (0.61–2.13) | 1.04 (0.69–1.55)‡ | 2.20 (1.43–3.38)†‡¶‖†† |
| Surgery | 0.43 (0.24–0.77)†‡§ | 0.82 (0.47–1.43) | 0.75 (0.51–1.11)‡ | 0.43 (0.27–0.67)†§**†† |

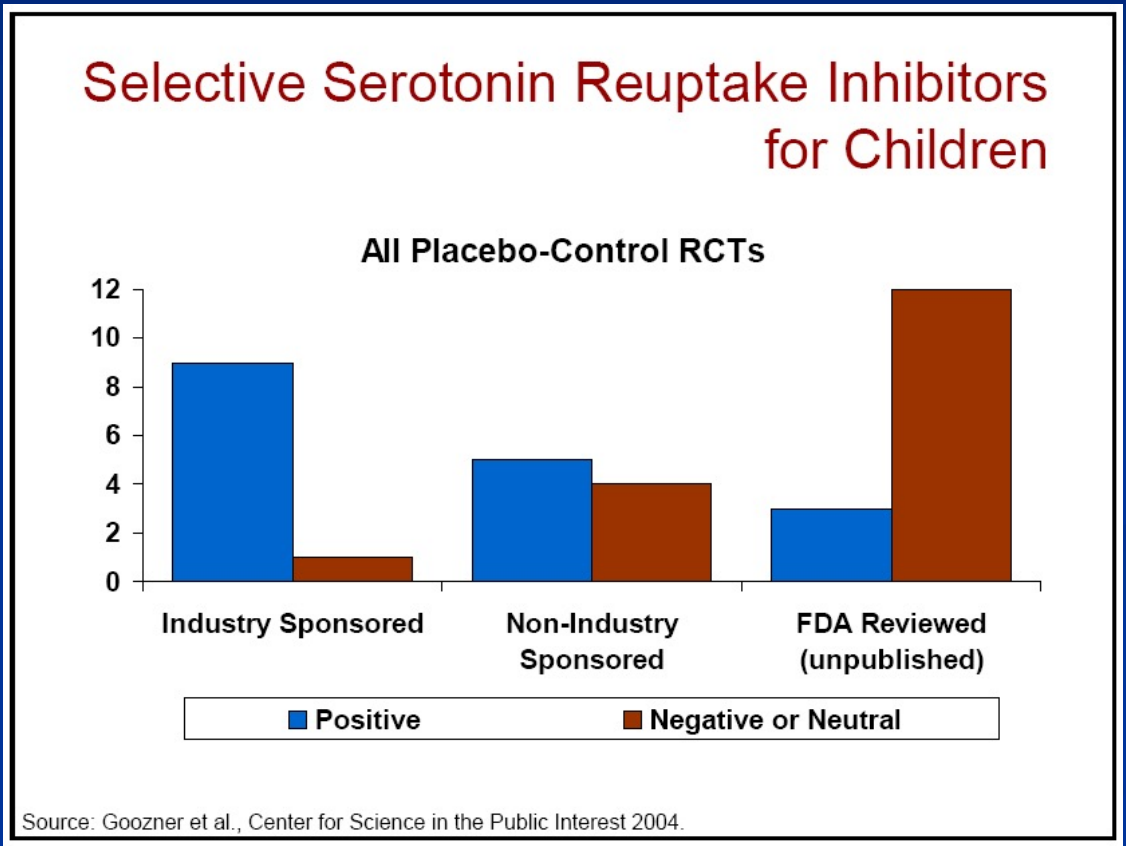
Academie Nieuwezorg

# COIs: Reporting and publication bias




Academie Nieuwezorg

# COIs: Reporting and publication bias




Academie Nieuwezorg

JAMA®

Online article and related content current as of October 1, 2008.

# Reported Outcomes in Major Cardiovascular Clinical Trials Funded by For-Profit and Not-for-Profit Organizations: 2000-2005

Paul M Ridker; Jose Torres

*JAMA*. 2006;295(19):2270-2274 (doi:10.1001/jama.295.19.2270)

http://jama.ama-assn.org/cgi/content/full/295/19/2270

**Table 2.** Proportion of Trials Significantly Favoring Newer Treatments Over Standard of Care

| Trials | No./Total % | | | *P* for Trend |
|---|---|---|---|---|
| | Not-for-Profit (n = 104) | Not-for-Profit and For-Profit (n = 62) | For-Profit (n = 137) | |
| All | 51/104 (49.0) | 35/62 (56.5) | 92/137 (67.2) | .005 |
| Clinical end points | 19/55 (34.6) | 24/44 (54.6) | 64/96 (66.7) | <.001 |
| Drug | 17/43 (39.5) | 24/46 (54.4) | 74/113 (65.5) | .002 |
| Device | 4/8 (50.0) | 9/13 (69.2) | 14/17 (82.4) | .07 |

Academie Nieuwezorg

JAMA®

Online article and related content current as of September 28, 2008.

## Everyone's a Little Bit Biased (Even Physicians)

Daylian M. Cain; Allan S. Detsky

*JAMA*. 2008;299(24):2893-2895 (doi:10.1001/jama.299.24.2893)

http://jama.ama-assn.org/cgi/content/full/299/24/2893

They believe their opinions are truly their own; but that is not the point. The question is: how did they come to have those beliefs?

Academie Nieuwezorg

# The Unintended Consequences of Conflict of Interest Disclosure

George Loewenstein, PhD
Sunita Sah, MD, PhD
Daylian M. Cain, PhD

JAMA, February 15, 2012

However, perhaps the most significant likely pitfall of disclosure is not its effects on the quality of advice received by individual patients or its inapplicability to many serious conflicts of interest, but the likelihood of a kind of moral licensing on the part of the profession as a whole—the rationalization that, with disclosure, the profession has dispensed with its obligation to deal with conflicts of interest. Conflicts of interest, including fee-for-service arrangements, are at the heart of the astronomical increases in health care costs in the United States, and transparency is no substitute for more substantive reform.

Academie Nieuwezorg

# distrust?

- Stats and bias
- Conflicts of interest: still highly relevant
- Outright fraud

Academie Nieuwezorg

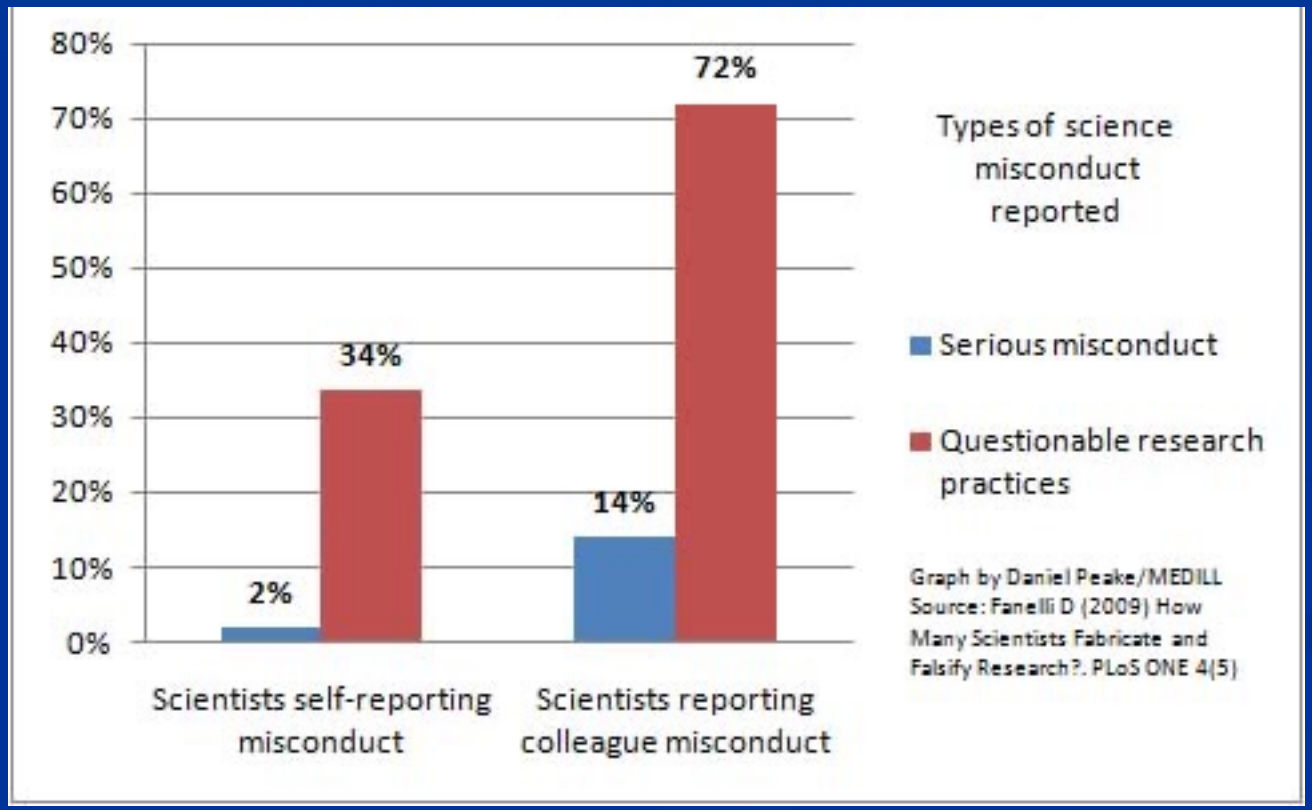
OPEN ACCESS Freely available online

PLoS one

# How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data

**Daniele Fanelli***

INNOGEN and ISSTI-Institute for the Study of Science, Technology & Innovation, The University of Edinburgh, Edinburgh, United Kingdom

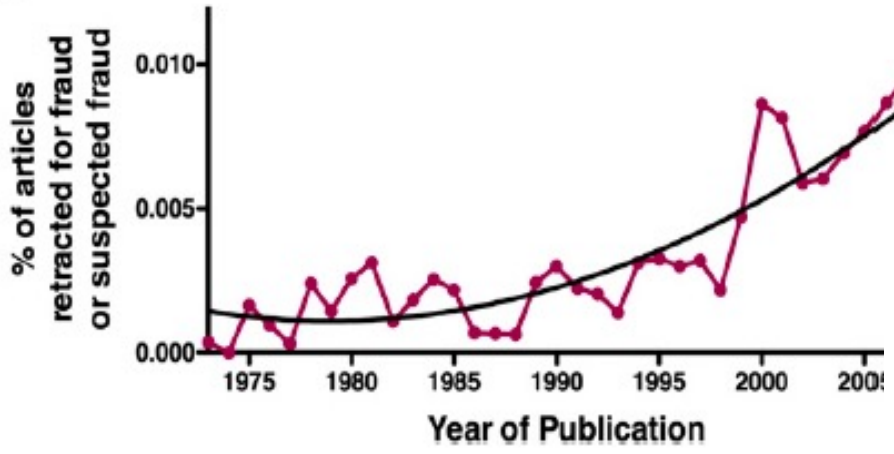


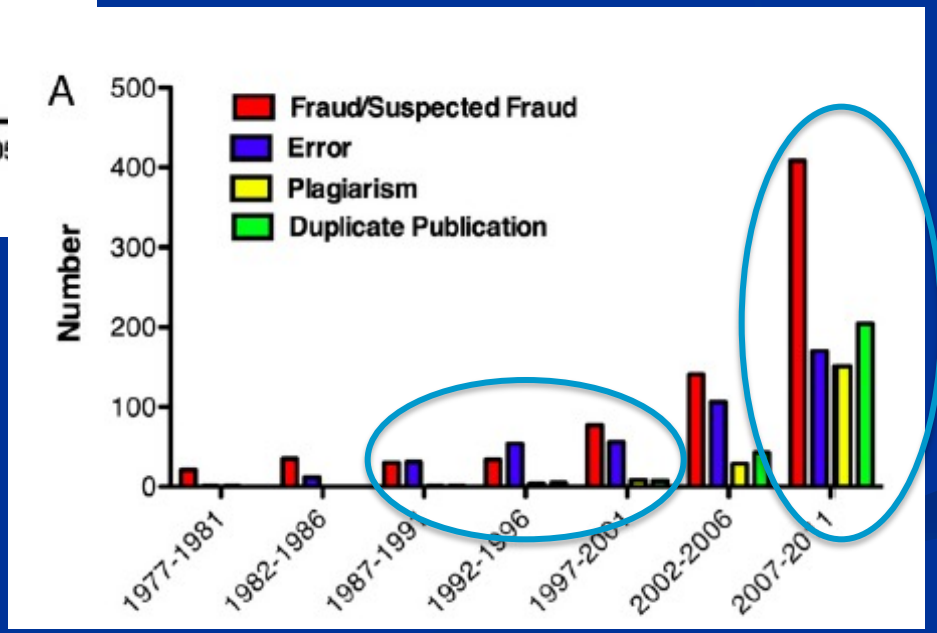

PLoS-One 2009;4:e5738

Academie Nieuwezorg

# Misconduct accounts for the majority of retracted scientific publications

Ferric C. Fang[a,b,1], R. Grant Steen[c,1], and Arturo Casadevall[d,1,2]

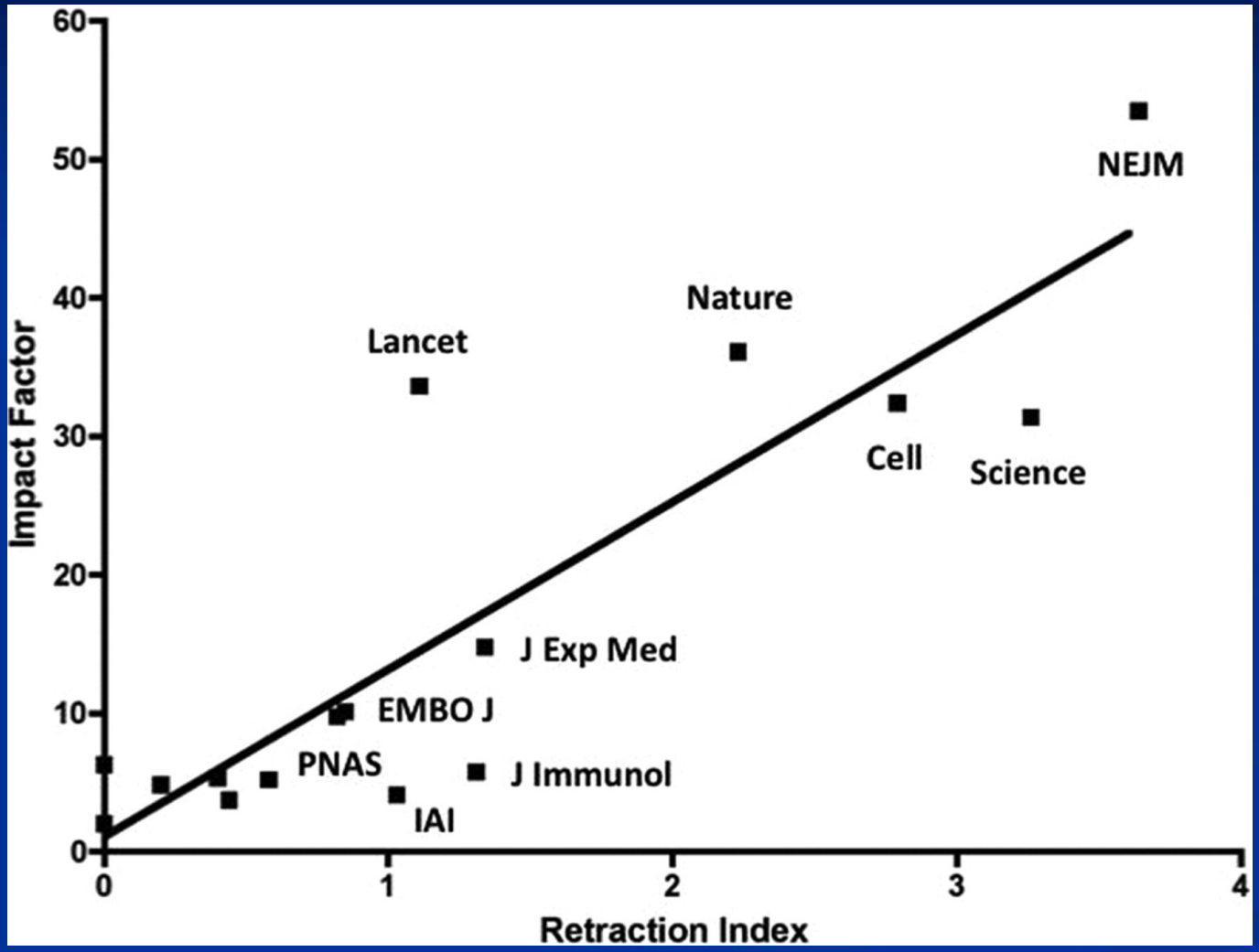



PNAS 2012;109:17028

Academie Nieuwezorg

Correlation between impact factor and retraction index.






Journals.ASM.org | Copyright © American Society for Microbiology. All Rights Reserved.

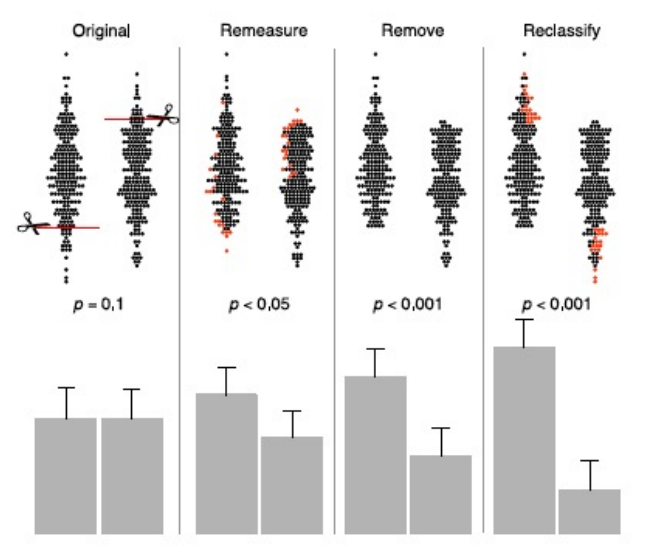
# Hoe stout moet je zijn?

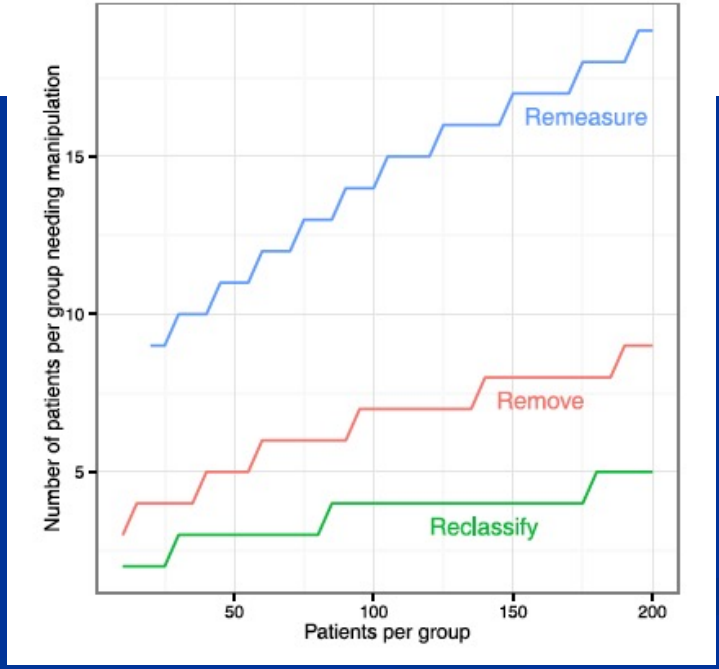
OPEN ACCESS Freely available online

PLOS ONE

## Why Even More Clinical Research Studies May Be False: Effect of Asymmetrical Handling of Clinically Unexpected Values

Matthew James Shun-Shin*, Darrel P. Francis




PLoS One 2013;8:e65323

Academie Nieuwezorg

# distrust?

- Stats and bias
- Conflicts of interest: still highly relevant
- Outright fraud: percentages…

Academie Nieuwezorg

# Unreliability: the readout is irreproducibility

## Raise standards for preclinical cancer research

C. **Glenn Begley** and **Lee M. Ellis** propose how methods, publications and incentives must change if patients are to benefit.

53 landmark basic science papers in hematology/oncology

**6/53 (11%) could be reproduced**

**REPRODUCIBILITY OF RESEARCH FINDINGS**

Preclinical research generates many secondary publications, even when results cannot be reproduced.

| Journal impact factor | Number of articles | Mean number of citations of non-reproduced articles* | Mean number of citations of reproduced articles |
|---|---|---|---|
| >20 | 21 | 248 (range 3–800) | 231 (range 82–519) |
| 5–19 | 32 | 169 (range 6–1,909) | 13 (range 3–24) |

VOL 483 | NATURE | 533

Academie Nieuwezorg

# Problems: What's old, what's new

- Oud:
  - IJdelheid
  - Bias / vooroordeel (*)
  - Publicatiebias (*?)
  - Statistische onkunde (*)
- Nieuw
  - Toenemende publicatiegekte
  - Overbevolking research
  - Hypercompetitie funding
  - Meet-gekte

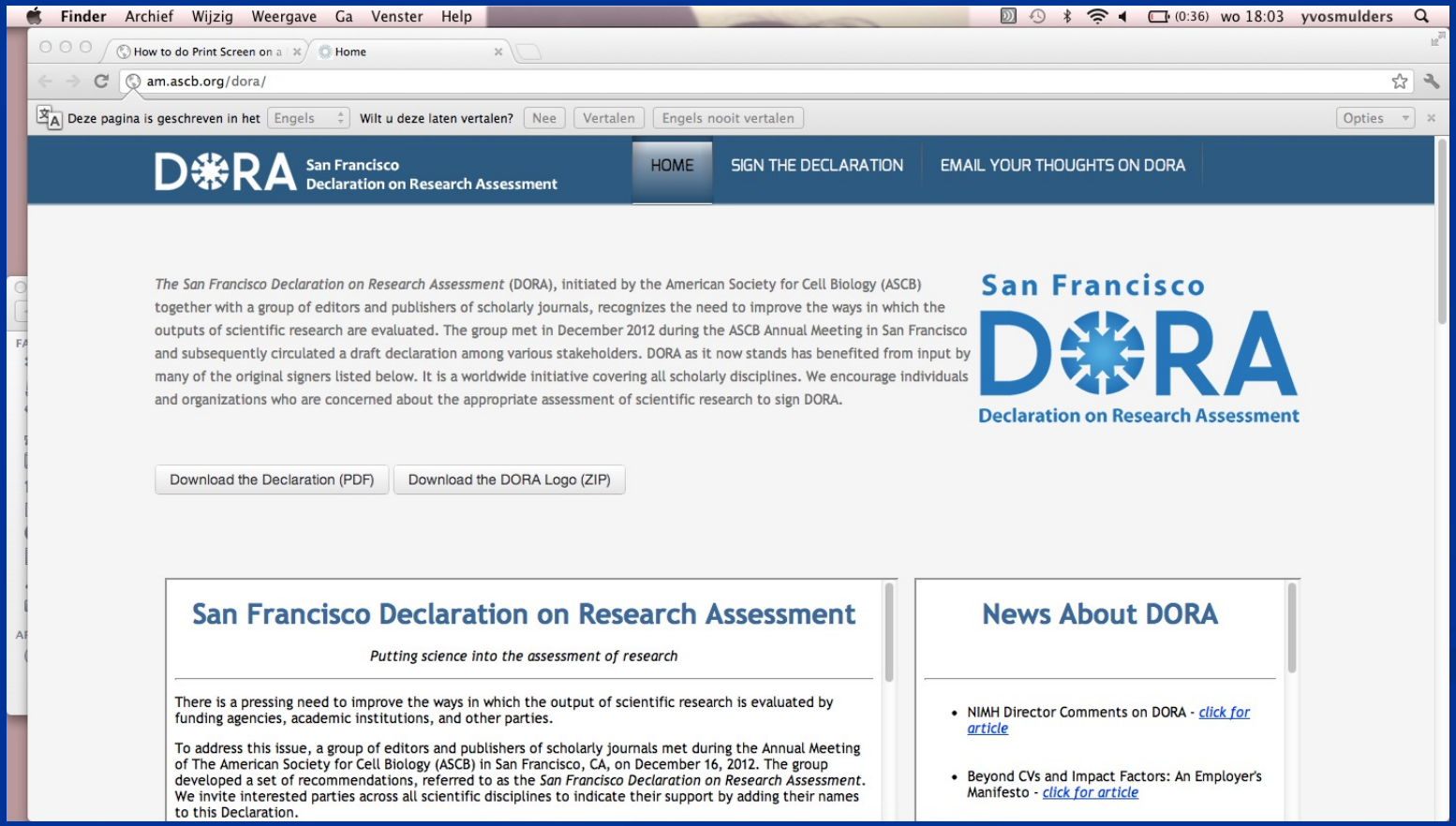
*: subtiele tekenen verbetering

**Waarheidslust maakt plaats voor scoringsdrift**

Academie Nieuwezorg

# Strong support for improvement

- Stop the impact-frenzy: JIF en H-index



# The Leiden Manifesto for research metrics

Use these ten principles to guide research evaluation, urge **Diana Hicks**, **Paul Wouters** and colleagues.

Data are increasingly used to govern science. Research evaluations that were once bespoke and performed by peers are now routine and reliant on metrics[1]. The problem is that evaluation is now led by the data rather than by judgement. Metrics have proliferated: usually well intentioned, not always well informed, often ill applied. We risk damaging the system with the very tools designed to improve it, as evaluation is increasingly implemented by organizations without knowledge of, or advice on, good practice and interpretation.

Before 2000, there was the Science Citation Index on CD-ROM from the Institute for Scientific Information (ISI), used by experts for specialist analyses. In 2002, Thomson Reuters launched an integrated web platform, making the Web of Science database widely accessible. Competing citation indices were created: Elsevier's Scopus (released in 2004) and Google Scholar (beta version released in 2004). Web-based tools to easily compare institutional research productivity and impact were introduced, such as InCites (using the Web of Science) and SciVal (using Scopus), as well as software to analyse individual citation profiles using Google Scholar (Publish or Perish, released in 2007).

In 2005, Jorge Hirsch, a physicist at the University of California, San Diego, proposed the *h*-index, popularizing citation counting for individual researchers. Interest in the journal impact factor grew steadily after 1995 (see 'Impact-factor obsession').

Lately, metrics related to social usage ▸

23 APRIL 2015 | VOL 520 | NATURE | 429

San Francisco DORA Declaration on Research Assessment

"Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions."

Academie Nieuwezorg

# SAMENVATTING

- Epidemiologie is mooi uitgangspunt, maar meer niet
- Aggregatieniveau van ‘maat nemen’ moet omhoog
- Waarheidsgehalte van onderzoeksuitkomst inschatten vereist heel veel kennis en deskundigheid
- Belangen en publicatiedruk zijn een probleem

Academie Nieuwezorg