

Lowering the P Value Threshold

Citation for published version (APA):

Janssens, A. C. J. W., & Penders, B. (2018). Lowering the P Value Threshold. *JAMA-Journal of the American Medical Association*, 320(9), 936-937. <https://doi.org/10.1001/jama.2018.8725>

Document status and date:

Published: 04/09/2018

DOI:

[10.1001/jama.2018.8725](https://doi.org/10.1001/jama.2018.8725)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Author Affiliations: Department of Respiratory and Critical Care Medicine, Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China.

Corresponding Author: Huahao Shen, MD, Key Laboratory of Respiratory Diseases of Zhejiang Province, Department of Respiratory and Critical Care Medicine, Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310009, China (huahaoshen@zju.edu.cn).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

1. Sobieraj DM, Weeda ER, Nguyen E, et al. Association of inhaled corticosteroids and long-acting β -agonists as controller and quick relief therapy with exacerbations and symptom control in persistent asthma: a systematic review and meta-analysis. *JAMA*. 2018;319(14):1485-1496. doi:10.1001/jama.2018.2769
2. Lundborg M, Wille S, Bjermer L, et al. Maintenance plus reliever budesonide/formoterol compared with a higher maintenance dose of budesonide/formoterol plus formoterol as reliever in asthma: an efficacy and cost-effectiveness study. *Curr Med Res Opin*. 2006;22(5):809-821. doi:10.1185/030079906X100212
3. Ställberg B, Ekström T, Neij F, et al; SHARE Trial Group. A real-life cost-effectiveness evaluation of budesonide/formoterol maintenance and reliever therapy in asthma. *Respir Med*. 2008;102(10):1360-1370. doi:10.1016/j.rmed.2008.06.017
4. Pavord ID, Jeffery PK, Qiu Y, et al. Airway inflammation in patients with asthma with high-fixed or low-fixed plus as-needed budesonide/formoterol. *J Allergy Clin Immunol*. 2009;123(5):1083-1089. doi:10.1016/j.jaci.2009.02.034
5. Egger M, Smith GD. Bias in location and selection of studies. *BMJ*. 1998;316(7124):61-66. doi:10.1136/bmj.316.7124.61

In Reply This systematic review and meta-analysis¹ addressed 1 of 6 research questions that comprised a larger project funded by the Agency for Healthcare Research and Quality in partnership with the National Heart, Lung, and Blood Institute. The protocol was registered with PROSPERO (CRD42016047985) and the final report is publicly available.²

Prior to publication, the draft report was posted for public comment and underwent peer review. This project was designed and conducted by a team without related conflicts of interest and no biases toward a particular result. As such, to our knowledge, no unpublished data were intentionally removed from any of the analyses.

Regarding the analysis comparing SMART with a higher dose of inhaled corticosteroids and LABA controller therapy with an end point of exacerbation risk, the 2 studies in question were excluded from the pooled estimate for valid reasons. First, the study by Ställberg et al³ included patients with both the same and higher inhaled corticosteroid doses in the comparator group; therefore, it was not pooled with studies that examined solely higher dosing.

Second, the study by Pavord et al⁴ did not specifically provide the number of study participants in each group who experienced a severe exacerbation; rather, they reported the time to first exacerbation or mean rate of severe exacerbations per patient-year. Therefore, the results shown in Figure 3 represented the best available evidence at the time for the outcome of exacerbation risk.

Diana M. Sobieraj, PharmD
William L. Baker, PharmD

Author Affiliations: Department of Pharmacy Practice, University of Connecticut School of Pharmacy, Storrs.

Corresponding Author: William L. Baker, PharmD, Department of Pharmacy Practice, University of Connecticut School of Pharmacy, 69 N Eagleville Rd, Unit 3092, Storrs, CT 06269 (william.baker_jr@uconn.edu).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

1. Sobieraj DM, Weeda ER, Nguyen E, et al. Association of inhaled corticosteroids and long-acting β -agonists as controller and quick relief therapy with exacerbations and symptom control in persistent asthma: a systematic review and meta-analysis. *JAMA*. 2018;319(14):1485-1496. doi:10.1001/jama.2018.2769
2. Sobieraj DM, Baker WL, Weeda ER, et al. Intermittent inhaled corticosteroids and long-acting muscarinic antagonists for asthma. <https://effectivehealthcare.ahrq.gov/topics/asthma-pharmacologic-treatment/research-2017/>. Accessed May 20, 2018.
3. Ställberg B, Ekström T, Neij F, et al; SHARE Trial Group. A real-life cost-effectiveness evaluation of budesonide/formoterol maintenance and reliever therapy in asthma. *Respir Med*. 2008;102(10):1360-1370. doi:10.1016/j.rmed.2008.06.017
4. Pavord ID, Jeffery PK, Qiu Y, et al. Airway inflammation in patients with asthma with high-fixed or low-fixed plus as-needed budesonide/formoterol. *J Allergy Clin Immunol*. 2009;123(5):1083-1089. doi:10.1016/j.jaci.2009.02.034

Lowering the *P* Value Threshold

To the Editor Dr Ioannidis¹ proposed lowering the *P* value threshold to .005 for medical studies. This proposal builds on the suggestion by Benjamin et al,² who were concerned that a *P* value threshold of .05 contributed to the nonreproducibility of studies through a high false-positive rate. However, the proposal for a lower *P* value threshold may provoke misunderstanding of the use of *P* values and thereby hinder innovation.

Several risk factors influence the gap between actual and expected false-positive rates.² However, biased estimates can also result from selection bias, information bias, or confounding bias, with false-positive findings more likely to be published.

In addition, failed replication occurs in cases of unknown heterogeneity among study populations. Yet simply reducing the *P* value threshold does not attenuate these problems. Instead, it may increase the possibility of false-negative results, impede the development of potential therapeutic agents at an early phase, increase needed sample sizes, lessen study feasibility, and aggravate *P* value hacking and data manipulation. Thus, the proposal to reduce the *P* value threshold addresses only superficial problems rather than the essential contributing factors.

We believe that resolving this issue begins with study design and supervision. First, potential bias should be considered during study design. Random sampling ensures the representativeness of research samples. Random allocation avoids potential confounding bias. Repeating experiments exposes coincidental results and ensures the reproducibility of findings.

Second, we prefer study replication over a lowered significance threshold for reducing the risk of false-positive results. Although this would require a larger number of participants, the combined false-positive rate of the original and replication studies would not exceed $.05 \times .05 = .0025$, and replication would yield lower risk and improved reproducibility.

Third, study protocol repository and registration systems should be promoted widely for accessibility to the scientific community. Modifications to the study protocol should be submitted to and approved by the system. Public accessibility to original experimental records also would improve reproducibility.

Fourth, the definition of the *P* value and its threshold should be treated objectively.^{3,4} Effect size should be considered ahead of statistical significance, and both should be used for better decision making. Different *P* value thresholds may be appropriate for different research fields (eg, .05 for traditional analysis, 5×10^{-8} for genome-wide research) and study objectives (.10 or .05 for screening phase, .05 or .01 for confirmation phase). New approaches, such as the Bayes and causal inference methods, also should be considered.

Yongyue Wei, PhD
Feng Chen, PhD

Author Affiliations: State Key Laboratory of Reproductive Medicine, Nanjing Medical University School of Public Health, Nanjing, China.

Corresponding Author: Feng Chen, PhD, Department of Biostatistics, School of Public Health, Nanjing Medical University, 101 Longmian Ave, Nanjing 211166, China (fengchen@njmu.edu.cn).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Chen reported being the president of the Statistical Theory and Methodology Committee of the Chinese Health Information Association. No other disclosures were reported.

- Ioannidis JPA. The proposal to lower *P* value thresholds to .005. *JAMA*. 2018; 319(14):1429-1430. doi:10.1001/jama.2018.1536
- Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. *Nat Hum Behav*. 2018;2(1):6-10. doi:10.1038/s41562-017-0189-z
- Kyriacou DN. The enduring evolution of the *P* value. *JAMA*. 2016;315(11):1113-1115. doi:10.1001/jama.2016.2152
- Wasserstein RL, Lazar NA. The ASA's statement on *P*-values: context, process, and purpose. *Am Stat*. 2016;70(2):129-133. doi:10.1080/00031305.2016.1154108

To the Editor The Viewpoint by Dr Ioannidis discussed the recent proposal by a coalition of 72 methodologists to lower the statistical significance threshold to .005.^{1,2} Ioannidis acknowledged limitations of this proposal; however, his Viewpoint did not discuss any financial implications of this change.¹ Powering randomized clinical trials (RCTs) and prospective observational studies at an a level of .005 would translate into considerably higher costs associated with larger sample sizes.

The increased costs associated with powering prospective studies at the .005 level could have serious repercussions, especially for publicly funded trials. First, in the current hypercompetitive environment in which many deserving studies do not get funded, these cost increases would result in an even higher number of clinically relevant studies that remain unfunded. In the absence of unlimited resources to support research, this proposal inevitably raises the question of whether limited funds should be spent on a few highly powered studies vs supporting a greater number of studies powered at the .05 level.

Second, this change would affect RCTs supporting new drug approvals. In addition to incentivizing the selection of sur-

rogate outcomes as end points, as acknowledged by Ioannidis,¹ the increased financial risk would motivate manufacturers to adopt more conservative approaches when making decisions regarding whether to support new molecules through the clinical phase. This would ultimately lead to delays in innovation and access to new therapies.

Third, lower *P* value thresholds would be particularly problematic in the design of studies for rare diseases, given the added difficulty in recruitment.

Although we commend efforts to mitigate the misuse of *P* values, we believe that lowering the statistical significance threshold would first require a broader debate that accounts for its consequences on access to research findings and availability of new therapies.

Inmaculada Hernandez, PharmD, PhD
Walid F. Gellad, MD, MPH
Chester B. Good, MD, MPH

Author Affiliations: Department of Pharmacy and Therapeutics, University of Pittsburgh School of Pharmacy, Pittsburgh, Pennsylvania (Hernandez); Division of General Internal Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania (Gellad); Insurance Services Division, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania (Good).

Corresponding Author: Inmaculada Hernandez, PharmD, PhD, Department of Pharmacy and Therapeutics, School of Pharmacy, University of Pittsburgh, 3501 Terrace St, Pittsburgh, PA 15261 (inh3@pitt.edu).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

- Ioannidis JPA. The proposal to lower *P* value thresholds to .005. *JAMA*. 2018; 319(14):1429-1430. doi:10.1001/jama.2018.1536
- Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. *Nat Hum Behav*. 2018;2(1):6-10. doi:10.1038/s41562-017-0189-z

To the Editor Changing the *P* value threshold to .005 does not solve the issue of how to interpret *P* values.¹ It simply shifts the burden of proof to require a more extreme difference.

Uncertainty regarding the potential benefit of a therapy that demonstrated a clinically meaningful but now statistically non-significant result would increase.² The demonstration of the reproducibility of experiments, a cornerstone of the scientific method, would be diminished and restricted to those therapies showing large differences (together with large sample sizes). Clinicians may feel abandoned.

To illustrate, a study of long-term survival of patients receiving intraperitoneal chemotherapy showed a clinical benefit with a median survival of 5.2 years vs 4.3 years in the control group, yet the *P* value was only .04.³ Under the proposed threshold, the result would not be statistically significant, yet clinicians would consider the difference clinically important.

The real problem is with the scientific interpretation of the statistical results and not the choice of significance level.⁴ It makes sense to place more effort on explaining the clinical relevance of a result rather than simply the statistical meaning. The consequences of lowering the threshold include how to interpret harm in the safety and subgroup analyses with non-significant interaction tests.

For example, the change in the standard treatment for patients with *K-ras* wild-type metastatic colorectal cancer might

have been missed because the P value for interaction for a differential treatment effect on survival between the *K-ras* mutation and wild type was only .01.⁵

Lowering the P value threshold would be detrimental to interpreting signal-finding associations in multivariable analyses and systematic reviews and yield little scientific gain and potentially major scientific loss.

Val GebSKI, MStat
Karen Byth, PhD, DIC

Author Affiliations: National Health and Medical Research Council Clinical Trials Centre, University of Sydney, Sydney, Australia.

Corresponding Author: Val GebSKI, MStat, University of Sydney, NHMRC Clinical Trials Centre, Level 5 MFB Building, 92-94 Parramatta Rd, Camperdown, New South Wales, Australia 2050 (val@ctc.usyd.edu.au).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

- Ioannidis JPA. The proposal to lower P value thresholds to .005. *JAMA*. 2018; 319(14):1429-1430. doi:10.1001/jama.2018.1536
- Shakespeare TP, GebSKI VJ, Veness MJ, Simes J. Improving interpretation of clinical studies by use of confidence levels, clinical significance curves, and risk-benefit contours. *Lancet*. 2001;357(9265):1349-1353. doi:10.1016/S0140-6736(00)04522-0
- Tewari D, Java JJ, Salani R, et al. Long-term survival advantage and prognostic factors associated with intraperitoneal chemotherapy treatment in advanced ovarian cancer: a gynecologic oncology group study. *J Clin Oncol*. 2015;33(13):1460-1466. doi:10.1200/JCO.2014.55.9898
- Resnick B. What a nerdy debate about p -values shows about science—and how to fix it. <https://www.vox.com/science-and-health/2017/7/31/16021654/p-values-statistical-significance-redefine-0005>. Accessed April 13, 2018.
- Karapetis CS, Khambata-Ford S, Jonker DJ, et al. *K-ras* mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med*. 2008;359(17):1757-1765. doi:10.1056/NEJMoa0804385

To the Editor P values are a common measure in most observational studies and RCTs. Researchers have long attached undue importance to significant P values despite the fact that P values by themselves have no clinical interpretation. The recommendation by Dr Ioannidis¹ to replace the arbitrary a level of .05 as the threshold of statistical significance with a lower value is laudable but begs for elaboration.

Statistical methods that assess sampling uncertainty, such as P values, have been the foundation for quantitative medical research since World War II.² However, P values are not often replicated in similar studies, and their definition is not directly associated with reproducibility.³

Scientific findings are not valid if they cannot be reproduced and are highly variable. Although other methodological issues relate to reproducibility, the P value is arguably at the root of the problem, given its wide variability and irreproducibility from study to study.⁴ The irreproducibility of P values was demonstrated using a simulation with data from a published RCT.⁵ The probability of attaining another statistically significant P value varied widely on replication.

Statistical power alone determined the distribution of P values and varied with sample size and effect size. Given the variability in replicated P values, lowering the α level to .005 or even .001 would make little difference in interpreting and applying the results clinically.

Categorizing findings as either statistically significant or not is a false dichotomy. Confidence intervals (CIs) share some of the weaknesses of P values. Of the means replicated in the simulation, 85.4% fell within the original 95% CI. However, the advantage of CIs is that they are more easily understood. Values outside the CI are not likely to exist in the population. When interpreted out of context, using P values without CIs can be misleading and potentially lead to biased inferences from clinical studies.

Medicine needs a more critical understanding of statistical power and inferences. Optimal statistical reasoning should consider P values in conjunction with CIs while contextualizing the findings in the clinical context.

Paul Barach, MD, MPH
Ronald L. Thomas, PhD
Steven E. Lipshultz, MD

Author Affiliations: Wayne State University School of Medicine, Children's Hospital of Michigan, Detroit, Michigan.

Corresponding Author: Paul Barach, MD, MPH, Carmen and Ann Adams Department of Pediatrics, Wayne State University School of Medicine, Children's Hospital of Michigan, Children's Research Center of Michigan, 3901 Beaubien, Detroit, MI 48202 (pbarach@gmail.com).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

- Ioannidis JPA. The proposal to lower P value thresholds to .005. *JAMA*. 2018; 319(14):1429-1430. doi:10.1001/jama.2018.1536
- Fidler F, Thomason N, Cumming G, Finch S, Leeman J. Editors can lead researchers to confidence intervals, but can't make them think: statistical reform lessons from medicine. *Psychol Sci*. 2004;15(2):119-126. doi:10.1111/j.0963-7214.2004.01502008.x
- Wasserman RL, Lazar NA. The ASA's statement on p -values: context, process, and purpose. *Am Stat*. 2016;70(2):129-133. doi:10.1080/00031305.2016.1154108
- Goodman SN. Toward evidence-based medical statistics, 1: the P value fallacy. *Ann Intern Med*. 1999;130(12):995-1004. doi:10.7326/0003-4819-130-12-199906150-00008
- Thomas RL, Barach PR, Wilkinson JD, Farooqi AA, Lipshultz SE. The danger of relying on the interpretation of p -values in single studies: irreproducibility of results from clinical studies. *Prog Pediatr Cardiol*. 2017;44:57-62. doi:10.1016/j.ppedcard.2017.02.002

To the Editor Dr Ioannidis argued that the proposal to lower the P value threshold for claiming statistical significance from .05 to .005 is a temporary solution to improve statistical inference and suggests additional training in methods and statistics for the scientific workforce.^{1,2}

To people who do not understand complex statistics, the P value threshold of .05 is a rule of thumb for claiming statistical significance. The rule can be safely applied because the threshold is .05 regardless of the underlying statistics, with a few exceptions. We agree with Ioannidis that improving statistical literacy may be unachievable² but believe that replacing 1 rule of thumb with 2 others can aid researchers in assessing scientific claims.

First, results vary with the methods. Statistical estimates may vary with differences in the selection of the study population, length of follow-up, and the measurement of the variables.³ The likelihood of such differences is relevant when

studies are performed using data that are not specifically collected for the research question at hand such as from electronic health records or biobanks. The variables in these data collections may not be measured ideally, and relevant confounders may be missing. If the results may be different with optimal study design and data collection, the *P* values will vary. Knowing that methodology affects the estimates and the *P* values prevents taking them too literally.

Second, each study is only 1 piece of the puzzle. Scientific claims and knowledge in the health sciences primarily follow from inductive reasoning that allows scientists to infer future instances from observed instances, across studies, through generalizations.⁴ The results of a single study represent only 1 piece of evidence, and the results of other studies may make a claim more or less likely. These other studies can be similar, aiming at replication, or different, contributing experimental evidence from, for example, animals or cells.⁵ Only the combined evidence determines the likelihood of the claim being true.

Science thrives on high-quality data, statistics, and reasoning. Ioannidis' aim to improve statistical literacy is a long-term project worthy of support. Strengthening scientific reasoning through 2 simple rules that discourage overreliance on the *P* value might be an immediate and effective substitute.

A. Cecile J. W. Janssens, PhD
Bart Penders, PhD

Author Affiliations: Department of Epidemiology, Emory University, Atlanta, Georgia (Janssens); Care and Public Health Research Institute, Maastricht University, Maastricht, the Netherlands (Penders).

Corresponding Author: A. Cecile J. W. Janssens, PhD, Department of Epidemiology, Rollins School of Public Health, Emory University, 1518 Clifton Rd NE, Atlanta, GA 30322 (cecile.janssens@emory.edu).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

1. Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. *Nat Hum Behav*. 2018;2(1):6-10. doi:10.1038/s41562-017-0189-z
2. Ioannidis JPA. The proposal to lower *P* value thresholds to .005. *JAMA*. 2018; 319(14):1429-1430. doi:10.1001/jama.2018.1536
3. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
4. Greenland S. Induction versus Popper: substance versus semantics. *Int J Epidemiol*. 1998;27(4):543-548. doi:10.1093/ije/27.4.543
5. Munafò MR, Davey Smith G. Robust research needs many lines of evidence. *Nature*. 2018;553(7689):399-401. doi:10.1038/d41586-018-01023-3

In Reply Drs Wei and Chen list several approaches that can be used to minimize bias and increase the reproducibility and utility of future studies. There is no reason why all these efforts cannot be pursued. Currently there is a rich discussion about ways to improve research.^{1,2} Changing the *P* value threshold will not hinder such improvements. Moreover, these proposed improvements cannot salvage studies already performed, whereas changing the *P* value threshold may better calibrate the interpretation of past literature.

Dr Hernandez and colleagues are concerned about higher costs associated with larger required sample sizes, a higher number of unfunded trials given limited public funds, and a

delay in access to new therapies. Their reasoning assumes that the clinical trials agenda is optimized and rationally driven with careful, clinically relevant, and statistically solid power calculations. This is far from the reality; most trials are notoriously underpowered, spuriously powered for irrelevant outcomes, heavily biased, or a combination of these. Instead of having 1 million mostly useless trials, a focus on fewer, better, more conclusive ones should help.

Regardless, sometimes clinical and implementation decisions may need to be adopted with only “suggestive” rather than “statistically significant” evidence. For rare diseases, this situation may be more common, but the terminology (“suggestive”) is then appropriate: it conveys that less is known than would be desired. In extremely rare conditions with minimal evidence, decisions may be made even with a *P* value greater than .05, trying to optimally balance the consequences and cost of false-positive and false-negative results.³

I do not share the concerns of Wei and Chen and Hernandez and colleagues about slowed innovation. Better decision rules lead to more efficient innovation, with better discrimination of what matters and what does not, and with fewer pursuits of false leads.

Mr Gebski and Dr Byth suggest that many clinically meaningful differences would be dismissed. Of course, interpreting the results of clinical studies should focus primarily on their clinical rather than their statistical significance. However, most of the effects in the range of *P* values between .005 and .05 in the past literature are not clinically significant. I acknowledged in the Viewpoint that exceptions do exist.⁴

For “suggestive” statistical significance with clear clinical significance, the latter should have priority. The 2 examples that Gebski and Byth provide as major failures of the *P* value threshold less than .005 demonstrate, in fact, that the rule works. The study of long-term survival of patients receiving intraperitoneal chemotherapy in ovarian cancer shows a *P* value of .002 for the survival benefit,⁵ not .04 as they claim. For *K-ras* and colorectal cancer, the *P* value of .01 is for survival (“suggestive”) but the *P* value is less than .001 (“statistically significant”) for progression-free survival.⁶

I agree with Dr Barach and colleagues regarding CIs and I favor contextualizing the findings in the clinical context. The simulation study adds to the literature about how *P* values are misinterpreted. However, their claim that “values outside the [95%] CI are not likely to exist in the population” is optimistic. It holds true only if multiple assumptions are met such as lack of any biases. Use of 95% CIs suffers from some, but not all, of the problems of *P* value thresholds less than .05.

I am sympathetic with the 2 additional rules of thumb that Drs Janssens and Penders propose to sensitize more clinical researchers, clinicians, and users of the scientific literature. However, having to resort to oversimplified rules of thumb perpetuates the current concept that widespread statistical illiteracy and innumeracy should be taken for granted. Scientific reasoning needs solid methodological training. Rules of thumb (including $P < .005$) are mostly tempering measures.

John P. A. Ioannidis, MD, DSc

Author Affiliation: Meta-Research Innovation Center at Stanford, Stanford University, Stanford, California.

Corresponding Author: John P. A. Ioannidis, MD, DSc, Department of Medicine and Meta-Research Innovation Center at Stanford, Stanford University, 1265 Welch Rd, Stanford, CA 94305 (jioannid@stanford.edu).

Conflict of Interest Disclosures: The author has completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and reported having authored 2 articles proposing a change in the *P* value threshold.

1. Ioannidis JP. How to make more published research true. *PLoS Med*. 2014;11(10):e1001747. doi:10.1371/journal.pmed.1001747
2. Munafò MR, Bishop DV, Button KS, et al. A manifesto for reproducible science. *Nat Hum Behav*. 2017;1:0021. doi:10.1038/s41562-016-0021
3. Ioannidis JPA, Hozo I, Djulbegovic B. Optimal type I and type II error pairs when the available sample size is fixed. *J Clin Epidemiol*. 2013;66(8):903-910.e2. doi:10.1016/j.jclinepi.2013.03.002
4. Ioannidis JPA. The proposal to lower *P* value thresholds to .005. *JAMA*. 2018;319(14):1429-1430. doi:10.1001/jama.2018.1536
5. Tewari D, Java JJ, Salani R, et al. Long-term survival advantage and prognostic factors associated with intraperitoneal chemotherapy treatment in advanced ovarian cancer: a gynecologic oncology group study. *J Clin Oncol*. 2015;33(13):1460-1466. doi:10.1200/JCO.2014.55.9898
6. Karapetis CS, Khambata-Ford S, Jonker DJ, et al. *K-ras* mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med*. 2008;359(17):1757-1765. doi:10.1056/NEJMoa0804385

CORRECTION

Incorrect Medication Reported: In the News From the Food and Drug Administration item entitled "Generic for Opioid Use Disorder,"¹ published in the July 17, 2018, issue of *JAMA*, an incorrect medication was reported. In the second paragraph, the second sentence should have read as follows: "Buprenorphine, naltrexone, and methadone are FDA-approved to help patients who misuse opioids." This article was corrected online.

1. Voelker R. Generic for opioid use disorder. *JAMA*. 2018;320(3):228.

Error in Figure: In the Original Investigation entitled "Effect of Loading Dose of Atorvastatin Prior to Planned Percutaneous Coronary Intervention on Major Adverse Cardiovascular Events in Acute Coronary Syndrome: The SECURE-PCI Randomized Clinical Trial"¹ published in the April 3, 2018, issue of *JAMA*,

an error occurred in a figure. In Figure 3, the last 2 row labels were switched. The second-to-last row should have been labeled "Bare-metal stent only" and the last row should have been labeled "≥1 Drug-eluting stent." This article was corrected online.

1. Berwanger O, Santucci EV, de Barros E Silva PGM, et al; SECURE-PCI Investigators. Effect of loading dose of atorvastatin prior to planned percutaneous coronary intervention on major adverse cardiovascular events in acute coronary syndrome: the SECURE-PCI randomized clinical trial. *JAMA*. 2018;319(13):1331-1340. doi:10.1001/jama.2018.2444

Guidelines for Letters

Letters discussing a recent *JAMA* article should be submitted within 4 weeks of the article's publication in print. Letters received after 4 weeks will rarely be considered. Letters should not exceed 400 words of text and 5 references and may have no more than 3 authors. Letters reporting original research should not exceed 600 words of text and 6 references and may have no more than 7 authors. They may include up to 2 tables or figures but online supplementary material is not allowed. All letters should include a word count. Letters must not duplicate other material published or submitted for publication. Letters not meeting these specifications are generally not considered. Letters being considered for publication ordinarily will be sent to the authors of the *JAMA* article, who will be given the opportunity to reply. Letters will be published at the discretion of the editors and are subject to abridgement and editing. Further instructions can be found at <http://jamanetwork.com/journals/jama/pages/instructions-for-authors>. A signed statement for authorship criteria and responsibility, financial disclosure, copyright transfer, and acknowledgment and the ICMJE Form for Disclosure of Potential Conflicts of Interest are required before publication. Letters should be submitted via the *JAMA* online submission and review system at <https://manuscripts.jama.com>. For technical assistance, please contact jama-letters@jamanetwork.org.

Section Editor: Jody W. Zylke, MD, Deputy Editor.