

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



LSHTM Research Online

Pocock, SJ; Stone, GW; (2016) The Nature of the P Value. The New England journal of medicine, 375 (22). pp. 2205-2206. ISSN 0028-4793 DOI: <https://doi.org/10.1056/NEJMc1612970>

Downloaded from: <http://researchonline.lshtm.ac.uk/3399655/>

DOI: <https://doi.org/10.1056/NEJMc1612970>

Usage Guidelines:

Please refer to usage guidelines at <https://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license: Copyright the publishers

<https://researchonline.lshtm.ac.uk>

identify potential new collaborators.² Nevertheless, under the terms of approval of our retrospective study by the institutional review board, we are not permitted to release patient-level data. Our retrospective study includes not only patients in different clinical trials but also patients who received “standard of care” treatment that was not determined by a specific clinical protocol. Different consents apply to the different populations included in the analysis; therefore, in order to perform our retrospective study, we applied for and received specific approval from the institutional review board under internal protocol 2281 of the Fred Hutchinson Cancer Research Center.

Furthermore, Hourigan and colleagues request disclosure of patient-level data pertaining to a large number of individual patient-related factors, some of which were not collected at the time of the chart review or included in the analysis. Although we agree that some of these factors might be worth examining, to collect, analyze, and release such information would require time and another institutional review board application.

In response to the other questions by Hourigan et al., bone marrow biopsies with cytogenetic

analysis were performed before transplantation in all the patients with myelodysplasia. In our system, we do not have a way of determining how many patients are considered for transplantation but do not undergo transplantation. The study by Walter et al. included transplants from both related and unrelated donors, and a total of 152 patients in that study were also involved in our study. With the goal of sharing data among various institutions in the future, we are willing to submit a new institutional review board application and collaborate with Hourigan and colleagues or others.

Filippo Milano, M.D., Ph.D.

Frederick R. Appelbaum, M.D.

Colleen Delaney, M.D.

Fred Hutchinson Cancer Research Center

Seattle, WA

fmlano@fredhutch.com

Since publication of their article, the authors report no further potential conflict of interest.

1. Brunstein CG, Petersdorf EW, DeFor TE, et al. Impact of allele-level HLA mismatch on outcomes in recipients of double umbilical cord blood transplantation. *Biol Blood Marrow Transplant* 2016;22:487-92.

2. Longo DL, Drazen JM. Data sharing. *N Engl J Med* 2016;374:276-7.

DOI: 10.1056/NEJMc1612872

The Nature of the P Value

TO THE EDITOR: Pocock and Stone (Sept. 8 issue) provide an excellent primer on evaluations of clinical trials and their applications to practice.¹ However, early on in their article they make a common but erroneous statement regarding the nature of the P value: “A P value of 0.05 carries a 5% risk of a false positive result.” This misconception is found throughout statistical and medical literature and was recently addressed in the following statement, published earlier this year by the American Statistical Association: “P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.”² A P value is an expression of the probability of getting results at least as extreme as what was observed, under the assumption that the null hypothesis is true.^{2,3} The P value as it was originally intended was simply meant to serve as a guide

for evaluating whether an experiment was worth repeating.

Daniel Hu, Pharm.D.

Providence St. Peter Hospital

Olympia, WA

daniel.hu@providence.org

No potential conflict of interest relevant to this letter was reported.

1. Pocock SJ, Stone GW. The primary outcome is positive — is that good enough? *N Engl J Med* 2016;375:971-9.

2. Wasserstein RL, Lazar NA. The ASA’s statement on p-values: context, process, and purpose. *Am Stat* 2016;70:129-33 (<http://dx.doi.org/10.1080/00031305.2016.1154108>).

3. Nuzzo R. Scientific method: statistical errors. *Nature* 2014;506:150-2.

DOI: 10.1056/NEJMc1612970

THE AUTHORS REPLY: We appreciate Hu’s comments on the meaning of P values. Of course, we endorse the statement from the American Statis-

tical Association, which is compatible with the view of the original creators of the P value — R.A. Fisher and others — more than a century ago. (See Pocock et al.¹ for a historical account that addresses some common misperceptions about P values.)

We disagree with Hu's claim that our statement — "A P value of 0.05 carries a 5% risk of a false positive result" when "there is no true difference between treatments" — is erroneous. According to the null hypothesis, $P < 0.05$ will occur 5% of the time.

We also respectfully disagree with Hu's last sentence, in which he states that "The P value as it was originally intended was simply meant to serve as a guide for evaluating whether an experiment was worth repeating." Fisher, in his seminal text, clearly described the P value as a

tool to be used to determine the strength of evidence contradicting the null hypothesis, recognizing that the smaller the P value, the lower the likelihood of a false positive result.²

Stuart J. Pocock, Ph.D.

London School of Hygiene and Tropical Medicine
London, United Kingdom
stuart.pocock@lshtm.ac.uk

Gregg W. Stone, M.D.

Columbia University Medical Center
New York, NY

Since publication of their article, the authors report no further potential conflict of interest.

1. Pocock SJ, McMurray JJV, Collier TJ. Making sense of statistics in clinical trial reports: part 1 of a 4-part series on statistics for clinical trials. *J Am Coll Cardiol* 2015;66:2536-49.

2. Fisher RA. *Statistical methods for research workers*. Edinburgh: Oliver and Boyd, 1925.

DOI: 10.1056/NEJMc1612970

TP53 and Histone H3.3 Mutations in Triple-Negative Lower-Grade Gliomas

TO THE EDITOR: The genomic landscape of lower-grade gliomas has recently been investigated in "multi-omic" analyses that have revealed molecular subgroups with prognostic relevance.¹ In particular, robust biomarkers, including 1p/19q codeletion (deletion of chromosome arms 1p and 19q), *IDH* mutation, and *TERT* promoter mutation, have been used to classify international cohorts of gliomas into five principal molecular groups with clinical relevance.² In the 2016 World Health Organization (WHO) classification of central nervous system tumors,³ *IDH* and 1p/19q are used as biomarkers to classify gliomas.

We evaluated tumor samples obtained from 459 adults (mean age, 44 years; median age, 43 years; range, 18 to 79) with grade II or III diffuse gliomas defined according to the molecular scheme (1p/19q codeletion and *IDH* and *TERT* mutations) described by Eckel-Passow et al.² A total of 30.3% of these gliomas were triple-positive, 9.2% had *IDH-TERT* co-mutations, 28.1% had *IDH* mutations only, 17.4% were triple-negative, 5.9% had *TERT* mutations only, and 9.1% had other genetic combinations. The principal molecular subgroups had prognostic significance across the cohort; this significance was indepen-

dent of the patient's age and the histologic type and grade of the tumor (see Fig. S1A and Table S1 in the Supplementary Appendix, available with the full text of this letter at NEJM.org).

The frequency of triple-negative gliomas was higher in our Chinese cohort than in the cohort described by Eckel-Passow et al. (17.4% vs. 7.0%).² We further examined this molecular subgroup for substratification of prognostic value. The substratification of clinical risk on the basis of the histologic grade of triple-negative gliomas was more robust than predictions of risk on the basis of histologic type ($P < 0.001$) (see Fig. S1B in the Supplementary Appendix).

TP53 was sequenced in 21 triple-negative gliomas for which there was sufficient tissue for analysis; survival data were available for 20 patients (Fig. 1B). A significant proportion of patients (43%) had gliomas with *TP53* mutations; these patients had shorter overall survival than those with wild-type *TP53* ($P = 0.03$) (Fig. 1A).

H3.3-K27M, another biomarker used to categorize gliomas in the WHO 2016 classification,³ was present in 19% of the triple-negative gliomas and was detected in the patients with reduced survival ($P = 0.02$) (Fig. 1A). Notably, all