Qeios

Comments on "The roles, challenges, and merits of the *p* value" by Chén et al.

Hening Huang
Teledyne RD Instruments (retired)
San Diego, CA 92127, USA
Email: heninghuang1@gmail.com

Summary

Chén et al. recently published a systematic review of the *p* value produced by null hypothesis significance tests (NHSTs) in *Patterns* (2023, **4**(12), 100878, ISSN 2666-3899, https://doi.org/10.1016/j.patter.2023.100878). We argue that their paper does not reveal the actual meaning of the *p* value in real-world problems, and their view on the *p* value is another form of common misconceptions about the *p* value. This commentary focuses on the *p* value produced by the two-sample *z*-test and explores its meaning. We argue that the *p* value is not an appropriate probabilistic measure in scientific decision-making; the exceedance probability or gain-probability is an appropriate probabilistic measure and can and should be used as an alternative to the *p* value.

Keywords exceedance probability; gain-probability; probabilistic measure; *p*-value; *z*-test

Introduction

The long-lasting debate about the validity of null hypothesis significance testing (NHST) (or simply hypothesis testing) and its produced *p*-values continues (e.g. Heckelei 2023, Aurbacher et al. 2024). On one hand, many scientists have suggested retiring or abandoning statistical significance and *p* values (e.g. Amrhein et al. 2019, McShane et al. 2018, Halsey 2019, Wasserstein and Lazar 2016, Wasserstein et al. 2019) and replacing significance testing with estimation statistics (e.g. Claridge-Chang and Assam 2016, Berner and Amrhein 2022, Huang 2023a). *Basic and Applied Social Psychology* has officially banned the NHST procedures since 2015 (Trafimow and Marks 2015). Fourteen physiotherapy journals that are the members of the International Society of Physiotherapy Journal Editors (ISPJE) have advised researchers to expect manuscripts to use estimation methods instead of NHSTs (Elkins et al. 2022). Moreover, many authors have called for statistics reform (e.g. Wagenmakers et al. 2011, Haig 2016, Colling and Szűcs 2021). The 'New Statistics' (Cumming (2014, Cumming and Calin-Jageman 2024) is seen as a form of statistics reform. Very recently, Trafimow et al. (2024) proposed using a new two-step process comprising the APP (a prior procedure) and gain-probability analyses and diagrams to replace the traditional two-step process comprising the power analysis and NHST. On the other hand, some authors defend NHST and *p* values (e.g. Verhulst 2016, Benjamini et al. 2021, Hand 2022, Lohse 2022, Chén et al. 2023).

Chén et al. (2023) recently published a review paper that provides a systematic examination of the *p* value from its roles and merits to its misuses and misinterpretations. Chén et al. (2023) argue that the *p* value and hypothesis testing form a useful probabilistic decision-making system, but the interpretation of the *p* value must be contextual, taking into account scientific questions, experimental design, and statistical principles. Moreover, Chén et al. (2023) believe that the *p*

https://doi.org/10.32388/SVHLGA.2

value will continue to play an important role in hypothesis-testing-based scientific enquiries, whether in its current form or modified formulations.

Correct interpretation of the *p* value is crucial for the debate about the validity of the *p* value-based hypothesis testing. We agree with Chén et al.'s view that "the interpretation of the *p* value must be contextual, considering the scientific question, experimental design, and statistical principles." However, we argue that their paper does not reveal the actual meaning of the *p* value. Although Chén et al. (2023) correctly mentioned common misconceptions about the *p* value, including that "the *p* value measures the probability that the research hypothesis is true" and that "the *p* value measures the probability that observed data are due to chance," they review the *p* value to be a value "with which one assigns probabilistic belief about the property [of the population] and decides whether to reject the hypothesis the hypothesis." We argue that their view of the *p* value is merely another form of misconceptions and does not capture the actual meaning of the *p* value in practical applications.

Chén et al. (2023) used NHST to develop their discussion about the *p* value. However, they did not specify which NHST procedure produced the *p* value in their discussion. We argue that the interpretation of a *p* value must be tied to a specific NHST procedure that produced it. In other words, the actual meaning of the *p* value cannot be revealed without examining the specific problem we want to address.

In this commentary, we focus on the *p* value produced by the two-sample *z*-test and explore its meaning. In the following sections, we first discuss the definition of the *p* value given by Chén et al. (2023), and then discuss the meaning of the *p* value produced by the two-sample *z*-test and why the *p* value is not an appropriate probabilistic measure and what are alternatives. Finally, we give conclusion and recommendation.

On the definition of the *p* value

Chén et al. (2023) defined the *p* value as follows: "the *p* value is the tail probability calculated using a test statistic." Under the hypothesis testing paradigm, the test statistic, such as the Z statistic or T statistic, is a standardized effect size that is assumed to follow the standard normal distribution or a *t*-distribution. However, it is important to note that standardized effect sizes are dimensionless; they do not have the physical units of the quantity of interest in practice. Schäfer (2023) argued that standardized effect sizes bear a high risk for misinterpretation. Baguley (2009) stated, "For most purposes simple (unstandardized) effect size is more robust and versatile than standardized effect size." In real-world applications, our domain knowledge about a quantity of interest is related to the physical units of that quantity. It is easier for practitioners to assess the practical significance [true importance] of effects using the physical units than the dimensionless standardized effect sizes (Huang 2023a).

We argue that the definition of the *p* value as the tail probability calculated using a test statistic (a standardized effect size) is the root cause of two problems with the *p* value. First, it is not clear what the *p* value as the tail probability really means in practical problems. As a result, the *p* value can be easily misinterpreted. Common misconceptions about the *p* value include that "the *p* value measures the probability that the research hypothesis is true" and that "the *p* value measures the probability that observed data are due to chance," as stated by Chén et al. (2023).

Second, the *p* value can be easily hacked through "*N*-chasing," a term coined by Stansbury (2020), because the *p* value decreases monotonically as sample size increases (Chén et al. 2023). *N*-chasing guarantees the statistical significance at any pre-specified threshold, even if the actual effect (or unstandardized effect size) is very small and has no practical significance [true importance] (Huang 2023a). Chén et al. (2023) considered *p*-hacking to be a paradox. However, we argue that this paradox stems from the intrinsic property of the *p* value. Chén et al. (2023) offered several suggestions to avoid *p*-hacking, including "… consider sample size and effect size during experimental plans." On the other hand, they stated,

> Indeed, given unlimited resources, most people may prefer studies with very large sample sizes because they feel larger sample studies are more reliable than smaller trials. Here, we do not advocate against large-sample studies (which have many advantages, as we see below); rather, we argue that one should treat the p value contextually and avoid being that aggressive scientist.

We argue that their suggestions cannot help solve the *p*-hacking problem, because there is nothing to stop scientists from using large samples, which is actually preferred, whenever possible, in any study. Therefore, the *p*-hacking problem caused by *N*-chasing cannot be solved unless the *p* value-based hypothesis testing is abandoned.

The meaning of the *p* value produced by the two-sample *z*-test

As mentioned in the introduction, the interpretation of a *p* value must be tied to the specific NHST procedure that produces it. In this section, we consider the *p* value produced by the two-sample *z*-test and explore its meaning.

Suppose that two samples (two datasets) $\{x_{1,1}, x_{1,2}, \ldots, x_{1,n1}\}$ and $\{x_{2,1}, x_{2,2}, \ldots, x_{2,n2}\}$ are randomly drawn from two independent normal distributions $X_1 \sim N(\mu_1, \sigma_1)$ and $X_2 \sim N(\mu_2, \sigma_2)$, respectively, where n_1 and n_2 are the sample sizes. Neither μ_1 nor μ_2 is known, but σ_1 and σ_2 are known. Let $\bar{x}_{1,D}$ and $\bar{x}_{2,D}$ denote the calculated sample means. The *z*-score for the two-sample equal-variance *z*-test is written as

$$z_p = \frac{\bar{x}_{1,D} - \bar{x}_{2,D}}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}. \quad (1)$$

One-tailed z-test

We first consider the one-tailed *z*-test for the null: the absolute effect (i.e. the difference between two means) is greater than zero. Assuming that $\bar{x}_{1,D} - \bar{x}_{2,D} > 0$, $z_p > 0$. The one-tailed *p* value can be calculated as

$$p_{\text{one-tailed}} = \Pr(Z < -z_p) = \Phi(-z_p), \quad (2)$$

where $\Phi(.)$ is the cumulative probability function of the standard normal distribution $Z \sim N(0,1)$, and *Z* is the standardized effect size (statistic) written as

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\bar{x}_{1,D} - \bar{x}_{2,D})}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad (3)$$

where $\bar{X}_1$ and $\bar{X}_2$ are the sample means (statistics) that are normally distributed: $\bar{X}_1 \sim N(\bar{x}_{1,D}, \frac{\sigma_1}{\sqrt{n_1}})$ and $\bar{X}_2 \sim N(\bar{x}_{2,D}, \frac{\sigma_2}{\sqrt{n_2}})$ respectively.

Indeed, as Eq. (2) indicates, $p_{\text{one-tailed}}$, is the left tail probability of the standardized effect size distribution. However, the probability statement, Eq. (2), does not tell us what $p_{\text{one-tailed}}$ actually means in practical problems.

To explore the actual meaning of $p_{\text{one-tailed}}$, we substitute the expressions for Z and z_p into Eq. (2). Then, Eq. (2) can be rewritten as (Huang 2022)

$$p_{\text{one-tailed}} = \Pr\left(\left[Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\bar{x}_{1,D} - \bar{x}_{2,D})}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right] < \left[-z_p = -\frac{\bar{x}_{1,D} - \bar{x}_{2,D}}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right]\right), \tag{4}$$

which is the same as (Huang 2022)

$$p_{\text{one-tailed}} = \Pr(\bar{X}_1 - \bar{X}_2 < 0) = \Pr(\bar{X}_1 < \bar{X}_2) \tag{5}$$

Note that $\Delta\bar{X} = \bar{X}_1 - \bar{X}_2$ is the unstandardized effect size (statistic). Therefore, $p_{\text{one-tailed}}$ is the estimated probability that the sample mean $\bar{X}_1$ is smaller than the sample mean $\bar{X}_2$.

When the population variances are unknown and estimated using the sample variances s_1^2 and s_2^2, according to the central limit theorem, the sample means are normally distributed: $\bar{X}_1 \sim N(\bar{x}_{1,D}, \frac{s_1}{c_{4,n_1}\sqrt{n_1}})$ and $\bar{X}_2 \sim N(\bar{x}_{2,D}, \frac{s_2}{c_{4,n_2}\sqrt{n_2}})$, respectively, where $c_{4,n}$ is the bias correction factor for the sample standard deviation (Huang 2022). $p_{\text{one-tailed}}$ can still be estimated using Eq. (5).

Two-tailed z-test

Now consider the two-tailed z-test for the null: the effect (i.e. the difference between two means) is zero. The two-tailed p value can be calculated as

$$p_{\text{two-tailed}} = 1 - [\Pr(-z_p < Z < z_p)] = \Pr(Z < -z_p) + 1 - \Pr(Z > z_p) \tag{6}$$
$$= \Phi(-z_p) + [1 - \Phi(z_p)] = 2\Phi(-z_p) = \psi_z,$$

where ψ_z is called the probability of compatibility (Huang 2023b). Therefore, the two-tailed p vale $p_{\text{two-tailed}}$ has the same meaning as the compatibility probability ψ_z, which measures the degree of compatibility between the two estimated sampling distributions $\bar{X}_1$ and $\bar{X}_2$.

Why the *p* value is not an appropriate probabilistic measure and what are alternatives?

Chén et al. (2023) cited David Hume's view that "all knowledge degenerates into probability" and argued that probability guides scientific research. We agree with their view that probabilities or probabilistic measures play an important role in scientific researches. However, we argue that an appropriate probabilistic measure must be independent of sample size so that it cannot be hacked through *N*-chasing. Apparently, the *p* value is not an appropriate probabilistic measure because it can be easily hacked through *N*-chasing.

As shown in the previous section, the *p* value produced by a two-sample *z*-test is actually a probabilistic measure of the difference (or compatibility) between the two estimated *sampling distributions of means*. It is important to note that, when dealing with sampling distributions, some evidence or properties of the observed data are confounded with sample size and therefore cannot be properly uncovered. For example, according to the central limit theorem, as sample size increases, the sampling distribution of means will become a normal distribution, even if the samples come from a nonnormal distribution, making it impossible to reveal the underlying distribution of the observed data. Therefore, the *p* value of a two-sample *z*-test based on sampling distributions, is not helpful for extracting evidence from or exploring the properties of the observed data.

In many practical applications, we actually need a probabilistic measure for the difference between two (estimated) population distributions. Exceedance probability (EP) is such a probabilistic measure (Huang 2022), which is briefly described below.

Consider the effect size $\Delta X = X_1 - X_2$, which is a random variable. We assume that the probability density function of ΔX is available (e.g. estimated using the observed data). The exceedance probability of ΔX against a specified value Δx_{EP} is defined as (Huang 2022)

$$\mathrm{EP}(\Delta x_{\mathrm{EP}}) = \Pr(\Delta X > \Delta x_{\mathrm{EP}}) = 1 - \Pr(\Delta X \leq \Delta x_{\mathrm{EP}}) \qquad (7)$$

The meaning of $\mathrm{EP}(\Delta x_{\mathrm{EP}})$ is that EP% of the ΔX values will be greater than Δx_{EP}. For example, $\mathrm{EP}(\Delta x_{75} = 5)$ means that 75% of the ΔX values will be greater than $\Delta x_{75} = 5$. In other words, if we draw random samples x_1 and x_2 from the populations X_1 and X_2 respectively, 75% of the x_1 samples will be greater than the x_1 samples by at least $\Delta x_{75} = 5$. If we assume that X_1 represents the "treatment" group and X_2 represents the "control" group. There is a 75% chance that a randomly picked person from the "treatment" group will score at least $\Delta x_{75} = 5$ higher than a randomly picked person from the "control" group. Therefore, Δx_{EP} provides a probabilistic measure of the effect size ΔX; it is called the probabilistic effect size (PES) (Huang 2022).

The concept of exceedance probability is essentially the same as the concept of gain-probability (G-P) proposed by Trafimow et al. (2022). A diagram of the exceedance probability $\mathrm{EP}(\Delta x_{\mathrm{EP}})$ will be the same as the Type A G-P diagram of Trafimow et al. (2022). The specified value Δx_{EP} is the same as the critical value in the Type A G-P diagram. Trafimow et al. (2022, 2024) also proposed the Type B G-P diagram, which provides vertical bars representing probabilities of a value falling within particular intervals. The G-P analysis can provide fine-grained information about the probabilities of various degrees of gain or loss (Trafimow et al. 2024).

In the special case $\Delta x_{\mathrm{EP}} = 0$,

$$\mathrm{EP}(0) = \Pr(\Delta X > 0) = \Pr(X_1 > X_2) \qquad (8)$$

The meaning of $\mathrm{EP}(0)$ is essentially the same as the meaning of the common language effect size (CLES) (McGraw and Wong 1992), the probability of superiority (PS) (Vargha and Delaney 2000, Grissom and Kim 2001), or the area under the receiver operating characteristic (AUC) (Huang 2022). CLES can be considered to be an approximation of $\mathrm{EP}(0)$ (Huang 2022).

Unlike the p value, which depends on sample size, the exceedance probability EP(0) (or CLES or PS) is independent of sample size. Therefore, the exceedance probability EP(0) (or CLES or PS) cannot be hacked through *N*-chasing. In addition, it is worth mentioning that the concept of exceedance probability and its analysis have been used in some engineering fields such as environmental protection and water quality control (e.g. U.S. EPA 1991, Di Toro 1984, and Huang and Fergen 1995).

Furthermore, from a philosophical perspective, we argue that a fundamental principle of scientific inductive reasoning is that scientific claims must be based on statistical inference and domain knowledge about the *population* properties (or population distribution) of the quantity under consideration (e.g. effect size). Indeed, a basic idea of statistical inference is to infer *population* properties (e.g. population parameters) from the observed data. However, the p value produced by a NHST procedure is not a population property; it is a property of *sampling distributions* that depend on sample size. Therefore, making scientific claims based on p values violates the fundamental principle of scientific inductive reasoning. In contrast, the exceedance probability or gain-probability (and G-P diagrams) is a population property that does not depend on sample size. It can and should be used as an alternative to p values, in conjunction with domain knowledge, to make scientific claims.

Conclusion and recommendation

Scientists do need an appropriate probabilistic measure in scientific decision-making. We argue that, according to the fundamental principle of scientific inductive reasoning, an appropriate probabilistic measure must be a *population* property (or property of *population distributions*) that is independent of sample size, so that it cannot be hacked through *N*-chasing. The p value is not an appropriate probabilistic measure because it is a property of *sampling distributions* that depend on sample size and can be easily hacked through *N*-chasing. The problem of p-hacking caused by *N*-chasing cannot be solved unless the p value-based hypothesis testing is abandoned. The exceedance probability or gain-probability (and G-P diagrams) is a population property that does not depends on sample size. It should be used as an alternative to p values in scientific decision-making.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Amrhein V, Greenland S, and McShane B 2019 Retire statistical significance *Nature* **567** 305-307

Aurbacher J, Bahrs E, Banse M, Hess S, Hirsch S, Hüttel S, Latacz-Lohmann U, Mußhoff O, Odening M, and Teuber R 2024 Comments on the p-value debate and good statistical practice, *German Journal of Agricultural Economics* **73**(1) 1-3

Benjamini Y, De V R, Efron B, Evans S, Glickman M, Graubard B I, He X, Meng X-L, Reid N, Stigler S M, Vardeman S B, Wikle C K, Wright T, Young L J and Kafadar K 2021 ASA President's Task Force Statement on Statistical Significance and Replicability *Harvard Data Science Review* **3**(3) https://doi.org/10.1162/99608f92.f0ad0287

Baguley T 2009 Standardized or simple effect size: what should be reported? *Br J Psychol.* **100**(Pt 3) 603-17 doi: 10.1348/000712608X377117 Epub 2008 Nov 17 PMID: 19017432
Berner D and Amrhein V 2022 Why and how we should join the shift from significance testing to estimation *J Evol Biol.* **35**(6) 777-787 doi: 10.1111/jeb.14009. Epub 2022 May 18. PMID: 35582935; PMCID: PMC9322409. https://onlinelibrary.wiley.com/doi/10.1111/jeb.14009
Chén O Y, Bodelet J S, Saraiva R G, Phan H, Di J, Nagels G, Schwantje T, Cao H, Gou J, Reinen J M, Xiong B, Zhi B, Wang X, and de Vos M, 2023 The roles, challenges, and merits of the p value *Patterns* **4**(12) 100878, ISSN 2666-3899, https://doi.org/10.1016/j.patter.2023.100878
Claridge-Chang A and Assam P 2016 Estimation statistics should replace significance testing *Nat Methods* **13** 108–109 https://doi.org/10.1038/nmeth.3729
Colling L J and Szűcs D 2021 Statistical Inference and the Replication Crisis, *Review of Philosophy and Psychology* **12** 121–147 https://doi.org/10.1007/s13164-018-0421-4
Cumming G 2014 The new statistics: why and how *Psychological Science* **25**(1) 7-29 DOI: 10.1177/0956797613504966
Cumming G and Calin-Jageman R 2024 *Introduction to the New Statistics Estimation, Open Science, and Beyond* 2nd edition ISBN 9780367531508 Routledge
Di Toro D M 1984 Probability model of stream quality due to runoff *Journal of Environmental Engineering* ASCE **110**(3) 607-628
Elkins MR, Pinto RZ, Verhagen A, Grygorowicz M, Söderlund A, Guemann M, Gómez-Conesa A, Blanton S, Brismée JM, Ardern C, Agarwal S, Jette A, Karstens S, Harms M, Verheyden G, Sheikh U. 2022 Statistical inference through estimation: recommendations from the International Society of Physiotherapy Journal Editors *European Journal of Physiotherapy* **24**(3) 129-133 DOI: 10.1080/21679169.2022.2073991
Environment protection agency (EPA) 1991 *Technical support document for water quality-based toxics control,* Office of Water, Washington, DC, EPA/505/2-90-001
Grissom R J and Kim J J 2001 Review of assumptions and problems in the appropriate conceptualization of effect size *Psychol Methods* **6**(2) 135-46 doi: 10.1037/1082-989x.6.2.135. PMID: 11411438
Haig B D 2016 Tests of statistical significance made sound *Educational and Psychological Measurement* **77** 489–506. https://doi.org/10.1177/0013164416667981
Halsey L G 2019 The reign of the *p*-value is over: what alternative analyses could we employ to fill the power vacuum? *Biology Letters* **15**(5) 20190174 https://doi.org/10.1098/rsbl.2019.0174
Hand D J 2022 Trustworthiness of Statistical Inference *Journal of the Royal Statistical Society Series A*: *Statistics in Society* 1**85**(1) 329–347 https://doi.org/10.1111/rssa.12752
Heckelei T, Hüttel S, Odening M and Rommel J 2023 The p-value debate and statistical (Mal) practice–implications for the agricultural and food economics community. *German Journal of Agricultural Economics* **72**(1) 47-67 https://doi.org/10.30430/gjae.2023.0231
Huang H 2022 Exceedance probability analysis: a practical and effective alternative to t-tests *Journal of Probability and Statistical Science* **20**(1) 80-97
Huang H 2023a Statistics reform: practitioner's perspective (preprint) *ResearchGate* https://www.researchgate.net/publication/373551061_Statistics_reform_practitioner's_perspective
Huang H 2023b Probability of net superiority for comparing two groups or group means *Lobachevskii Journal of Mathematics* **44**(11) 42-54

Huang H and Fergen R E 1995 Probability-domain simulation - A new probabilistic method for water quality modeling *WEF Specialty Conference "Toxic Substances in Water Environments: Assessment and Control"* (Cincinnati, Ohio, May 14-17, 1995)

Lohse K 2022 In Defense of Hypothesis Testing: A Response to the Joint Editorial From the International Society of Physiotherapy Journal Editors on Statistical Inference Through Estimation *Physical Therapy*, **102**(11) 118 https://doi.org/10.1093/ptj/pzac118

McGraw K O and Wong S P 1992 A common language effect size statistic *Psychological Bulletin* **111**(2) 361–365 https://doi.org/10.1037/0033-2909.111.2.361

McShane B B, Gal D, Gelman A, Robert C P, and Tackett J L 2018 Abandon statistical significance *The American Statistician* **73** DOI: 10.1080/00031305.2018.1527253

Schäfer T 2023 On the use and misuse of standardized effect sizes in psychological research OSF Preprints June 7 doi:10.31219/osf.io/x8n3h

Stansbury D 2020 *p*-Hacking 101: *N* Chasing *The Clever Machine* https://dustinstansbury.github.io/theclevermachine/p-hacking-n-chasing

Trafimow D and Marks M 2015 Editorial *Basic and Applied Social Psychology* **37** 1-2

Trafimow D, Hyman M R, Kostyk A, Wang Z, Tong T, Wang T, and Wang C 2022 Gain-probability diagrams in consumer research *International Journal of Market Research* **64**(4) 470–483 https://doi.org/10.1177/14707853221085509

Trafimow D, Tong T, Wang T, Choy S T B, Hu L, Chen X, Wang C, and Wang Z 2024 Improving Inferential Analyses Pre-Data and Post-Data *Psychological Methods* (to be published)

Vargha A and Delaney H D 2000 A critique and improvement of the CL common language effect size statistic of McGraw and Wong *Journal of Educational and Behavioral Statistics* **25** 101–132 doi: 10.3102/10769986025002101

Verhulst B 2016 In defense of *p* values *AANA J.* **84**(5) 305-308 PMID: 28366961 PMCID: PMC5375179

Wagenmakers E-J, R. Wetzels D B and Maas H L J van der 2011 Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem *Journal of Personality and Social Psychology* **100** 426–432 https://doi.org/10.1037/a0022790

Wasserstein R L and Lazar N A 2016 The ASA's statement on *p*-values: context, process, and purpose, *The American Statistician* **70** 129-133 DOI:10.1080/00031305.2016.1154108

Wasserstein R L, Schirm A L, and Lazar N A 2019 Moving to a world beyond "$p < 0.05$" *The American Statistician* **73**:sup1 1-19 DOI: 10.1080/00031305.2019.1583913