MEDICINA CLÍNICA Y SOCIAL
Med. clín. soc. 2022;6(3):103-108. e-ISSN: 2521-2281
SETIEMBRE-DICIEMBRE. https://doi.org/10.52379/mcs.v6i3.254

ORIGINAL

Factors associated with knowledge of p-value in a sample of medical students

Factores asociados al conocimiento del valor p en una muestra de estudiantes de medicina

Víctor Juan Vera-Ponce[1,2], Jenny Raquel Torres-Malca[1,2], Fiorella E. Zuzunaga-Montoya[1,3], Jhony A. De La Cruz-Vargas[1], Liliana Cruz-Ausejo[3]

[1] Universidad Ricardo Palma, Universidad Ricardo Palma, Perú.
[2] Universidad Tecnológica de Perú, Lima, Perú.
[3] Universidad Científica del Sur, Lima, Perú.

Received: 25/05/2022
Revised: 10/06/2022
Accepted: 16/06/2022

Corresponding author

Dr. Victor Vera-Ponce
Instituto de Investigaciones de Ciencias Biomédicas, Perú.
victor.vera@urp.edu.pe

Conflicts of interests

The authors declare that there is no conflict of interest.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ABSTRACT

Introduction: The role of p-value knowledge for clinical practice is elemental; however, insufficient evidence on this is found in health science students. **Objective:** To determine the factors associated with p-value knowledge in human medical students. **Methods:** Analytical cross-sectional study. Application of a virtual survey to human medicine students from different faculties in Peru. **Results:** 54.69% had sufficient knowledge of p-value. The multivariate analysis found a statistically significant association with having sufficient knowledge on this topic in those who were between 6th to 9th semester (APr: 1.118; 95% CI 1.051 - 1.412; p=0.009) and medical internship (APr: 1.234; 95% CI 1.073 - 1.418; p=0.003); taking an external course in biostatistics, epidemiology or research (APr: 1.420; 95% CI 1.227 - 1.643; p<0.001); having read 6 to 12 articles per year (APr: 1.353; 95% CI 1.196 - 1.530; p<0.001) and more than 12 articles per year (APr: 1.590; 95% CI 1.313 - 1.967; p<0.001); and publishing at least one scientific article (APr: 1.397; 95% CI 1.199 - 1.628; p<0.001) or more than one (APr:1.424; 95% CI 1.196 - 1.696; p<0.001). **Conclusion:** It was found that the academic semester, having taken an external course, having read more than 6 articles per year and having published at least one scientific article are independently associated with having greater understanding of this topic.

Keywords: knowledge; association; statistics and numerical data; students.

RESUMEN

Introducción: El conocimiento del rol de valor-p para la práctica clínica es fundamental; sin embargo, la evidencia científica de éste en los estudiantes de ciencias de la salud no es suficiente. **Objetivo:** Determinar los factores asociados al conocimiento sobre el valor-p en estudiantes de medicina humana.
Métodos: Estudio transversal analítico. Se aplicó una encuesta virtual a estudiantes de medicina humana de distintas facultades de medicina del Perú. **Resultados:** El 54.69% tuvo un conocimiento suficiente sobre el rol del valor-p. El análisis multivariado encontró asociaciones estadísticamente significativas con tener conocimiento suficiente en este tema en aquellos que se encontraban entre el 6° y 9° semestre (APr: 1.118; 95% CI 1.051 - 1.412; p=0.009), eran internos de medicina (APr: 1.234; 95% CI 1.073 - 1.418; p=0.003); haber llevado un curso externo de bioestadística, epidemiología o investigación (APr: 1.420; 95% CI 1.227 - 1.643; p<0.001); leer entre 6 y 12 artículos científicos por año (APr: 1.353; 95% CI 1.196 - 1.530; p<0.001), leer más de 12 artículos por año (APr: 1.590; 95% CI 1.313 - 1.967; p<0.001); y haber publicado al menos un artículo científico (APr: 1.397; 95% CI 1.199 - 1.628; p<0.001) o más de uno (APr:1.424; 95% CI 1.196 - 1.696; p<0.001).
Conclusión: Los hallazgos mostraron que el semestre académico, haber llevado un curso externo, leído más de 6 artículos por años y publicado al menos un artículo científico se asocian independientemente con tener un mayor entendimiento de este tópico.

Palabras clave: Conocimiento; asociación; estadística y datos numéricos; estudiantes.

How to cite this article: Vera-Ponce VJ, Torres-Malca JR, Zuzunaga-Montoya FE, De la Cruz-Vargas JA, Cruz-Ausejo L. Factors associated with knowledge of p-value in a sample of medical students. Med. clín. soc. 2022;6(3):103-108

INTRODUCTION

The p-value is a topic of fundamental importance in epidemiology and research, despite its controversies and questioning (1). Not only is it part of the statistical inference process based on hypothesis testing, at a critical level, p-value expresses the degree of comparability between the null hypothesis and the data, being specific the p-value is the probability associated with contrast statistics when the null hypothesis is true (1,2). Moreover, it also allows a reflective process for decision-making in health and the critical analysis of scientific articles (2).

However, studies in different parts of the world have found deficiencies in the knowledge of this. Horton et al. (3,4) mention that health professionals have greater difficulties in understanding statistical methods, while Andreu et al. (3) report that the prevalence, in Argentina, of low knowledge about p-values in doctors and therapists is 63%. Whilst in Peru, the Araoz-Melgarejo et al study shows that insight into statistical analyzes is low (5).

Although most of the work is focused on knowing if the health professionals knew biostatistics (6-10), it should be known how the understanding of the p-value, as a biostatistics' tool, is independent, especially in undergraduate students. For this reason, the objective of this research is to determine the factors associated with knowledge about p-value in a sample of human medicine students.

METHODS

Study design: Cross-sectional analytical study based on the analysis of a virtual questionnaire distributed from September 1, 2021, to October 1, 2021. Population, sample and eligibility criteria: The population was made up of 1192 medical students of both sexes, belonging to faculties of human medicine in Peru. Those who agreed to participate in the study and those who reported residing in the country were included. Those who were in the first, second and third cycle of the career (by standardization, due to the probability of not having taken the biostatistics course), those under 18 years of age, and those who did not adequately complete the questionnaire questions were excluded. Consecutive non-probabilistic sampling was carried out.

Variable definition: The questionnaire contained three groups of questions: The first part consisted of 8 sociodemographic questions that were age; sex; academy semester; external course in epidemiology, biostatistics or research; reading of scientific articles; type of university; if is the author of an article published this year and number of articles published. The second section consisted of 11 nominal questions (True/False/Don't know) about p-values. If the answer was correct, a point was awarded, while if it was incorrect, none was awarded. The result of this was categorized dichotomously, grouped into "sufficient knowledge" (≥ 6 points) vs "insufficient knowledge" (< 6 points).

Data collection and procedure: Given the national situation (COVID-19 pandemic), it was decided to collect the information virtually. The questionnaire was designed in Google Form. The test lasted about 10 minutes per person. This was distributed through the online survey on Facebook and Whatsapp, to contact university medical students.

Statistical analysis: Statistical analyzes were performed with STATA version 17.0 software. For the descriptive analysis, the qualitative variables were summarized in absolute and relative frequencies; while the quantitative variable was presented in the form of median and interquartile ranges, due to the non-normal distribution evaluated by histogram. In the bivariate analysis, the chi-square test was performed. Finally, a generalized linear model of the Poisson family with robust variance was performed to obtain the crude prevalence ratio (PRc) and adjusted (PRa) for the covariates mentioned above. It was considered statistically significant with the p-value <0.05 and the 95% confidence interval (95% CI) was presented.

Ethical aspects: Informed consent was given to all participants. The information obtained did not violate the privacy and integrity of the study participants, since they were filled out anonymously. The procedures complied with the ethical standards of the Council for International Organizations of Medical Sciences (CIOMS) and the Declaration of Helsinki.

RESULTS

A total of 1192 students were surveyed. The 57.55% were female, while 54.87% were between the 10th and 14th cycle. Only 28.86% belonged to a public university. 35.99% took an external course in biostatistics, epidemiology or research and 54.69% presented sufficient knowledge about the p-value. Regarding the bivariate analysis, no statistically significant association was found with age (p=0.156) and university (p=0.098). The rest of the characteristics and analyzes can be seen in the first column of Table 1.

TABLE 1. CHARACTERISTICS AND BIVARIATE ANALYSIS OF P-VALUE KNOWLEDGE IN PERUVIAN MEDICAL STUDENTS (N=1192).

	Knowledge of p-value			
Characteristics	Insufficient	Sufficient	p*
	n (%)	n (%)	n (%)	
Gender				
Female	686 (57,55)	292 (42,57)	394 (57,43)	0,023
Male	506 (42,45)	249 (49,21)	257 (50,79)	
Categorized age				
18 - 24 years old	715 (59,88)	312 (43,64)	403 (56,36)	0,156
25 years and older	479 (40,12)	229 (47,81)	250 (52,19)	
Academic semester				
4th to 5th semester	256 (21,48)	138 (53,91)	118 (46,09)	< 0,001
6th to 9th semester	282 (23,66)	96 (34,04)	186 (65,96)	
10th to 14th semester	654 (54,87)	307 (46,94)	347 (53,06)	
University				
Public	344 (28,86)	169 (49,13)	175 (50,87)	0,098
Private	848 (71,14)	372 (43,87)	476 (56,13)	
External course in biostatistics, epidemiology or research				
No	763 (64,01)	447 (58,58)	316 (41,42)	< 0,001
Yes	429 (35,99)	94 (21,91)	335 (78,09)	
Number of articles read in the year				
Up to 5 articles	442 (37,08)	252 (57,01)	190 (42,9)	< 0,001
6 to 12 articles	432 (36,24)	164 (37,96)	268 (62,04)	
More than 12 articles	318 (26,68)	125 (39,31)	193 (60,69)	
Number of articles published				
No	714 (59,90)	427 (59,80)	287 (40,20)	< 0,001
One	375 (31,46)	92 (24,53)	283 (75,47)	
More than 1	103 (8,64)	22 (21,36)	81 (78,64)	

*Analysis performed with the chi-square test of independence.

Overall, the question with the most correct answers was about the concept of p-value as probability (72.73%; 95% CI 70.13% - 75.19%), while the question with the least correct answers was about the interpretation of the p-value in a clinical analysis (13.92%; 95% CI 12.07% - 16.01%). The rest of the responses can be seen in Table 2.

TABLE 2: PERCENTAGE OF CORRECT ANSWERS TO THE KNOWLEDGE QUESTIONS IN P-VALUES (N=1192).

Question	Objective	Correct % (CI 95%)
1	The p-value is a probability	72,73 (70,13 - 75,19)
2	A non-significant p-value (p> 0.05) indicates that the null hypothesis is true.	61,99 (70,13 - 75,19)
3	A non-significant p-value (p> 0.05) indicates that the treatment effect under analysis is not clinically important.	56,69 (50,85 - 56,51)
4	A non-significant p-value (p> 0.05) indicates that both treatments are similar.	39,43 (36,69 - 42,24)
5	p-value indicates the probability that the null hypothesis is true given the results of our study	64,18 (61,41 - 66,85)
6	A non-significant p-value (p> 0.05) indicates that we should accept the null hypothesis.	54,53 (51,69 - 57,34)
7	If we obtain a significant p-value (p<0.05), we should reject the null hypothesis	48,32 (45,49 - 51,16)
8	The p-value obtained (p = 0.02) indicates the probability of obtaining similar results if the same study is repeated with a similar sample.	21,39 (19,15 - 23,82)
9	A statistically significant result (p <0.05) indicates that the treatment effect under analysis is clinically significant	47,73 (44,91 - 50,58)
10	The p value observed in our study was significant (p = 0.02). This confirms that the treatment effect was greater than that observed in a similar study with a p value = 0.04	13,92 (12,07 - 16,01)
11	Which of the following statements is the definition of a p-value?	55,96 (53,12 - 58,76)

Table 3 shows the multivariate analysis of each factor associated with knowledge of biostatistics. The variables used for adjustment were gender, categorized age, academic cycle, external course, type of university, number of articles read, and number of articles published. A statistically significant association was found for those between the 6th and 9th cycle (APr: 1.118; 95% CI 1.051 - 1.412; p=0.009) and medical internship (APr: 1.234; 95% CI 1.073 - 1.418; p=0.003). Taking an external course in biostatistics, epidemiology or research (APr: 1420; 95% CI 1227 - 1643; p<0.001); having read 6 to 12 articles per year (APr: 1353; 95% CI 1196 - 1530; p < 0.001) and more than 12 articles per year (PRa: 1590; 95% CI 1313 - 1967; p < 0.001); and publish at least one scientific article (PRa: 1.397; 95% CI 1.199 - 1.628; p<0.001) or more than one (APr:1.424; 95% CI 1.196 - 1.696; p<0.001).

TABLE 3. SIMPLE AND MULTIPLE POISSON REGRESSION ANALYSIS OF FACTORS ASSOCIATED WITH P-VALUE KNOWLEDGE IN PERUVIAN MEDICAL STUDENTS (N=1192).

	Bivariate analysis			Multivariable regression		
	CPr	CI 95%	p	APr	CI 95%	p
Gender						
Female	Ref.			Ref.		
Male	0,884	0,794 - 0,984	0,025	0,909	0,824 - 1,004	0,061
Categorized age						
18 to 24 years old	Ref.			Ref.		
25 years and older	0,926	0,832 - 1,031	0,160	0,930	0,842 - 1,027	0,153
Academic semester						
4th to 5th semester	Ref.			Ref.		
6th a 9th semester	1,431	1,223 - 1,674	< 0,001	1,118	1,051 - 1,412	0,009
10th to 14th semester	1,151	0,989 - 1,339	0,068	1,234	1,073 - 1,418	0,003
University						
Public	Ref.			Ref.		
Private	1,103	0,979 - 1,244	0,107	1,076	0,959 - 1,206	0,212
External course in biostatistics, epidemiology or research						
No	Ref.			Ref.		
Yes	1,885	1,709 - 2,080	< 0,001	1,420	1,227 - 1,643	< 0,001
Number of articles read in the year						
Up to 5 articles	Ref.			Ref.		
6 to 12 articles	1,443	1,267 - 1,644	< 0,001	1,353	1,196 - 1,530	< 0,001
More than 12 articles	1,412	1,228 - 1,623	< 0,001	1,359	1,196 - 1,543	< 0,001
Number of articles published						
No	Ref.			Ref.		
One	1,877	1,688 - 2,088	< 0,001	1,397	1,199 - 1,628	< 0,001
More than 1	1,956	1,710 - 2,239	< 0,001	1,424	1,196 - 1,696	< 0,001

Adjusted for: sex, age categorized, academic year, external course, type of university, number of articles read and number of articles published. CPr: Crude prevalence ratio. APr: Adjusted prevalence ratio. 95% CI: Confidence interval at 95%.

DISCUSIÓN

It was evidenced that a little more than half understood the biostatistical results reported in the medical literature. This is the first study that reports the level of knowledge of the p-value for Peruvian medical students. Numerous investigations have studied statistical knowledge in human medicine students and medical residents (11-13), but few have focused solely on the p-value, since in a circumscribed way, knowledge of this has been evaluated with only one or two questions, undervaluing or overvaluing this topic.

The percentage of students with sufficient knowledge was slightly higher than 50%. Although other studies have found that the values of ignorance of this topic revolve around 60% (3,14,15), In general, the concepts

of biostatistics in our environment are low (5). Furthermore, the question that had the fewest correct answers was the one that requested an interpretation of the p-value, despite the fact that the majority knew the concept of the p-value as probability. Works such as Lecoutre et al. (16) and Badenes-Ribera (14) have found that there is a weakness in the interpretation of this. This also means that the lack of knowledge of the value p is more frequent to the interpretation than to the theoretical concept. This may be because when health professionals read scientific articles, they do not usually apply an interpretation beyond applying the p < alpha rule to know if there are statistically significant differences or not (17).

Both age and gender were not associated with knowledge about p-value. The first should not, since the gender difference should not cause a greater or lesser knowledge of it, and it is consistent with what was found in other works (3,6). In the case of the second, this may be due to the fact that it does not matter that they are older, if they have not studied the subject, they do not have to know this well (3,16). No association was found between public and private universities, which also coincided with the work of Andreu et al. (3).

With higher academic semesters, there have been several courses that request the need to understand the concepts of p-value, as in clinical courses, where published cases are discussed, so it allows you to learn more about the subject as you finish your degree. The Araoz-Melgarejo study had the same result when it evaluated biostatistical knowledge in medical students (5).

Taking an external course in research allows the student to be instructed in the subject in a more focused way, first because it would not be a compulsory course, being voluntary on the part of the student, and, secondly, because these courses focus on deeper statistical topics (18). The study by Andreu et al. (3) found that the lack of training in scientific research methodology increased the probability of having less knowledge on this subject. Other investigations also found the same, although they were aimed at knowledge in biostatistics in a general way (6,19,20). Furthermore, as more scientific articles are read and manuscripts are published, the understanding of the p-value is much greater. These results coincide with the works of Andreu et al. (3), that reading less than 6 articles per year increased the probability of having less knowledge on this topic.

This study has limitations. First, the questionnaire was sent through internet media to obtain a good number of undergraduate students. Second, since a probabilistic sampling has not been carried out, it is likely that it will not be representative at the national level; however, given the characteristics that may be similar among students, some inference could finally be made. Thirdly, the cut-off was arbitrarily grouped to define with what score it can be said that you have sufficient knowledge and the opposite. However, the authors considered that the median plus one would be an adequate score, in addition to considering that a dichotomized value is of greater analytical understanding.

ACKNOWLEDGEMET

We would like to thank every Medicine student who filled this questionnaire, for their time and contribution to science.

AUTHORS CONTRIBUTIONS

All the authors participated in the conception of the idea, methodology design, synthesis of evidence, writing and final approval of this manuscript.

REFERENCIAS

1. Mendoza C. El Valor P en Epidemiología. Revista Chilena de Salud Pública. 2006;10(1):47–51. URL.
2. Kain Z, MacLAren J. Valor de p inferior a 0,05: ¿qué significa en realidad? Pediatrics. 2007;63(3):118–120. URL.
3. Andreu MF, Ballve LPD, Verdecchia DH, Monzón AM, Carvalho TD de. Is the p-value properly interpreted by critical care professionals? Online survey. Rev Bras Ter Intensiva. 2021;33(1):88–95. https://doi.org/10.5935/0103-507X.20210009
4. Horton NJ, Switzer SS. Statistical methods in the journal. N Engl J Med. 2005;353(18):1977–1979. https://doi.org/10.1056/NEJM200511033531823
5. Araoz-Melgarejo VA, Espinoza BM, Quiñones-Laveriano DM, Cruz-Vargas JADL. Basic knowledge and attitudes towards biostatistics in sixth- and seventh-year medical students. Turkish Journal of Computer and Mathematics Education (TURCOMAT). 2021;12(14):1286–1302. URL.
6. Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. JAMA. 2007;298(9):1010–1022. https://doi.org/10.1001/jama.298.9.1010
7. Polychronopoulou A, Eliades T, Taoufik K, Papadopoulos MA, Athanasiou AE. Knowledge of European orthodontic postgraduate students on biostatistics. Eur J Orthod. 2011;33(4):434–440. https://doi.org/10.1093/ejo/cjq098
8. Taylor RS, Reeves BC, Ewings PE, Taylor RJ. Critical appraisal skills training for health care professionals: a randomized controlled trial [ISRCTN46272378]. BMC Med Educ. 2004;4(1):30. https://doi.org/10.1186/1472-6920-4-30
9. Ganasegeran K, Ch'ng ASH, Jamil MFA, Looi I. Clinicians' Perceived Understanding of Biostatistical Results in the Medical Literature: A Cross-Sectional Study. Medicina (Kaunas). 2019;55(6):E227. https://doi.org/10.3390/medicina55060227
10. Msaouel P, Kappos T, Tasoulis A, Apostolopoulos AP, Lekkas I, Tripodaki E-S, et al. Assessment of cognitive biases and biostatistics knowledge of medical residents: a

multicenter, cross-sectional questionnaire study. Med Educ Online. 2014;19:23646. https://doi.org/10.3402/meo.v19.23646

11. Alzahrani SH, Aba Al-Khail BA. Resident physician's knowledge and attitudes toward biostatistics and research methods concepts. Saudi Med J. 2015;36(10):1236–1240. https://doi.org/10.15537/smj.2015.10.11842
12. Best AM, Laskin DM. Oral and maxillofacial surgery residents have poor understanding of biostatistics. J Oral Maxillofac Surg. 2013;71(1):227–234. https://doi.org/10.1016/j.joms.2012.03.010
13. Couture F, Nguyen D-D, Bhojani N, Lee JY, Richard PO. Knowledge and confidence level of Canadian urology residents toward biostatistics: A national survey. Can Urol Assoc J. 2020;14(10):E514–19. https://doi.org/10.5489/cuaj.6495
14. Badenes-Ribera L, Frias-Navarro D, Iotti B, Bonilla-Campos A, Longobardi C. Misconceptions of the p-value among Chilean and Italian Academic Psychologists. Front Psychol. 2016;7:1247. https://doi.org/10.3389/fpsyg.2016.01247
15. Badenes-Ribera L, Frías-Navarro D, Monterde-i-Bort H, Pascual-Soler M. Interpretation of the p value: A national survey study in academic psychologists from Spain. Psicothema. 2015;27(3):290–295. https://doi.org/10.7334/psicothema2014.283
16. Lecoutre M-P, Poitevineau J, Lecoutre B. Even statisticians are not immune to misinterpretations of Null Hypothesis Tests. International Journal of Psychology. 2003;38(1):37–45. https://doi.org/10.1080/00207590244000250
17. Lakens D. The Practical Alternative to the p Value Is the Correctly Used p Value. Perspect Psychol Sci. 2021;16(3):639–48. https://doi.org/10.1177/1745691620958012
18. OJEDA HS MM. Problems and challenges of teaching biostatistics to medical students and professionals. Medical Teacher. 1999;21(3):286–288. https://doi.org/10.1080/01421599979545
19. Chima SC, Nkwanyana NM, Esterhuizen TM. Impact of a short biostatistics course on knowledge and performance of postgraduate scholars: Implications for training of African doctors and biomedical researchers. Niger J Clin Pract. 2015;18 Suppl:S62-70. https://doi.org/10.4103/1119-3077.170818
20. Kiekkas P, Panagiotarou A, Malja A, Tahirai D, Zykai R, Bakalis N, et al. Nursing students' attitudes toward statistics: Effect of a biostatistics course and association with examination performance. Nurse Educ Today. 2015;35(12):1283–1288. https://doi.org/10.1016/j.nedt.2015.07.005