

Size effect in observational studies in Public Oral Health: importance, calculation and interpretation

Flávia Martão Flório (<https://orcid.org/0000-0001-7742-0255>)¹

Luciane Zanin (<https://orcid.org/0000-0003-0218-9313>)¹

Leônidas Marinho dos Santos Júnior (<https://orcid.org/0000-0002-4685-906X>)¹

Marcelo de Castro Meneghim (<https://orcid.org/0000-0003-2673-3627>)²

Gláucia Maria Bovi Ambrosano (<https://orcid.org/0000-0001-5950-6370>)²

Abstract *The objective of this study was to analyze the scientific literature in public oral health regarding calculation, presentation, and discussion of the effect size in observational studies. The scientific literature (2015 to 2019) was analyzed regarding: a) general information (journal and guidelines to authors, number of variables and outcomes), b) objective and consistency with sample calculation presentation; c) effect size (presentation, measure used and consistency with data discussion and conclusion). A total of 123 articles from 66 journals were analyzed. Most articles analyzed presented a single outcome (74%) and did not mention sample size calculation (69.9%). Among those who did, 70.3% showed consistency between sample calculation used and the objective. Only 3.3% of articles mentioned the term effect size and 24.4% did not consider that in the discussion of results, despite showing effect size calculation. Logistic regression was the most commonly used statistical methodology (98.4%) and Odds Ratio was the most commonly used effect size measure (94.3%), although it was not cited and discussed as an effect size measure in most studies (96.7%). It could be concluded that most researchers restrict the discussion of their results only to the statistical significance found in associations under study.*

Key words *Statistical data interpretation, Observational study, Bias*

¹ Departamento de Saúde Coletiva, Faculdade São Leopoldo Mandic.

R. José Rocha Junqueira
13, Swift. 13045-755
Campinas SP Brasil.
flaviaflorio@yahoo.com

² Faculdade de Odontologia de Piracicaba, Universidade Estadual de Campinas.
Piracicaba SP Brasil.

Introduction

Effect size is a descriptive measure that allows the discussion of results in terms of the magnitude of the effect of intervention or study factor¹, and it is recommended that this value be reported and interpreted by researchers in their scientific articles².

Taken together, effect size and statistical significance allow the true significance to be assessed without a possible misleading effect of the sample size^{3,4}, which can occur when only statistical significance is taken into account⁵. In this way, it is possible to describe and analyze the observed effects, since large but not statistically significant effects suggest that future studies need greater test power (larger sample size), while small but significant effects due to the large sample size, must be taken into account and discussed, thus avoiding overestimation of the observed effect³.

The level of significance still dominates the preference of researchers when discussing data found, even though the debate about the need for its proper interpretation⁶ is not new, since the real meaning and isolated interpretations of p-values may be accompanied by misinterpretation⁷. In observational studies, much more than in randomized trials, bias and confusion can suppress the assumption that there is only a 5% probability that the observed effect is seen by chance when in reality there is no effect, since by definition, in this type of study, there is no intervention and exposure may not be the only potential explanation for differences observed in results⁵.

Hypothesis tests are applied to control the probabilities of errors when rejecting or not a hypothesis. However, when analyzed in isolation, the results of these tests only inform the probability of the result found to be by chance and frequently, results with lower probability p-values (e.g. $p < 0.001$) are erroneously interpreted as having stronger effect than those with higher p-values (e.g. $p < 0.05$)⁸. Determining the magnitude of the effect of interest and the accuracy of estimating the magnitude of this effect⁹ are fundamental aspects to be considered when weighing the clinical or practical importance of the results, and for this purpose, the analysis of effect sizes and confidence intervals must be considered^{4,9}.

Cohen presents and classifies the effect sizes for different statistical methodologies^{10,11}, being commonly presented as the standardized mean

difference (Cohen's *d* or Hedges' *g*) or as the strength of association (Pearson's *r*) between two groups or variables¹². Cohen^{10,11} also provided guidelines for the interpretation of these values based on the notion that an average effect should be noticeable to the naked eye of a careful observer: values of 0.20; 0.50 and 0.80 for Cohen's *d* and Hedges' *g* and 0.10; 0.30 and 0.50 for the correlation coefficient are commonly considered, respectively, as indicative of small, medium and large effects, which represent the manifestation of the phenomenon evaluated in the population.

The effect size depends on the result obtained and the population of interest and therefore, it is suggested that the classification of the effect size distribution should be analyzed in each of the study areas¹².

In the area of public oral health, investigations often seek to identify association between risk or protective factors for diseases or clinical measures. In this case, the measures that quantify the magnitude of this association are usually expressed by odds ratio (OR), prevalence ratio (PR) or relative risk (RR), depending on the study design and type of variables under study¹³, the first two measures being indicated for cross-sectional observational studies, with OR also indicated in case-control studies and RR indicated for longitudinal studies. These measures are considered non-standardized effect size statistics, as they indicate the direction and strength of association between exposure variables and the outcome.

For OR, which is the effect size index most commonly used to demonstrate increase or decrease in the chance of disease in epidemiological studies, the authors determined that, for disease rate of 1% in the unexposed group, the reference limits that reflect "weak association" (Cohen's $d=0.20$); "moderate association" (Cohen's $d=0.50$) or "strong association" (Cohen's $d=0.80$) are ORs of 1.68, 3.47 and 6.71, respectively. Considering disease rate of 5% in unexposed people, the corresponding reference limits are 1.52, 2.74 and 4.72⁸.

Thus, the objective of the present study was to analyze and discuss a section of the specific scientific literature in the area of public oral health regarding calculation, presentation and discussion of the effect size on the results of observational studies. In addition, the study aimed to detail the calculations and interpretation of effect size measures that can be used in articles in this area.

Methods

Type of study and ethical considerations

This is an observational, retrospective study with theoretical discussion. As this is a study with data collected from public domain databases, there was no need for ethical evaluation.

Search strategy, selection of journals and studies

In January 2020, a search was carried out in electronic databases considering the period from January 2015 to December 2019. Searches took place considering articles published with open and free access in MEDLINE - via PubMed using MeSH terms (Medical Subject Headings): (oral health) OR (dentistry) AND (logistic models) AND (analysis regression) AND free full text[sb] AND "last 5 years"[PDat])). All observational studies found were included.

Study variables

Two calibrated examiners performed the search for articles and by consensus, they were assisted in cases of doubts or disagreements by a third examiner, who collected and analyzed the following information in the selected articles:

Basic information: Journal; Year of publication.

About the study: Type of study; Research objective; Sample size; Number of variables; Detailing of outcomes; Instruments used in data collection; Presence or absence of sample statistical calculation; Parameters used for sample size calculation; Consistency of sample calculation with the research objective; Statistical methodology used; Did the study mention the term effect size?; Effect size presentation and, if so: which measure was presented, what is the value of the minimum significant effect size, if the minimum significant value was medium or large and not significant, was this finding discussed?; if the minimum significant value was small and significant, was it discussed?; Did the study consider the effect size found in the conclusion section?

Effect size calculation and classification

The effect size measures found in articles were detailed in terms of their concepts, calculations and interpretations.

Effect size in the norms of journals in the area

A search was carried out in the norms of journals responsible for the publication of three or more articles selected for the present search, seeking in the guidelines to authors the presence of recommendation for effect size presentation.

Results

Description of studies

A total of 123 articles from 66 journals were included in the study, of which 9.8% (12) were published in 2015; 17.1% (21) in 2016; 30.1% (37) in 2017; 25.2% (31) in 2018 and 17.9% (22) in 2019.

Table 1 presents a summary of the main characteristics of analyzed articles. It was observed that most of selected studies had a single outcome (74%), did not report sample size calculation (69.9%) and among those that did, in 70.3% of them, there was consistency between sample size calculation and the study objective. For articles in which these consistencies were not observed, there is a common error in calculating the sample size for the purpose of estimating prevalence in studies with the objective of measuring association.

Table 2 presents the statistical methodologies used in studies and the way in which results are presented. It was observed that logistic regression was the most commonly used statistical methodology and the effect sizes of associations were represented in articles mainly by the odds ratio, which in turn, had small magnitudes and little discussed in most articles. It was also observed that only 3.3% of articles mentioned the term effect size and 24.4% did not consider, despite having calculated, the effect size in the discussion of results.

Effect size in the norms of journals in the area

Table 3 presents the results of the search in the journals' norms regarding the presence of recommendation for the effect size presentation in manuscripts. It was observed that together, they published 50.3% of the evaluated production and only 2 of the 10 journals mentioned effect size presentation in the guidelines to authors.

Table 1. Characteristics of the articles analyzed (January 2015 to December 2019, MEDLINE database - via PubMed).

Characteristic	Categories	Frequencies	%
Number of outcomes	1	91	74.0
	2	25	20.3
	3	3	2.4
	4 or more	4	3.3
Sample calculation presentation	No	86	69.9
	Yes	37	30.1
Consistency of sample calculation with the question and design	No	11	29.7
	Yes	26	70.3

Source: Authors.

Table 2. Methodological characteristics of research in the area of Public Health (January 2015 to December 2019, MEDLINE Base - via PubMed).

Characteristic	Categories	Frequencies	%
Statistical methodologies used	Linear regression	1	0.8
	Logistic Regression	121	98.4
	Poisson Regression	1	0.8
Effect sizes	Odds ratio	116	94.3
	Relative risk	2	1.6
	Prevalence ratio	1	0.8
	Not presented	4	3.2
Minimum significant OR (effect size) in cases where OR are presented	Up to 1.5	54	55.1
	1.5-2.0	16	16.3
	2.0-2.5	12	12.2
	2.5-3.0	7	7.1
	>3.0	9	9.2
When the minimum significant OR was medium or large, it was discussed	No	3	33.3
	Yes	6	66.7
When the minimum significant OR was small, it was discussed	No	12	44.4
	Yes	15	55.6
Mentions effect size	No	119	96.7
	Yes	4	3.3
Takes effect size into account in the conclusion?	No	30	24.4
	Yes	93	75.6

Source: Authors.

Analysis of effect sizes presented

As a way of presenting the effect sizes used in articles in the area, the concepts of odds ratio (OR), relative risk (RR) and prevalence ratio (PR) and their calculations based on simulated data are detailed.

Odds ratio (OR)

OR with respective confidence intervals can be estimated from the coefficients of logistic regression models.

To exemplify the calculation and facilitate the measurement interpretation, simulated data presented in Table 4 were used. Two cross-sectional studies were simulated to evaluate the association between the consumption of sweetened beverages and caries experience in children, with similar results, but different sample sizes and logistic regression analysis was used to estimate OR.

Considering data from simulation 1, the sample size used was 64. Despite OR being 2.15, the confidence interval is wide due to the small sample size (95%CI: 0.66-6.95) and the associa-

Table 3. Journals with more articles evaluated and recommendations on the effect size presentation according to guidelines to authors. (January 2015 to December 2019, MEDLINE Base - via PubMed).

Journal	Frequency	%	Recommendation
Braz Oral Res	16	13.0	Not included
Cien Saude Colet	11	8.9	Not included
PLoS One	9	7.3	<ul style="list-style-type: none"> • Reporting of statistical methods: If appropriate, provide sample sizes, along with a description of how they were determined. <i>If a sample size calculation was performed, specify the inputs for power, effect size and alpha.</i> Where relevant, report the number of independent replications for each experiment • Reporting of statistical results (Reporting parameters): Test statistics (F/t/r) and associated degrees of freedom should be provided. <i>Effect sizes and confidence intervals should be reported where appropriate.</i> If percentages are provided, the numerator and denominator should also be given
Int J Environ Res Public Health	7	5.7	Not included
Int J Equity Health	4	3.3	Not included
Aust Dent J.	3	2.4	Not included
BMC Public Health	3	2.4	Not included
BMJ Open	3	2.4	<ul style="list-style-type: none"> • Results: main results with (for quantitative studies) 95% confidence intervals and, where appropriate, the exact level of statistical significance and the number need to treat/harm. Whenever possible, state absolute rather than relative risks
J Appl Oral Sci.	3	2.4	Not included
J Formosan Med Assoc.	3	2.4	Not included

Source: Authors.

tion was not statistically significant ($p=0.3211$). The result of the same study was then simulated (Simulation 2); however, with larger sample size ($n=632$). It was observed that results were similar, that is, OR was 2.11, but with 95% CI of 1.44-3.08 and in this case, the association was statistically significant ($p=0.0001$).

In both cases, OR is close to two, but depending on the sample size, there is a change in the amplitude of the confidence interval and in the statistical significance. In simulation 2, it was observed that children who consumed sweetened beverages were 2.11 (95%CI: 1.44-3.08) times more likely of experiencing dental caries. To understand what this significant chance represents, in the group of children who did not consume sweetened beverages, 172 children had caries experience; therefore, the likelihood of having caries experience in this group is $172/75=2.29$. Likewise, the likelihood of having caries experience in children who consume sweetened beverages is $319/66=4.83$. The relationship between these

two likelihoods (4.83/2.29) results in odds ratio of 2.11.

When OR is significantly greater than one, the category under study is more likely of having the event than the reference category.

Prevalence Ratio (PR)

PR with the respective confidence intervals can be estimated from Negative Binomial and Poisson regression models.

Table 4 also presents the results of simulation 2, calculating this measure of association in substitution to OR. It was observed that in the group of children with caries experience, the prevalence of schoolchildren who consumed sweetened beverages is 1.19 times higher than in the group of children without caries experience. In the group of children with caries experience, the prevalence of children who do not consume sweetened beverages is 69.6% and 82.9% for those who do. Calculating the relationship between the two prevalences (82.9%/69.6%) leads to prevalence ratio

of 1.19. The greater the distance between PR in relation to PR=1 (both for more and for less), the greater the effect size for this variable.

When PR is significantly greater than one, the category under study has higher prevalence of the event than the reference category.

Relative Risk (RR)

This measure of association can only be calculated in longitudinal cohort studies¹³ and, therefore, represents the relative risk of developing the outcome in exposed in relation to unexposed ones. RR with the respective confidence intervals can be estimated from Negative Binomial and Poisson regression models.

While PR is the ratio between two prevalences, RR is the ratio between two incidences. As an example, in a simulated study (Table 5), the impact of caries experience on oral health-related quality of life was evaluated. From the negative binomial regression analysis, RRs were estimated. For caries experience, RR was 1.50 (95%CI: 1.04-2.17), $p=0.0204$. In this case, the interpretation is that the presence of caries is associated with a 50% increase in the impact of oral health on quality of life. In the same way as OR and PR, the greater the distance between RR in relation to RR=1 (both for more and for less), the greater the effect size for this variable.

In the group without caries experience, the risk of having worse quality of life is $30/100=30\%$. In the group with caries experience, the risk of having worse quality of life is $45/100=45\%$. So, the relative risk= $45\%/30\%=1.5$. That is, children with caries experience are 1.5 times more likely of having worse quality of life.

Discussion

The present study reinforces the fact that although the literature in the area of statistics for a long time exposes the need and importance of presenting and discussing the effect size in articles, it was observed that only 3.3% of evaluated articles mentioned in their texts the term effect size and 24.4% did not consider the effect size in the discussion of results, despite having calculated it. Since 1925, Fisher proposed that researchers add the correlation rate or η (eta) to the significance of the analysis of variance (ANOVA), that is, the effect size, representing the strength of association between independent and dependent variables¹⁴.

Although slower than necessary, there has been pressure from scientific journals on researchers so that effect sizes are reported and interpreted in articles^{3,15}. Among journals analyzed in the literature review carried out in this study, those that published three or more of the included studies were selected, which together accounted for more than 50% of selected articles, and it was found that only 20% of them explicitly suggested reporting the effect size of articles in their guidelines. This finding is in agreement with a previous study that identified that only a small portion of journals from different areas explicitly recommended in the standards to authors calculating the magnitude of the effect size¹⁵.

There is a great deal of misunderstanding in literature about the correct definition of the effect size, which is sometimes incorrectly used. Kelley and Preacher¹⁶ propose a definition for the effect size and discuss it based on three partic-

Table 4. Example of the use of odds ratio (OR) or prevalence ratio (PR) in the analysis of the association between consumption of sweetened beverages and caries experience in children (simulated data).

Consumption of sweetened beverages	Frequency	Caries experience		(95%CI)	p-value
		No	Yes		
Odds ratio - Simulation 1					
No	25	8 (30.0%)	17 (68.8%)	Reference	
Yes	39	7 (16.9%)	32 (81.8%)	2.15 (0.66-6.95)	0.3211
Odds ratio - Simulation 2					
No	247	75 (30.4%)	172 (69.6%)	Reference	
Yes	385	66 (17.1%)	319 (82.9%)	2.11 (1.44-3.08)	0.0001
Prevalence Ratio - Simulation 2					
No	247	75 (30.4%)	172 (69.6%)	Reference	
Yes	385	66 (17.1%)	319 (82.9%)	1.19 (1.08-1.31)	<0.0001

Source: Authors.

Table 5. Use of Relative Risk (RR) in the analysis of the association between caries experience and oral health-related quality of life (simulated data).

Caries experience	Frequency	Quality of life		PR (95%CI)	p-value
		Worse	Better		
No	100	30 (30.0%)	70 (70.0%)	Reference	
Yes	100	45 (45.0%)	55 (55.0%)	1.50 (1.04-2.17)	0.0204

Source: Authors.

ularities (dimension, measure/index and value). According to these authors, the effect size can be presented with a statistic that estimates the magnitude of the effect (for example, correlation coefficient=0.3) or with a qualitative interpretation of this statistic (median correlation), which must take into account the practical applicability of the finding. Also according to the authors, effect size is often linked to the idea of substantive significance (for example, practical, clinical, medical or managerial importance), which can be understood as the degree to which stakeholders (scientists, professionals, politicians, managers, consumers, decision makers, the general public, etc.) would consider a discovery important and worthy of attention and possibly of action.

In this context, the exclusive use of the significance level to analyze and discuss the findings is not enough^{4,5} as it only informs whether the research result is due to the analyzed effect or to chance (sample variability). Practical significance informs whether the results are useful in the real world and is analyzed by the effect size found, being essential to draw the attention of researchers to the need to analyze the effect sizes found in their publications^{10,11}. In addition, previously observed effect sizes can serve as a basis for calculating power, for estimating the appropriate sample size in further studies^{1,3,17}, for understanding the study results in the context of previous studies, in addition to facilitating the incorporation of their results in future meta-analyses, which are very relevant as a standard method of quantitative review in biology⁹.

According to Kirk¹⁴, the magnitude of the effect can be classified into three categories: a) measure of the strength of associations, b) measure of the effect size (typically standardized difference between means), c) other measures.

Most articles that present and/or discuss effect size use ANOVA, t test and calculate the effect according to Cohen^{10,11}, but as verified in the present study, these statistical methodologies are rarely used in articles in the area of Public

Oral Health and very little is said about effect size when using logistic regression analysis, which is a statistical methodology used in 98.4% of articles.

In the present study, it was observed that articles present the effect size by measuring the strength of associations between variables, since in 94.3% of selected articles, the presentation of the odds ratio (OR) was verified, corroborating Chen *et al.*⁸, who reported that this is probably the most commonly used effect size index in epidemiological studies because it reflects the chances of a successful or desired outcome in the intervention group in relation to the chances of a similar outcome in the control group¹⁵.

Breaugh¹⁸ highlights some misconceptions about effect size estimates and introduces a series of effect size measures that, according to the author, depending on the research context and target population, can better communicate the importance of the relationship between two variables. In the case of dichotomous variables, there is a limitation in the use of phi as an effect size measure (ϕ is a measure commonly used as effect size in 2 x 2 contingency table analysis) because its possible amplitude is affected by the variable distribution. In certain areas such as medicine, it is common for a risk ratio to be reported as a measure of effect size. In this context, many statisticians have suggested reporting OR as a measure of effect, rather than the risk ratio or the phi coefficient, as seen in articles evaluated in the present study.

A desirable property of OR is that its possible range of values is not influenced by the marginal distributions of variables. It was observed in the present study that the vast majority of articles evaluated (96.7%) presented OR as a measure of the effect size, but 24.4% did not take this value into account when discussing the results and conclusions, which leads to the understanding that the authors have based the discussion and conclusion of their works only on p-values. In articles evaluated, 8.9% made conclusions based on significant association without mentioning

that the effect size was small. In addition, 33.3% of articles concluded that the association was not significant, without mentioning that the OR was medium or large, that is, the sample was probably small and other studies need to be carried out with larger samples. Therefore, the authors have to take this into account and these two important information, that is, the p-value and the effect size, in this case the degree of association (OR).

Chen *et al.*⁸ present OR classification into small, medium and large according to the probabilities being compared and Durlak¹⁵ presents a guide for the selection, calculation and interpretation of effect sizes. In this study, different types of commonly used effect sizes are discussed.

Ferguson¹⁹ recommends small, medium, and large odds ratio effect sizes of 2.0; 3.0 and 4.0, but recommends caution in their use, as they are not “anchored” to the Pearson correlation coefficient. Although many authors have pointed to problems with ϕ as a measure of association and encourage the use of odds ratios as alternative, effect size recommendations for odds ratios do not generally exist. The authors demonstrate the relationship between ϕ and odds ratio and recommend chance effect sizes derived from Cohen’s work. For a 1:1 allocation ratio, odds ratio of 1.22; 1.86 and 3.00 correspond to small, medium, and large effect sizes.

Thus, the effect size (substantive significance) complements the statistical significance and one measure does not replace the other, and must be analyzed in a complementary way, so that a step towards scientific veracity is taken. Ialongo²⁰ presents an introduction and a guide for the reader interested in the use of effect size estimation and emphasizes that evidence can be quantified by hypothesis tests, which represent the probability (or p-value) by which it is likely to consider the observation shaped by chance (the

so-called “null hypothesis”) rather than by the phenomenon (the so-called “alternative hypothesis”). The size at which the p-value is considered small enough to exclude the effect of chance corresponds to statistical significance. So, when the researcher obtains a non-significant result, two possibilities must be considered: the first is that there is no phenomenon and only the effect of chance is being observed, and the second is that the phenomenon exists, but its effect is small and confused with the effect of chance.

It is in the second possibility that the issue about the importance of presenting the phenomenon when it actually exists arises, quantifying it by calculating the effect size, that is, how large (or small) is the expected effect produced by the phenomenon in relation to the observation through which it is intended to be detected. For this reason, researchers should be encouraged to present the effect size in their work, particularly reporting it whenever the p-value is mentioned.

Among the study limitations, it is noteworthy that a literature review was carried out to contextualize the theme and the frequencies presented apply only to this review. Despite this, the results presented here allowed performing a theoretical discussion on the subject, providing an opportunity to understand that the report and discussion of the effect size in studies should be done as a routine and that reviewers and editors of scientific journals should pay attention to their report and appropriate discussion.

It could be concluded in the present study that most researchers restricted the discussion of their results only to the statistical significance found in the tested associations and journals do not explicitly indicate the need to present the magnitude of effects and the need to consider it in the discussion of results and conclusion of the study.

Collaborations

FM Flório, GMB Ambrosano and MC Meneghim devised the study. FM Flório and L Zanin conducted the study and writing of the manuscript. FM Flório, L Zanin, LM Santos Junior, GMB Ambrosano and MC Meneghim contributed with the final and critical review of the manuscript.

References

1. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol* 2013; 4:863.
2. Wilkinson L, Task Force on Statistical Inference. Statistical methods in psychology journals: Guidelines and explanations. *Am Psychol* 1999; 54:594-604.
3. Lindenau JDR, Guimarães LSP. Calculando o tamanho de efeito no SPSS. *Rev HCPA* 2012; 32(3):363-381.
4. Espírito Santo H, Daniel F. Calcular e apresentar tamanhos do efeito em trabalhos científicos (1): As limitações do $p < 0,05$ na análise de diferenças de médias de dois grupos. *Rev Port Invest Comport Soc* 2015; 1(1):3-16.
5. Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Stat Med* 2014; 33(2):209-218.
6. Baker M. Statisticians issue warning over misuse of P values. *Nature* 2016; 531(7593):151.
7. Gigerenzer G. Statistical Rituals: The Replication Delusion and How We Got There. *Adv Methods Pract Psychol Sci* 2018; 1(2):198-218.
8. Chen H, Cohen P, Chen S. How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Commun Stat Simul Comput* 2010; 39(4):860-864.
9. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc* 2007; 82(4):591-605.
10. Cohen J. *Statistical power analysis for the behavioral sciences*. 2ª ed. Mahwah: Lawrence Erlbaum Associates; 1988.
11. Cohen J. A power primer. *Psychol Bull* 1992; 112:155-159.
12. Brydges CR. Effect Size Guidelines, Sample Size Calculations, and Statistical Power in Gerontology. *Innov Aging* 2019; 3(4):igz036.
13. Papaléo CLM. *Estimação de risco relativo e razão de prevalência com defecho binário*. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2009.
14. Kirk RE. Practical significance: A concept whose time has come. *Edu Psychol Measuram* 1996; 56:746-759.
15. Durlak JA. How to select, calculate, and interpret effect sizes. *J Pediatr Psychol* 2009; 34(9):917-928.
16. Kelley K, Preacher KJ. On Effect Size. *Psychol Methods* 2012; 17(2):137-152.
17. Olivier J, Bell ML. Effect sizes for 2x2 contingency tables. *PLoS One* 2013; 8(3):e58777.
18. Breaugh JA. Effect Size Estimation: Factors to Consider and Mistakes to Avoid. *Journal of Management* 2003; 29(1) 79-97.
19. Ferguson CJ. An effect size primer: A guide for clinicians and researchers. *Prof Psychol Res Pract* 2009; 40(5):532-538.
20. Ialongo C. Understanding the effect size and its measures. *Biochem Med (Zagreb)* 2016; 26(2):150-163.

Article submitted 25/04/2022

Approved 12/08/2022

Final version submitted 14/08/2022

Chief editors: Romeu Gomes, Antônio Augusto Moura da Silva