JCSM
Journal of Clinical Sleep Medicine

LETTERS TO THE EDITOR

The Real Role of Sensitivity, Specificity and Predictive Values in the Clinical Assessment

Response to Raphael. Error in calculation of predictive values in paper on screening for sleep bruxism. *J Clin Sleep Med* 2016;12(2):277.

Marcelo Palinkas, DDS, MSc, PhD[1]; Graziela De Luca Canto, DDS, MSc, PhD[2,3]; Laíse Angélica Mendes Rodrigues, DDS, MSc, PhD[1]; César Bataglion, DDS, MSc, PhD[5]; Selma Siéssere, DDS, MSc, PhD[1]; Marisa Semprini, DDS, MSc, PhD[1]; Simone Cecilio Hallak Regalo, DDS, MSc, PhD[1]

[1]Department of Morphology, Physiology and Basic Pathology, Ribeirão Preto School of Dentistry, University of São Paulo, São Paulo, Brazil; [2]Department of Dentistry, Federal University of Santa Catarina, Florianopolis, Brazil; [3]Department of Dentistry, University of Alberta, Edmonton, Canada; [4]Department of Restorative Dentistry, Ribeirão Preto School of Dentistry, University of São Paulo, São Paulo, Brazil

We would like to thank Dr. Raphael, because her letter[1] published in this issue of the *Journal of Clinical Sleep Medicine* provides us with an opportunity to emphasize points already highlighted in our article.

The main goal of our study was to evaluate the diagnostic ability of signs and symptoms of sleep bruxism (SB) according to the criteria of the American Academy of Sleep Medicine (AASM) and a diagnostic classification system proposed by international experts to assess SB.

The validity of a diagnostic test is determined as the ability of a test to tell who have the disease from who do not. For this purpose, two components are calculated: sensitivity and specificity. Sensitivity is the ability to correctly identify those who have the disease, while specificity is the ability to correctly identify those who do not have the disease.[2] In order to calculate sensitivity and specificity, it is required that patients be identified by another test which provides a more permanent result, is often more sophisticated, more invasive, and more expensive, named gold standard.[2] In our study, the polysomnography (PSG) was considered the gold standard for SB assessment.

The evaluation of the validity of a diagnostic test is usually performed on selected contexts as well, with two equally-numbered groups of patients—one with the disease, one without it—as this is an efficient way of describing sensitivity and specificity.[3] Having that in mind, we selected a control group with the same number of patients as the group of patients with the disease for the sample, then we calculated the sensitivity and the specificity of each tested diagnostic criteria. We know that sensitivity and specificity are characteristics of the diagnostic test, although the predictive value is also influenced by the prevalence.[3] That is why when we want to evaluate the discriminatory capacity of a particular diagnostic test, we calculate sensitivity and specificity, even though the predictive values are clinically more useful.[2] In our study, besides the main measures to determine accuracy, we chose to present additional analyses reporting predictive values of each test, which were calculated based on the prevalence of our sample and not from the literature, despite the limitations. However, our conclusion was mainly based on sensitivity and specificity.

The ideal situation is that a diagnostic test is highly sensitive and highly specific, but as this is not always possible, we seek a counterbalance between sensitivity and specificity.[3] Despite having a good specificity (82%), in our study the AASM criteria did not show sensitivity (58%) high enough to replace the gold standard (PSG) in the diagnosis of SB. However, analyzing sensitivity and specificity, it is possible to state that AASM criteria can be used as an effective screening tool.

In addition to that, another good way of expressing the relationship between sensitivity and specificity is with the ROC curve, where the overall accuracy of the test can be seen in the area under the curve; the larger the area, the better the test, i.e. tests with good discriminatory power have their curve in the upper left corner. The ROC curves are especially useful for comparing possible alternative tests for the same diagnosis.[3] By calculating the odds ratio (OR), the diagnostic test performance information is summarized in a single number rather than two, facilitating the comparison between different tests.[3] To clarify these points, we presented the ROC curve (Figure 4) and OR (Table 1), as well as the forest plot (Figure 3). In these figures, it was possible to compare the various tests and identify that the items "muscle fatigue" and " temporal headache" can also be used as a screening tool and assert that the AASM criteria are the best screening tool when compared to "lock jaw" and "mandibular muscle pain." It is clear that the usefulness of the screening methods depends on the context in which they are used. A screening program is more productive and effective if targeted to a high-risk population, because when performed in an entire population for an uncommon disease, it can be expensive and detect only a few cases.[2] Our study suggests that the AASM criteria can be used as an effective screening tool to ensure a better selection of patients for referral to PSG. We know that the results of any test should be interpreted in the context of disease prevalence in the population where it originated. Health care professionals should be familiar with the mathematical relationships between the properties of diagnostic

tests and the information that they provide to different clinical situations. This includes the characteristics of the clinical situation, such as signs, symptoms, and risk of disease, and they influence the probability of finding a disease.[3]

Finally, the relative complexity and high costs associated with overnight PSG as the gold standard approach employed for diagnosing the vast majority of sleep disorders has spurred the quest for alternative diagnostic methods.[4] The development of less expensive, simple, and reliable screening tools that permit precise screening of at-risk populations is essential. If accurate identification of those subjects with or without definitive disease is accomplished using such simplified and less onerous tools, then timely access to clinical care would be possible to a large sector of the population.[5] A recent study agrees with our results that the current use of self-reported SB failed to significantly predict the presence or absence of either moderate or severe SB as assessed by PSG.[6]

Although a single set of simple criteria cannot capture the complexity of such a sleep movement disorder, it may be reasonable to combine the criteria evaluated in this study into a clinical algorithm that could form the basis for a new diagnostic system combining intelligent methodologies and clinical insights. Our study is only an initial point; future research is needed to clarify this topic and provide us with better tools with acceptable diagnostic test accuracy.[7]

CITATION

Palinka M, De Luca Canto G, Rodrigues LA, Bataglion C, Siéssere S, Semprini M, Regalo SC. The real role of sensitivity, specificity and predictive values in the clinical assessment. *J Clin Sleep Med* 2016;12(2):279–280.

REFERENCES

1. Raphael KG. Error in calculation of predictive values in paper on screening for sleep bruxism. J Clin Sleep Med 2016;12:277.
2. Gordis L. Epidemiology, 3rd ed. Philadelphia, PA: Elsevier/Saunders, 2004.
3. Fletcher RH, Fletcher SW, Wagner EH. Clinical epidemiology: the essentials, 3rd ed. Porto Alegre: Artmed, 2003.
4. Gozal D. Serum, urine, and breath-related biomarkers in the diagnosis of obstructive sleep apnea in children: is it for real? Curr Opin Pulm Med 2012;18:561–7.
5. De Luca Canto G, Pacheco-Pereira C, Aydinoz S, Major PW, Flores-Mir C, Gozal D. Biomarkers associated with obstructive sleep apnea: a scoping review. Sleep Med Rev 2015;23:28–45.
6. Raphael KG, Janal MN, Sirois DA, et al. Validity of self-reported sleep bruxism among myofascial temporomandibular disorder patients and controls. J Oral Rehabil 2015;42:751–8.
7. Palinkas M, De Luca Canto G, Rodrigues LA, et al. Comparative capabilities of clinical assessment, diagnostic criteria, and polysomnography in detecting sleep bruxism. J Clin Sleep Med 2015;11:1319–25.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication January, 2016
Accepted for publication January, 2016
Address correspondence to: Graziela De Luca Canto, Departamento de Odontologia-Centro de Ciências da Saúde, Campus Reitor João David Ferreira Lima, Florianópolis - Santa Catarina Brazil, CEP: 88040-900; Email: graziela.canto@ufsc.br

DISCLOSURE STATEMENT

The authors have indicated no financial conflicts of interest.