Internal construct validity of the Brazilian version of a tool for assessing the population’s knowledge of human papillomavirus

Validade interna do construto da versão brasileira de um instrumento para aferição do conhecimento sobre o papilomavirus humano

Fabiana Schuelter Trevisol, Rodrigo Dias Nunes, Gabriel Oscar Cremona Parma, André Luciano Manoel, Clávison Martinelli Zapelini, Daisson José Trevisol, Jefferson Traebert

ABSTRACT: Objective: To verify the internal construct validity of the Brazilian Portuguese version of a tool for measuring the general population’s knowledge of human papillomavirus (HPV). Materials and methods: A cross-culturally adapted Brazilian Portuguese version of a measurement tool originally designed for English speaking populations was administered to 330 adults in Tubarão, Santa Catarina, Southern Brazil. After examining the overall suitability of the method, we performed investigations based on the item response theory and exploratory factor analysis. Results: Ten of the 29 items presented a low contribution to the construct and were excluded from subsequent analysis. The factor analysis yielded three factors, which explained approximately 51% of the variance variability. A different arrangement from the original measurement tool was found: general HPV knowledge, with six items; HPV vaccination knowledge, with five items; HPV transmission and testing knowledge, with eight items. Conclusion: The Brazilian Portuguese version under study presented a different behavior from the original measurement tool, but proved to be a reliable and valid instrument in assessing the Brazilian population’s knowledge about HPV.

Keywords: Papillomaviridae. Questionnaires. Validation studies.
INTRODUCTION

The persistent human papillomavirus (HPV) infection is the main cause of cervical cancer and is associated with neoplasms in other tissues. Cervical cancer is one of the most common causes of cancer-related deaths in women in low- and middle-income countries1.

According to recent estimates by the HPV Information Center2, cervical cancer is the third most frequent among women aged 15–44 years in Brazil. Infection with HPV types 16 and 18, which are responsible for approximately 70% of cervical cancer cases, can be prevented through vaccination3. However, vaccine coverage is low among adolescents, the target population of HPV immunoprophylaxis in Brazil4. Some authors have argued that the low HPV vaccine uptake in Brazil may be due to fear of adverse reactions (following media reports of neurological symptoms in clusters of girls in Brazil), parental vaccine hesitancy, and/or logistical challenges to vaccinating adolescents at health care centers5,6.

Among the main restrictive factors of preventive practices regarding the cervical cancer is the lack of knowledge about the disease and its prevention, especially in adolescent populations, considered the most vulnerable group for acquiring HPV infection7. Media advertisements, health professionals, and parents are the most frequent sources of information, but doubts on the accuracy of the message transmitted remain8. We believe that by measuring the knowledge about HPV, it will be possible to identify the need to expand the information on the virus, diagnosis, treatment, and prevention.

There are several published studies on the knowledge about HPV in different population strata, in many countries9–13. However, the lack of an instrument with the same theoretical
construct to gauge this knowledge prevents the comparability of data. Waller et al.\textsuperscript{14} proposed a scale to measure lay knowledge of HPV in the United States, the United Kingdom, and Australia. The questionnaire consists of 29 items with three subscales — general HPV knowledge, HPV vaccination knowledge, and HPV testing knowledge.

Recently, we proposed and published a cross-culturally adapted Brazilian Portuguese version of that tool. The details of the cross-cultural adaptation process can be found in Manoel et al.\textsuperscript{15}. The preliminary model showed satisfactory reliability; however, the construct validity still needs to be established. The type of validation that allows verifying whether an instrument actually measures what it proposes to measure is called construct validity, given by the ability of a test to measure a theoretical trait or construct, thus validating a background theory\textsuperscript{14}. In fact, it checks whether the observations that supported the arguments correspond to the theoretical parameters of the investigated subject. Construct validity is not limited to the measurement. Its broader purpose focuses on the validation of the theory on which it relied for the construction of the instrument\textsuperscript{16}.

In 2014, the Brazilian National Health System launched the HPV immunization campaign, but the goals were not reached\textsuperscript{1}. Possibly, the lack of general public knowledge has raised social and ethical discussions that could explain it, at least partially. Public understanding of preventive measures and consequences of HPV infection can play an important role in contributing to the achievement of these goals. Thus, it is crucial to have a valid instrument for measuring HPV knowledge that can be administered to different segments of the population.

This study aimed to verify the internal construct validity of the Brazilian Portuguese version of a tool for measuring the general population’s knowledge of HPV. We expect that a validated tool can help determine how education measures can be implemented.

**MATERIALS AND METHODS**

**STUDY DESIGN AND PARTICIPANTS**

This is a cross-sectional study. The 29 items of the cross-culturally adapted Brazilian Portuguese version\textsuperscript{15} of the instrument proposed to measure lay knowledge of HPV in English speaking countries\textsuperscript{14} were self-administered by 330 adults. The sample comprised parents of 9–15-year-old adolescents, recruited from public and private schools in the municipality of Tubarão, Santa Catarina, Southern Brazil, and community health agents from the same municipality, who had not been previously trained in HPV screening. The sample was non-probabilistic, as previously described\textsuperscript{15,17}.

An informed consent form, complying with standards of the Declaration of Helsinki, was delivered to the participants, along with the HPV knowledge survey instrument. The Human Research Ethics Committee of Universidade do Sul de Santa Catarina, Brazil, approved the research project, under protocol number 734,735.
DATA ANALYSIS

We assessed the psychometric properties of the Brazilian Portuguese version of the questionnaire at the item level based on the item response theory (IRT). Items that did not fit the IRT model, evaluated by the chi-square test, were excluded from the subsequent analysis. For the remained items, we used classical reliability and exploratory factor analysis (EFA) with principal component analysis (PCA) to assess internal consistency and determine possible factors or subscales.

Questions 1 (Have you ever heard of HPV?), 2 (Have you ever heard of HPV vaccination?), and 3 (Have you ever heard of HPV testing?) were excluded from the analysis because their purpose was solely to enable the participant to continue responding to each of the three proposed sections.

IRT is a modern method recommended by the Patient-Reported Outcomes Measurement Information System (PROMIS) group to test the item loading. We analyzed the 29 items to identify the discriminatory power of the instrument and of each item alone. The analysis was performed in the open software R 3.30.

Cronbach’s alpha was used to assess internal consistency reliability after a previous analysis of the overall suitability of the dataset. A matrix between each pair of questions was examined through Pearson’s linear correlation. We also carried out the Kaiser-Meyer-Olkin (KMO) and Bartlett’s sphericity tests. All tests were performed in the software IBM SPSS version 18.0 (IBM Corp., Armonk, USA) for statistical analysis.

The PCA to define the number of factors involved two analyses: using the Kaiser-Guttman criterion, which considers factors with eigenvalues greater than or very close to one ($\lambda \geq 1$); performing a scree plot test to observe the eigenvalues.

Based on the Kaiser-Guttman criterion, we assumed that the loadings between factors with eigenvalues smaller than one and the original variables should be low, given that they would have had higher correlations with previously extracted factors with higher eigenvalues. Therefore, the original questions that shared small variance percentages with the other questions had their factor loadings raised in a single factor. The analysis of communalities allowed us to verify if any question did not share a significant percentage of variance with the defined factors.

We used the Varimax method to minimize the number of questions that presented high loadings in a given factor by redistributing the loadings and maximizing the shared variance in factors with smaller eigenvalues. Finally, we carried out a theoretical evaluation according to the information required by the item and performed a regrouping based on the possibilities offered by the statistical model.

RESULTS

The sample consisted mainly of women (87.6%), and the mean age was 41.3 ± 8.8 years. More details about the sample are available in Manoel et al. and Manoel et al.
ITEM RESPONSE THEORY

The one-dimensionality of the instrument allowed us to proceed with IRT. Three classic models were created:

- rasch;
- two-parameter logistic (ltm);
- three-parameter logistic (tpm).

They were analyzed through ANOVA. Compared to rasch, the ltm model presented better results [Akaike Information Criterion (AIC): ltm 9900.3 < rasch 10084.0; Bayesian Information Criterion (BIC): ltm 10120.6 < rasch 10194.2; p < 0.001]. The second analysis showed similar results between ltm and tpm (AIC: tpm 9881.3 < ltm 9900.3; BIC: tpm 10211.8 > ltm 10120.7; p < 0.001). Both models (ltm or tpm) could be valid, but we selected the two-parameter logistic, since it is the most classic model.

Figure 1 shows the item trace lines. They were expressed numerically by the discrimination and difficulty parameters. IRT demonstrated that the item 1e had the strongest
discrimination (3.4) while the item 1q (0.0) had the lowest contribution. Otherwise, item 3f presented the highest difficulty level (1.7) while the lowest was identified in item 1q (-175.9). We excluded ten items with poor discrimination power (1g, 1k, 1m, 1q, 2c, 2e, 3a, 3b, 3d, and 3f) and performed a new IRT analysis. All items with a p-value greater than or equal to 0.05 remained for the EFA.

### RELIABILITY AND FACTOR ANALYSIS

The instrument demonstrated an overall Cronbach’s alpha index of 0.80. The Cronbach’s alpha index if each item was deleted from the subscale remained similar to the overall value. The assessment of the overall suitability of the dataset for EFA showed that the relationship between the number of subjects interviewed and the number of questions was 11.4, which represented a favorable condition. The correlation matrix of the variables presented Pearson correlation coefficient values above 0.2. The overall adequacy of the EFA for the data set showed a score of 0.82 in the KMO statistic, evidencing a correlation between the variables. Bartlett’s test presented a significance level of p < 0.001.

The scree plot indicated the existence of two to five factors, depending on the slope of the curve points (Figure 2). According to the Kaiser-Guttman criterion, the extraction of five factors was responsible for explaining 51.38% of the total variance of the instrument,
with eigenvalues greater than or close to one (Table 1). PCA was performed with two, three, four, and five factors to identify which configuration could better explain the theoretical affinity between items. We considered that, usually, the indicated difference between dimensions should be greater than 2.5 to allow selecting the next factor.

The best configuration was obtained with three factors. The rotation of the factors by the Varimax method sought to minimize the number of variables with high factor loadings in a factor and maximize the variation between the weights of each main component. The component matrix, after orthogonal rotation, aimed to maximize the factor loadings, so that each variable was associated with only one factor, simplifying the interpretation of these factors. All values lower than 0.2 were eliminated due to a weak correlation. The rotated factor loading matrix showed that factor 1 included items related to “general information,” factor 2 to “vaccination,” and factor 3 to “transmission and diagnosis.” Items 1c, 1e, and 1j demonstrated loadings for factors 1 and 3, but although the highest loading belonged to factor 1, they were nearest to factor 3 according to the theoretical analysis. The same occurred to item 2b, with loading in factors 2 and 3, but better proximity to factor 2. Table 2 shows how the items were effectively related to the factors. The sequence order followed the factor and value of correlation (above 0.2) of each item.

**DISCUSSION**

The results pointed to keeping the latent construct, and the Brazilian Portuguese version might retain 19 of the 29 items, with a different arrangement from the originally proposed instrument: general HPV knowledge, with six items; HPV vaccination knowledge, with five items; HPV transmission and testing knowledge with eight items. The first dimension comprised questions related to “general HPV knowledge” as proposed by Waller et al. The second also consisted of questions from the original “HPV testing knowledge.” However, the

<table>
<thead>
<tr>
<th>Component</th>
<th>Square summation of the extracted factor loadings</th>
<th>Square summation of rotated factor loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Percent variance</td>
</tr>
<tr>
<td>1</td>
<td>4.29</td>
<td>22.57</td>
</tr>
<tr>
<td>2</td>
<td>1.73</td>
<td>9.10</td>
</tr>
<tr>
<td>3</td>
<td>1.50</td>
<td>7.90</td>
</tr>
<tr>
<td>4</td>
<td>1.21</td>
<td>6.40</td>
</tr>
<tr>
<td>5</td>
<td>1.03</td>
<td>5.43</td>
</tr>
</tbody>
</table>
third dimension of the Brazilian Portuguese instrument involved a combination of two questions from the original “HPV vaccination knowledge” and six questions from the original “general HPV knowledge.” Nevertheless, we emphasize that ten items presented factor loadings in two dimensions and two other items in three dimensions, indicating certain non-specificity. It forced the researchers to arbitrate their distribution among the dimensions based on the theoretical status, respecting the statistical results of each loading on the factor. The practical need for separating the items clearly by dimension remains. On the other hand, probably the most important practical result of this study was providing a shorter and valid version with good internal consistency, indicating good reliability.

Table 2. Rotated factor loading matrix with three factors by the Varimax method, rearranged and eliminating scores lower than 0.2.

<table>
<thead>
<tr>
<th>Items</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>HPV can cause cervical cancer</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>HPV is very rare</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>A person could have HPV for many years without knowing it</td>
<td>0.63</td>
</tr>
<tr>
<td>HPV always has visible signs or symptoms</td>
<td>0.53</td>
</tr>
<tr>
<td>HPV can cause HIV/AIDS</td>
<td>0.51</td>
</tr>
<tr>
<td>HPV can be cured with antibiotics</td>
<td>0.31</td>
</tr>
<tr>
<td>Girls who had an HPV vaccine do not need a Pap test when they are</td>
<td>0.74</td>
</tr>
<tr>
<td>older</td>
<td></td>
</tr>
<tr>
<td>The HPV vaccine requires three doses</td>
<td>0.66</td>
</tr>
<tr>
<td>Someone who had an HPV vaccine cannot develop cervical cancer</td>
<td>0.65</td>
</tr>
<tr>
<td>HPV vaccines are most effective if given to people who have never had</td>
<td>0.57</td>
</tr>
<tr>
<td>sex</td>
<td></td>
</tr>
<tr>
<td>HPV vaccines protect against genital warts</td>
<td>0.23</td>
</tr>
<tr>
<td>HPV can be transmitted by genital skin-to-skin contact</td>
<td>0.58</td>
</tr>
<tr>
<td>When you have an HPV test, you get the results on the same day</td>
<td>0.53</td>
</tr>
<tr>
<td>Most sexually active people will get HPV at some point in their lives</td>
<td>0.47</td>
</tr>
<tr>
<td>There are many types of HPV</td>
<td>0.43</td>
</tr>
<tr>
<td>An HPV test can tell you how long you have had an HPV infection</td>
<td>0.38</td>
</tr>
<tr>
<td>HPV can be transmitted during sexual intercourse</td>
<td>0.40</td>
</tr>
<tr>
<td>Men cannot get HPV</td>
<td>0.33</td>
</tr>
<tr>
<td>Having many sexual partners increases the risk of contracting HPV</td>
<td>0.29</td>
</tr>
</tbody>
</table>

HPV: human papillomavirus; HIV: human immunodeficiency virus; AIDS: acquired immunodeficiency syndrome.
It is noteworthy that the Brazilian public health system does not use the HPV test as a screening tool, which limits the knowledge about the diagnosis even for health professionals. In addition, Brazil recently modified the vaccination schedule to two anti-HPV doses, while other countries maintained the three-dose vaccination schedule.

The low anti-HPV vaccination coverage observed in Brazil is mainly due to hesitation and resistance from many parents who are still deciding on the health care of their children. When considering the child’s age and the presumed time frame until sexual exposure, parents usually underestimate the child’s susceptibility to acquiring sexually transmitted infections or even to developing cancer in the future. Another relevant factor would be the connection between HPV and sexual activity. This issue has been a major source of religious controversy surrounding the HPV vaccine. Several religious groups and parents have expressed concern that the vaccine would be a trigger for promiscuity and early sexual life among adolescents. Furthermore, it has been argued that religious norms that regulate the sexual activity of unmarried women make the HPV vaccine unnecessary.

Once a valid instrument is available to gauge the general knowledge of the population about HPV and aspects related to vaccination, transmission, and diagnosis, it is possible to plan and develop more appropriate educational strategies. There are several differences between subgroups, as well as in sociodemographic status, including age, ethnicity, maternal schooling, healthcare coverage, and health providers’ recommendations about HPV. Indeed, one study demonstrated that young people were unaware that the HPV vaccine could be given to males. Authors have suggested improving the discussion with health care providers about this issue, as some social barriers still need to be overcome. One of the principal concerns regarding the HPV diagnosis is the knowledge generated by the information that follows the diagnosis per se.

The search for the partner responsible for the transmission and the need to determine when the infection has occurred are questions that remain unanswered. The emotional stress generated by the diagnosis and non-scientific sources of information can contribute to negative results in the progress of the disease or even in the possibility of transmission. Therefore, an effective instrument able to measure HPV knowledge is crucial to focus on preventive programs and health policies.

The existence of an instrument to determine the knowledge about HPV allows safer choices and more reliable strategies for the dissemination of this knowledge in a specific population. It is important to remember that teachers and administrators will be on the front lines of these programs, coordinated with health providers and parents. Also, there will always be anti-campaigns to defeat.

Among the limitations of this study are the predominantly female sample and the HPV-related problems they might have experienced that are not known, which theoretically have the potential to influence the results. Such a hypothesis would imply the need to incorporate a generic question about the issue, even if the interviewee were male, since he could have experienced the problem, personally or in his family. The greater the heterogeneity of the sample, the more representative of the target population it would be.
Nonetheless, further studies should include a discriminatory analysis between groups with and without a history of personal or family problems related to HPV infection.

Lastly, we underline that at the time of data collection, the vaccination schedule included three doses. Currently, the country has a two-dose regimen. Thus, we suggest changing question 2.2 of the final version of the instrument in Portuguese based on the current vaccination schedule.

**CONCLUSION**

We can conclude that the Brazilian Portuguese version under study presented a different behavior from the original measurement tool. A shorter version with 19 questions proved to be a reliable and valid instrument in assessing the Brazilian population’s knowledge about HPV. In this scenario, it is possible to infer that the chances of success of health promotion and disease prevention actions would improve.

**REFERENCES**

2. Human Papillomavirus (HPV) Information Center. Complementary data on cervical cancer prevention [Internet]. [accessed on Apr. 29, 2019]. Available at: www.hpvcentre.net


Received on: 01/23/2019
Revised on: 06/10/2019
Accepted on: 08/05/2019

Authors’ contribution: Fabiana Schuelter Trevisol: concept and design; data collection; data analysis; writing of the manuscript; approval of the final version. Rodrigo Dias Nunes: data analysis; writing of the manuscript; approval of the final version. Gabriel Oscar Cremona Parma: data analysis; writing of the manuscript; approval of the final version. André Luciano Manoel: concept and design; data collection; data analysis; writing of the manuscript; approval of the final version. Clávison M. Zapelini: concept and design; data collection; data analysis; writing of the manuscript; approval of the final version. Jefferson Traebert: data analysis; writing of the manuscript; approval of the final version.

© 2020 Associação Brasileira de Saúde Coletiva
This is an open access article distributed under the terms of the Creative Commons license.