3º ENCONTRO CONEXÃO MULHERES E ECONOMIA - CM&E

RESUMO EXPANDIDO

Área Temática: Economia da Saúde

CERVICAL CANCER IN THE BRAZILIAN STATES

Samanda[1]

Resumo: Cervical cancer is one of the most prevalent cancers in Brazil, with a highly heterogeneous distribution across its Federal Units. The aim of this study was to identify groupings by Federal Unit based on health variables that could represent the three pillars of actions proposed by the World Health Organization (WHO) for the elimination of cervical cancer using unsupervised machine learning techniques. After selecting seven health variables, Principal Component Analysis (PCA) was applied, allowing for the structural reduction of the data into two factors. The PCA outputs were used as inputs for hierarchical clustering analysis, resulting in the identification of four clusters. Cluster 1 (MG and PR) was the one closest to the target recommended by the WHO, with the highest percentages of vaccination coverage, diagnosis, and treatment. Conversely, Cluster 4 (AC, AM, AP, MA, and PA) represented the extreme opposite, with the poorest results in vaccination coverage and treatment. Cluster 3 (AL, CE, PB, PE, PI, and SE) showed the worst results in diagnosis variables. Clusters 3 and 4 suggest difficulties in accessing healthcare services correlated with poverty. Cluster 2 (BA, DF, ES, GO, MS, MT, RJ, RN, RO, RR, RS, SC, SP, and TO) differed from Cluster 3 due to its lower poverty index.

Palavras-chave: Principal component analysis; Cluster analysis; Uterine cervical neoplasms.

1 INTRODUÇÃO (OU APRESENTAÇÃO)

In Brazil, according to estimates from INCA (2022), approximately 17,000 new cases of cervical cancer are diagnosed annually, which represents about 15 cases per 100,000 women. However, there are significant differences between regions and Federal Units (UFs) of Brazil. The estimated age-adjusted incidence rates of cervical cancer vary from 12.9 in the Southeast region to 20.5 per 100,000 women in the Northern region (INCA, 2022).

Although it is one of the most common cancers among women in Brazil, cervical cancer, also known as cervical cancer, can be 100% prevented or controlled through primary prevention actions such as vaccination against the Human Papillomavirus (HPV), secondary prevention

[1] Samanda Silva da Rosa. Doutora em Economia do Desenvolvimento. Pontifícia Universidade Católica do Rio Grande do Sul. Samanda.srosa@gmail.com

(screening and treatment of pre-cancerous lesions), and tertiary prevention (early diagnosis and treatment of invasive cervical cancer) (INCA, 2016). Due to its importance in terms of public health, in August 2020, during the World Health Assembly, the World Health Organization (WHO) launched the global strategy to accelerate the elimination of cervical cancer (Seventy-third World Health Assembly, 2020). According to this strategy, all countries should achieve and maintain an incidence rate of fewer than four cases of cervical cancer per 100,000 women to achieve the elimination of cervical cancer (WHO, 2020). To make this goal possible, the 90-70-90 actions were proposed, resting on three pillars: a) vaccination: 90% of girls should be vaccinated with the HPV vaccine by the age of 15; b) screening: 70% of women should undergo a high-performance screening test by the age of 35, and again at the age of 45; c) treatment: 90% of women with pre-cancerous lesions and 90% of women with invasive cancer should receive appropriate treatment (WHO, 2020).

HPV infection is a well-established cause of cervical cancer (De Sanjose et al., 2018). Virtually 100% of cases are related to HPV infection. Among the more than 40 types of HPV that infect the genital area, types 16 and 18 are considered to have a high oncogenic risk and are responsible for about 70% of all cervical cancers (Bruni et al., 2021; De Sanjose et al., 2018). In terms of primary prevention, HPV vaccines that prevent infection by HPV types 16 and 18 have been available on the market since 2006 (WHO, 2017). In Brazil, the HPV vaccine was introduced in 2014 in the national immunization program for girls aged 9 to 14, representing a significant ally in reducing the incidence of cervical cancer in Brazil (SOARES, 2023).

2 MÉTODO (OU OPÇÕES METODOLÓGICAS)

Through a literature review, seven quantitative variables were selected to represent indicators related to the three pillars of actions in the fight against cervical cancer (Table 1). Subsequently, unsupervised machine learning techniques were employed, specifically Principal Component Analysis for structural reduction followed by hierarchical clustering (FÁVERO; FÁVERO, 2017).

Table 1. List of indicators selected as variables and the data baseline year.

Indicadores	Ano
HPV first-dose vaccine coverage at 15 years old	2021
Proportion of women aged 25 to 64 years who reported having had a cytological examination in the last 3 years	2019

Proportion of women aged 25 to 64 years who reported never having had a cervical cytological examination	2019
Proportion of cases with treatment initiation > 60 days after diagnosis	2022
Cervical cancer incidence adjusted per 100,000 women	2022
Cervical cancer mortality adjusted per 100,000 women	2022
Percentage of poverty	2021

Source: Original research data.

The Principal Component Analysis (PCA) is based on creating uncorrelated factors from the linear combination of the original variables. The correlation matrix P was defined using the variables from the database, as shown in Eq. (1), and the Pearson correlation coefficient was calculated between variable pairs as per Eq. (2) (FÁVERO; FÁVERO, 2017).

$$p = \begin{pmatrix} 1 & P_{12} & \cdots & P_{1k} \\ P_{21} & 1 & \cdots & P_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ P_{k1} & P_{k2} & \cdots & 1 \end{pmatrix} \quad (1)$$

$$p = \frac{\sum_{i=1}^{n}(x_{1i}-\bar{x}_1).(x_{2i}-\bar{x}_2)}{\sqrt{\sum_{i=1}^{n}(x_{1i}-\bar{x}_1)^2}.\sqrt{\sum_{i=1}^{n}(x_{2i}-\bar{x}_2)^2}} \quad (2)$$

For the clustering analysis, the outputs of PCA were used as inputs. A dissimilarity matrix was created using the factors generated in the Principal Component Analysis, employing the Euclidean distance measure between clusters and a hierarchical agglomeration scheme with complete linkage method. Since the clusters were formed based on the factors, there was no need to apply the Z-Score because the factors had already been standardized during the PCA.

A dendrogram, which is a graphical representation of hierarchical agglomeration, was created to determine the optimal number of clusters. ANOVA was performed to check if at least one cluster had a statistically different mean from the others (p-value < 0.05).

3 RESULTADOS E DISCUSSÃO

The values of the variables analyzed in this study by Federal Unit (UF) can be seen in Table 2.

Table 2. Variables analyzed in the study by federal unit

UF	CV1[(1)]	Papa3[(2)]	No Papa[(3)]	IncColo [(4)]	MortColo[(5)]	Poverty[(6)]	Tto60[(7)]
AC	43,1	82,1	6,7	15,4	10,9	45,5	72,4
AL	74,3	71,3	15,6	18,5	7,5	50,4	47,8
AP	64,7	76,1	13,3	26,7	16,9	46,8	67,4
AM	76,3	85,0	13,3	31,7	16,5	51,4	67,7
BA	65,9	81,5	4,8	11,8	5,0	47,3	57,7
CE	83,9	72,6	8,5	14,0	6,5	45,9	36,2
DF	77,7	79,6	8,4	11,1	6,2	15,7	44,9

ES	79,0	83,4	2,8	9,4	6,5	27,2	37,2
GO	72,9	75,8	8,0	9,1	5,8	22,5	50,7
MA	68,0	77,8	9,1	21,1	9,6	57,9	61,4
MT	83,1	79,7	6,3	11,1	6,3	20,2	47,1
MS	74,4	83,3	3,6	17,7	6,2	21,0	27,2
MG	76,6	79,7	6,7	7,7	3,3	25,3	41,4
PA	62,0	76,5	10,6	19,5	9,5	46,9	68,0
PB	81,1	66,5	15,8	10,5	6,5	47,2	62,6
PR	90,3	85,6	5,0	9,8	4,8	17,6	20,8
PE	74,1	77,1	9,0	12,1	5,7	50,3	39,8
PI	72,0	75,0	9,8	15,2	7,2	45,8	25,3
RJ	65,2	86,1	4,2	11,8	4,9	22,8	55,6
RN	64,6	78,4	7,6	12,1	5,5	42,9	41,5
RS	71,6	84,8	4,9	7,1	5,3	13,5	38,0
RO	67,5	80,7	5,1	16,4	6,0	31,7	44,9
RR	81,3	80,7	8,0	13,3	12,8	46,2	66,7
SC	83,9	85,8	3,0	17,2	4,7	10,2	40,7
SP	83,9	85,3	3,8	7,6	3,5	17,7	51,9
SE	74,7	74,0	9,1	13,9	6,7	48,2	37,5
TO	76,3	75,4	6,5	16,8	7,6	33,6	64,0

Source: Original research data.

Note: (1) First-dose vaccination coverage [CV1]; (2) Cervical cytology in the last three years [Papa3]; (3) Never had cervical cytology [NãoPapa]; (4) Cervical cancer incidence per 100,000 [IncColo]; (5) Cervical cancer mortality per 100,000 [MortColo]; (6) Percentage of poverty [Pobreza]; (7) Time between diagnosis and treatment exceeding 60 days [Tto60].

While the analysis of the variable "proportion of women aged 25 to 64 years who reported having had a cytological examination in the last three years," which represents the second pillar of the WHO's strategy for cervical cancer elimination, appears to meet the recommended goal of 70%, it is important to note that it was not possible to extract the coverage of examinations performed within the Brazilian Unified Health System (SUS) from the Cancer Information System (SISCAN) since it has not been fully implemented nationwide (SISCAN, 2022). Therefore, population surveys with samples exclusively from capitals and the Federal District were used, which may not be representative of the reality in the states. The rate of over 15% of women reporting never having undergone cytological screening in AL and PB is concerning.

The variable "proportion of cases with treatment initiation exceeding 60 days after diagnosis" was the only indicator that could be extracted from the Department of Informatics of the Unified Health System (DATASUS) to represent the third pillar of the WHO's strategy (cervical cancer treatment), with an average of 48.8% of women waiting more than 60 days to start treatment, ranging from 20.8% (PR) to 72.4% (AC).

4 CONCLUSÃO (OU CONSIDERAÇÕES FINAIS)

This set of unsupervised machine learning techniques allowed for the structural reduction of seven health variables related to cervical cancer into two factors and the identification of four groupings of Federal Units (UFs). The UFs' groupings based on health variables can serve as an important tool for decision-making in targeted actions and resource allocation for one or more of the pillars recommended for the elimination of cervical cancer, such as promoting health through intensified vaccination campaigns for eligible young individuals to receive the HPV vaccine, organized cervical cancer screening for women aged 25 to 64, and improving the healthcare network for cancer treatment.

REFERÊNCIAS

DE SANJOSE, Silvia; BROTONS, Maria; PAVON, Miguel Angel. The natural history of human papillomavirus infection. **Best practice & research Clinical obstetrics & gynaecology**, v. 47, p. 2-13, 2018.

FÁVERO, Luiz; FÁVERO, Patrícia. **Análise de dados: técnicas multivariadas exploratórias com SPSS e STATA**. Elsevier Brasil, 2017.

BRASIL. INSTITUTO NACIONAL DE CÂNCER JOSÉ ALENCAR GOMES DA SILVA. Coordenação de Prevenção e Vigilância. Divisão de Detecção Precoce e Apoio à Organização de Rede. **Diretrizes brasileiras para o rastreamento do câncer do colo do útero**, 2016. Disponível em:< https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/diretrizesparaorastreamentodocanc erdocolodoutero_2016_corrigido.pdf>. Acesso em: 01 outubro 2023.

INCA. Estimativa 2023: incidência do Câncer no Brasil. 2022. Disponível em: https://www.gov.br/inca/pt-br/assuntos/cancer/numeros/estimativa. Acesso em: 25 setembro 2023.

SOARES, Larissa; COSSE, Luisa; FERNANDES, Juliana. ASSISTÊNCIA DE ENFERMAGEM NA PREVENÇÃO DO PAPILOMAVÍRUS HUMANO EM ADOLESCENTES (ENFERMAGEM). **Repositório Institucional**, v. 1, n. 1, 2023.

Seventy-third world health assembly. Global strategy to accelerate the elimination of cervical cancer as a public health problem and its associated goals and targets for the period 2020–2030. 2020 Disponível em: < https://apps.who.int/gb/ebwha/pdf_files/WHA73/A73_R2-en.pdf>. Acesso em: 01 outubro 2023.

Sistema de Informação do Câncer [SISCAN]. 2022. Histo do colo Disponível em:< http://tabnet.datasus.gov.br/cgi/webtabx.exe?siscan/histo_residbr.def>. Acesso em: 06 outubro 2023.

WORLD HEALTH ORGANIZATION et al. Human papillomavirus vaccines: WHO position paper, May 2017–Recommendations. **Vaccine**, v. 35, n. 43, p. 5753-5755, 2017.

WORLD HEALTH ORGANIZATION et al. **Global strategy to accelerate the elimination of cervical cancer as a public health problem**. World Health Organization, 2020.