

# SELEÇÃO DE CRITÉRIOS DIAGNÓSTICOS PARA O MESOTELIOMA MALIGNO POR MEIO DA ANÁLISE DE COMPONENTES PRINCIPAIS

Ilias De Musis<sup>1</sup>  
Lenita de Souza Ferreira<sup>2</sup>  
Kuang Hongyu<sup>3</sup>

**RESUMO:** O mesotelioma maligno (MM) é um tipo de câncer agressivo que atinge o mesotélio, em principal a pleura. Sua principal causa é a exposição ao amianto, relacionando-se fortemente com atividades laborais. Diagnosticar o MM pode ser um desafio devido aos seus sintomas inespecíficos e semelhanças com outras doenças. No contexto de registros médicos, a análise multivariada pode ser utilizada para a otimização do diagnóstico. O presente estudo emprega a análise de componentes principais (ACP), utilizando registros de saúde de pacientes da Turquia. A aplicação da ACP visa o descarte de variáveis menos relevantes, simplificando o processo diagnóstico enquanto minimizando a perda de informação. Foram obtidos 17 componentes principais, dos quais 10 tiveram autovalores abaixo de 1. Isso sugere 10 variáveis para descarte, por apresentarem maiores coeficientes de ponderação com os componentes de menor variância. Com base nesses resultados, sugere-se o uso das seguintes variáveis em futuros estudos: duração da exposição ao amianto, duração dos sintomas, contagem de leucócitos pleural, albumina sérica, albumina pleural, glicose sanguínea e proteína C-reativa.

**Palavras-chave:** Amianto, análise multivariada, câncer, descarte de variáveis, pleura.

## SELECTION OF DIAGNOSTIC CRITERIA FOR MALIGNANT MESOTELIOMA THROUGH MAIN COMPONENT ANALYSIS

**ABSTRACT:** Malignant mesothelioma (MM) is an aggressive type of cancer that targets the mesothelium, primarily the pleura. Its main cause is asbestos exposure, strongly linked to occupational activities. Diagnosing MM can be challenging due to its non-specific symptoms and resemblance to other diseases. In the context of medical records, multivariate analysis may be used for diagnostic optimization. This study employs Principal Component Analysis (PCA), using hospital reports from patients in Turkey. The application of PCA aims to discard less relevant variables, simplifying the diagnostic process while minimizing information loss. 17 principal components were identified, of which 10 had eigenvalues below 1. This suggests 10 variables for disposal, as they have the highest weighting coefficients with the components of lower variance. Based on these results, the use of the following variables in future studies is recommended: duration of asbestos exposure, duration of symptoms, pleural white blood cell count, serum albumin, pleural albumin, blood glucose and C-reactive protein.

**Keywords:** Asbestos, Multivariate analysis, cancer, discarding variables, pleura.

---

<sup>1,2</sup> Graduandos em Estatística da Universidade Federal de Mato Grosso, Campus Universitário de Cuiabá – ilias.musis@sou.ufmt.br

<sup>3</sup> Prof. Dr. Adjunto do Departamento de Estatística. Universidade Federal de Mato Grosso, Av. Fernando Corrêa da Costa, nº 2367, Bairro Boa Esperança. CEP: 78060-900, Cuiabá, MT, Brasil - prof.kuang@gmail.com

## INTRODUÇÃO

O mesotelioma maligno (MM) é uma neoplasia altamente agressiva do mesotélio, tecido que reveste determinados órgãos, sendo a pleura detentora de 81% dos casos, seguida do peritônio (15%), pericárdio (4%) e, em menor frequência, da túnica vaginal (Brasil, 2020; INCA, 2012). Na pleura, apesar de ser uma doença multifatorial associada também a causas genéticas, sua principal causa está associada à exposição ao asbesto, mais conhecido como amianto. Esta relação faz do MM um tipo de câncer com forte associação ao trabalho, bem como à exposição ambiental e ao risco indireto por contato com materiais contaminados com a fibra. Por esses motivos, é mais frequente em homens com idade acima de 50 anos (Hung *et al.*, 2020). Tem um período de latência de 20 a 50 anos entre a exposição e o aparecimento da doença (Lacourt *et al.*, 2012; Petrof *et al.*, 2020).

Os sintomas são inespecíficos e comuns às doenças pleurais, como dispneia, dor torácica, tosse, perda de peso, astenia, fadiga e, eventualmente, abaulamento no tórax, dificultando a identificação, e o diagnóstico prescinde de dados da história e da exposição ocupacional ou ambiental. Também é frequentemente confundido com outros tipos de metástase pleurais e com câncer de pulmão (Brasil, 2020). Terra Filho, Freitas e Nery (2006) ressaltam a dificuldade da distinção entre o mesotelioma e o adenocarcinoma, sendo necessária a utilização de diversos métodos diagnósticos aprofundados, como a análise de marcadores imuno-histoquímicos.

O asbesto é uma fibra mineral composta de cálcio e magnésio, abundante na natureza e foi amplamente utilizado em inúmeras aplicações industriais. Seus usos variam desde a produção de telhas e caixas d'água até componentes para as indústrias bélica, aeroespacial e naval (Castro; Giannasi; Novello, 2003). O risco associado à exposição ao amianto é intensificado pela durabilidade de suas fibras, não facilmente decompostas por fagócitos. Uma vez inaladas, essas fibras podem se deslocar do pulmão para os linfonodos regionais, tecidos mesoteliais, pleurais e peritoneais, e outros órgãos, e são capazes de induzir tumores (Suzuki; Yuen, 2002).

A incidência do MM na população mundial é de um a dois casos por milhão de habitantes ao ano, com grande variação regional atrelada ao consumo de amianto. Contudo, sua mortalidade é alarmante, com mais de 80% dos afetados sendo levados a óbito no primeiro ano após o diagnóstico. A ausência de um diagnóstico preciso tende a acarretar uma progressão mais grave da doença (Hajj *et al.*, 2021; INCA, 2012). Regiões onde não existem medidas regulatórias apropriadas na exposição e uso do amianto concentram mais mortes por ano, como é o caso do Reino Unido, onde é estimado um número de 90.000 mortes relacionadas ao amianto no ano 2050. No Brasil, apesar do uso e comercialização desta substância ter sido restrito em 2017, é esperado que o pico de óbitos relacionados ao amianto no país se dê entre 2021 e 2026 (Barrera R; Chavarría G; Morales F, 2010; Miranda; Santos; Azeredo, 2020).

A exposição ambiental ao amianto é comum na região sudeste da Anatólia, Turquia e em certas partes da Grécia. Nessas regiões utilizam-se solos contendo amianto para fins de isolamento térmico e impermeabilização de telhados e paredes, especialmente como componente do reboco de paredes caídas (Metintas *et al.*, 2017). A exposição ambiental ocorre através da inalação de solo carregado de amianto e resulta em exposição contínua desde o nascimento para os habitantes das áreas rurais (Sezgi *et al.*, 2014). Neste cenário, foi disponibilizado um conjunto de dados de registros de saúde de 324 pacientes provenientes da região sudeste da Turquia com sintomas compatíveis com o MM para análise (Tanrikulu; Er, 2016).

Prontuários médicos contêm uma vasta gama de informações; enquanto algumas são imprescindíveis para a aceitação ou descarte de uma suspeita clínica, outras podem ser comuns em diversos agravos de saúde, dificultando o processo de diagnóstico. Ao abordar enfermidades específicas, como o mesotelioma maligno, a capacidade de identificar e priorizar informações médicas torna-se fundamental. Distinguir quais destes dados estão diretamente relacionados ao MM é crucial para diagnósticos precisos e, diante da grande quantidade de dados, pode ser uma tarefa laboriosa, exigindo abordagens estatísticas sofisticadas.

Chicco e Rovelli (2019) salientam que, em contraste com os sintomas, características clínicas oferecem dados quantitativos que facilitam o diagnóstico do mesotelioma pleural. Modelos existentes que projetam a sobrevivência do paciente utilizam características clínicas como tempo desde o diagnóstico, contagem de plaquetas, hemoglobina e estágio da doença.

A análise multivariada emerge, neste contexto, como uma abordagem de grande valor. Ao lidar com múltiplas variáveis simultaneamente, ela permite identificar padrões, correlações e fatores dominantes dentro de um conjunto de dados. Especificamente, a Análise de Componentes Principais (ACP) é uma técnica de redução de dimensionalidade que transforma um grande número de variáveis possivelmente correlacionadas em um número menor de componentes não correlacionados, chamados de componentes principais. Esses componentes representam combinações lineares das variáveis originais e são ordenados de forma que o primeiro componente capture a maior parte da variabilidade nos dados, o segundo componente capture a segunda maior parte, e assim por diante (Jolliffe; Cadima, 2016).

Com base nesse princípio, variáveis que possuem maior correlação com os componentes principais que menos representam a variabilidade dos dados tendem a ser menos importante para explicar a variância total. Assim, seu descarte pode ser considerado, otimizando a análise (Barbosa *et al.*, 2006). Paiva *et al.* (2010) defendem que, ao lidar com uma grande quantidade de variáveis, é vantajoso descartar aquelas que são redundantes ou que dificultam a mensuração para tornar as análises mais eficientes, levando a uma redução de custo e tempo das pesquisas.

Neste estudo, investiga-se a aplicação da análise de componentes principais para a redução da dimensionalidade do conjunto de dados de saúde relacionados ao mesotelioma na Turquia, conforme descrito por Abdullah Tanrikulu e Orhan Er (2016). Essa abordagem visa identificar e priorizar as variáveis mais relevantes para o diagnóstico do mesotelioma maligno, uma condição cujo reconhecimento precoce é crucial para o tratamento e prognóstico dos pacientes. Ao reduzir a dimensionalidade do conjunto de dados sem uma perda significativa de informação, a ACP pode otimizar processos de diagnóstico, tornando-os mais eficientes e precisos.

## MATERIAL E MÉTODOS

### Conjunto de dados

Para as análises executadas neste trabalho, foram utilizados dados hospitalares reais de pacientes da Faculdade de Medicina da Universidade de Dicle, na Turquia, disponibilizados por Tanrikulu e Er (2016). O banco de dados inclui registros de saúde de 324 pacientes, dos quais 96 foram diagnosticados com mesotelioma, 29,63% do total.

Informações sobre as variáveis do conjunto de dados foram extraídas das descrições originalmente propostas pelos curadores (Er *et al.*, 2012) e complementadas por detalhes apresentados em análise subsequente de Chicco e Rovelli (2019). Para os intuítos deste artigo, serão estudadas as variáveis descritas no Quadro 1.

**Quadro 1: Descrição das variáveis utilizadas.**

<b>Código</b>	<b>Denominação</b>	<b>Descrição</b>
AGE	Idade	Idade do paciente, em anos
DAE	Duração da exposição ao amianto	Tempo de exposição ambiental ao amianto, em anos
DS	Duração dos sintomas	Intervalo de tempo em que o paciente apresenta sintomas, em anos
WB	Contagem de leucócitos pleural	Quantidade de leucócitos no líquido pleural
WBC	Contagem de leucócitos	Quantidade e qualidade dos leucócitos no sangue
PLT	Contagem de plaquetas	Quantidade de plaquetas no sangue
SEDIM	Sedimentação	Rapidez com que os eritrócitos se assentam em um tubo de ensaio em uma hora.
LDH	Desidrogenase láctica sanguínea	Nível de enzima que auxilia na produção de energia no corpo.
ALP	Fosfatase alcalina	Nível de enzima associada à detecção doença hepática ou distúrbios ósseos.
TP	Proteína total	Quantidade total de proteína presente no soro.
ALB	Albumina sérica	Nível de albumina no sangue.
GLUC	Glicose sanguínea	Quantidade de glicose presente em uma amostra sanguínea.
PLD	Desidrogenase láctica pleural	Nível de enzima que indica se o líquido pleural é exudato ou transudato.
PP	Proteína pleural	Nível de proteína no líquido pleural, utilizado para classificar efusões como transudatas ou exudatas.
PA	Albumina pleural	Nível de albumina no líquido pleural.
PG	Glicose pleural	Nível de glicose no líquido pleural, sendo níveis baixos potencialmente indicativos de infecção ou malignidade.
CRP	Proteína C-reativa	Nível de proteína reagente de fase aguda, frequentemente elevado em pacientes com mesotelioma pleural.

**Fonte: Elaborado pelos autores com base em Chicco e Rovelli (2019).**

## Análise estatística

Os componentes principais são determinados por meio da diagonalização de matrizes simétricas positivas semi-definidas. A análise de componentes principais (ACP) é elogiada por sua capacidade de adaptação, sendo identificada como a técnica estatística multivariada de maior uso em praticamente todas as áreas científicas, conforme apontado por Abdi e Williams (2010). Sua grande versatilidade é associada à acessibilidade dos cálculos, com a disponibilidade de uma vasta gama de *softwares* capazes de realizar as operações matriciais necessárias (Hongyu; Sandanielo; Oliveira Junior, 2016).

Esta técnica apresenta algumas vantagens notáveis, sendo uma das mais relevantes a sua capacidade de lidar com o problema da multicolinearidade. Multicolinearidade refere-se à presença de correlações altas entre as variáveis independentes, tornando de difícil discernimento o efeito isolado de cada uma e as reais relações entre elas (Tabachnick; Fidell, 2013). Ao transformar o conjunto original de variáveis intercorrelacionadas em um novo conjunto de variáveis não correlacionadas, a ACP oferece uma solução amplamente empregada na análise de dados, conforme apontado por Jolliffe (2010).

Adicionalmente, a ACP é instrumental na redução da dimensionalidade dos dados. Por meio desta técnica, é possível sintetizar uma grande quantidade de variáveis em poucos componentes principais que descrevem a maior parte da variação dos dados. Estes componentes são ortogonais, ou seja, perpendiculares entre si, assegurando que cada um explique uma porção única e independente da variação total dos dados (Mardia; Kent; Bibby, 1979).

Os componentes principais podem ser derivados de matrizes de correlação ou covariâncias. Conforme Mingoti (2005), quando há grande variação nas variâncias, devido a diferentes unidades de medida, é apropriado usar a matriz de covariâncias das variáveis normalizadas. Esse método é equivalente à obtenção de componentes principais por meio da matriz de correlação (Jobson, 1992).

Dadas as variáveis  $X_1, X_2, \dots, X_p$  para  $n$  indivíduos ou unidades experimentais, forma-se uma matriz de dados  $\mathbf{X}(n \times p)$ :

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$$

Seja um conjunto de  $p$  variáveis  $X_1, X_2, \dots, X_p$  com médias  $\mu_1, \mu_2, \dots, \mu_p$  e variância  $\sigma_1^2, \sigma_2^2, \sigma_3^2, \dots, \sigma_p^2$ , respectivamente. Sendo a covariância entre a  $i$ -ésima e  $k$ -ésima variável definida por  $\sigma_{ik}$ , sua matriz de covariância  $\mathbf{\Sigma}$  é dada por:

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11}^2 & \cdots & \sigma_{1p}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{n1}^2 & \cdots & \sigma_{nn}^2 \end{bmatrix}$$

Obtendo-se os pares de autovalores e autovetores  $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$ , em que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ , associados a  $\mathbf{\Sigma}$ , o  $i$ -ésimo componente principal é definido por (Johnson; Wichern, 2007):

$$CP_i = e_i' \mathbf{X} = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p, \quad i=1, \dots, p.$$

Dessa forma, os componentes principais são deduzidos a partir das  $p$  variáveis contidas no vetor  $\mathbf{X}$ . A abordagem adotada para obtê-los envolve projetar as coordenadas das

variáveis originais em um plano, maximizando a distância entre elas, resultando na maximização da variabilidade de  $\mathbf{CP}_i$ . A variância de  $\mathbf{CP}_i$  pode ser descrita como:

$$\text{Var}(\mathbf{CP}_i) = \text{Var}(\mathbf{e}'_i \mathbf{X}) = \mathbf{e}'_i \text{Var}(\mathbf{X}) \mathbf{e}_i = \mathbf{e}'_i \boldsymbol{\Sigma} \mathbf{e}_i, \quad i = 1, \dots, p.$$

Ao aplicar a decomposição espectral da matriz  $\boldsymbol{\Sigma}$ , dada por  $\boldsymbol{\Sigma} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'$ , em que  $\mathbf{P}$  é composta pelos autovetores de  $\boldsymbol{\Sigma}$  em suas colunas e  $\boldsymbol{\Lambda}$  a matriz diagonal de autovalores de  $\boldsymbol{\Sigma}$ , temos (Hongyu; Sandanielo; Oliveira Junior, 2016; Jobson, 1992):

$$\text{tr}(\boldsymbol{\Sigma}) = \text{tr}(\mathbf{P}\boldsymbol{\Lambda}\mathbf{P}') = \text{tr}(\boldsymbol{\Lambda}\mathbf{P}'\mathbf{P}) = \text{tr}(\boldsymbol{\Lambda}\mathbf{I}) = \text{tr}(\boldsymbol{\Lambda}) = \sum_{i=1}^p \lambda_i$$

e

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{bmatrix}.$$

Além disso,  $\text{tr}(\boldsymbol{\Sigma})$  é definida pela soma dos elementos da diagonal, de forma que:

$$\text{tr}(\boldsymbol{\Sigma}) = \sum_{i=1}^p \sigma_{ii} = \sum_{i=1}^p \lambda_i$$

Temos, então, que a variabilidade total contida nas variáveis originais é equivalente à variabilidade total contida nos componentes principais. Cada componente do vetor  $\mathbf{e}_i$  também deve ser examinado. A magnitude de cada elemento indica a relevância da variável para o  $i$ -ésimo componente principal (Johnson; Wichern, 2007).

A influência de cada componente principal é frequentemente expressa em percentuais. O nível de contribuição de cada componente, representado pela proporção da variância total explicada por esse, pode ser calculado da seguinte forma (Jolliffe, 2010):

$$C_k = \frac{\text{Var}(\mathbf{Z}_i)}{\sum_{i=1}^p \text{Var}(\mathbf{Z}_i)} \cdot 100 = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \cdot 100 = \frac{\lambda_i}{\text{tr}(\boldsymbol{\Sigma})} \cdot 100$$

Esse valor é levado em consideração para a seleção de componentes principais a serem mantidos. No geral, é recomendado que a variância acumulada explicada pelos componentes escolhidos esteja entre 70% e 90% (Johnson; Wichern, 2007; Jolliffe, 2010).

Uma estratégia frequentemente adotada para a retenção de fatores é critério de Kaiser (1960), que determina que os componentes principais para os quais o autovalor  $> 1$ , explicam a maioria da variação dos dados originais e, por isso, devem ser mantidos. Este critério é fundamentado na ideia que um componente que não oferece uma explicação de variância superior à de uma única variável original não é eficaz na redução de dimensionalidade.

Um complemento visual a essa abordagem é o *scree plot*. Introduzido e nomeado por Cattell (1966), esse gráfico ilustra a magnitude dos autovalores em relação ao número de componentes. Ao analisá-lo, busca-se o ponto em que a trajetória da curva tende a se nivelar, indicando quantos componentes reter. É considerado um método mais subjetivo de seleção de componentes e deve ser complementado por outros critérios (Jolliffe, 2010; Neisse; Hongyu, 2016).

De acordo com Regazzi e Cruz (2020), no contexto de descarte de variáveis, aquelas com coeficientes de maior magnitude no componente principal de menor autovalor são consideradas de menor importância e passíveis de descarte em estudos futuros. Este processo

de descarte se alinha ao método B2 descrito por Jolliffe (1972). Nesse, após realizar a análise de componentes principais e decidir pela retenção de  $n$  variáveis, uma variável é associada a cada um dos últimos componentes principais ( $p - n$ ), considerados menos importantes. Para cada componente considerado, é escolhida a variável com o coeficiente mais elevado, contanto que essa não tenha sido previamente associada a outro componente. Em seguida, as variáveis escolhidas são descartadas.

Todas as análises deste estudo foram conduzidas por meio do *software* estatístico R (R Core Team, 2023), com o auxílio do ambiente *RStudio* (Posit team, 2023). Foram utilizados recursos dos pacotes *FactoMineR* (Lê; Josse; Husson, 2008) e *Factoextra* (Kassambara; Mundt, 2020).

## RESULTADOS E DISCUSSÕES

Com base nos resultados obtidos pela técnica dos componentes principais, os respectivos autovalores e porcentagens da variância explicada por cada um estão apresentados na Tabela 1. A análise de componentes principais gerou 17 componentes principais, dos quais os 7 primeiros foram responsáveis por explicar 62,27% da variância dos dados.

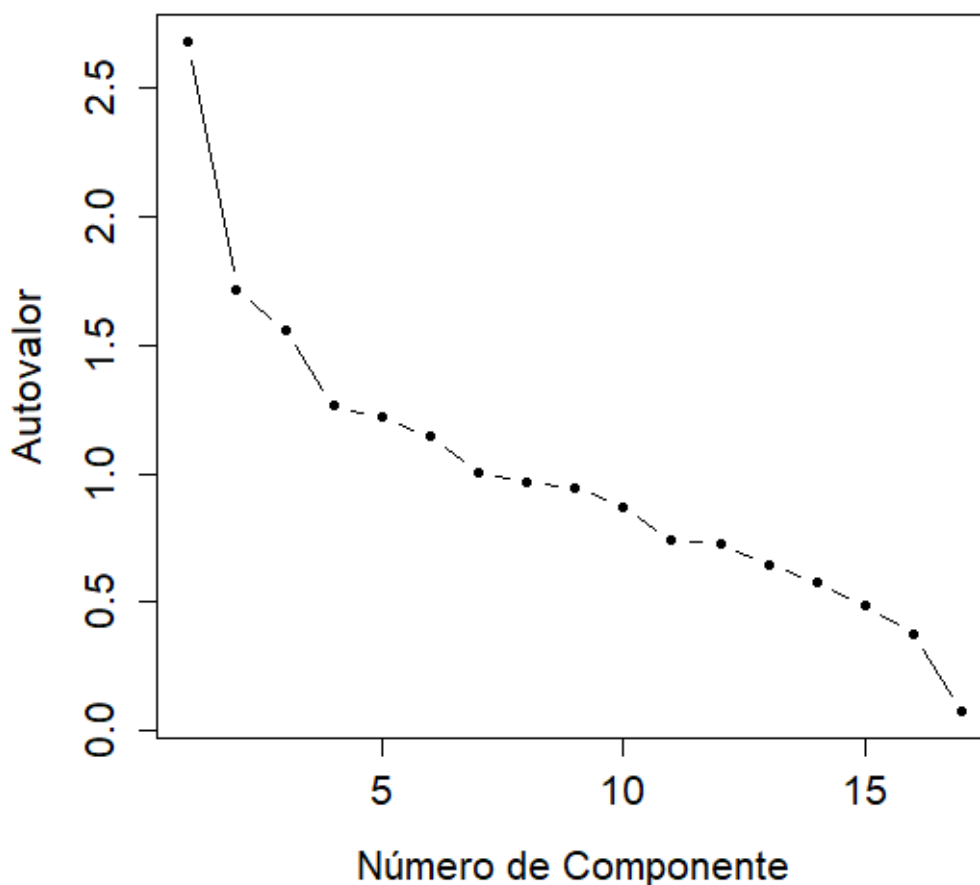
Leite *et al.* (2009), ao analisarem características de desempenho e de carcaça de codornas utilizando a técnica de componentes principais, verificaram que quatro componentes principais foram suficientes para explicar 75% da variação total dos dados. Em estudo de Meira *et al.* (2013), referente a seleção de características de cavalos, foram utilizados seis componentes principais, explicando 78,57% da variância dos dados. Ao estudar propriedades do solo e plantas na agricultura de precisão, Ferraz *et al.* (2019) observaram que os primeiros 7 componentes principais explicavam 88,98% da variância total.

**Tabela 1: Componentes principais (CP), autovalores e porcentagem da variância explicada e proporção acumulada (%) pelos componentes.**

Componente Principal	Autovalor*	% da Variância	% Acumulada da Variância
CP1	<b>2,68</b>	15,76	15,76
CP2	<b>1,72</b>	10,09	25,85
CP3	<b>1,56</b>	9,17	35,02
CP4	<b>1,26</b>	7,43	42,45
CP5	<b>1,22</b>	7,17	49,62
CP6	<b>1,14</b>	6,73	56,35
CP7	<b>1,01</b>	5,92	62,27
CP8	0,97	5,68	67,94
CP9	0,95	5,57	73,52
CP10	0,87	5,11	78,63
CP11	0,74	4,36	82,99
CP12	0,73	4,28	87,27
CP13	0,64	3,78	91,05
CP14	0,58	3,40	94,44
CP15	0,49	2,89	97,33
CP16	0,38	2,21	99,54
CP17	0,08	0,46	100,00

\* Valores em negrito indicam autovalores superiores a 1.

Figura 1: Scree plot dos autovalores dos componentes principais.



Conforme observado na Tabela 1, dos 17 componentes principais encontrados, 10 (58,82%) apresentaram autovalores  $< 1$ , sendo então passíveis de descarte de acordo com o critério de Kaiser (1960). O *scree plot*, ilustrado na **Erro! Fonte de referência não encontrada.**, corrobora com essa interpretação, sugerindo que a informação relevante foi capturada pelos componentes retidos.

Conforme a metodologia descrita por Regazzi e Cruz (2020), as variáveis mais fortemente associadas com cada um dos CPs não significativos (CP8 a CP17) são identificadas para descarte. Este princípio está alinhado com a ideia de que, considerando um componente com pequeno autovalor como de menor relevância, pode-se inferir que a variável que o domina tem pouca importância para a variância total dos dados (Mardia; Kent; Bibby, 1979).

As 11 variáveis com os maiores coeficientes de ponderação a partir do último componente principal correspondem aos valores destacados em negrito na Tabela 2. Portanto, as variáveis sugeridas para descarte são, em ordem de menor importância para explicação da variação total: Proteína pleural (PP), Glicose pleural (PG), Proteína total (TP), Fosfatase alcalina (ALP), Desidrogenase lática sanguínea (LDH), Sedimentação (SEDIM), Contagem de plaquetas (PLT), Contagem de leucócitos (WBC) e Idade (AGE).



**Tabela 2: Coeficientes de ponderação das variáveis com os 10 últimos componentes principais.**

Variáveis	CP8	CP9	CP10	CP11	CP12	CP13	CP14	CP15	CP16	CP17
AGE	0,00	0,09	0,20	-0,02	0,11	<b>-0,54</b>	0,37	-0,08	-0,12	0,03
DAE	0,00	0,22	-0,29	-0,03	-0,07	0,50	0,39	0,01	-0,03	-0,01
DS	0,07	0,21	-0,08	0,08	0,43	0,28	0,25	0,03	-0,10	-0,02
WB	0,44	0,40	0,44	0,21	-0,29	0,08	0,06	0,16	-0,08	-0,03
WBC	<b>0,66</b>	-0,33	-0,34	-0,07	-0,26	-0,18	0,01	-0,03	-0,02	-0,04
PLT	-0,08	0,02	-0,01	<b>-0,78</b>	-0,04	0,03	0,04	0,35	-0,02	0,04
SEDIM	0,45	0,24	-0,27	-0,12	<b>0,58</b>	-0,10	0,02	-0,09	-0,03	0,01
LDH	-0,19	0,15	-0,08	0,14	0,00	-0,35	<b>0,53</b>	-0,29	-0,34	-0,01
ALP	-0,14	0,01	<b>-0,58</b>	0,28	-0,15	-0,05	0,25	-0,08	0,25	-0,01
TP	0,00	0,28	0,13	-0,21	0,08	0,10	0,10	<b>-0,60</b>	0,18	0,03
ALB	-0,02	0,23	-0,14	0,27	0,12	-0,14	0,02	0,56	-0,21	-0,06
GLUC	-0,18	-0,39	0,09	0,10	0,42	0,00	0,11	0,15	0,01	-0,04
PLD	0,19	<b>-0,45</b>	0,23	0,18	0,15	0,37	0,09	-0,15	-0,29	-0,02
PP	0,00	-0,13	0,02	-0,07	-0,07	-0,02	0,04	-0,10	-0,03	<b>-0,71</b>
PA	0,05	-0,16	0,00	0,02	-0,06	-0,01	0,01	-0,03	-0,07	0,70
PG	-0,06	-0,03	-0,18	-0,21	-0,09	0,02	0,05	-0,12	<b>-0,74</b>	-0,02
CRP	-0,16	0,15	-0,10	0,10	-0,21	0,21	0,52	-0,06	-0,28	0,03

Por outro lado, para estudos subsequentes, recomenda-se a consideração das seguintes variáveis: Duração da exposição ao amianto (DAE), Duração dos sintomas (DS), Contagem de leucócitos pleural (WB), Albumina sérica (ALB), Albumina pleural (PA), Glicose sanguínea (GLUC) e Proteína C-reativa (CRP).

A seleção de variáveis determinada pela análise implica em uma forte associação entre a duração da exposição ao amianto e o desenvolvimento do mesotelioma maligno. Notavelmente, dado o uso de informações de pacientes provenientes da região sudeste da Turquia, onde o amianto é prevalente por motivos naturais e industriais, é esperado que a duração da exposição seja maior que a média encontrada em outros contextos geográficos (Bayram *et al.*, 2019; Metintas *et al.*, 2017). A especificidade do cenário turco oferece detalhes sobre essa variável e corrobora descobertas similares em pesquisas de escopo global. A associação entre o tempo de exposição ao amianto e o mesotelioma maligno é indicada em diversas pesquisas (Lacourt *et al.*, 2012; Loomis; Richardson; Elliott, 2019; Pass *et al.*, 2005).

Em estudo conduzido por Chicco e Rovelli (2019), as variáveis de contagem de plaquetas, proteína pleural e idade foram identificadas como sendo de significativa relevância na predição do mesotelioma. No presente estudo, tais variáveis foram consideradas redundantes na análise de componentes principais e, por consequência, descartadas. Apesar dessas discrepâncias, ambos os estudos concordaram quanto à relevância de diversas variáveis, destacando-se particularmente a duração da exposição ao amianto, duração dos sintomas, proteína C-reativa e albumina pleural. Essa concordância reforça a relevância desses indicadores na investigação da doença.

Em revisão sistemática da literatura de 3484 artigos únicos sobre a doença, Santos *et al.* (2022) selecionaram dezessete estudos que preenchiam os critérios de inclusão adotados. Em oito, datados de 1969 a 2020, foi possível determinar o período de intervalo da primeira exposição. O tempo de exposição foi mais prolongado nos estudos que consideraram dados ocupacionais e não ocupacionais combinados, totalizando 28 anos, enquanto nos estudos puramente ocupacionais e para-ocupacionais, a média de exposição foi de 13 e 12 anos,

respectivamente. Essas informações indicam a relevância tanto da duração quanto a compreensão dos tipos de exposição para o estudo dos fatores relacionados ao mesotelioma.

Embora a duração dos sintomas também tenha revelado associação com a doença, os sintomas inicialmente surgem de forma gradual e inespecífica, o que frequentemente leva a erros no diagnóstico. O período entre o início dos sintomas e o diagnóstico do mesotelioma pleural pode se estender por anos (Zhang *et al.*, 2015).

No acompanhamento de 2 pacientes com mesotelioma pleural maligno em estágio 4, Baran (2016) fornece uma descrição da evolução temporal desde os primeiros sintomas da doença, diagnóstico, até o óbito. O período observado compreendeu 1 ano e 2 meses para o primeiro paciente e 8 anos para o segundo.

Apesar da relevância da variável na análise, existe uma lacuna na literatura sobre a relação entre a contagem de plaquetas no líquido pleural e o diagnóstico do mesotelioma em específico. Em contraste, a contagem global de leucócitos elevada observada no líquido pleural é um achado frequente em derrames pleurais exsudativos resultantes de efusões malignas da pleura (Arora; Boster, 2023; Cunha, 1983).

As dosagens bioquímicas são exames complementares e por si só não são decisivas no diagnóstico de mesotelioma maligno, mas auxiliam na caracterização da amostra. A razão de 0,5 entre o teor de glicose sanguínea e o teor de glicose pleural, por exemplo, é encontrada em efusões pleurais exsudativas. Da mesma forma, a diferença entre a albumina sérica e a albumina no líquido pleural inferior a  $0,6 \pm 0,4$  g/dl (Comar *et al.*, 2008), apesar da presença de hipoalbuminemia no líquido pleural, é um achado frequente no MPM (Carreiro *et al.*, 1986), o que corrobora os resultados do estudo.

Diversos artigos mencionam a elevação dos níveis de proteína C-reativa em estudos de relato de caso de MPM (Ghanim *et al.*, 2012). A combinação de proteína C-reativa em níveis elevados e albumina encontrada em resposta inflamatória está associada à sobrevida de pacientes e prenuncia o pior prognóstico (Hajj *et al.*, 2021; Takamori *et al.*, 2018).

## CONCLUSÃO

Com base nos resultados obtidos, a análise de componentes principais se mostrou efetiva para a seleção de variáveis. Do total de 17 componentes principais identificados, 7 foram considerados de relevância significativa, possibilitando a exclusão de 10 variáveis. Dessa forma, 58,82% das variáveis foram consideradas redundantes ou apresentaram por baixa variabilidade.

Recomenda-se a inclusão das seguintes variáveis para pesquisas futuras relacionadas ao diagnóstico do mesotelioma maligno: duração da exposição ao amianto, duração dos sintomas, contagem de leucócitos pleural, albumina sérica, albumina pleural, glicose sanguínea e proteína C-reativa.

## REFERÊNCIAS BIBLIOGRÁFICAS

ABDI, H.; WILLIAMS, L. J. Principal component analysis. **Wiley Interdisciplinary Reviews: Computational Statistics**, [s. l.], v. 2, n. 4, p. 433–459, 2010. DOI 10.1002/wics.101. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/wics.101>. Acesso em: 2 jun. 2023.

ARORA, R. D.; BOSTER, J. **Malignant Pleural Effusion**. Treasure Island, FL: StatPearls Publishing, 2023. *E-book*. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK574541/>. Acesso em: 19 out. 2023.

BARAN, F. D. P. **O itinerário terapêutico revelado pelo familiar da pessoa com mesotelioma maligno: estudo de casos múltiplos**. 2016. Dissertação (Mestrado em Enfermagem) - Universidade Federal do Paraná, Curitiba, 2016. Disponível em: <https://hdl.handle.net/1884/45767>. Acesso em: 17 out. 2023.

BARBOSA, L. *et al.* Avaliação de características de qualidade da carne de suínos por meio de componentes principais. **Revista Brasileira de Zootecnia**, [s. l.], v. 35, n. 4, p. 1639–1645, 2006. DOI 10.1590/S1516-35982006000600009. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1516-35982006000600009&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-35982006000600009&lng=pt&tlng=pt). Acesso em: 21 set. 2023.

BARRERA R, R.; CHAVARRÍA G, J.; MORALES F, J. Mesotelioma maligno: Experiencia clínico-patológica de 247 casos. **Revista chilena de enfermedades respiratorias**, Santiago, v. 26, n. 3, 2010. DOI 10.4067/S0717-73482010000300003. Disponível em: [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0717-73482010000300003&lng=en&nrm=iso&tlng=en](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-73482010000300003&lng=en&nrm=iso&tlng=en). Acesso em: 23 set. 2023.

BARRERA R, R.; CHAVARRÍA G, J.; MORALES F, J. Mesotelioma maligno: Experiencia clínico-patológica de 247 casos. **Revista chilena de enfermedades respiratorias**, Santiago, v. 26, n. 3, 2010. DOI 10.4067/S0717-73482010000300003. Disponível em: [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0717-73482010000300003&lng=en&nrm=iso&tlng=en](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-73482010000300003&lng=en&nrm=iso&tlng=en). Acesso em: 23 set. 2023.

BAYRAM, M. *et al.* Asbestos-related Diseases in Turkey: Caused Not Only by Naturally Occurring Fibers but Also by Industrial Exposures. **American Journal of Respiratory and Critical Care Medicine**, New York, v. 199, n. 5, p. 656–659, 2019. DOI 10.1164/rccm.201810-1922LE. Disponível em: <https://www.atsjournals.org/doi/10.1164/rccm.201810-1922LE>. Acesso em: 19 out. 2023.

BRASIL. Ministério da Saúde. Secretaria de Ciência, Tecnologia, Inovação e Insumos Estratégicos em Saúde. **Diretrizes brasileiras para o diagnóstico do mesotelioma maligno de pleura**. Brasília, DF: Ministério da Saúde, 2020. Relatório de recomendação. Disponível em: [https://sbpt.org.br/portal/wp-content/uploads/2020/12/Diretrizes-Brasileiras\\_Diagnostico\\_MMP.pdf](https://sbpt.org.br/portal/wp-content/uploads/2020/12/Diretrizes-Brasileiras_Diagnostico_MMP.pdf). Acesso em: 19 out. 2023.

CARREIRO, L. *et al.* Mesoteliomas pleurais malignos. **Acta Médica Portuguesa**, Lisboa, v. 7, n. 2, p. 87–92, 1986.

CASTRO, H.; GIANNASI, F.; NOVELLO, C. A luta pelo banimento do amianto nas Américas: uma questão de saúde pública. **Ciência & Saúde Coletiva**, [s. l.], v. 8, n. 4, p. 903–911, 2003.

DOI 10.1590/S1413-81232003000400013. Disponível em:  
[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-81232003000400013&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232003000400013&lng=pt&tlng=pt). Acesso em: 23 set. 2023.

CATTELL, R. B. The Scree Test For The Number Of Factors. **Multivariate Behavioral Research**, [s. l.], v. 1, n. 2, p. 245–276, 1966. DOI 10.1207/s15327906mbr0102\_10. Disponível em: [http://www.tandfonline.com/doi/abs/10.1207/s15327906mbr0102\\_10](http://www.tandfonline.com/doi/abs/10.1207/s15327906mbr0102_10). Acesso em: 18 out. 2023.

CHICCO, D.; ROVELLI, C. Computational prediction of diagnosis and feature selection on mesothelioma patient health records. **PLOS ONE**, [s. l.], v. 14, n. 1, 2019. DOI 10.1371/journal.pone.0208737. Disponível em: <https://dx.plos.org/10.1371/journal.pone.0208737>. Acesso em: 24 set. 2023.

COMAR, S. R. *et al.* ANÁLISE CITOLÓGICA DO LÍQUIDO PLEURAL NO HOSPITAL DAS CLÍNICAS DA UNIVERSIDADE FEDERAL DO PARANÁ (UFPR). **Estudos de Biologia**, Curitiba, v. 30, n. 70/72, 2008. DOI 10.7213/reb.v30i70/72.22799. Disponível em: <https://periodicos.pucpr.br/index.php/estudosdebiologia/article/view/22799>. Acesso em: 17 out. 2023.

CUNHA, A. L. **Derrame pleural**: análise laboratorial e histopatológica de 58 casos num hospital geral. 1983. Trabalho de Conclusão de Curso (Bacharelado em Medicina) - Centro de Ciências da Saúde, Universidade Federal de Santa Catarina, Florianópolis, 1983. Disponível em: <https://repositorio.ufsc.br/handle/123456789/117465>. Acesso em: 19 out. 2023.

ER, O. *et al.* An approach based on probabilistic neural network for diagnosis of Mesothelioma's disease. **Computers & Electrical Engineering**, [s. l.], v. 38, n. 1, p. 75–81, 2012. DOI 10.1016/j.compeleceng.2011.09.001. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0045790611001261>. Acesso em: 6 out. 2023.

FERRAZ, G. A. S. *et al.* Principal components in the study of soil and plant properties in precision coffee farming. **Agronomy Research**, [s. l.], v. 17, n. 2, p. 418–429, 2019. DOI 10.1515/AR.19.114. Disponível em: <https://dspace.emu.ee/xmlui/handle/10492/4828>. Acesso em: 19 out. 2023.

GHANIM, B. *et al.* Pretreatment Serum C-Reactive Protein Levels Predict Benefit From Multimodality Treatment Including Radical Surgery in Malignant Pleural Mesothelioma: A Retrospective Multicenter Analysis. **Annals of Surgery**, [s. l.], v. 256, n. 2, p. 357–362, 2012. DOI 10.1097/SLA.0b013e3182602af4. Disponível em: <https://journals.lww.com/00000658-201208000-00025>. Acesso em: 19 out. 2023.

HAIJ, G. N. M. *et al.* Malignant pleural mesothelioma: an update. **Jornal Brasileiro de Pneumologia**, [s. l.], 2021. DOI 10.36416/1806-3756/e20210129. Disponível em: <http://www.jornaldepneumologia.com.br/details/3621/en-US/malignant-pleural-mesothelioma--an-update>. Acesso em: 24 set. 2023.

HONGYU, K.; SANDANIELO, V. L. M.; OLIVEIRA JUNIOR, G. J. Análise de Componentes Principais: Resumo Teórico, Aplicação e Interpretação. **E&S Engineering and Science**, Mato Grosso, v. 5, n. 1, p. 83–90, 2016. DOI 10.18607/ES201653398. Disponível em:

<https://periodicoscientificos.ufmt.br/ojs/index.php/eng/article/view/3398>. Acesso em: 2 jun. 2023.

HUNG, Y. P. *et al.* Molecular characterization of diffuse malignant peritoneal mesothelioma. **Modern Pathology**, [s. l.], v. 33, n. 11, p. 2269–2279, 2020. DOI 10.1038/s41379-020-0588-y. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0893395222004483>. Acesso em: 24 set. 2023.

INSTITUTO NACIONAL DE CÂNCER José de Alencar Gomes da Silva. **Diretrizes para a vigilância do câncer relacionado ao trabalho**. Rio de Janeiro, RJ: INCA, 2012. *E-book*. Disponível em: <https://www.inca.gov.br/publicacoes/livros/diretrizes-para-vigilancia-do-cancer-relacionado-ao-trabalho>. Acesso em: 24 set. 2023.

JOBSON, J. D. **Volume 2: Categorical and Multivariate Methods**. Ned. New York: Springer New York, 1992.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 6. ed. Upper Saddle River, N.J: Pearson Prentice Hall, 2007.

JOLLIFFE, I. T. Discarding Variables in a Principal Component Analysis. I: Artificial Data. **Applied Statistics**, [s. l.], v. 21, n. 2, p. 160, 1972. DOI 10.2307/2346488. Disponível em: <https://www.jstor.org/stable/10.2307/2346488?origin=crossref>. Acesso em: 17 out. 2023.

JOLLIFFE, I. T. **Principal component analysis**. 2. ed. New York: Springer, 2010. (Springer series in statistics).

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, [s. l.], v. 374, n. 2065, 2016. DOI 10.1098/rsta.2015.0202. Disponível em: <https://royalsocietypublishing.org/doi/10.1098/rsta.2015.0202>. Acesso em: 7 jun. 2023.

KAISER, H. F. The Application of Electronic Computers to Factor Analysis. **Educational and Psychological Measurement**, [s. l.], v. 20, n. 1, p. 141–151, 1960. DOI 10.1177/001316446002000116. Disponível em: <http://journals.sagepub.com/doi/10.1177/001316446002000116>. Acesso em: 6 out. 2023.

KASSAMBARA, A.; MUNDT, F. **Factoextra**: Extract and Visualize the Results of Multivariate Data Analyses. [S. l.], 2020. Disponível em: <https://CRAN.R-project.org/package=factoextra>. Acesso em: 2 jun. 2023.

LACOURT, A. *et al.* Temporal patterns of occupational asbestos exposure and risk of pleural mesothelioma. **European Respiratory Journal**, [s. l.], v. 39, n. 6, p. 1304–1312, 2012. DOI 10.1183/09031936.00005111. Disponível em: <http://erj.ersjournals.com/lookup/doi/10.1183/09031936.00005111>. Acesso em: 23 set. 2023.

LÊ, S.; JOSSE, J.; HUSSON, F. FactoMineR: An R Package for Multivariate Analysis. **Journal of Statistical Software**, [s. l.], v. 25, n. 1, 2008. DOI 10.18637/jss.v025.i01. Disponível em: <http://www.jstatsoft.org/v25/i01/>. Acesso em: 2 jun. 2023.

LEITE, C. D. S. *et al.* Avaliação de características de desempenho e de carcaça de codornas de corte por meio da análise de componentes principais. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, [s. l.], v. 61, n. 2, p. 498–503, 2009. DOI 10.1590/S0102-09352009000200030. Disponível em:

[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-09352009000200030&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-09352009000200030&lng=pt&tlng=pt). Acesso em: 19 out. 2023.

LOOMIS, D.; RICHARDSON, D. B.; ELLIOTT, L. Quantitative relationships of exposure to chrysotile asbestos and mesothelioma mortality. **American Journal of Industrial Medicine**, [s. l.], v. 62, n. 6, p. 471–477, 2019. DOI 10.1002/ajim.22985. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/ajim.22985>. Acesso em: 19 out. 2023.

MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. **Multivariate analysis**. 1. ed. London: Academic Press, 1979. (Probability and mathematical statistics).

MEIRA, C. T. *et al.* Seleção de características morfofuncionais de cavalos da raça Mangalarga Marchador por meio da análise de componentes principais. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, [s. l.], v. 65, n. 6, p. 1843–1848, 2013. DOI 10.1590/S0102-09352013000600036. Disponível em:

[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-09352013000600036&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-09352013000600036&lng=pt&tlng=pt). Acesso em: 17 out. 2023.

METINTAS, S. *et al.* Asbestos knowledge and awareness level in central part of Anatolia. **International Journal of Occupational and Environmental Health**, [s. l.], v. 23, n. 3, p. 243–249, 2017. DOI 10.1080/10773525.2018.1470789. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/10773525.2018.1470789>. Acesso em: 23 set. 2023.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2005.

NEISSE, A. C.; HONGYU, K. Aplicação de componentes principais e análise fatorial a dados criminais de 26 estados dos EUA. **E&S Engineering and Science**, [s. l.], v. 5, n. 2, p. 105–115, 2016. DOI 10.18607/ES201654354. Disponível em: <https://periodicoscientificos.ufmt.br/ojs/index.php/eng/article/view/4354>. Acesso em: 19 out. 2023.

PAIVA, A. L. D. C. *et al.* Análise de componentes principais em características de produção de aves de postura. **Revista Brasileira de Zootecnia**, [s. l.], v. 39, n. 2, p. 285–288, 2010. DOI 10.1590/S1516-35982010000200009. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1516-35982010000200009&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-35982010000200009&lng=pt&tlng=pt). Acesso em: 21 set. 2023.

PASS, H. I. *et al.* Asbestos Exposure, Pleural Mesothelioma, and Serum Osteopontin Levels. **New England Journal of Medicine**, [s. l.], v. 353, n. 15, p. 1564–1573, 2005. DOI 10.1056/NEJMoa051185. Disponível em: <http://www.nejm.org/doi/abs/10.1056/NEJMoa051185>. Acesso em: 19 out. 2023.

PETROF, O. *et al.* On the impact of residential history in the spatial analysis of diseases with a long latency period: A study of mesothelioma in Belgium. **Statistics in Medicine**, [s. l.], v. 39, n. 26, p. 3840–3866, 2020. DOI 10.1002/sim.8697. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/sim.8697>. Acesso em: 24 set. 2023.

Posit team. **RStudio**: Integrated Development Environment for R. Versão 2023.06.0. Boston, MA: Posit Software, PBC, 2023. Disponível em: <http://www.posit.co/>. Acesso em: 19 out. 2023.

R Core Team. **R**: A Language and Environment for Statistical Computing. Versão 4.3.0. Vienna, Austria: R Foundation for Statistical Computing, 2023. Disponível em: <https://www.R-project.org/>. Acesso em: 19 out. 2023.

REGAZZI, A. J.; CRUZ, C. D. **Análise multivariada aplicada**. Viçosa: Universidade Federal de Viçosa, 2020. v. 2 *E-book*. Disponível em: [ftp://ftp.ufv.br/dbg/biodata/An%20E1lise%20Multivariada%20Aplicada%202020%20Adair\\_Co sme.pdf](ftp://ftp.ufv.br/dbg/biodata/An%20E1lise%20Multivariada%20Aplicada%202020%20Adair_Co sme.pdf). Acesso em: 18 out. 2023.

SANTOS, C. *et al.* Asbestos Exposure and Malignant Pleural Mesothelioma: A Systematic Review of Literature. **Portuguese Journal of Public Health**, [Portugal], v. 40, n. 3, p. 188–202, 2022. DOI 10.1159/000527971. Disponível em: <https://www.karger.com/Article/FullText/527971>. Acesso em: 17 out. 2023.

SEZGI, C. *et al.* Oxidative Status and Acute Phase Reactants in Patients with Environmental Asbestos Exposure and Mesothelioma. **The Scientific World Journal**, [s. l.], v. 2014, p. 1–5, 2014. DOI 10.1155/2014/902748. Disponível em: <http://www.hindawi.com/journals/tswj/2014/902748/>. Acesso em: 23 set. 2023.

SUZUKI, Y.; YUEN, S. R. Asbestos Fibers Contributing to the Induction of Human Malignant Mesothelioma. **Annals of the New York Academy of Sciences**, New York, v. 982, n. 1, p. 160–176, 2002. DOI 10.1111/j.1749-6632.2002.tb04931.x. Disponível em: <https://nyaspubs.onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.2002.tb04931.x>. Acesso em: 23 set. 2023.

TABACHNICK, B. G.; FIDELL, L. S. **Using multivariate statistics**. 6. ed. Boston: Pearson, 2013.

TAKAMORI, S. *et al.* The C-Reactive Protein/Albumin Ratio is a Novel Significant Prognostic Factor in Patients with Malignant Pleural Mesothelioma: A Retrospective Multi-institutional Study. **Annals of Surgical Oncology**, [s. l.], v. 25, n. 6, p. 1555–1563, 2018. DOI 10.1245/s10434-018-6385-x. Disponível em: <http://link.springer.com/10.1245/s10434-018-6385-x>. Acesso em: 19 out. 2023.

TANRIKULU, A.; ER, O. Mesothelioma's disease data set. UCI Machine Learning Repository, 2016. DOI 10.24432/C5903T. Disponível em: <https://archive.ics.uci.edu/dataset/351>. Acesso em: 24 set. 2023.

TERRA FILHO, M.; FREITAS, J. B. P. D.; NERY, L. E. Doenças asbesto-relacionadas. **Jornal Brasileiro de Pneumologia**, [s. l.], v. 32, n. 2, p. S48–S53, 2006. DOI 10.1590/S1806-37132006000800009. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1806-37132006000800009&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1806-37132006000800009&lng=pt&tlng=pt). Acesso em: 23 set. 2023.

ZHANG, WEIQUN *et al.* Advances in the diagnosis, treatment and prognosis of malignant pleural mesothelioma. **Annals of Translational Medicine**, [s. l.], v. 3, n. 13, 2015. Disponível em: <https://atm.amegroups.org/article/view/7018>.