

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

DISSERATAÇÃO DE MESTRADO

POTENCIALIDADES DA ESPECTROMETRIA NO  
INFRAVERMELHO E QUIMIOMETRIA PARA CARACTERIZAÇÃO  
DE PERFIS DE MEDICAMENTOS APREENDIDOS

DANIELLE KOCHENBORGER JOHN

Porto Alegre, Novembro de 2020.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

DANIELLE KOCHENBORGER JOHN

**POTENCIALIDADES DA ESPECTROMETRIA NO INFRAVERMELHO E  
QUIMIOMETRIA PARA CARACTERIZAÇÃO DE PERFIS DE  
MEDICAMENTOS APREENDIDOS**

Dissertação apresentada como requisito parcial para a  
obtenção do grau de Mestre em Química

Prof. Dr. Marco Flôres Ferrão  
Orientador

Porto Alegre, Novembro de 2020.

*"A vida não é fácil para nenhum de nós. Mas, e daí? Devemos ter perseverança e, acima de tudo, confiança em nós mesmos. Devemos acreditar que somos dotados para algo e que isso deve ser conquistado."*

Marie Curie

## **Agradecimentos**

Ao meu orientador, Marco Flôres Ferrão, pela paciência, ensinamentos em quimiometria, dedicação e amizade durante este tempo tumultuado.

À UFRGS, UNISC, IFRS e Polícia Federal pela estrutura, equipamentos e amostras necessárias para a realização do presente trabalho.

Ao CNPq pelo auxílio financeiro.

Aos meus colegas do LAQIA, meu muito obrigada pelo acolhimento e risadas. Sempre teremos Nona Deusa.

Aos meus pais e irmão, Davi, Rosane e Dimitri, pelo cuidado que têm por mim, pelo suporte nas horas complicadas e pelo amor a qualquer hora. Amo vocês.

À minha tia Madalena e minha prima Melina, pelo acolhimento, conversas, conselhos, incentivos e sororidade.

Ao meu namorado, Elias, pelo suporte emocional, carinhos, companheirismo e revisões de texto que foram essenciais durante essa jornada acadêmica.

Muito obrigada a todos!

## Sumário

Lista de Figuras.....	vi
Lista de Tabelas.....	ix
Lista de Abreviaturas.....	x
Resumo.....	xii
Abstract.....	xiii
1 Introdução.....	14
2 Revisão bibliográfica.....	16
2.1 Química forense.....	16
2.2 Medicamento irregulares.....	16
2.2.1 Medicamentos tipo-anfetamina.....	20
2.2.1.1 Anorexígenos.....	21
2.2.2 Antidepressivos.....	24
2.2.3 Anabolizantes.....	25
2.3 Inteligência forense.....	27
2.4 Espectroscopia no infravermelho.....	29
2.4.1 Espectroscopia no infravermelho médio.....	31
2.4.1.1 Reflexão total atenuada.....	31
2.4.2 Espectroscopia no infravermelho próximo.....	32
2.5 Quimiometria.....	33
2.5.1 Métodos de reconhecimentos de padrões não supervisionados ..	34
2.5.1.1 Análise de componentes principais.....	35
2.5.1.2 Análise de agrupamentos por métodos hierárquicos.....	36
2.5.2 Métodos supervisionados de reconhecimento de padrões.....	37
2.5.2.1 K-enésimos vizinhos mais próximos.....	37
2.5.2.2 <i>Soft independent modeling of class analogy</i> .....	38

2.5.2.3	Análise discriminante pelo método de quadrados mínimos parciais	39
2.6	Espectroscopia no infravermelho e quimiometria	40
3	Materiais e métodos	43
3.1	Amostras	43
3.2	Espectroscopia de infravermelho médio	43
3.3	Espectroscopia no infravermelho próximo	44
3.4	Análise multivariada	44
4	Resultados e discussão	45
4.1	Infravermelho médio	45
4.1.1	Análise exploratória	46
4.1.2	Métodos supervisionados	58
4.2	Infravermelho próximo	65
4.2.1	Análise exploratória	67
4.2.2	Métodos supervisionados	72
4.3	Aplicação na inteligência forense	78
5	Conclusão	80
	Referências bibliográficas	81

## Lista de Figuras

<b>Figura 1.</b> Produtos fora do padrão e falsificados reportados à OMS de 2013 a 2017. ....	18
<b>Figura 2.</b> Fluxograma para a identificação do nível de risco do uso de medicamentos contrafeitos.....	19
<b>Figura 3.</b> Estrutura base de compostos anfetamínicos.....	201
<b>Figura 4.</b> Estruturas moleculares dos adulterantes cafeína e lidocaína. ....	212
<b>Figura 5.</b> Estrutura molecular da sibutramina, um anorexígeno. ....	223
<b>Figura 6.</b> Estrutura molecular do femproporex. A linha curva indica o ponto de clivagem para a liberação da anfetamina. ....	234
<b>Figura 7.</b> Estrutura molecular da benzocaína. ....	245
<b>Figura 8.</b> Estrutura molecular do antidepressivo fluoxetina. ....	256
<b>Figura 9.</b> Estrutura molecular da testosterona. ....	27
<b>Figura 10.</b> Principais etapas do processo iterativo de inteligência forense aplicado na análise de contrafeitos. ....	29
<b>Figura 11.</b> Espectro eletromagnético, com destaque para a região do infravermelho.....	290
<b>Figura 12.</b> Tipos de vibrações moleculares (+ indica movimento em direção ao leitor e – para longe). ....	301
<b>Figura 13.</b> Análise simplificada por ATR.....	323
<b>Figura 14.</b> Etapas para a aplicação de quimiometria. Etapas diretamente relacionadas a aplicação quimiométricas estão destacadas. ....	345
<b>Figura 15.</b> Classificação de amostra desconhecida em uma das classes por kNN. ....	39
<b>Figura 16.</b> Limites das hipercaixas dos modelos SIMCA e amostra desconhecida. ....	390
<b>Figura 17.</b> Representação do modelo PLS-DA para duas classes.....	401
<b>Figura 18.</b> Espectros de infravermelho médio das amostras sem pré-processamento.....	456
<b>Figura 19.</b> Espectros de ATR-FTIR para amostras de fármacos: a) espectros brutos da região espectral de 1.800-750 $\text{cm}^{-1}$ e b) média dos espectros pós-tratamentos SNV e 1ª derivada de Savitzky-Golay com janela de 13 pontos e polinômio de 2ª ordem.....	47

<b>Figura 20.</b> Análise de agrupamentos por métodos hierárquicos dos espectros obtidos por ATR-FTIR. ....	48
<b>Figura 21.</b> Gráfico de escores da análise de componentes principais da PC1xPC2.....	49
<b>Figura 22.</b> Gráfico de escores da análise de componentes principais da PC1xPC3.....	490
<b>Figura 23.</b> Gráfico de pesos da primeira componente principal da PCA da região espectral entre 1.800 e 750 $\text{cm}^{-1}$ .....	501
<b>Figura 24.</b> Gráfico de pesos da segunda componente principal da PCA da região espectral entre 1.800 e 750 $\text{cm}^{-1}$ .....	512
<b>Figura 25.</b> Gráfico de pesos da terceira componente principal da PCA da região espectral entre 1.800 e 750 $\text{cm}^{-1}$ .....	523
<b>Figura 26.</b> Média das amostras 8-11 e 12-20, da Classe 3.....	534
<b>Figura 27.</b> Análise de agrupamentos por métodos hierárquicos dos espectros obtidos por ATR-FTIR com a “pseudo Classe 5”, em azul escuro. ....	545
<b>Figura 28.</b> Espectros brutos com destaque para a região de fingerprint utilizada na criação dos modelos (1.800 – 750 $\text{cm}^{-1}$ ) e região selecionada por i-PCA (1.600 – 1.567 $\text{cm}^{-1}$ ). ....	556
<b>Figura 29.</b> Dendrograma resultante da HCA do intervalo espectral selecionado por i-PCA (1.600 – 1.567 $\text{cm}^{-1}$ ).....	57
<b>Figura 30.</b> Gráfico de escores da PC1 e PC2 referente ao intervalo espectral selecionado por i-PCA (1.600 – 1.567 $\text{cm}^{-1}$ ).....	57
<b>Figura 31.</b> Gráfico de escores da PC1 e PC2 referente ao intervalo espectral selecionado por i-PCA (1.600 – 1.567 $\text{cm}^{-1}$ ) com a adição das amostras da Classe 5. ....	58
<b>Figura 32.</b> Taxa de erro na validação cruzada para a determinação do número de vizinhos mais próximos do modelo kNN.....	590
<b>Figura 33.</b> Taxa de erro e de amostras não atribuídas na validação cruzada para a determinação do número de variáveis latentes do modelo PLS-DA. ....	623
<b>Figura 34.</b> Classificação das amostras por PLS-DA utilizando a faixa espectral de 1.600 a 1.567 $\text{cm}^{-1}$ .....	645
<b>Figura 35.</b> Espectros de infravermelho próximo das amostras sem pré-processamento.....	67



<b>Figura 36.</b> Espectros de NIR para amostras de fármacos: a) espectros brutos da região espectral de 7.500 a 4.000 $\text{cm}^{-1}$ e b) média dos espectros pós-tratamentos MSC e 1ª derivada de Savitzky-Golay com janela de 13 pontos e polinômio de 2ª ordem.....	67
<b>Figura 37.</b> Análise de agrupamentos por métodos hierárquicos dos espectros obtidos por NIR. ....	68
<b>Figura 38.</b> Análise de agrupamentos por métodos hierárquicos dos espectros obtidos por NIR, com a Classe 5 em azul escuro.....	69
<b>Figura 39.</b> Gráfico de escores da análise de componentes principais da PC1xPC2.....	690
<b>Figura 40.</b> Espectros tratados com destaque para a região selecionada por i-PCA (4.438 – 4.330 $\text{cm}^{-1}$ ).....	701
<b>Figura 41.</b> Dendrograma resultante da HCA do intervalo espectral selecionado por i-PCA (4.438 – 4.330 $\text{cm}^{-1}$ ).....	701
<b>Figura 42.</b> Gráfico de escores da PC1 e PC2 referente ao intervalo espectral selecionado por i-PCA (4.438 – 4.330 $\text{cm}^{-1}$ ).....	712
<b>Figura 43.</b> Dendrograma resultante da HCA do intervalo espectral selecionado por i-PCA (4.438 – 4.330 $\text{cm}^{-1}$ ) com a adição das amostras da Classe 5. ....	712
<b>Figura 44.</b> Gráfico de escores da PC1 e PC2 referente ao intervalo espectral selecionado por i-PCA (4.438 – 4.330 $\text{cm}^{-1}$ ) com a adição das amostras da Classe 5. ....	723
<b>Figura 45.</b> Taxa de erro na validação cruzada para a determinação do número de vizinhos mais próximos do modelo kNN.....	734
<b>Figura 46.</b> Taxa de erro e de amostras não atribuídas na validação cruzada para a determinação do número de variáveis latentes do modelo PLS-DA. ....	756
<b>Figura 47.</b> Classificação das amostras por PLS-DA utilizando a faixa espectral de 4.438 a 4.330 $\text{cm}^{-1}$ .....	78
<b>Figura 48.</b> Processo iterativo da aplicação de métodos quimiométricos na inteligência forense. ....	790

## Lista de Tabelas

<b>Tabela 1.</b> Regiões espectrais do infravermelho.....	290
<b>Tabela 2.</b> Uso de quimiometria na análise de medicamentos e drogas adulteradas.....	412
<b>Tabela 3.</b> Amostras analisadas e respectivos princípios ativos encontrados por GC-MS. ....	434
<b>Tabela 4.</b> Acurácia, especificidade e sensibilidade do modelo kNN para amostras analisadas por ATR-FTIR utilizando a faixa 1.600-1.567 $\text{cm}^{-1}$ . ....	590
<b>Tabela 5.</b> Acurácia, especificidade e sensibilidade do modelo SIMCA para amostras analisadas por ATR-FTIR utilizando a faixa 1.600-1.567 $\text{cm}^{-1}$ . ....	601
<b>Tabela 6.</b> Acurácia, especificidade e sensibilidade do modelo PLS-DA para amostras analisadas por ATR-FTIR utilizando a faixa 1.600-1.567 $\text{cm}^{-1}$ . ....	634
<b>Tabela 7.</b> Acurácia, especificidade e sensibilidade do modelo kNN para amostras analisadas por NIR utilizando a faixa 4.438 a 4.330 $\text{cm}^{-1}$ .....	745
<b>Tabela 8.</b> Acurácia, especificidade e sensibilidade do modelo SIMCA para amostras analisadas por NIR utilizando a faixa 4.438 a 4.330 $\text{cm}^{-1}$ .....	745
<b>Tabela 9.</b> Acurácia, especificidade e sensibilidade do modelo PLS-DA para amostras analisadas por NIR utilizando a faixa 4.438 a 4.330 $\text{cm}^{-1}$ .....	77

## Lista de Abreviaturas

ANVISA - Agência Nacional de Vigilância Sanitária

ATR - Reflexão Total Atenuada (do inglês, "*Attenuated Total Reflection*")

ATR-FTIR – Espectroscopia De Infravermelho Médio Com Transformada De Fourier Com Reflexão Total Atenuada (do inglês, "*Fourier Transform Infrared Spectroscopy With Attenuated Total Reflection*")

DSC - Calorimetria Diferencial De Varredura (do inglês, "*Differential Scanning Calorimetry*")

FIR - Infravermelho Distante (do inglês, "*Far Infrared*")

FTIR - Espectroscopia No Infravermelho Por Transformada De Fourier (do inglês, "*Fourier Transform Infrared*")

GC-MS - Cromatografia Gasosa Acoplada A Espectrometria De Massas (do inglês, "*Gas Chromatography Mass Spectrometry*")

HCA - Análise De Agrupamentos Por Métodos Hierárquicos (do inglês, "*Hierarchical Clusters Analysis*")

HPLC-DAD-MS/MS - Cromatografia Líquida De Alta Eficiência Com Detector De Arranjo De Diodos Acoplada À Espectrometria De Massas Sequencial (do inglês, "*High-Performance Liquid Chromatograph Coupled With On-Line Mass Spectrometry*")

HPLC-MS - Cromatografia Líquida De Alta Eficiência Acoplada Ao Espectrômetro De Massas (do inglês, "*High-Performance Liquid Chromatograph Coupled With Mass Spectrometry*")

IFA – Ingrediente farmacêutico ativo

i-PCA - Análise De Componentes Principais Por Intervalo (do inglês, "*Interval Principal Components Analysis*")

ISRS - Inibidor Seletivo Da Receptação Da Serotonina

kNN - K-enésimos Vizinhos Mais Próximos (do inglês, "*K- Nearest Neighbors*")

LC-DAD - Cromatografia Líquida Com Detector De Arranjo De Diodos (do inglês, "*Liquid Chromatography With Photodiode Array Detection*")

LC-HRMS - Cromatografia Em Fase Líquida Acoplada A Espectrometria De Massas De Alta Resolução (do inglês, "*Liquid Chromatography Coupled High Resolution Mass Spectrometry*")

LSD - ácido lisérgico (do alemão, "*Lysergsäurediethylamid*")

MDMA - Metilenodioximetanfetamina

MIR - Infravermelho Médio (do inglês, "*Mid Infrared*")

MSC – Correção Multiplicativa De Espalhamento (do inglês, "*Multiplicative Scatter Correction*")

NBOME - 2-(4-iodo-2,5-dimetoxifenil)-N-[(2-methoxyphenyl)metil]etanamina

NIR - Infravermelho Próximo (do inglês, "*Near Infrared*")

OMS - Organização Mundial de Saúde

PC - Componentes Principais (do inglês, "*Principal Components*")

PCA - Análise De Componentes Principais (do inglês, "*Principal Component Analysis*")

PLS-DA - análise discriminante por mínimos quadrados parciais (do inglês, "*Partial Least Squares Discriminant Analysis*")

SIMCA – do inglês, "*Soft Independent Modeling of Class Analogy*"

SNV – Padronização Normal De Sinal (do inglês, "*Standard Normal Variate*")

SSFFC - Produtos Médicos De Qualidade Inferior / Adulterados / Falsamente Rotulados / Falsificados / Contrafeitos (do inglês, "*Substandard / Spurious / Falsely Labelled / Falsified / Counterfeit Medical Products*")

SWGDRUG – do inglês, "*Scientific Working Group for the Analysis of Seized Drugs*"

VC - Validação Cruzada

VL - Variáveis Latentes

## Resumo

A Organização Mundial da Saúde estima que cerca de 10% dos medicamentos comercializados mundialmente são falsificados ou fora das especificações, principalmente em países cujo os índices de desenvolvimento são relativamente baixos. Esses medicamentos de baixa qualidade podem causar sérios danos à saúde de seus consumidores, visto que não existe um controle de qualidade adequado em sua produção. Este trabalho teve como objetivo desenvolvimento de um método rápido e simples de análise dos perfis dos medicamentos apreendidos pela Polícia Federal do Rio Grande do Sul através da espectroscopia no infravermelho com transformada de Fourier por reflexão total atenuada (ATR-FTIR) e no infravermelho próximo (NIR) associada a quimiometria. Para isso, foram aplicados métodos de análise multivariada não supervisionados e supervisionados HCA, PCA, kNN, SIMCA e PLS-DA em dados de 42 duas amostras de cinco classes de medicamentos, previamente analisadas por GC-MS. Com o método i-PCA, foi selecionada a região espectral de interesse tanto nos dados de ATR-FTIR quanto de NIR. As análises exploratórias HCA e PCA indicaram que amostras que formavam uma das classes, pois apresentavam a mesma mistura de compostos (lidocaína e cafeína), foram produzidas em diferentes condições, sendo, provavelmente, de fabricantes e/ou lotes distintos. O método de classificação por PLS-DA foi o mais eficiente no conjunto de dados de ATR-FTIR e o SIMCA com os dados de NIR. Portanto, métodos espectroscópicos ATR-FTIR e NIR associados a quimiometria permitem o desenvolvimento de técnicas de triagem rápidas, de baixo custo e sem preparo de amostras, auxiliando na detecção dos perfis de medicamentos com suspeita de falsificação.

Palavras-chave: quimiometria, medicamentos falsificados, espectroscopia no infravermelho, ATR-FTIR, NIR.

## **Abstract**

The World Health Organization estimates that about 10% of medicines sold worldwide are counterfeit or out of specification, mainly in countries whose development rates are relatively low. Low-quality drugs can cause severe damage to the health of their consumers since there is no adequate quality control in their production. This work aimed to develop a fast and easy method for analyzing the profiles of drugs seized by the Federal Police of Rio Grande do Sul through infrared spectroscopy with Fourier transform by attenuated total reflection (ATR-FTIR) and in the near-infrared (NIR) associated with chemometrics. For this, unsupervised and supervised multivariate analysis methods HCA, PCA, kNN, SIMCA, and PLS-DA were applied to data from 42 samples from five classes of drugs, previously determined by identifying the active principles by GC-MS of each one of the samples. With the iPCA method, a spectral region of interest was selected in both the ATR-FTIR and NIR data. The exploratory analyzes HCA and PCA indicated that samples that formed one of the classes, as they presented the same active ingredients (lidocaine and caffeine), were produced under different conditions, probably from distinct manufacturers or batches. The exploratory analyzes HCA and PCA indicated that samples that formed one of the classes, as they presented the same mixture of compounds (lidocaine and caffeine), were produced under different conditions, probably from different manufacturers or batches. The PLS-DA classification method was the most efficient in both the ATR-FTIR and NIR data sets. Therefore, ATR-FTIR and NIR spectroscopic methods associated with chemometrics allow the development of fast, low-cost screening techniques without sample preparation, assisting the detection of drug profiles with suspected counterfeiting.

*Keywords:* chemometrics, counterfeit medicines, infrared spectroscopy, ATR-FTIR, NIR.

## 1 Introdução

Nas últimas décadas, o acesso a medicamentos tem colaborado com a melhora na qualidade de vida ao reduzir taxas de mortalidade e morbidades. Apesar dos benefícios causados pelo desenvolvimento de novas tecnologias na indústria de fármacos, seu uso indiscriminado pela população tem impacto, muitas vezes, irreversível. O comércio *online* e ilegal de compostos farmacêuticos é uma das portas de entrada para falsificações que, em sua maioria, possuem pouco ou nenhum controle em seu processo de fabricação<sup>1</sup>.

As diferentes formas de falsificação tornam a identificação desses medicamentos bastante complexas visto que o produto pode: não possuir o ingrediente farmacêutico ativo (IFA), conter o princípio ativo errado ou possuir doses diferentes das indicadas ou mais de um fármaco ativo. Além disso, o IFA presente pode não ser autorizado por agências reguladoras ou ter sido produzido de diferentes formas <sup>2</sup>. O reconhecimento desses medicamentos falsificados torna-se de suma importância, uma vez que essas alterações no produto podem causar danos sérios à saúde de seus consumidores <sup>1</sup>.

Análises cromatográficas, necessárias para essa averiguação, são onerosas e consomem tempo e tornam-se um atraso em investigações e dificulta apreensões. Neste cenário, buscam-se metodologias analíticas rápidas e de fácil compreensão que possibilitem a identificação das falsificações e também forneçam informações que possam contribuir para o rastreamento dos produtores ilícitos. A aplicação de diferentes técnicas de torna-se uma possibilidade prática, visto que, a espectroscopia de infravermelho não é destrutiva e não exige preparo de amostras, reduzindo tempo e custo de análise. Essas análises aliadas a modelos quimiométricos facilitam a interpretação dos dados adquiridos através da explicação de comportamentos e previsão para classes de amostras<sup>3,4</sup>. A quimiometria associa estatística e matemática a dados químicos, tendo como objetivos acentuar as informações químicas relevantes. Métodos quimiométricos não supervisionados são técnicas apropriadas para o reconhecimento de padrões em amostras de origem desconhecida, pois revela tendências de comportamento entre elas, podendo indicar a proveniência das mesmas. Já as técnicas supervisionadas possuem como principal função indicar

a qual grupo as substâncias pertencem, o que auxiliaria o processo de investigação policial, principalmente *in situ*<sup>5</sup>.

O presente trabalho utilizou dados no infravermelho próximo e médio em conjunto de métodos quimiométricos exploratórios e supervisionados na identificação de padrões e classificação de amostras de medicamentos falsificados apreendidos pela Polícia Federal com o objetivo no desenvolvimento de metodologias de triagem rápidas e de fácil aplicação, que possibilitem o auxílio à inteligência forense no rastreamento dos produtores ilícitos.



## 2 Revisão bibliográfica

### 2.1 Química forense

A aplicação da ciência a exercício da lei consiste em uma extensa área conhecida como ciência forense. Esse grande grupo interdisciplinar engloba conteúdos de física, biologia, informática, farmácia, medicina, química, entre vários outros, e possui a finalidade de identificar e quantificar objetos de interesse judicial <sup>6,7</sup>. No Brasil, o Código de Processo Penal <sup>8</sup> determina que as ciências forenses, junto aos seus peritos, são essenciais em casos de processos penais, como estabelecidas pelos Art. 158 e Art. 159.

Art. 158. Quando a infração deixar vestígios, será indispensável o exame de corpo de delito, direto ou indireto, não podendo supri-lo a confissão do acusado.

Art. 159. O exame de corpo de delito e outras perícias serão realizados por perito oficial, portador de diploma de curso superior.

A prova pericial é de suma importância na resolução de crimes, visto que é através da perícia que se pode encontrar as conclusões necessárias em uma investigação. Entre as diferentes áreas periciais, destaca-se a química, que engloba grande parcela das ciências forenses, através de testes presuntivos, como a caracterização de drogas de abuso apreendidas, identificação de disparo de arma de fogo, além de análises confirmatórias realizadas em laboratórios de perícia, por meio de equipamentos analíticos. As principais classes analisadas por peritos de química são: drogas de abuso, suplementos alimentares, medicamentos, agrotóxicos, resíduos de explosivos, combustíveis entre outras <sup>9</sup>.

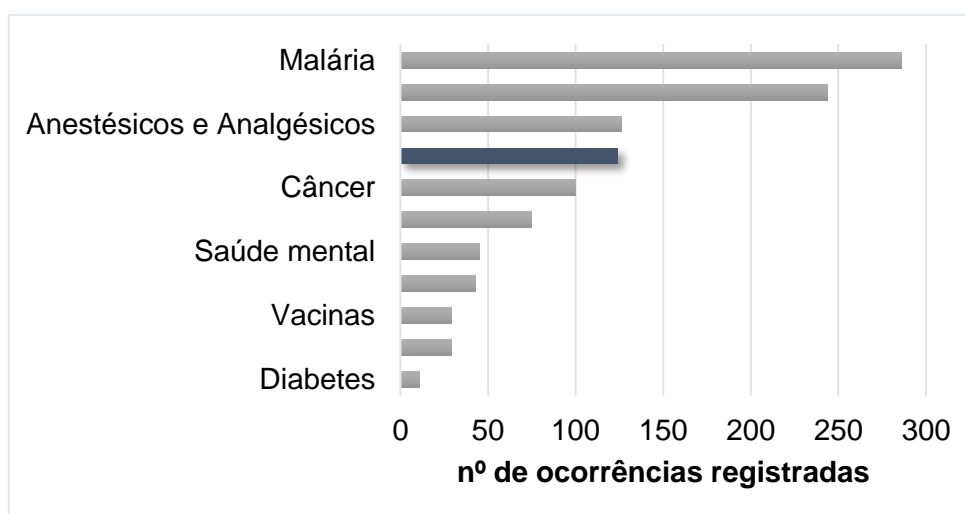
### 2.2 Medicamento irregulares

A Organização Mundial de Saúde (OMS) estima que cerca de 10% de todo medicamento comercializado mundialmente é falsificado ou fora da especificação (*substandard*). Medicamentos irregulares são definidos como “produtos médicos de qualidade inferior / adulterados / falsamente rotulados / falsificados / contrafeitos” (*Substandard / Spurious / Falsely labelled / Falsified /*

*Counterfeit medical products – SSFFC*)<sup>1</sup>. Os SSFFC podem ser classificados como:

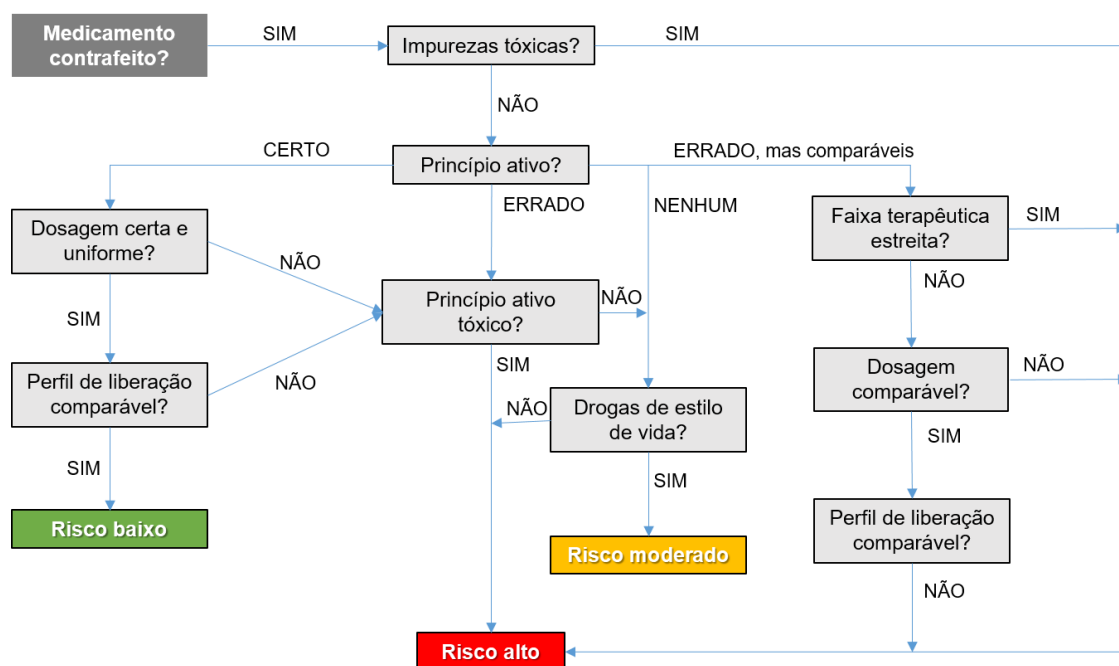
- a. Produtos médicos fora das especificações, considerados de qualidade inferior, que são autorizados, porém não cumprem padrões e especificações exigidas;
- b. Medicamentos não registrados e/ou licenciados, que não passaram por avaliações e aprovações necessárias impostas pelos órgãos reguladores;
- c. Medicamentos falsificados, que possuem composição, origem ou identidade fraudadas<sup>1</sup>.

Medicamentos autênticos com composição diferente da original, são especialmente perigosos, visto que seu uso pode mascarar sinais clínicos que proporcionariam o tratamento incorreto do enfermo<sup>2</sup>. A identificação e denúncia do comércio destes produtos é de grande importância, devido aos efeitos prejudiciais causados pelo consumo de medicamentos irregulares. De 2013 a 2017, foram reportados à OMS apenas 1112 casos de venda de produtos falsificados ou fora do padrão no mundo todo. A distribuição destes casos quanto ao tipo de medicamento pode ser observada na Figura 1. Destacam-se os chamados produtos de estilo de vida (em azul), que são os anabolizantes, medicamentos para disfunção erétil, queda de cabelo e emagrecimento, visto que, diferentemente dos demais, são mais consumidos em países de primeiro mundo<sup>1</sup>.



**Figura 1.** Produtos fora do padrão e falsificados reportados à OMS de 2013 a 2017.

Koczwarra e Dressman (2017) <sup>10</sup> criaram um método para classificação de medicamentos contrafeitos de acordo com seu provável nível de risco ao usuário, demonstrado na Figura 2. Conforme a proposta, medicamentos de baixo risco são aqueles cujas falsificações atuam no organismo da mesma maneira que o medicamento genuíno, também chamados de falsificações profissionais. Medicamentos de risco moderado são, principalmente, os não essenciais, sem eficiência significativa, porém não apresentam efeitos tóxicos. Os medicamentos de alto risco são essenciais e não apresentam eficácia.



**Figura 2.** Fluxograma para a identificação do nível de risco do uso de medicamentos contrafeitos.

Através da internet, medicamentos são comercializados ilegalmente e sem o acompanhamento médico necessário<sup>11</sup>. Morelato e outros (2020) <sup>12</sup> realizaram um estudo no qual foram analisados dados referentes ao comércio de medicamentos prescritos em um grande criptomercado. Benzodiazepínicos e analgésicos opióides correspondem a maior parte das transações de medicamentos do website. O III Levantamento Nacional sobre o Uso de Drogas pela População Brasileira <sup>13</sup>, constatou que os benzodiazepínicos, opiáceos e anfetamínicos são as classes de medicamentos mais consumidas pelos brasileiros de forma diferente da prescrita, principalmente pela restrição de sua comercialização. A prevalência do uso desses medicamentos foi maior entre as mulheres, uma vez que existe uma maior pressão da sociedade para que se

encaixem em um padrão de beleza, o que leva ao consumo de medicamentos, principalmente anorexígenos anfetamínicos <sup>14</sup>.

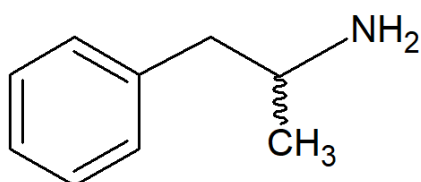
Watkins, Griffin e Gardner (2019)<sup>15</sup> encontraram diferentes concentrações e princípios ativos nas amostras de medicamentos obtidas na internet. Esta falta de qualidade dos produtos vendidos de forma online é perigosa e pode causar danos graves a saúde de seus consumidores. Arens e outros (2016) <sup>16</sup> relataram os efeitos do uso de alprazolam falsificado em oito pacientes que o adquiriram ilegalmente. Esse ansiolítico causou efeitos inesperados como sedação prolongada, fraqueza unilateral e sensação de formigamento, além de insuficiência renal. A causa desses efeitos colaterais foi a presença de fentanila e etizolam como princípios ativos do medicamento em substituição ao alprazolam. Essas substâncias também possuem efeitos analgésicos e ansiolíticos, porém com mecanismos de ação distintos. A fentanila é um anestésico de uso hospitalar 100 vezes mais potente que a morfina e está relacionado a um alto número de overdoses, principalmente nos Estados Unidos. Esse anestésico é comumente utilizado em falsificações de ansiolíticos e nas adulterações de heroína e cocaína <sup>17</sup>. Já o etizolam é um ansiolítico derivado de benzodiazepínicos. Seu uso é permitido em países como o Japão, Índia e Itália, porém, devido ao seu baixo custo e seus efeitos similares aos de ansiolíticos como o alprazolam, sendo também comumente utilizado em suas adulterações e está relacionado a um aumento de mortes causadas por consumo diferente da orientação médica pois é de fácil acesso <sup>18</sup>. A identificação desses pacientes resultou em uma investigação pelas agências policiais locais.

Tivura e colaboradores (2016) <sup>19</sup> avaliaram a qualidade de medicamentos antimaláricos vendidos em farmácias, drogarias e hospitais de Gana. Através da análise por Cromatografia Líquida de Alta Eficiência acoplada ao Espectrômetro de Massas (HPLC/MS), detectou-se que cerca de 35% das amostras apresentaram concentrações do princípio ativo fora das especificações, indicando que o controle de qualidade destes produtos era ineficiente. O índice de medicamentos falsificados e fora das especificações é maior em países em desenvolvimento pois não possuem recursos suficientes para a detecção em larga escala<sup>20</sup>.

Ao analisar diferentes cápsulas de uma mesma embalagem de sildenafil, através de Cromatografia Líquida de Alta Eficiência com Detector de Arranjo de Diodos acoplada à Espectrometria de Massas Sequencial (HPLC-DAD-MS/MS), Venhuis e colaboradores (2014) <sup>21</sup> encontraram dosagens diferentes do princípio ativo, além de pílulas que continham apenas os excipientes. Os medicamentos contrafeitos avaliados também possuíam tamanho e pesos diferentes do medicamento original. Essa situação demonstra que o preparo das amostras avaliadas tem grande influência na identificação de contrafeitos, uma vez que, no caso, poderia ocorrer um falso positivo para a presença de sildenafil. Empregando dados de Calorimetria Diferencial de Varredura (DSC) à análise hierárquica de clusters, Santos e colaboradores (2019) <sup>22</sup> identificaram amostras genuínas e contrafeitas de sildenafil e tadalafila, possibilitando o uso de diferentes tipos de dados para esse tipo de identificação.

### 2.2.1 Medicamentos tipo-anfetamina

As anfetaminas, (Figura 3), são estimulantes do sistema nervoso central capazes de gerar quadros de euforia e vigília. Foram primeiramente sintetizadas em 1886, mas somente na década de 1930 que foram utilizadas com finalidades terapêuticas para transtornos pulmonares como asma. Durante a Segunda Guerra Mundial, foi administrado à soldados para que pudessem se manter acordados e alertas <sup>23</sup>. Até os dias de hoje, medicamentos tipo-anfetamínicos como o metilfenidato são utilizados para o tratamento de doenças como a síndrome do déficit de atenção <sup>24</sup>.

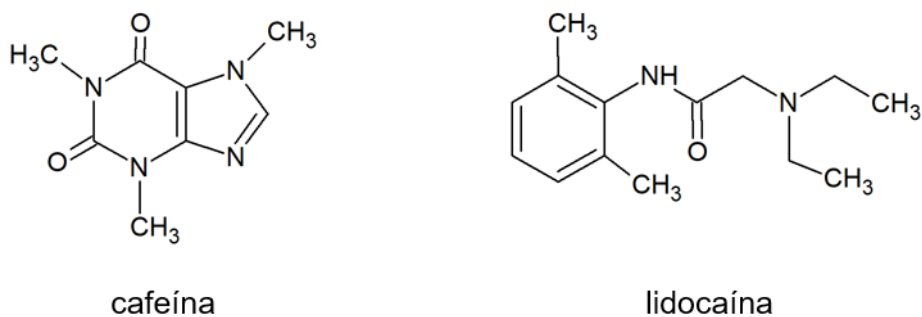


**Figura 3.** Estrutura base de compostos anfetamínicos.

No Brasil, as anfetaminas são classificadas como psicotrópicos e consideradas substâncias sujeitas a controle especial como indica a Portaria N<sup>o</sup> 344, de 12 de Maio de 1998 <sup>25</sup>. O uso de estimulantes tipo anfetamínicos

(popularmente conhecidos como “rebite”) é comum entre caminhoneiros, que possuem excessivas jornadas de trabalho, dirigindo por longas distâncias durante muitos dias seguidos sem fazer pausas adequadas para descanso, gerando assim maiores chances de acidentes no trânsito <sup>26</sup>. Estudos mostram que cerca de 5,8% dos motoristas brasileiros utilizam anfetamínicos enquanto trabalham, frequentemente combinada com outras drogas ilícitas como cocaína e canabinóides <sup>28</sup>. Entre as principais consequências do uso desse tipo de substância psicoativa é o efeito rebote, em que o usuário apresenta um estado de fadiga e sono, além de depressão. Esse tipo de efeito está diretamente relacionado a situações de risco no trânsito. Além disso, outros efeitos colaterais imediatos incluem taquicardia e aumento da temperatura corporal, além de perda de apetite <sup>29</sup>.

Substâncias como cafeína e lidocaína (Figura 4), dois dos principais adulterantes da cocaína, também são utilizados nas elaborações de rebites. A cafeína possui efeitos estimulantes similares aos da anfetamina, tornando-se uma alternativa barata. A lidocaína é um anestésico que possui efeitos semelhantes porém mais fortes que a cocaína e que pode causar, em baixas doses, problemas no sistema nervoso central e cardiovasculares<sup>30</sup>.



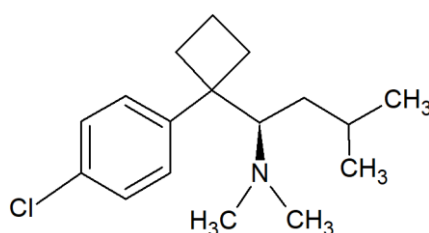
**Figura 4.** Estruturas moleculares dos adulterantes cafeína e lidocaína.

### 2.2.1.1 Anorexígenos

A OMS estima que cerca de 1,9 bilhões de adultos estão acima do peso ou obesos. Esta epidemia de obesidade é causadora de inúmeras comorbidades como doenças cardiovasculares, diabetes, derrame e alguns tipos de câncer. No Brasil, 55,7% da população está acima do peso, enquanto que 19,8% é considerada obesa, ou seja, com o índice de massa corporal acima de 30. Muitas

peessoas, buscando um tratamento, acabam por consumir medicamentos anorexígenos sem orientação médica, principalmente aqueles que os consomem para fins estéticos <sup>31-33</sup>.

A sibutramina (Figura 5) é um dos medicamentos para tratamento da obesidade mais consumidos no Brasil e está na lista de substâncias sujeitas a controle especial da Agência Nacional de Vigilância Sanitária (ANVISA) <sup>34</sup>. Este medicamento psicotrópico anorexígeno bloqueia a receptação de noradrenalina e serotonina, causando redução do apetite. Devido aos seus efeitos colaterais adversos, como infartos e derrames, seu uso é proibido em diversos países <sup>35</sup>.

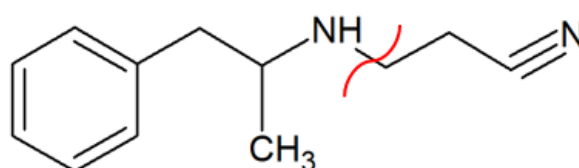


**Figura 5.** Estrutura molecular da sibutramina, um anorexígeno.

No Brasil, os medicamentos mazindol, femproporex e anfepramona foram retirados do mercado em 2011. O motivo da retirada destas substâncias inibidoras de apetite do tipo anfetamínicos foi uma revisão de segurança. Foi constatado que os riscos associados a tais medicamentos eram maiores que seus benefícios. Através da Resolução da Diretoria Colegiada nº133 de 2016, tais substâncias tiveram sua comercialização e dispensação legalizadas novamente, entrando na lista de substâncias sujeitas a controle especial junto à sibutramina.

As substâncias anorexígenas são muitas vezes alvo de adulteração, principalmente devido ao seu alto consumo entre pessoas que buscam a perda de peso de forma rápida. Sardela e colaboradores (2018) <sup>36</sup>, utilizando Cromatografia em Fase Líquida Acoplada a Espectrometria de Massas de Alta Resolução (LC-HRMS), analisaram exemplares de anfepramona adquiridas no comércio ilegal. Das amostras analisadas, não foi detectado o princípio ativo declarado em nenhuma delas, apenas os excipientes, indicando uma falsificação de risco moderado. Em outro estudo, realizado por Mariotti e outros (2013) <sup>37</sup>, amostras ilegalmente comercializadas de femproporex foram apreendidas pela

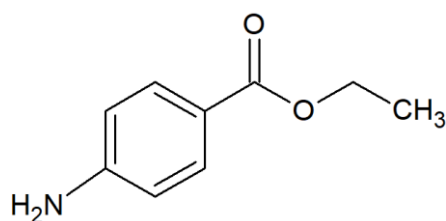
Polícia Federal. Foi identificada a presença de sibutramina através de análise por Cromatografia Gasosa acoplada a Espectrometria de Massas (GC-MS) em diversas amostras e com o uso da quimiometria aliada a dados de infravermelho médio foi possível afirmar que alguns dos medicamentos contrafeitos foram produzidos em um mesmo local. Medicamentos anorexígenos, em especial o femproporex, utilizado no tratamento a curto prazo para obesidade, também são empregados inadequadamente como estimulantes. Esse tipo de uso é comum uma vez que ocorre a biotransformação do femproporex em anfetamina no organismo através da clivagem da ligação nitrogênio-cianoetil <sup>38,39</sup>.



**Figura 6.** Estrutura molecular do femproporex. A linha curva indica o ponto de clivagem para a liberação da anfetamina.

Em outro estudo com apreensões da Polícia Federal <sup>40</sup>, amostras de sibutramina analisadas por GC-MS apresentaram benzocaína e cafeína como princípios ativos. A benzocaína (Figura 7) é comumente empregada como anestésico local, apesar de também ser utilizada em alguns tratamentos anti-obesidade, pois adormece nervos do estômago que são responsáveis pela sensação de fome <sup>41</sup>. Este fármaco não possui uma dose máxima segura estabelecida e pode causar metemoglobinemia, uma doença rara, porém bastante séria, que causa o desprendimento do átomo de ferro da hemoglobina do sangue prejudicando a oxigenação dos tecidos <sup>42</sup>. A cafeína também é um adulterante bastante comum em anorexígenos, uma vez que aumenta a massa do adulterado, possui propriedades estimulantes e é de fácil acesso, mas, em excesso, pode causar alterações de sono e humor além de taquicardia <sup>30</sup>.





**Figura 7.** Estrutura molecular da benzocaína.

Outro tipo comum de adulteração é a adição ilegal de substâncias anorexígenas em suplementos alimentares naturais devido a atividades farmacológicas. Os chamados medicamentos naturais promovem uma falsa segurança, uma vez que sua maioria não passa por um controle de qualidade rígido. Muitas dessas adulterações podem provocar casos graves de problemas cardiovasculares, dependência química e até mesmo psicose <sup>43</sup>.

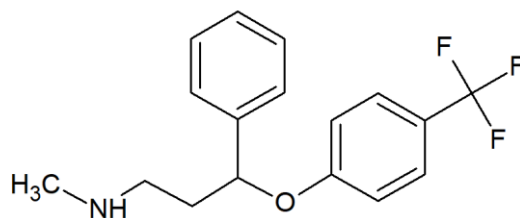
### **2.2.2 Antidepressivos**

Os antidepressivos são classes de medicamentos criadas para tratar depressão e estabilizar humor através de regulações nos chamados neurotransmissores, como dopamina, serotonina e norepinefrina. Outros tratamentos que utilizam esse tipo de medicamento incluem: transtornos alimentares, déficit de atenção, dores crônicas, incontinência urinária, distúrbios de sono, ejaculação precoce, urticária, entre outros <sup>44</sup>.

Os antidepressivos são drogas amplamente prescritas, uma vez que transtornos depressivos afetam cerca de 350 milhões de pessoas em todo mundo, principalmente na faixa etária entre 25 e 35 anos. Um dos medicamentos antidepressivos mais receitados pelos médicos é a fluoxetina, um inibidor seletivo da receptação da serotonina (ISRS) utilizado no combate a depressão desde a década de 1980 e, devido a menores efeitos colaterais, também é empregado no tratamento de outras desordens como bulimia, ansiedade e transtorno obsessivo compulsivo <sup>45</sup>.

A fluoxetina (Figura 8) é frequentemente encontrada como adulterante em medicamentos para perda de peso por possuir propriedades anorexígenas e ansiolíticas. Freitas e colaboradores (2019) <sup>46</sup> avaliaram medicamentos apreendidos pela Polícia Federal no Estado do Rio Grande do Sul. Aliando

técnicas de preparo de amostra utilizando micro-ondas, DSC e GC-MS, pôde-se observar a presença de antidepressivos em produtos que deveriam apresentar anfepramona, porém possuíam apenas fluoxetina como princípio ativo.



**Figura 8.** Estrutura molecular do antidepressivo fluoxetina.

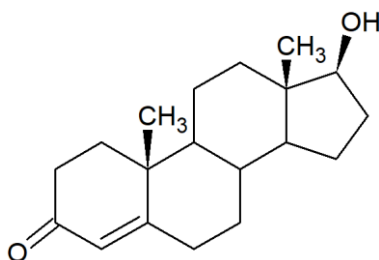
Neves e Caldas (2017) <sup>47</sup> caracterizaram diferentes compostos em suplementos alimentares para perda de peso apreendidos pela PF entre os anos de 2010 e 2016. Dentre as substâncias encontradas utilizando GC-MS, muitas amostras possuíam doses não declaradas no rótulo de sibutramina, cafeína e fluoxetina. Este tipo de adulteração também é comum em produtos rotulados como naturais ou herbais. Dastjerdi e outros (2018) <sup>48</sup> encontraram adulterações em suplementos herbais e dietéticos comercializados em lojas físicas e online. Por meio de análise de GC-MS, foram identificadas drogas sintéticas em 72% dos produtos analisados, destacando-se os antidepressivos fluoxetina e venlafaxina.

O uso de medicamentos e suplementos adulterados com fluoxetina, mesmo em doses baixas, pode causar efeitos colaterais como náusea, diarreia, perda de apetite, insônia, ansiedade e redução da libido. A síndrome serotoninérgica é comum em usuários de fluoxetina, caracterizada por sintomas como confusão mental, tremores, agitação, arritmia e hipertermia <sup>45</sup>.

### 2.2.3 Anabolizantes

Esteroides anabolizantes, ou esteroides anabólicos androgênicos, são versões sintéticas do hormônio masculino testosterona (Figura 9) ou derivados. Seu uso terapêutico é indicado em casos de hipogonadismo, impotência, perda muscular, endometriose, depressão, entre outros. A produção natural de testosterona no organismo masculino é cerca de 10 mg diários, promovendo o

crescimento de músculos e o aparecimento de características masculinas durante a puberdade <sup>49</sup>.



**Figura 9.** Estrutura molecular da testosterona.

Apesar de esteroides anabolizantes serem prescritos para o tratamento de diversos distúrbios, os mesmos acabam por serem utilizados de forma inadequada com objetivos estéticos e para aumento da performance esportiva. Em pesquisa realizada no Norte do Brasil <sup>50</sup>, observou-se o uso indiscriminado de tais substâncias entre frequentadores e professores de academias de ginástica. Os entrevistados relataram como principais efeitos colaterais observados: acne, aumento da agressividade e engrossamento da voz.

O uso não terapêutico de anabolizantes está frequentemente associado ao comércio ilegal dos mesmos. De acordo com Abrahin, Sousa e Santos (2014) <sup>51</sup>, estima-se que metade dos produtos anabolizantes apreendidos pela Polícia Federal não apresentam as substâncias descritas na embalagem, enquanto que 28,3% possuem apenas substâncias não declaradas.

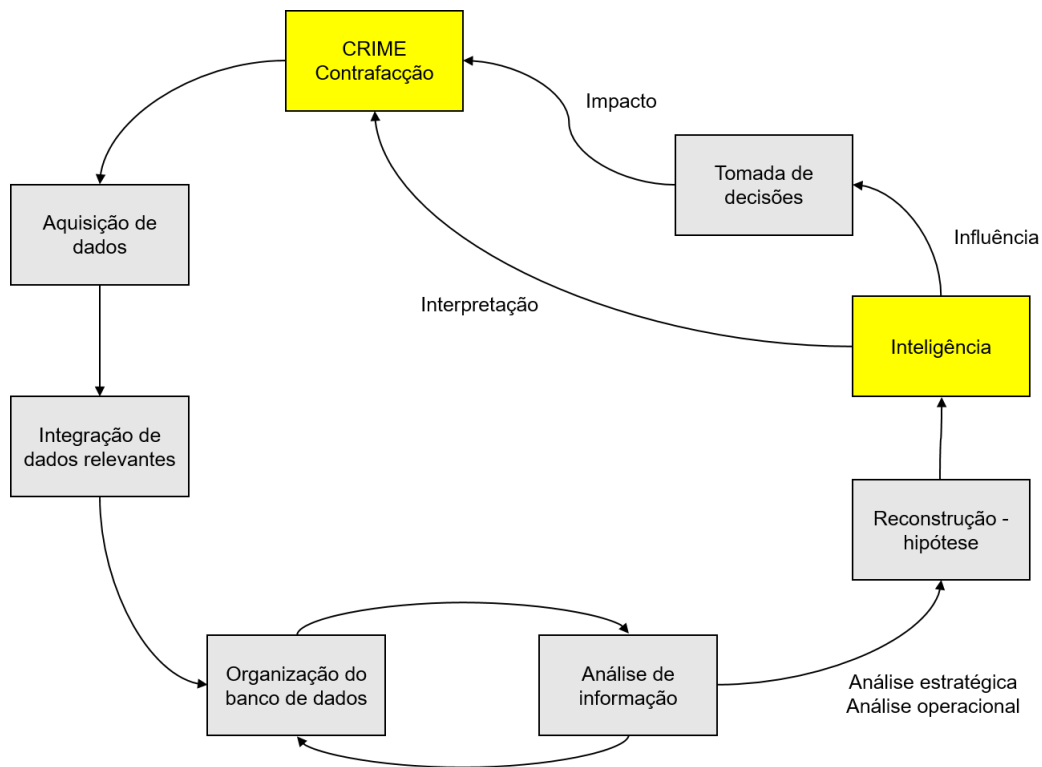
Neves e colaboradores (2017) <sup>52</sup>, utilizando Espectroscopia no Infravermelho por Transformada de Fourier (FTIR) associada à quimiometria, analisaram amostras autênticas e contrafeitas de Durateston, um repositório de testosterona. Ao criar um modelo de análise discriminante por mínimos quadrados parciais, observou-se que a técnica quimiométrica alcançou 100% de taxa de eficiência, sensibilidade e especificidade, provando-se um método mais efetivo e barato que os obtidos por GC-MS. Berneira e outros (2019) <sup>53</sup> analisaram ampolas de esteroides apreendidos por DSC, CG-MS e FTIR. Em cerca de 25% das amostras, somente o excipiente foi encontrado. Apenas em uma das amostras restantes o princípio ativo era correspondente ao descrito no rótulo, enquanto que nas demais o fármaco se encontrava bem abaixo da concentração indicada.

O uso de tais produtos sem o acompanhamento necessário pode gerar efeitos adversos graves. Dentre os principais efeitos colaterais destacam-se disfunção erétil, depressão, alteração de comportamento, problemas cardiovasculares e intoxicação no fígado. O consumo de anabolizantes associado a outras drogas e álcool também aumenta a agressividade e comportamentos violentos <sup>49</sup>.

### **2.3 Inteligência forense**

O conceito de inteligência consiste na produção de conhecimento e análise do mesmo, dando suporte a atividades relacionadas. Quando aplicada a dados de ciências forenses, é uma grande aliada em investigações policiais, uma vez que contribui de maneira estruturada na detecção de crimes repetitivos, além de apoiar modelos de policiamento e redução da criminalidade através da criação de banco de dados <sup>54</sup>.

A Figura 10 ilustra o processo iterativo de inteligência forense aplicada no combate de medicamentos falsificados. Inicialmente é imprescindível para o bom uso da inteligência forense que as informações analíticas sejam adquiridas e organizadas em um banco de dados estruturado, de forma que seja revisitado e alimentado constantemente com novas informações. Esse processo tem como objetivo dar uma visão atualizada do fenômeno em curso e permite examinar o risco e tomar decisões estratégicas e operacionais. Através desse estudo é possível interpretar o crime e reconstruí-lo <sup>55</sup>.



**Figura 10.** Principais etapas do processo iterativo de inteligência forense aplicado na análise de contrafeitos.

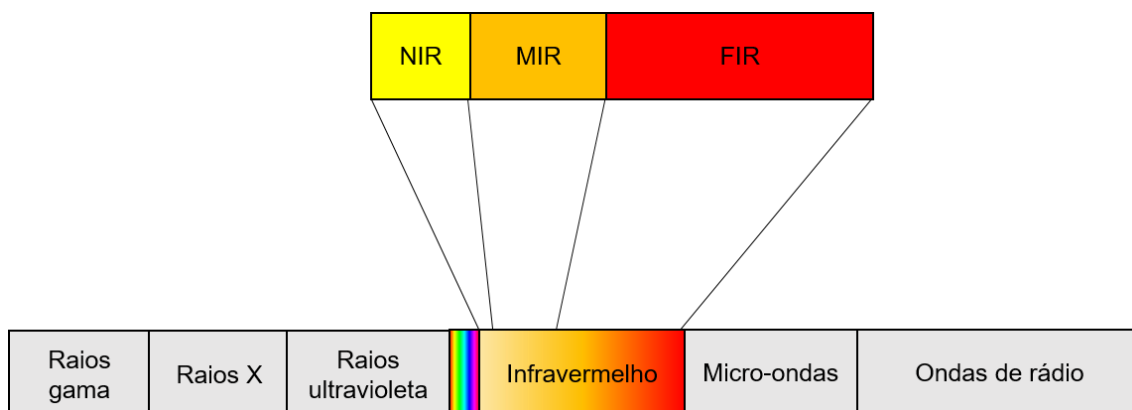
Dégardin e Roggo (2016) <sup>56</sup> utilizaram diferentes técnicas analíticas na análise de medicamentos contrafeitos. Eletroforese capilar, espectroscopia no infravermelho médio, próximo e ultravioleta-visível foram empregadas nas análises químicas, além da inspeção visual de caixas e comprimidos. Como resultados, conseguiram encontrar correlações entre diferentes medicamentos através das análises realizadas, constatando que seriam da mesma rede de falsificações.

A combinação de métodos analíticos também é um grande aliado na implementação da inteligência forense. Estudos que utilizam diferentes técnicas de espectroscopia molecular com o intuito de criar uma base de dados para a aplicação de técnicas quimiométricas a fim de caracterizar os perfis de diferentes medicamentos falsificados são comuns<sup>57-59</sup>. Desse modo, é possível identificar que apreensões distintas dessas falsificações possuíam o mesmo perfil químico, sendo, portanto de uma mesma fabricação. Esse tipo de aplicação, na inteligência forense, é capaz de agilizar e orientar ações para a combate de falsificadores.

## 2.4 Espectroscopia no infravermelho

A espectroscopia no infravermelho baseia-se na vibração dos átomos de uma molécula e é comumente utilizada em análises qualitativas e quantitativas, classificada como Categoria A pela SWGDRUG (2019) <sup>60</sup>, devido sua capacidade seletiva. O espectro de infravermelho é obtido através da passagem de radiação infravermelho pela amostra, que, ao absorve-la, aumenta a amplitude dos movimentos vibracionais das ligações moleculares, processos que podem ser quantificados <sup>61</sup>.

A região do infravermelho corresponde a intervalo do espectro eletromagnético situada entre a visível e micro-ondas, apresentado na Figura 11. A região do infravermelho divide-se em próximo (*near infrared* – NIR), médio (*mid infrared* – MIR) e distante (*far infrared* – FIR), conforme Tabela 1. Para uma molécula apresentar modos ativos no infravermelho suas ligações devem exibir mudanças no momento de dipolo durante seu movimento rotacional ou vibracional. Esta energia absorvida está relacionada aos modos vibracionais e rotacionais dos grupos funcionais presentes na molécula <sup>61</sup>.

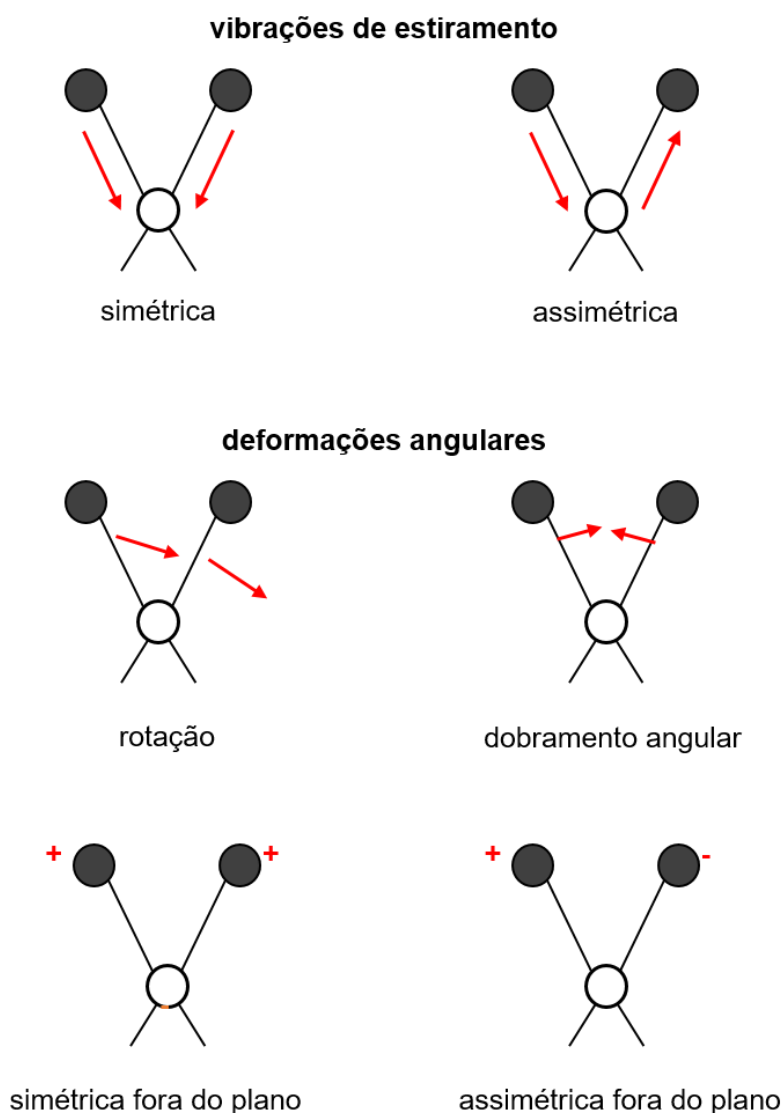


**Figura 11.** Espectro eletromagnético, com destaque para a região do infravermelho.

**Tabela 1.** Regiões espectrais do infravermelho.

Região	Nº de onda (cm <sup>-1</sup> )	Comprimento de onda (nm)
NIR	12.800 - 4.000	780 - 2.500
MIR	4000 - 400	2.500 - 5.000
FIR	200 - 10	5.000 - 100.000

As posições dos átomos em uma molécula flutuam continuamente devido a diferentes vibrações e rotações de suas ligações, dificultando a interpretação de espectros de moléculas com muitos átomos, uma vez que podem ocorrer interações entre os vários centros vibracionais. As vibrações moleculares são categorizadas em: estiramentos, que podem ser simétricos ou assimétricos, modificando a distância interatômica entre dois átomos; e deformações angulares, divididas em tesoura no plano, balanço no plano, torção fora do plano e sacudida fora do plano <sup>62</sup>, ilustradas.Figura 12.



**Figura 12.** Tipos de vibrações moleculares (+ indica movimento em direção ao leitor e – para longe).

## 2.4.1 Espectroscopia no infravermelho médio

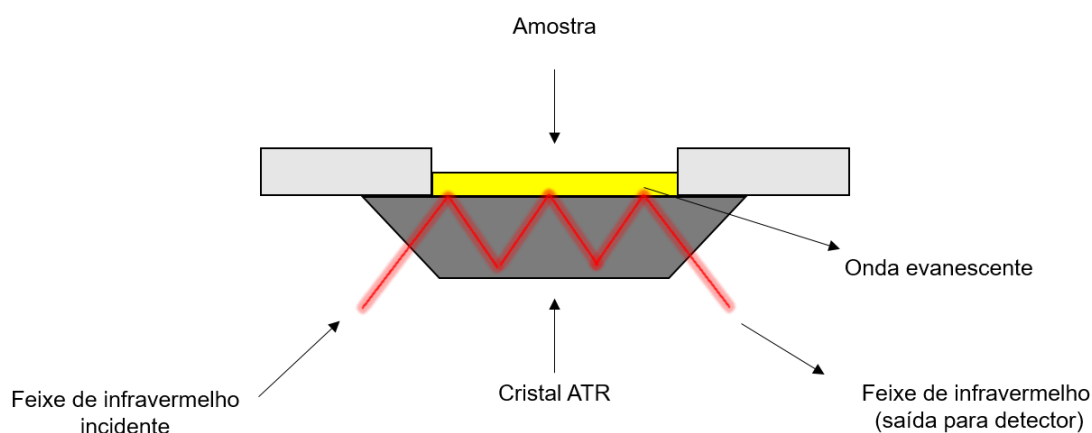
A faixa do infravermelho médio é a mais utilizada tanto em análises qualitativas quanto quantitativas. Estendendo-se de 4.000 até 670  $\text{cm}^{-1}$ , o MIR é uma ferramenta importante para a determinação de estruturas de espécies orgânicas e bioquímicas desde 1950, através do advento dos espectrofotômetros de feixe duplo, o que reduziu o tempo necessário para uma análise qualitativa <sup>62</sup>.

A região conhecida como impressão digital (do inglês, *fingerprint*) - entre 1.200 e 600  $\text{cm}^{-1}$  - é muito relevante na identificação de compostos, visto que até pequenas diferenças estruturais da molécula resultam em mudanças significativas de intensidade e distribuição da absorbância de seu espectro. A interpretação exata dos espectros nessa faixa é raramente possível devido sua complexidade. Entretanto, muitos compostos apresentam bandas características nessa região, o que é útil na identificação de moléculas e substâncias <sup>62</sup>.

### 2.4.1.1 Reflexão total atenuada

O método de reflexão total atenuada (*Attenuated Total Reflectance* - ATR) é baseado na reflexão interna da radiação infravermelha. Ao entrar em contato com materiais com altos índices de refração, como o ZnSe, a radiação infravermelha é totalmente refletida internamente. Em contato com a amostra, cujo o índice de refração é menor, parte desta radiação penetra na mesma, atenuando o feixe do comprimento de onda correspondente as bandas de absorção, surgindo então o espectro. Esta radiação penetrante que interage com a amostra é conhecida como onda evanescente<sup>62</sup>. A Figura 13 ilustra simplificadaamente uma análise por reflexão total atenuada.





**Figura 13.** Análise simplificada por ATR.

A profundidade da penetração da onda evanescente depende do comprimento de onda ( $\lambda_c$ ), índice de refração dos materiais (elemento de ATR e amostra,  $n_c$  e  $n_s$ , respectivamente), índice de refração do ar ( $n$ ), e o ângulo de incidência do feixe ( $\theta$ ), conforme Equação 1.

$$dp = \frac{\lambda_c}{2\pi n \sqrt{\left[ \sin^2 \theta - \left( \frac{n_s}{n_c} \right)^2 \right]}}$$

**Equação 1.** Profundidade da penetração da onda evanescente na amostra.

#### 2.4.2 Espectroscopia no infravermelho próximo

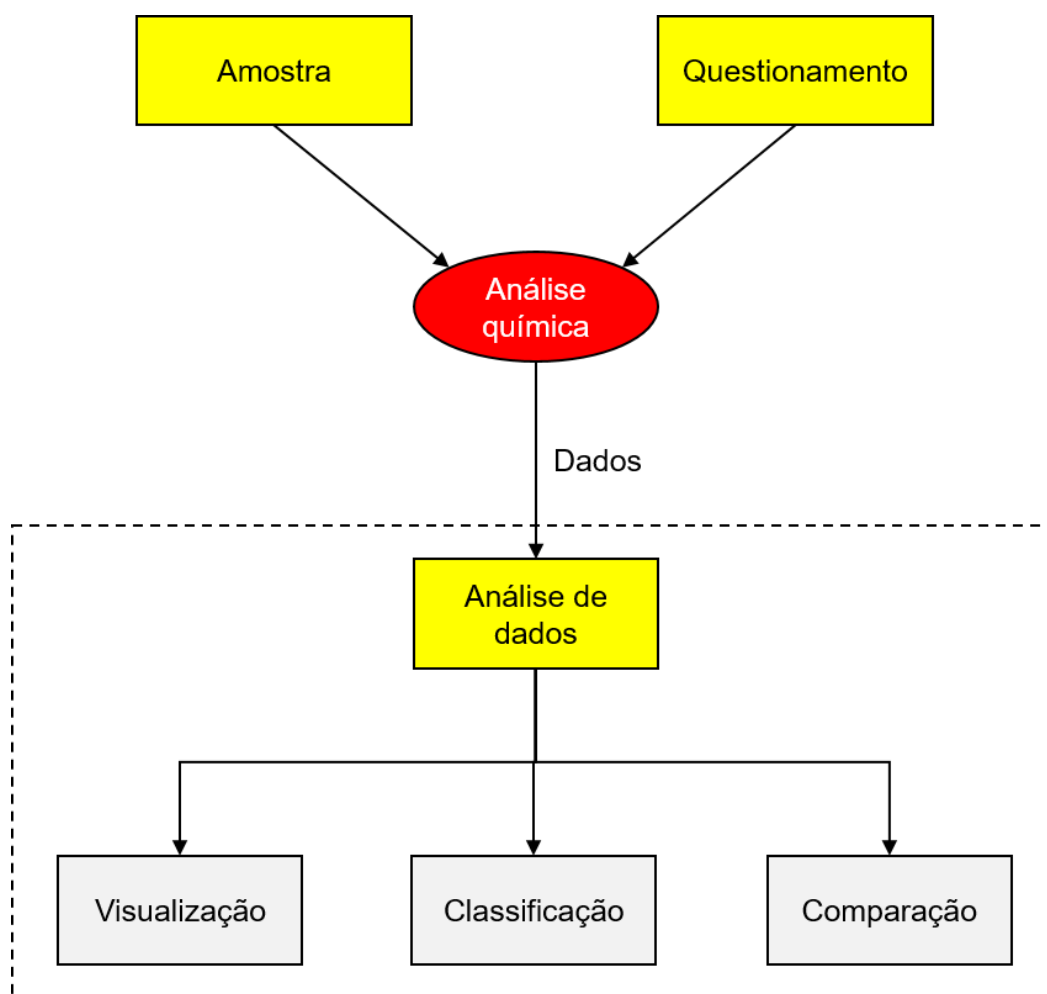
O infravermelho próximo consiste na região espectral de 13.000 a 4.000  $\text{cm}^{-1}$ . As bandas de absorção nesta região são combinações e sobretons fundamentais de estiramentos que ocorrem entre 3.000 e 1.700  $\text{cm}^{-1}$ , no infravermelho médio. As principais ligações observadas nesta faixa são as ligações presentes em grupos alifáticos e aromáticas (C-H), amins e amidas (N-H), além de hidroxila (O-H) <sup>62</sup>.

Comparado ao infravermelho médio, o NIR é utilizado em análises quantitativas com maior frequência quando associado a ferramentas quimiométricas. O uso da quimiometria é um grande aliado da técnica NIR, uma vez que esta região espectral apresenta *fingerprints* distintas de materiais orgânicos <sup>63</sup>.

## 2.5 Quimiometria

Com o avanço das técnicas analíticas instrumentais, um grande volume de dados é gerado. Estes dados possuem um extenso número de variáveis, como, por exemplo, cromatogramas ou espectros de infravermelho, onde cada tempo de retenção ou número de onda, respectivamente, corresponde a uma variável. Devido à complexidade dos dados, necessita-se de métodos estatísticos mais avançados para sua análise. Técnicas quimiométricas, portanto, são excelente escolhas, uma vez que são capazes de identificar padrões nos dados, classificar ou quantificar amostras <sup>64</sup>.

A quimiometria pode ser definida como uma disciplina que aplica métodos estatísticos e matemáticos a dados químicos com a finalidade de extrair o máximo de informação química, além de possibilitar o planejamento de experimentos e sua otimização <sup>64</sup>. A Figura 14 ilustra as etapas principais para uma aplicação quimiométrica. A escolha do método de análise está diretamente relacionada ao questionamento em relação as amostras. Após a coleta dos dados, realiza-se a análise dos mesmos, de forma que os resultados sejam também apresentados de acordo com as questões relacionadas. Esta apresentação pode ser visual, onde os resultados estatísticos são apresentados de forma gráfica. Já na classificação ocorre a investigação se uma amostra pertence ou não a uma classe pré-definida, enquanto que na comparação, amostras são avaliadas se possuem a mesma origem <sup>7</sup>.



**Figura 14.** Etapas para a aplicação de quimiometria. Etapas diretamente relacionadas a aplicação quimiométricas estão destacadas.

### 2.5.1 Métodos de reconhecimentos de padrões não supervisionados

Métodos não supervisionados de reconhecimento de padrões não requerem nenhum conhecimento sobre a identificação das amostras *a priori*, uma vez que são utilizados para identificar grupos de amostras baseando-se em características comuns. Estes métodos podem, portanto, ser empregados para encontrar similaridades e diferenças entre amostras de um conjunto de dados. Os métodos de análise exploratória mais utilizadas na quimiometria são análise de componentes principais e análise de agrupamentos por métodos hierárquicos

5,64.

### 2.5.1.1 Análise de componentes principais

A análise de componentes principais (PCA, do inglês: *Principal Component Analysis*) foi desenvolvida em 1930, na área de psicologia. A PCA consiste na projeção de dados multivariados de forma que a sua dimensão seja reduzida, sem que as relações entre as amostras sejam afetadas. Este método possui a capacidade de separar e ampliar as informações mais relevantes em um conjunto de dados, assim como evidenciar amostras atípicas (*outliers*) e detectar relações entre amostras e variáveis <sup>5</sup>. Na química, a PCA é um método comumente utilizado na análise de dados espectroscópicos e cromatográficos, no controle de qualidade de fármacos e na identificação de drogas de abuso <sup>3,4,65,66</sup>.

Através da PCA, a correlação significativa entre variáveis de um conjunto de dados permite que um número menor de novas variáveis seja encontrado de forma que descrevam a mesma informação presente nos dados originais. Para a redução do número de variáveis é realizada através uma compressão dos dados, combinando linearmente variáveis originais e agrupando aquelas que possuem informações semelhantes. As novas variáveis são chamadas de componentes principais (PCs – *Principal Components*), também conhecidas como autovetores. As PCs são não correlacionadas e ortogonais entre si, ou seja, cada PC possui informações distintas, sendo que a ordem de “importância” inicia com aquela que é mais representativa (PC1), oferecendo maior quantidade de informação, e assim por diante <sup>5</sup>.

Do ponto de vista matemático, a PCA pode ser descrita como a decomposição de uma matriz  $X(I \times J)$ , em duas matrizes como na Equação 2, uma representando os escores ( $T$ ) e outra os pesos ( $L$ ). Os escores indicam o grau de correlação entre as amostras, já os pesos revelam as relações entre as variáveis. A matriz  $E$  indica os resíduos, desvios entre as projeções e coordenadas originais <sup>5</sup>.

$$X = TL^T + E$$

**Equação 2.** Decomposição de matrizes na PCA.

### 2.5.1.2 Análise de agrupamentos por métodos hierárquicos

A análise de agrupamentos por métodos hierárquicos (HCA, do inglês: *Hierarchical Clusters Analysis*) é, assim como a PCA, um método não supervisionado utilizado no tratamento de dados químicos. A HCA tem sua origem na biologia, através da taxonomia numérica, sendo um método empregado para descobrir padrões naturais de comportamento das amostras. A principal função da HCA é agrupar amostras de modo que as pertencentes ao mesmo cluster sejam mais semelhantes do que amostras de outros grupos. Através de uma árvore hierárquica, denominada dendrograma, é possível visualizar o grau de similaridade entre objetos <sup>5</sup>.

A HCA é uma técnica aglomerativa, ou seja, considera cada amostra um único elemento, agrupando-as de forma sistemática por semelhança em um processo iterativo, até que ocorra a formação de um único grupo <sup>5</sup>. Na química, este método é rotineiramente empregado em dados espectroscópicos, visto a redução de dimensionalidade proporcionada pelo mesmo <sup>67,68</sup>.

A primeira etapa para a aplicação da técnica de HCA é a definição de um índice numérico para medir a distância entre amostras, uma vez que a HCA utiliza a proximidade entre objetos como característica de semelhança. A distância Euclidiana é a mais comum e utilizada em situações rotineiras. Expressa através da Equação 3, onde a distância Euclidiana entre pontos com coordenadas  $A=(a_1,a_2)$  e  $B=(b_1,b_2)$  no espaço multidimensional é apresentada usando o teorema de Pitágoras, onde  $x_{aj}$  e  $x_{bj}$  são valores numéricos da  $j$ -ésima coordenada de  $A$  e  $B$ , respectivamente <sup>5</sup>.

$$d_{AB} = \sqrt{(x_{a1} - x_{b1})^2 + (x_{a2} - x_{b2})^2 + \dots + (x_{aj} - x_{bj})^2}$$

#### Equação 3. Distância Euclidiana no espaço multidimensional

Os métodos de ligação entre os grupos podem ser: vizinho mais próximo, onde a distância entre dois grupos é dada pela maior proximidade entre objetos dos mesmos; vizinho mais distante, onde a distância entre os agrupamentos é dada pela maior distância entre os objetos dos grupos; média, onde ocorre uma média ponderada entre as amostras dos grupos; centroide, onde calcula-se o

ponto central e então a distância; e o método de Ward, que calcula a soma dos quadrados das distâncias do centroide médio entre cada um dos grupos, sendo que os dois grupos com menor acréscimo na soma quadrática total são unidos<sup>5</sup>.

## **2.5.2 Métodos supervisionados de reconhecimento de padrões**

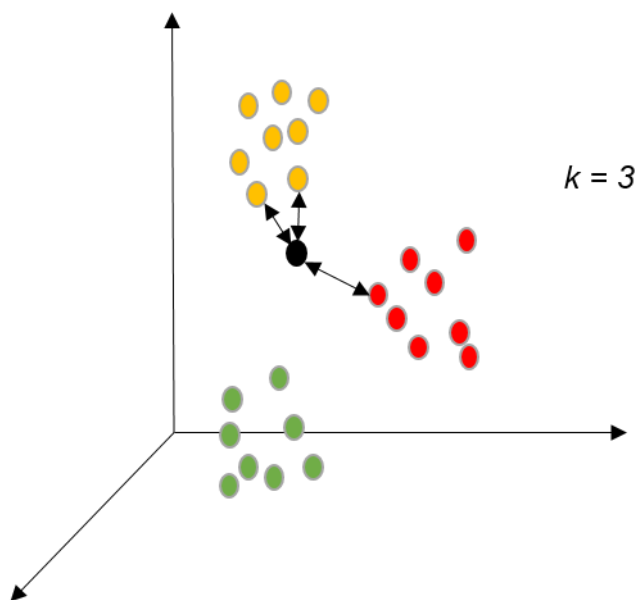
A análise supervisionada possibilita que amostras sejam categorizadas em diferentes classes previamente a aplicação do método. Por exemplo, é possível a classificação de acordo com tipo de medicamento, autenticidade, local de apreensão etc. Ou seja, ao contrário dos métodos de análise exploratória, são necessários conhecimentos *a priori* sobre cada amostra. Esse grupo de amostras, cuja propriedade de interesse é conhecida, é chamada de conjunto de treinamento. A partir desse conjunto de dados é construído o modelo de classificação. Devido a isso, esse conjunto de técnicas é denominado métodos supervisionados<sup>5</sup>.

Os métodos de classificação podem ser divididos em duas grandes categorias: paramétricos ou probabilísticos e não paramétricos ou não probabilísticos (determinísticos). Os métodos probabilísticos consideram que as variáveis satisfazem uma distribuição normal, usando a homogeneidade da matriz variância-covariância das classes para posicionar a região limitada de classificação. Já os métodos discriminantes baseiam-se na matriz dos dados **X** e no conhecimento prévio de classe de cada amostra, sem considerar medidas estatísticas<sup>5</sup>.

### **2.5.2.1 K-ésimos vizinhos mais próximos**

O kNN (do inglês: *k<sup>th</sup> Nearest Neighbors*) é um método supervisionado determinístico. A construção de um modelo de classificação kNN é realizada de forma que cada amostra do conjunto de treinamento seja excluída uma única vez e classificada de acordo com as amostras restantes. As distâncias entre as amostras do conjunto de treinamento e aquela que foi excluída são então calculadas, normalmente usando-se a distância Euclidiana (Equação 3). A amostra anteriormente excluída é classificada de acordo com a maioria dos votos

dos vizinhos mais próximos, sendo atribuída a classe mais votada. A Figura 15 ilustra o método kNN. Esse processo é aplicado a todas as amostras do conjunto de treinamento <sup>5</sup>.



**Figura 15.** Classificação de amostra desconhecida em uma das classes por kNN.

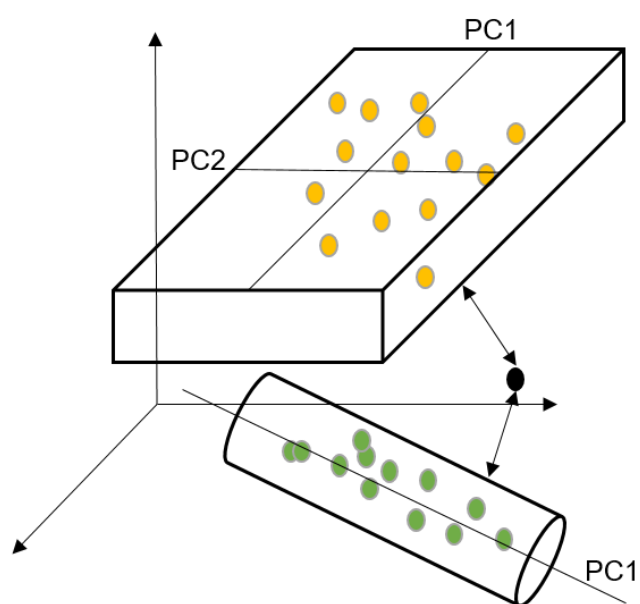
O método kNN possui muitas vantagens, sendo a principal delas ser muito simples. Por funcionar bem conjuntos de dados com número muito maior de variáveis que amostras, é bastante aconselhável para a classificação de dados espectroscópicos, ou quando uma das classes possui poucas amostras. Dentre as desvantagens destaca-se a incapacidade do método para detectar *outliers*, o que pode ser contornado com o uso prévio de PCA <sup>5</sup>.

### 2.5.2.2 *Soft independent modeling of class analogy*

O método SIMCA foi desenvolvido por Svante Wold nos anos 1970 <sup>69</sup>, sendo o primeiro pacote computacional da quimiometria e obtendo grande influência na química. Esta técnica assume que amostras parecidas possuem tendência de distribuição uniforme, sendo, portanto, probabilístico. No modelo SIMCA, uma PCA é feita para cada uma das classes do conjunto de treinamento. Para cada classe, um número de fatores (componentes principais) adequado é

determinado através de uma validação cruzada com o intuito de maximizar a capacidade preditiva do modelo<sup>5</sup>.

A definição das componentes principais para cada classe é a primeira etapa na construção de um modelo SIMCA, estabelecendo então os limites espaço-linha. Após, cada classe é delimitada no hiperplano complementar ortogonal, onde encontram-se os resíduos e, posteriormente, determina-se o tamanho da hipercaixa para cada classe. Quando um novo objeto é projetado no subespaço, comparam-se as distâncias entre o mesmo e cada classe, atribuindo-o aquela mais próxima<sup>5,70</sup>, como ilustrado na Figura 16.



**Figura 16.** Limites das hipercaixas dos modelos SIMCA e amostra desconhecida.

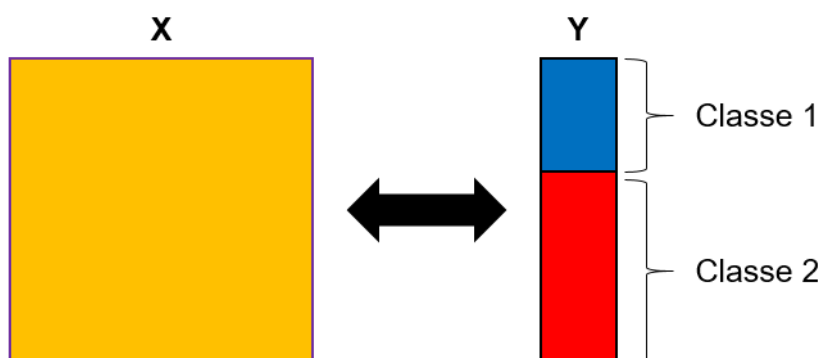
### 2.5.2.3 Análise discriminante pelo método de quadrados mínimos parciais

A análise discriminante pelo método de quadrados mínimos parciais (PLS-DA, do inglês: *Partial Least Squares Discriminant Analysis*) é uma variação do método de regressão PLS, introduzido por Wold, Sjöström e Eriksson (2001)<sup>71</sup>. Através de combinações lineares das variáveis originais, conhecidas como variáveis latentes (VLs), é possível obter uma visualização gráfica e entender os padrões dos dados e relações entre VLs de escores e pesos. O presente método é baseado no algoritmo PLS2, que opera utilizando diversas variáveis



dependentes Y. A principal diferença é que as variáveis, na PLS-DA, são qualitativas (classes), e não quantitativas, como numa regressão PLS tradicional <sup>72</sup>.

Na PLS-DA, posteriormente a definição das classes do conjunto de treinamento, o vetor de classes é desdobrado, e o algoritmo PLS2 aplicado nas respostas qualitativas multivariadas. PLS-DA retornará então com um valor  $y^{calc}$  para cada amostra, com valores entre 0 e 1. Se o valor for próximo a 0, então amostra não pertencerá a tal classe; já se o valor for próximo a um, indica que a amostra pertence a mesma. A Figura 17 representa um modelo PLS-DA para duas classes distintas. Um *threshold* (limite) pode ser definido para cada uma das classes, sendo que se  $y^{calc}$  for maior que o limite estabelecido para a classe, então a amostra será atribuída a tal, caso contrário não <sup>70,72</sup>.



**Figura 17.** Representação do modelo PLS-DA para duas classes.

## 2.6 Espectroscopia no infravermelho e quimiometria

Muitos estudos combinam técnicas espectroscópicas de NIR e ATR-FTIR, aliadas a ferramentas quimiométricas na identificação de medicamentos e drogas adulteradas, como observado na Tabela 2.

Deconinck e colaboradores (2018) <sup>77</sup>, ao utilizar dados de infravermelho médio em conjunto ao modelo PLS-DA, alcançaram 98,3% de acurácia na classificação do conjunto de teste ao diferenciar amostras adulteradas com compostos farmacêuticos e genuínas de suplementos alimentares naturais, além de obterem uma ideia inicial da natureza de seus adulterantes. Rodionova e

outros (2018) <sup>79</sup> empregaram uma variação mais recente do SIMCA, o DD-SIMCA <sup>83</sup>, e infravermelho próximo. O método atingiu 100% de acurácia na classificação de medicamentos anti-histamínicos genuínos e falsificados de alta qualidade. A técnica utilizada no estudo possui resultados que ultrapassam os encontrados quando testes de referência são empregados.

**Tabela 2.** Uso de quimiometria na análise de medicamentos e drogas adulteradas.

<b>Técnica analítica</b>	<b>Amostra</b>	<b>Método quimiométrico</b>	<b>Ref.</b>
<b>ATR-FTIR</b>	MDMA	PLS-DA	73
	Vildagliptina e metformina	PLSR	74
	Lactose	HCA e PCA	75
	Viagra e Cialis	Seleção de variáveis e kNN	76
	Suplementos alimentares	PLS-DA	77
	Ervas medicinais	SIMCA	78
<b>NIR</b>	Anti-histamínicos	DD-SIMCA	79
	Drogas e medicamentos	HCA, PCA e PLSR	40
	Ciprofloxacina	QDA	80
	Cocaína	PLSR e SIMCA	81
	Medicamentos	kNN, QDA, SVM e LDA	82

Deconinck e colaboradores (2018) <sup>77</sup>, ao utilizar dados de infravermelho médio em conjunto ao modelo PLS-DA, alcançaram 98,3% de acurácia na classificação do conjunto de teste ao diferenciar amostras adulteradas com compostos farmacêuticos e genuínas de suplementos alimentares naturais, além de obterem uma ideia inicial da natureza de seus adulterantes. Rodionova e colaboradores (2018) <sup>79</sup> empregaram uma variação mais recente do SIMCA, o DD-SIMCA <sup>83</sup>, e infravermelho próximo. O método atingiu 100% de acurácia na classificação de medicamentos anti-histamínicos genuínos e falsificados de alta qualidade. A técnica utilizada no estudo possui resultados que ultrapassam os encontrados quando testes de referência são empregados.

Correia e colaboradores (2018) <sup>40</sup> realizaram diversos estudos empregando um instrumento portátil de infravermelho próximo (microNIR), muito útil em

atividades forenses *in loco* com amostras de diferentes drogas sintéticas como cocaína, ecstasy, drogas planejadas (*designer drugs*) e medicamentos apreendidos pela polícia. Ao utilizar o modelo de regressão PLS para quantificar amostras de cocaína, não ocorreram diferenças significativas entre a metodologia proposta e resultados obtidos por Cromatografia líquida com detector de arranjo de diodos (LC-DAD). PCA e HCA foram capazes de discriminar entre ácido lisérgico (LSD) e derivados de feniletilaminas N-2-metoxibenzil substituídas (NBOMes), enquanto que os tabletes de ecstasy formaram cinco diferentes grupos diferenciados pela presença de (1) metilendioximetanfetamina (MDMA); (2) anfetamina, estimulantes tipo anfetamina, metanfetamina, MDMA 2C-B; (3) anfetamina e cafeína; (4) MDMA e (5) etilona e cetamina. Já as análises de PCA das amostras de medicamentos indicaram diferentes dosagens do fármaco ativo, mostrando falta de controle de qualidade em sua produção, além de diferentes princípios ativos, formando diferentes agrupamentos no gráfico de escores.

Brito e colaboradores (2020)<sup>76</sup> desenvolveram um método de seleção de variáveis para otimizar a classificação de amostras autênticas e contrafeitas de Viagra e Cialis através do método kNN. Os dados coletados em ATR-FTIR tiveram 0,51% de seus números de onda selecionados pelo modelo para Viagra, resultando em 100% de acurácia. Já para as amostras de Cialis, apenas 0,17% dos números de onda foram retidos, resultando também em 100% de acurácia.

### 3 Materiais e métodos

#### 3.1 Amostras

As amostras previamente pulverizadas de medicamentos foram obtidas junto a Polícia Federal do Estado do Rio Grande do Sul através de apreensões realizadas pela mesma. Esse grupo de amostras constitui-se de fármacos anabolizantes, anorexígenos, antidepressivos e estimulantes comercializados ilegalmente. No total, foram utilizadas 42 amostras nas análises de infravermelho médio e próximo, sendo que os compostos presentes foram previamente identificados por GC-MS, como apresenta a Tabela 3. As amostras da Classe 2 eram identificadas como sibutramina e as da Classe 3 como “Nobésio Forte”, uma marca de rebite conhecida entre os caminhoneiros. Nota-se que as algumas das amostras da Classe 4 também eram comercializadas como o anorexígeno sibutramina e as mesmas continham um composto em grande quantidade que não pode ser identificado durante a cromatografia. As amostras não identificadas eram ilegalmente comercializadas como diferentes compostos anabolizantes.

**Tabela 3.** Amostras analisadas e respectivos princípios ativos encontrados por GC-MS.

Classe	Nº de amostras	Compostos identificados
1	4	femproporex
2	3	benzocaína + cafeína
3	13	lidocaína + cafeína
4	6	fluoxetina
“5”	16	não identificado

#### 3.2 Espectroscopia de infravermelho médio

Os espectros de FTIR foram obtidos em um espectrofotômetro Perkin Elmer modelo Frontier acoplado ao acessório de ATR com cristal de seleneto de zinco (ZnSe). Os registros espectrais foram feitos em triplicata para cada amostra, na faixa de 4.000-750  $\text{cm}^{-1}$ , a temperatura ambiente, com resolução de 2  $\text{cm}^{-1}$  e 32 varreduras.

### **3.3 Espectroscopia no infravermelho próximo**

Os espectros de NIR foram adquiridos por reflexão em um espectrofotômetro Perkin Elmer, modelo Spectrum 400. Todas as amostras foram registradas em triplicata na faixa de 10.000 – 4.000  $\text{cm}^{-1}$ , a temperatura ambiente, com resolução de 2  $\text{cm}^{-1}$  e 32 varreduras.

### **3.4 Análise multivariada**

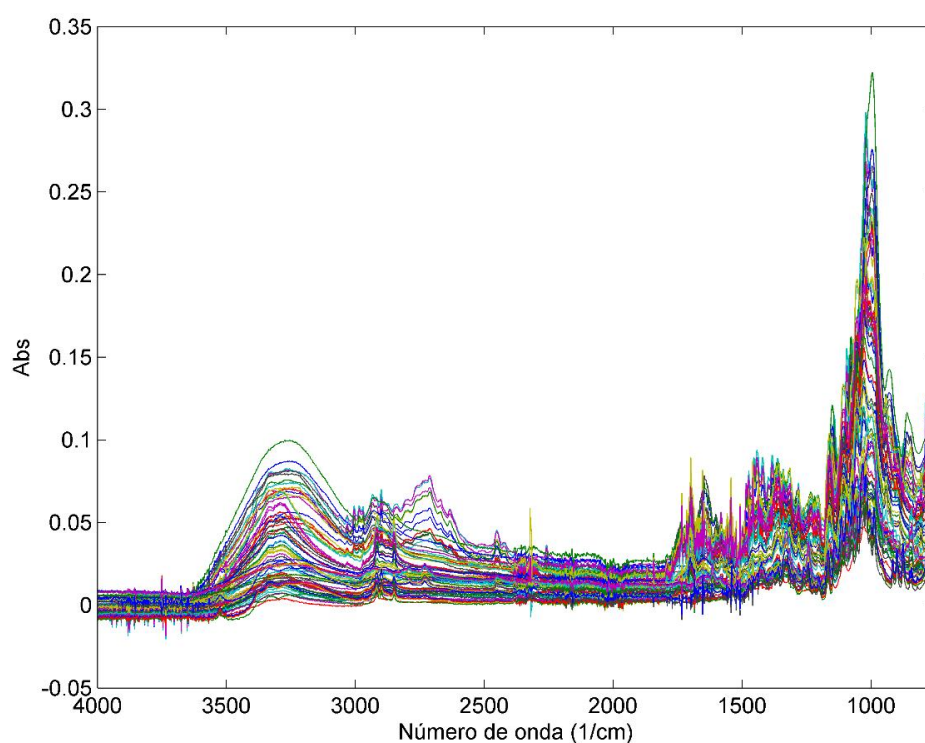
A análise multivariada das amostras dos fármacos foi realizada a partir dos espectros adquiridos por ATR-FTIR e NIR. Métodos não supervisionados (PCA e HCA) e supervisionados (kNN, SIMCA e PLS-DA) foram utilizados na identificação de padrões e classificação das amostras.

O software MATLAB R2012b (The MathWorks, Inc., Natick, Massachusetts, EUA) e toolboxes de métodos não supervisionados <sup>84</sup> e supervisionados <sup>72</sup> foram utilizados na construção dos modelos em um computador com processador Intel Core i7 2,00 GHz.

## 4 Resultados e discussão

### 4.1 Infravermelho médio

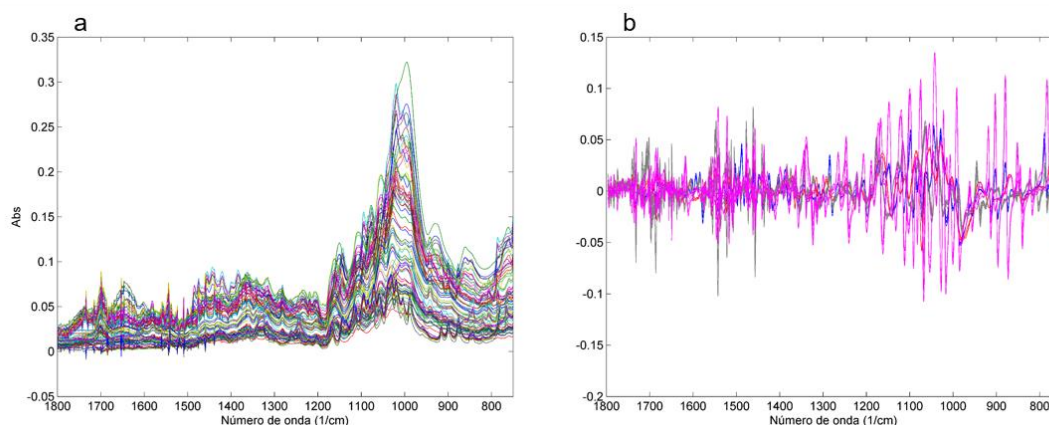
Os espectros das amostras analisadas por ATR-FTIR encontram-se na Figura 18. Por terem sido registradas diretamente, as amostras, além de apresentarem os princípios ativos, possuem também excipientes, o que aumenta a complexidade dos espectros obtidos, contribuindo para a identificação correta do fármaco e implementação da inteligência forense. Uma vez que cada produção apresenta características semelhantes, como tipo de excipiente e sua quantidade, pode haver uma indicação caso diferentes medicamentos fossem produzidos em um mesmo local ou pertencentes a um mesmo lote.



**Figura 18.** Espectros no infravermelho médio das amostras sem pré-processamento.

Para a criação dos modelos quimiométricos, foi selecionada a região de *fingerprint* de 1800 a 750  $\text{cm}^{-1}$ , como apresentado na Figura 19a. Os dados foram então pré-processados por SNV (*Standard Normal Variate*) para a correção de efeitos de espalhamento e tamanho de partícula sólida, posteriormente aplicada 1ª derivada de Savitzky-Golay com janela de 13 pontos e polinômio de 2ª ordem

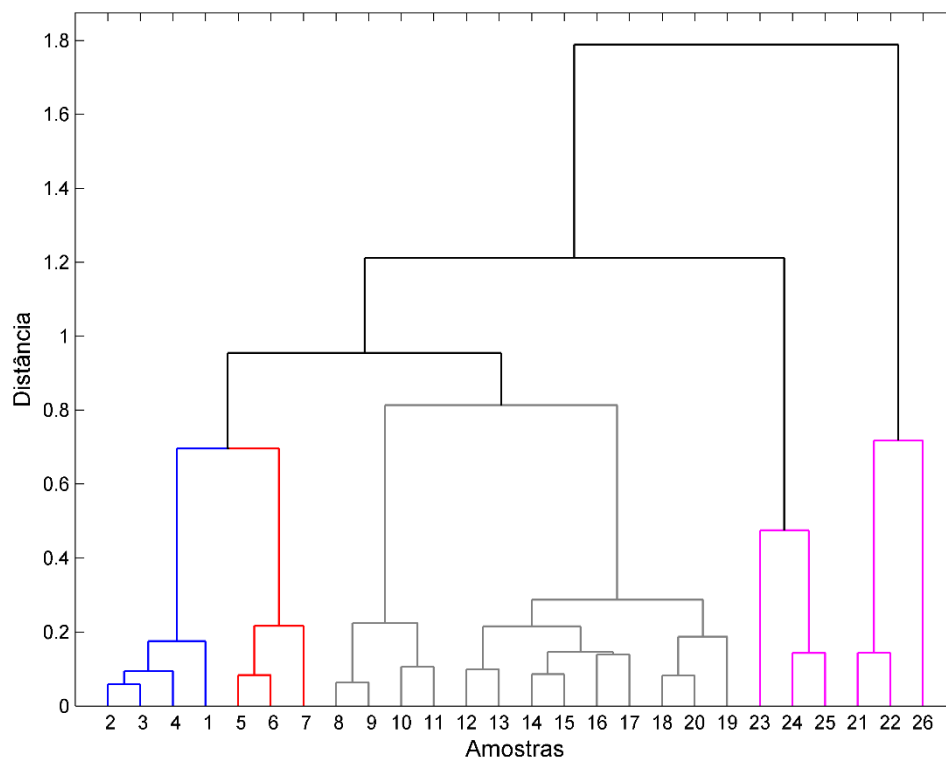
e ao final foram centrados na média. O resultado dos pré-processamentos podem ser visualizados na Figura 19b.



**Figura 19.** Espectros de ATR-FTIR para amostras de fármacos: a) espectros brutos da região espectral de 1.800-750  $\text{cm}^{-1}$  e b) média dos espectros pós-tratamentos SNV e 1ª derivada de Savitzky-Golay com janela de 13 pontos e polinômio de 2ª ordem.

#### 4.1.1 Análise exploratória

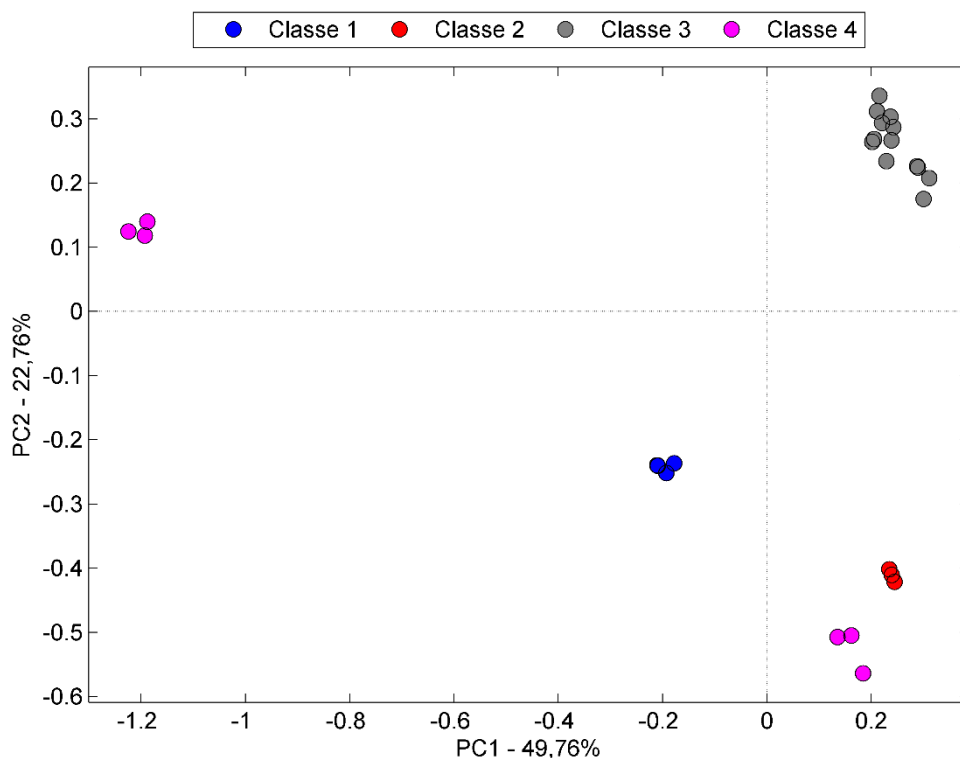
A análise de agrupamentos por métodos hierárquicos é uma técnica exploratória e foi elaborada com intuito de visualizar o comportamento geral inicial dos dados obtidos, uma vez que a HCA utiliza toda a informação proveniente dos espectros de infravermelho. O dendrograma visualizado na Figura 20 foi elaborado com a média dos espectros das amostras previamente pré-processadas. Observou-se que as amostras se agruparam de maneira ordenada de acordo com as classes as quais pertencem, formando quatro grandes grupos. Destaca-se a Classe 3, em cinza, que forma dois subgrupos. Já a Classe 4, em rosa, apresenta-se em dois grupos distintos entre si, apesar de possuírem o mesmo princípio ativo. Esta situação ocorre devido a presença de um composto que não pôde ser identificado por GC-MS em algumas amostras dessa classe, sugerindo que há uma diferença grande entre estas e aquelas que contém apenas fluoxetina como principal princípio ativo.



**Figura 20.** Análise de agrupamentos por métodos hierárquicos dos espectros obtidos por ATR-FTIR.

Na análise de componentes principais pode-se observar a dispersão das amostras de acordo com cada classe farmacêutica. A Figura 21 apresenta o gráfico de escores das duas primeiras componentes principais, representando 72,52% da variância espectral dos dados. Ao analisar PC1 x PC2, observa-se o agrupamento das amostras da Classe 1, que possuem femproporex como seu princípio ativo, nas partes negativas da PC1 e PC2. Amostras da Classe 2 se encontram na parte positiva da PC1 e negativa da PC2, enquanto que a Classe 3, os chamados rebites, encontra-se nas partes positivas da PC1 e PC2. Esta diferenciação entre classes é bastante relevante visto que ambas classes possuem cafeína em sua composição, porém também apresentam benzocaína e lidocaína (Classe 2 e 3, respectivamente).

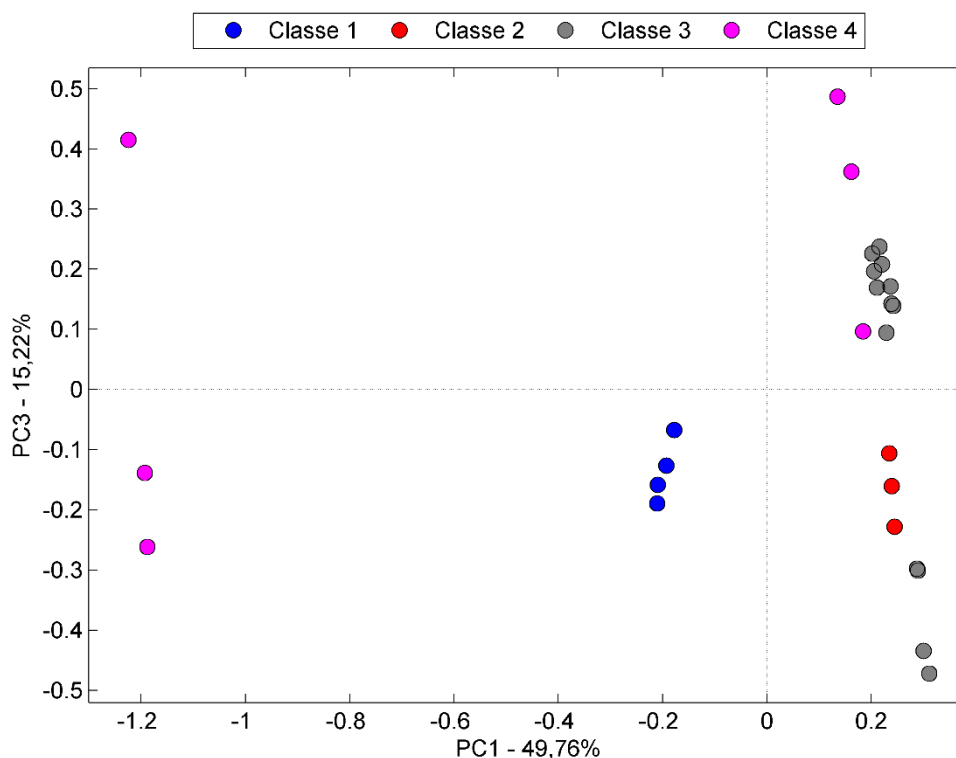




**Figura 21.** Gráfico de escores da análise de componentes principais para PC1xPC2.

Amostras da Classe 4 se subdividiram em 2 grupos: o primeiro na parte positiva da PC1 e o segundo na parte negativa. Através da análise de GC-MS, além de fluoxetina, foi detectado um composto desconhecido em grande quantidade cuja identificação não foi realizada, o que, indica uma falsificação grosseira do medicamento em questão, uma vez que estas amostras eram comercializadas como sibutramina.

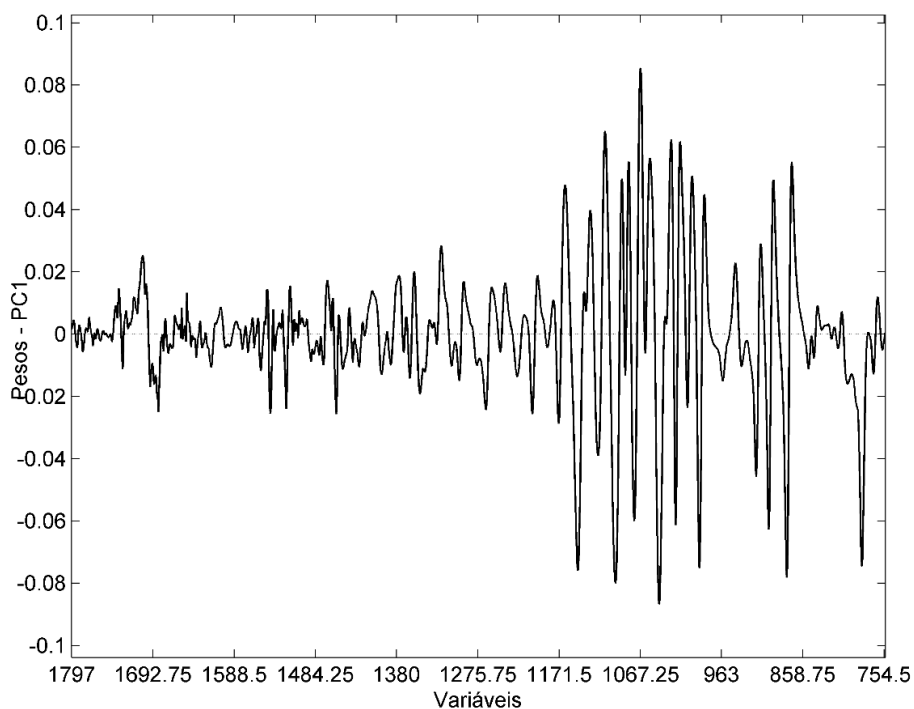
A Figura 22 apresenta o gráfico de escores da PCA das PC1 e PC3. Destaca-se a separação da Classe 3, onde algumas amostras se encontram na parte positiva da PC3, enquanto outras apresentam-se na parte negativa. A diferenciação intraclasses também foi observada previamente no dendrograma da HCA (Figura 20). Esta divisão ocorre, possivelmente, a uma diferença de fabricantes, pois todas as amostras pertencentes a esta classe possui os mesmos princípios ativos - lidocaína e cafeína - identificados através de GC-MS.



**Figura 22.** Gráfico de escores da análise de componentes principais para PC1xPC3.

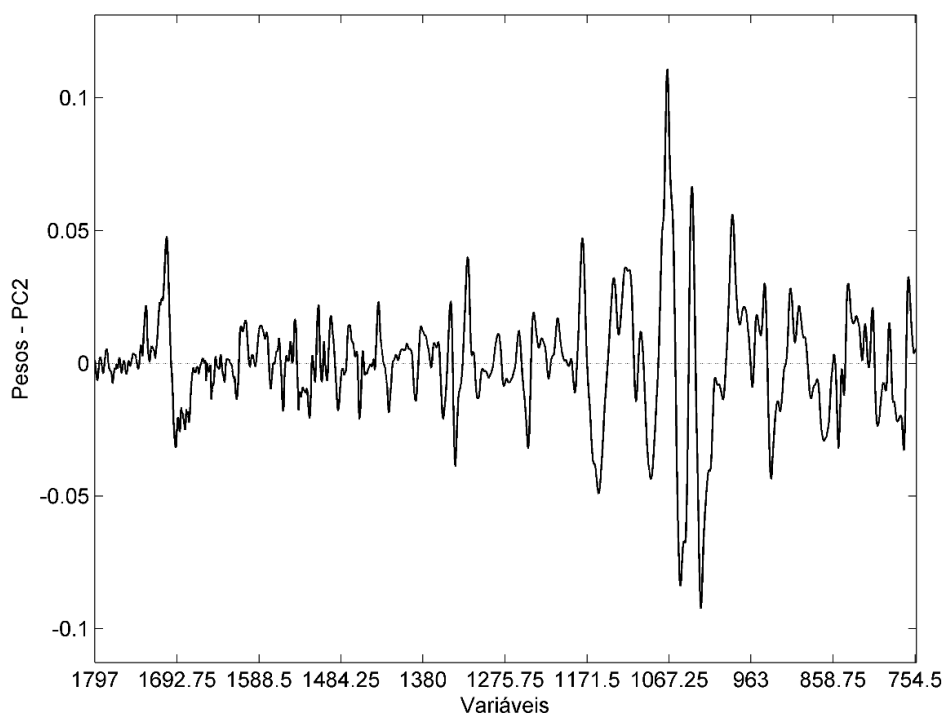
Analisando o gráfico de pesos, apresentados na Figura 23, observa-se que na primeira componente principal, os números de onda que possuem mais peso estão entre 1.200 e 850  $\text{cm}^{-1}$ , região que designa, sobretudo deformações angulares e axiais de ligações C-C, C-O, C-N e C-X. A parte negativa desta região é caracterizada, em especial pelo estiramento O-CH<sub>2</sub>-CH<sub>3</sub> (1.100  $\text{cm}^{-1}$ ) presentes na benzocaína, estiramento simétrico C-O (1.076 e 1.042  $\text{cm}^{-1}$ ) existente tanto no grupo éster da benzocaína quanto no éter na fluoxetina, e deformação fora do plano da ligação C-H (991  $\text{cm}^{-1}$ ), características de aromáticos, além de apresentar uma banda característica do grupo CF<sub>3</sub> em 782  $\text{cm}^{-1}$ . O grupo CF<sub>3</sub> está presente na fluoxetina, da Classe 4. A diferenciação dessa classe na PCA está possivelmente relacionada a proporção de fluoxetina no fármaco. Já na parte positiva da PC1, os pesos que mais contribuem para a distribuição apresentam-se nas bandas 1.066  $\text{cm}^{-1}$  (característico de compostos aromáticos), 1.035  $\text{cm}^{-1}$  (estiramento simétrico C-O-C presente em benzoatos,

presente na benzocaína) e  $870\text{ cm}^{-1}$  (deformação fora do plano C-H), relacionados a compostos aromáticos, que estão presentes em todas as classes.



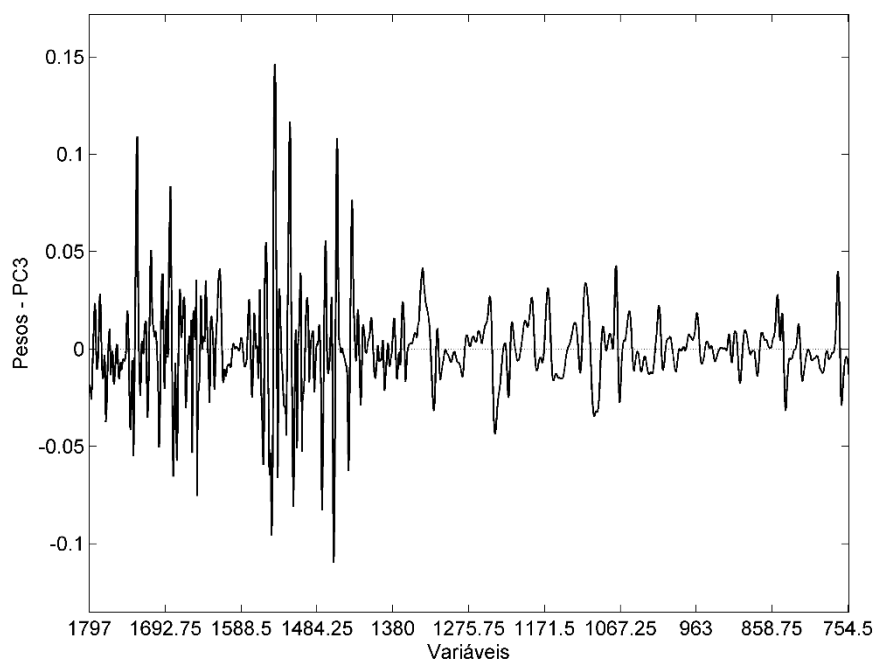
**Figura 23.** Gráfico de pesos da primeira componente principal da PCA da região espectral entre  $1.800$  e  $750\text{ cm}^{-1}$ .

O gráfico de pesos da PC2 na Figura 24 indica que as bandas que mais influenciam na parte positiva são àquelas relacionadas a anéis aromáticos, como em  $1.177\text{ cm}^{-1}$ , que está associada ao benzeno para-substituído, encontrado nas estruturas moleculares da benzocaína e fluoxetina. Além disso, observa-se deformações simétrica fora do plano  $\text{CH}_3$  em  $1070\text{ cm}^{-1}$ , balanço no plano C-H em  $1.037\text{ cm}^{-1}$  e deformação fora do plano C-H em  $985\text{ cm}^{-1}$  <sup>85</sup>.



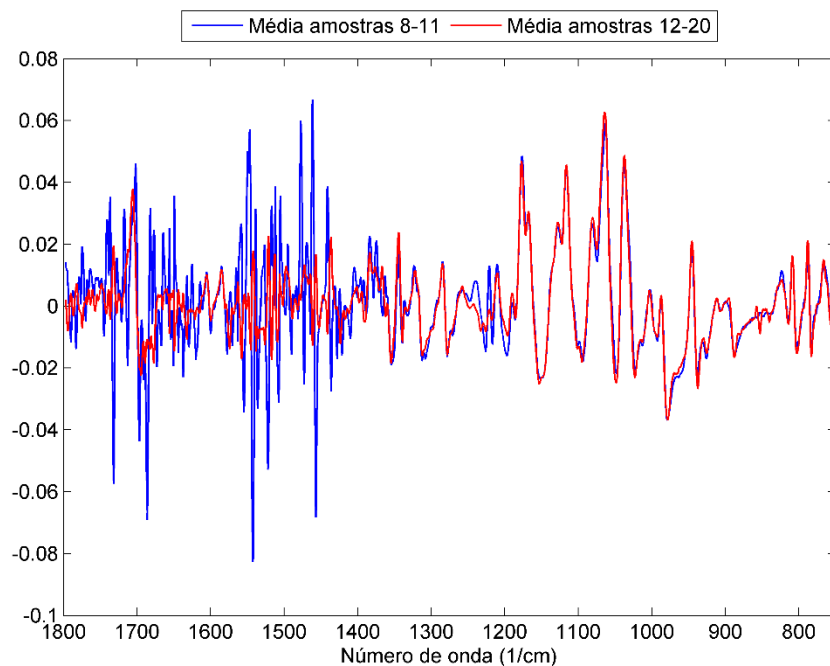
**Figura 24.** Gráfico de pesos da segunda componente principal da PCA da região espectral entre 1.800 e 750  $\text{cm}^{-1}$ .

Os pesos da PC3 (Figura 25) se comportam de forma diferente dos anteriores, uma vez que são maiores entre 1.700 e 1.450  $\text{cm}^{-1}$ . Esta região evidencia, na parte negativa, a carbonila C=O em 1.681  $\text{cm}^{-1}$  e a vibração C-N-H em 1.546  $\text{cm}^{-1}$  característica das amidas. A parte positiva caracteriza-se pela presença de bandas características de amidas cíclicas, referente a cafeína (1.731  $\text{cm}^{-1}$ )<sup>85</sup>; banda da ligação C-N, proveniente de uma amidina substituída também presente na cafeína (1.685  $\text{cm}^{-1}$ )<sup>86</sup>; estiramentos das ligações CH<sub>2</sub> presentes em carboidratos em 1.460  $\text{cm}^{-1}$ <sup>87</sup>. A presença de carboidratos nas amostras não é atípica, uma vez que muitos medicamentos os utilizam como excipientes. Uma quantidade maior ou menor de carboidratos como a celulose, lactose e amido pode explicar, principalmente, a separação intraclasse da Classe 3, assim como a concentração de cafeína.



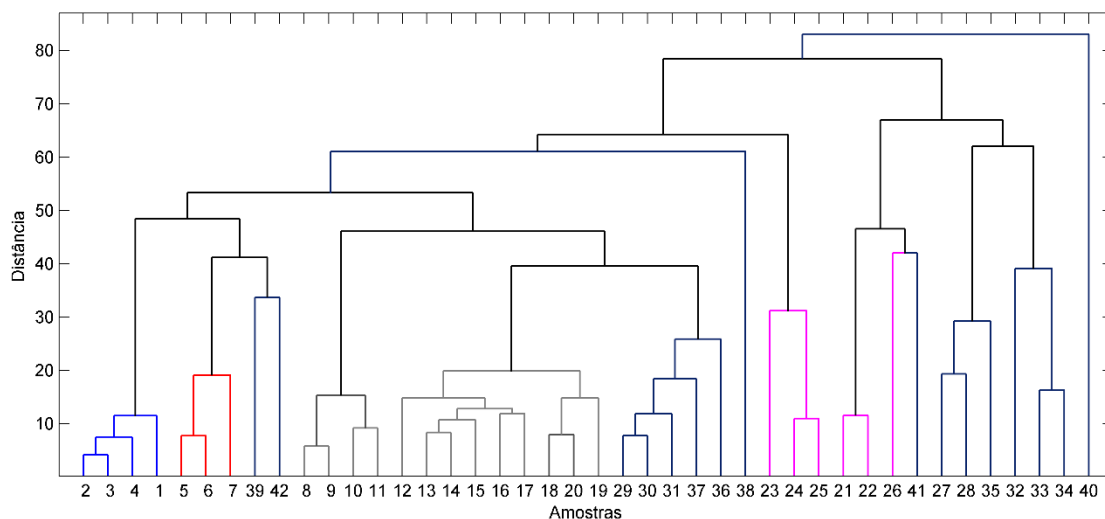
**Figura 25.** Gráfico de pesos da terceira componente principal da PCA da região espectral entre 1.800 e 750  $\text{cm}^{-1}$ .

Observando os espectros pré-processados médios das amostras da Classe 3 na Figura 26 evidencia-se essa diferença, uma vez que as amostras 8 a 11 (em azul) apresentam um perfil espectral com mais intensidade na região entre 1.800 e 1.400  $\text{cm}^{-1}$ . Essa distinção confirma que as amostras são provenientes de diferentes origens. Mariotti e colaboradores <sup>37</sup>, através da análise de componentes principais, discriminaram amostras cujo princípio ativo era o mesmo, indicando que eram de diferentes fabricantes. Também encontraram ligações entre amostras provenientes de diferentes apreensões, concluindo que foi utilizada a mesma formulação farmacêutica para sua fabricação.



**Figura 26.** Média das amostras 8-11 e 12-20, da Classe 3.

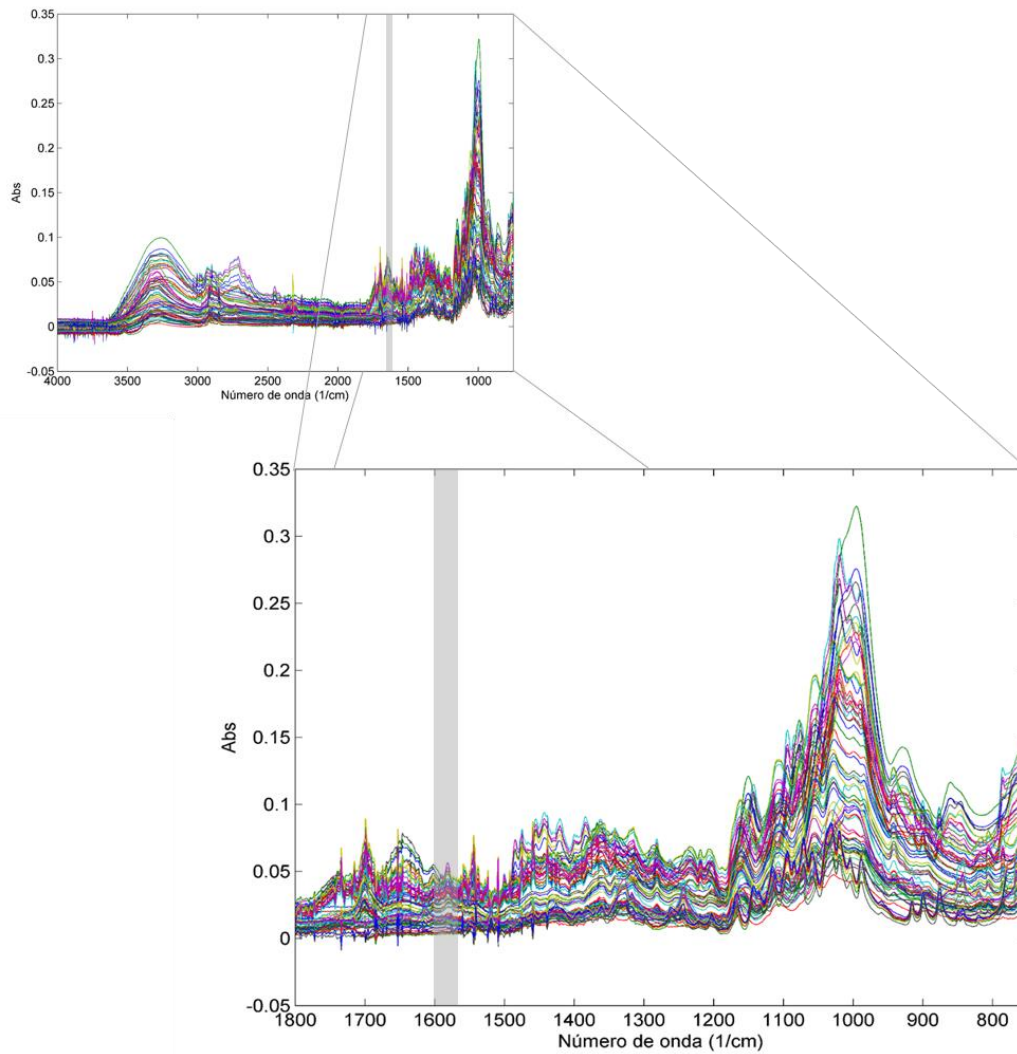
Com o intuito de compreender o comportamento das amostras cuja classificação por GC-MS não foi possível, as mesmas foram agrupadas em uma “pseudo Classe 5”. Utilizando dados da faixa espectral entre 1.800 e 750  $\text{cm}^{-1}$ , construiu-se uma HCA para observar o comportamento das novas amostras em relação as anteriores. Na Figura 27 nota-se que os medicamentos comercializados como anabolizantes, que formam a Classe 5, não se apresentam em um grupo uniforme, misturando-se com as demais amostras. Destaca-se a amostra 41, ilegalmente comercializada como estanozolol, um derivado da testosterona, que foi agrupado ao grupo de amostras que possuíam fluoxetina e uma substância desconhecida, indicando uma possível similaridade em sua composição.



**Figura 27.** Análise de agrupamentos por métodos hierárquicos dos espectros obtidos por ATR-FTIR com a “pseudo Classe 5”, em azul escuro.

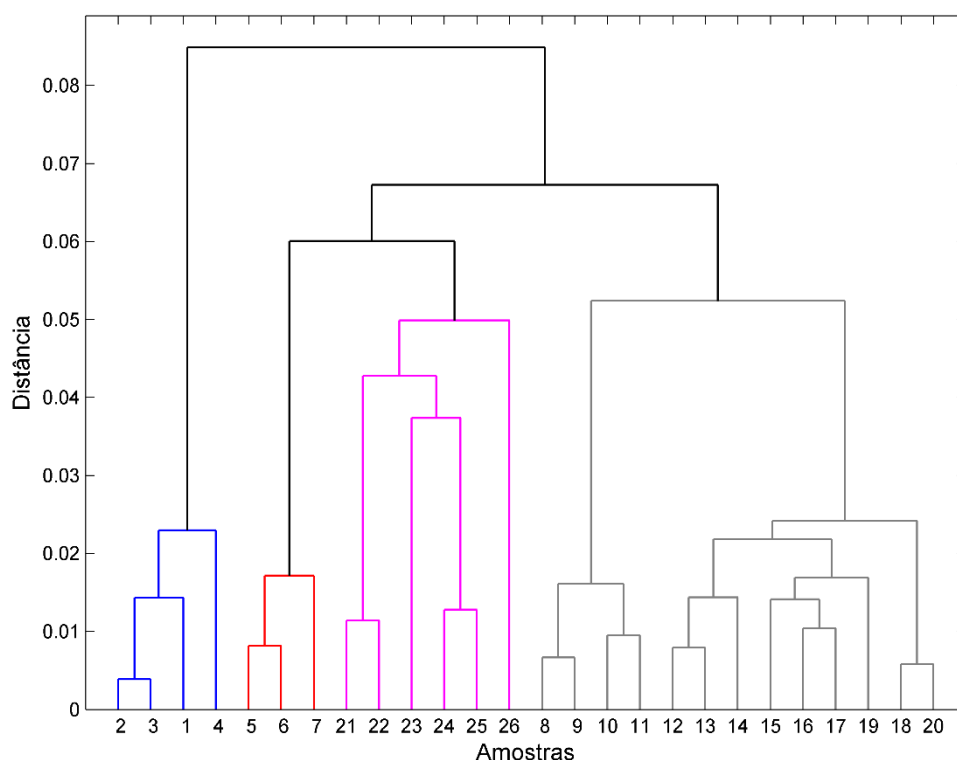
A iPCA é uma técnica similar a PCA, porém o algoritmo divide o espectro em regiões ( $i$ ), onde uma análise de componentes principais é realizada para cada intervalo. Por exemplo, em uma iPCA onde  $i=2$  as variáveis serão divididas em duas regiões e uma PCA será aplicada em cada uma delas separadamente. Para as amostras analisadas, o  $i$  foi de 2, 4, 8, 12 e 16 nos espectros de ATR-FTIR, utilizado a ferramenta iPCA do iToolbox para Matlab <sup>88</sup>.

Através da iPCA, selecionou-se o intervalo espectral que melhor separou as classes de acordo com seus princípios ativos. A Figura 28 apresenta a região espectral de  $1.600 - 1.567 \text{ cm}^{-1}$ . Esta região corresponde principalmente a ligações C=C de aromáticos e a deformação angular simétrica no plano da  $\text{NH}_2$  <sup>85</sup>. Os princípios ativos identificados por CG-MS femproporex, benzocaína, lidocaína e fluoxetina possuem anéis aromáticos em suas estruturas e apenas a benzocaína possui um  $\text{NH}_2$  em sua estrutura. O dendrograma apresentado na Figura 29 indica que esta região espectral pode agrupar de forma ordenada as amostras estudadas de forma semelhante e mais específica do que ao utilizar a faixa espectral completa.



**Figura 28.** Espectros brutos com destaque para a região de fingerprint utilizada na criação dos modelos (1.800 – 750  $\text{cm}^{-1}$ ) e região selecionada por iPCA (1.600 – 1.567  $\text{cm}^{-1}$ ).

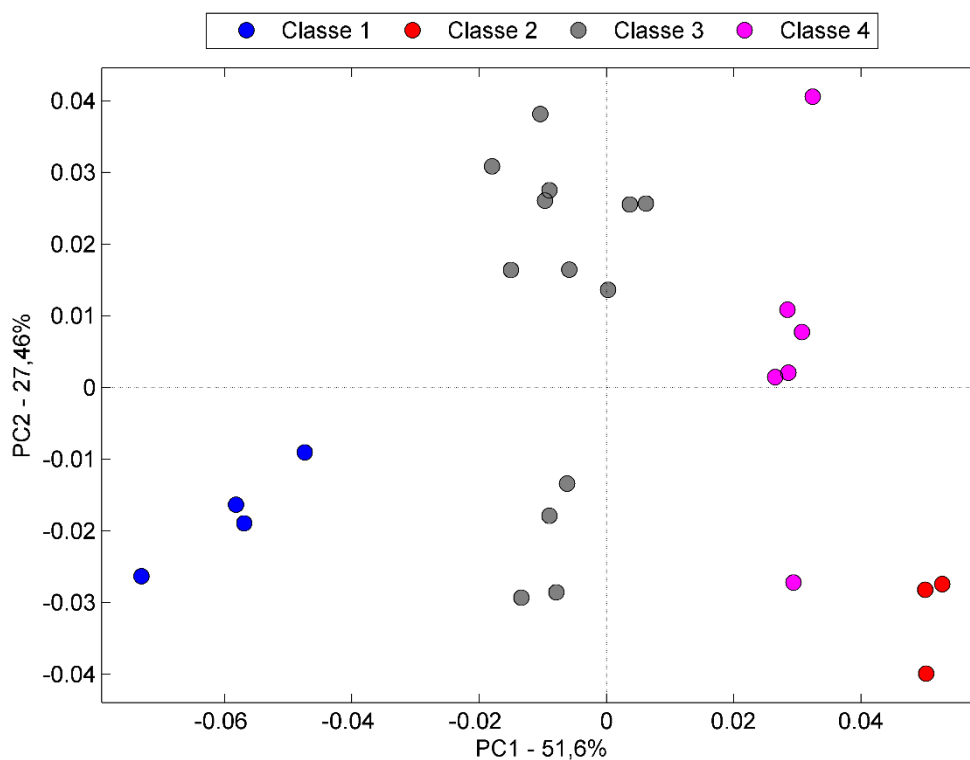




**Figura 29.** Dendrograma resultante da HCA do intervalo espectral selecionado por iPCA (1.600 – 1.567  $\text{cm}^{-1}$ ).

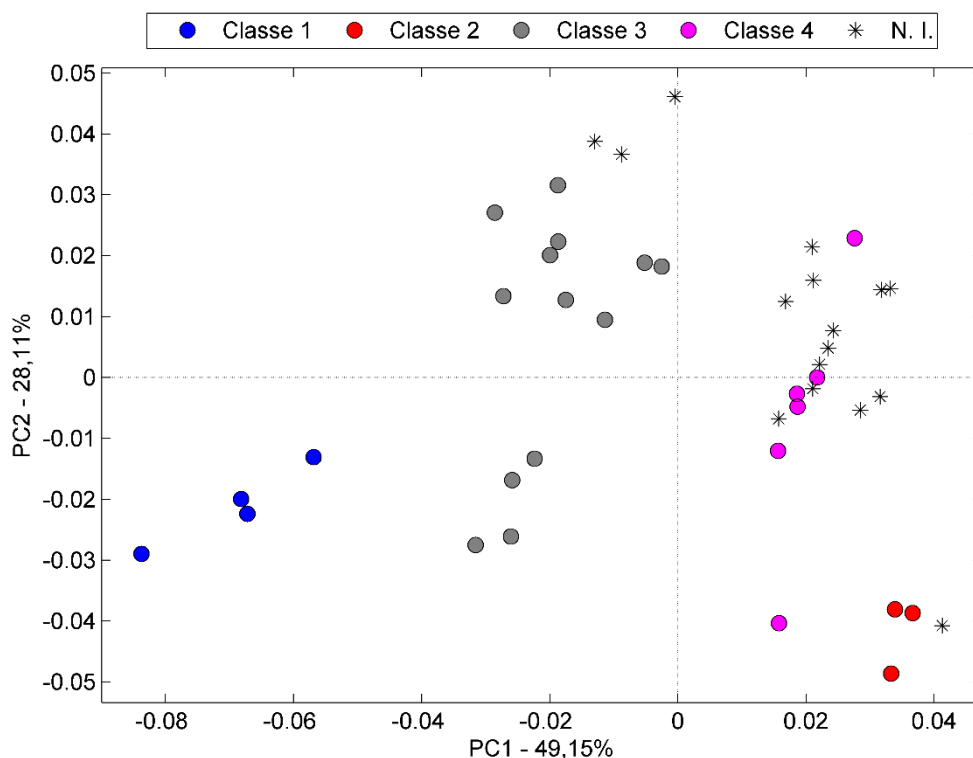
O gráfico de escores das duas primeiras componentes principais, que correspondem a 79,06% da variância dos dados, é apresentado na Figura 30. Observou-se que, utilizando apenas a região espectral selecionada, existe a tendência de formação de grupos de acordo com a classe farmacêutica de cada amostra ao longo da primeira componente principal. Os grupos formados são mais evidenciados do que aqueles formados utilizando toda a faixa espectral de 1.800 a 750  $\text{cm}^{-1}$ , em especial a Classe 4. Essa classe possui o mesmo princípio ativo, porém também apresenta um composto não identificado em algumas das amostras. A aproximação das amostras dessa classe no gráfico de escores indica que a região espectral selecionada está relacionada ao princípio ativo identificado femproporex. A Classe 3 está novamente dividida em dois subgrupos, indicando que formulações diferentes foram utilizadas no processo de fabricação dessas amostras. Isto demonstra que, possivelmente, as amostras, apesar de possuírem o mesmo princípio ativo, não são de um mesmo lote e/ou fornecedor. Ma e Wang (2015)<sup>89</sup>, ao utilizarem ATR-FTIR para a caracterização de medicamentos de diferentes fabricantes e lotes, obtiveram

resultados satisfatórios com apenas a PCA, conseguindo o produto final de dois fabricantes distintos, porém, não alcançaram sucesso na identificação de diferentes lotes.



**Figura 30.** Gráfico de escores para PC1xPC2 referente ao intervalo espectral selecionado por iPCA (1.600 – 1.567 cm<sup>-1</sup>).

Na Figura 31, as amostras não identificadas apresentaram um comportamento semelhante às anteriores, não indicando a presença de *outliers* ou a formação de grupos específicos. A tendência de formar grupos específicos indicaria que as amostras não identificadas apresentam composição distintas dos demais, uma vez que os grupos formados pelas demais amostras estão relacionados aos princípios ativos das mesmas.



**Figura 31.** Gráfico de escores da PC1xPC2 referente ao intervalo espectral selecionado por iPCA (1.600 – 1.567 cm<sup>-1</sup>) com a adição das amostras da Classe 5.

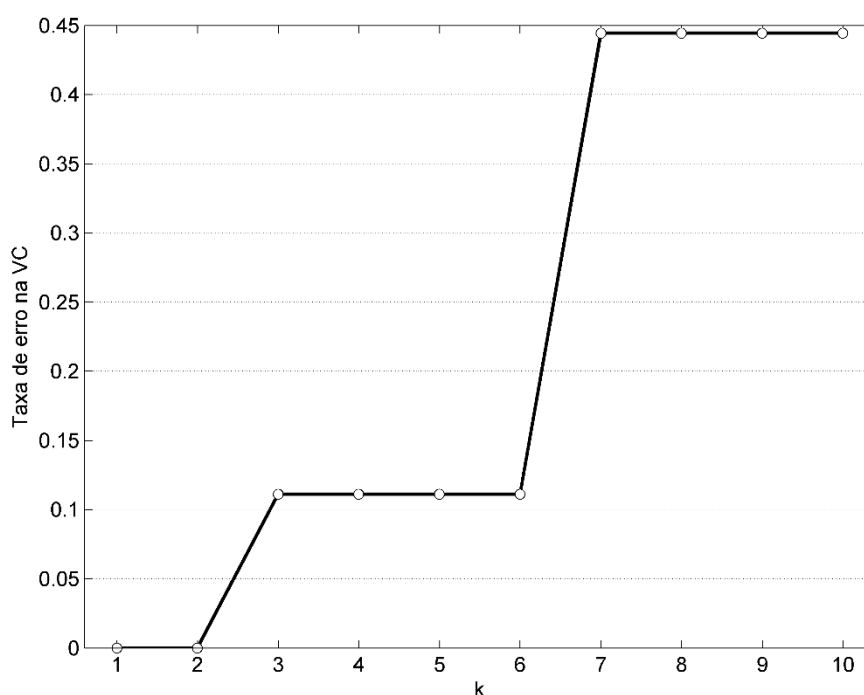
#### 4.1.2 Métodos supervisionados

As análises supervisionadas foram realizadas apenas com a faixa espectral selecionada por iPCA, de 1.600-1.567 cm<sup>-1</sup>. Devido ao baixo número de amostras da Classe 2, as mesmas não foram incluídas nos modelos. Todas as 23 identificadas previamente por CG-MS foram utilizadas no conjunto de treinamento. O modelo foi validado através da validação cruzada. A validação cruzada é uma etapa importante, principalmente em modelos cujo o número de amostras é relativamente baixo.

O método de classificação kNN é bastante simples comparado com os demais métodos supervisionados, sendo similar a HCA. Inicialmente, otimizou-se o método através da escolha do número de vizinhos mais próximos (k), através da construção de modelos de classificação aumentando k

progressivamente, realizando uma validação cruzada e avaliando a taxa de erro de encontrada.

Os resultados obtidos nessa etapa podem ser visualizados na Figura 32. O número de k escolhido foi 2, para que não ocorressem erros durante a previsão de novas amostras devido a classes de diferentes tamanhos. A Tabela 4 apresenta os resultados para a acurácia, sensibilidade e especificidade do modelo kNN proposto. O método obteve 100% de classificação correta na validação cruzada.



**Figura 32.** Taxa de erro na validação cruzada para a determinação do número de vizinhos mais próximos do modelo kNN.

**Tabela 4.** Acurácia, especificidade e sensibilidade do modelo kNN para amostras analisadas por ATR-FTIR utilizando a faixa 1.600-1.567  $\text{cm}^{-1}$ .

	AC <sup>1</sup>	Classe 1		Classe 3		Classe 4	
		ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>
<b>Treinamento</b>	100	100	100	100	100	100	100
<b>VC</b>	100	100	100	100	100	100	100

<sup>1</sup> Acurácia (%); <sup>2</sup> Especificidade (%); <sup>3</sup> Sensibilidade (%);

Com o intuito de direcionar futuras análises, as amostras que não passaram previamente por identificação por GC-MS foram previstas pelo modelo kNN. Das 16 amostras comercializadas como anorexígenos, 7 foram classificadas como semelhantes às da Classe 1, que possui femproporex, e 9 como semelhantes às da Classe 3, que contém uma mistura de lidocaína e cafeína. Destaca-se que o kNN é um método que não é capaz de identificar amostras anômalas e sempre classifica todas as amostras em alguma classe.

No modelo SIMCA foi utilizado o método *leave-one-out* na validação cruzada visando a escolha do menor número de componentes principais para cada classe de modo que a classificação fosse a mais eficiente possível. Dessa maneira, foi selecionada uma componente principal para as Classes 1 e 3, e duas para a Classe 4. A Tabela 5 apresenta as figuras de mérito para este modelo. Observa-se que a classificação foi 100% correta na validação cruzada. Porém, nesta última, duas amostras da Classe 4 não foram classificadas, representando 9% do total de amostras. A utilização do SIMCA na caracterização de medicamentos falsificados é bastante eficiente, embora seja uma técnica mais flexível que as demais, o que tem como consequência a não atribuição de amostras a nenhuma das classes, pois elas podem não pertencer as classes modeladas ou serem semelhantes a mais de uma. Custers e colaboradores (2015)<sup>3</sup> alcançaram resultados similares na identificação de diferentes princípios ativos em amostras de medicamentos para disfunção erétil, embora tenham utilizado um número de componentes principais muito maiores para a construção do modelo com dados de ATR-FTIR.

**Tabela 5.** Acurácia, especificidade e sensibilidade do modelo SIMCA para amostras analisadas por ATR-FTIR utilizando a faixa 1.600-1.567 cm<sup>-1</sup>.

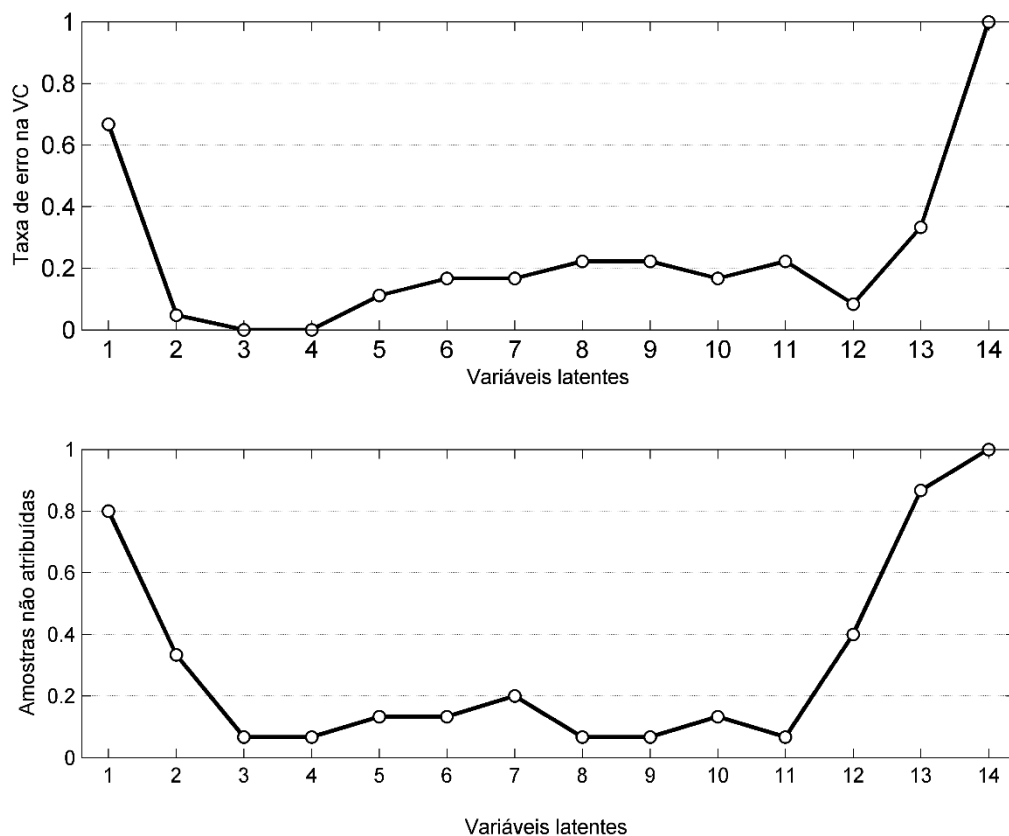
	AC <sup>1</sup>	Classe 1		Classe 2		Classe 3		Classe 4		N.A. <sup>4</sup>
		ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	
<b>Treinamento</b>	100	100	100	100	100	100	100	100	100	0
<b>VC</b>	100	100	100	100	100	100	100	100	100	2

<sup>1</sup> Acurácia (%); <sup>2</sup> Especificidade (%); <sup>3</sup> Sensibilidade (%); <sup>4</sup> n<sup>o</sup> de amostras que não foram atribuídas a nenhuma das classes

A combinação de infravermelho médio e SIMCA também pode ser utilizada como uma técnica de triagem em amostras desconhecidas, como apresentaram Deconink e colaboradores (2017) <sup>78</sup> ao classificar remédios tradicionais chineses. Amostras desconhecidas foram classificadas pelo modelo e posteriormente foram confirmadas através de GC-MS que pertenciam a classe classificada. Este tipo de aplicação tem como objetivo principal diminuir o tempo necessário para determinação da qualidade do produto.

Ao classificar as amostras não identificadas por CG-MS, o modelo SIMCA proposto não atribuiu para 15 amostras a nenhuma das classes, indicando que não pertenciam as mesmas. Apenas uma amostra fora classificada como pertencente à Classe 1. A ferramenta SIMCA realiza a classificação das amostras formando regiões de confiança de cada classe no espaço das componentes principais. Desse modo, ao realizar uma triagem dessas amostras, o indicativo seria que as mesmas não apresentam os compostos marcadores das classes modeladas, sendo necessário uma nova investigação com outras técnicas, como por exemplo as cromatográficas.

Para o modelo PLS-DA, devido ao baixo número de amostras, foi utilizado o método *leave-one-out* na validação cruzada para a escolha do número de variáveis latentes. Esta fase na elaboração de um modelo de classificação tem como finalidade a definição do número de variáveis latentes baseada na avaliação da magnitude dos erros de previsão. A escolha do número de variáveis latentes é de grande importância uma vez que é responsável pela quantidade de informação que será utilizada no modelo. Se o número de VLS for muito alto ocorre um sobreajuste, de modo que o ruído instrumental é inserido no modelo, fazendo com que os resultados das calibrações sejam supermodelados, não condizendo com a realidade. Já se o número de variáveis latentes for muito pequeno, ocorre um subajuste, fazendo com que o modelo tenha baixa capacidade preditiva. A Figura 33 mostra os resultados da validação cruzada para a escolha do número de variáveis latentes utilizadas no modelo PLS-DA. Para tal, foram utilizadas 3 variáveis latentes.



**Figura 33.** Taxa de erro e de amostras não atribuídas na validação cruzada para a determinação do número de variáveis latentes do modelo PLS-DA.

O modelo foi avaliado de acordo com a sua especificidade e sensibilidade, conforme Tabela 6. Observa-se que para todas as classes, tanto a especificidade quanto a sensibilidade atingiram valores expressivos de 100%, indicando a classificação correta das amostras. Comparado com o método SIMCA, o PLS-DA não possui o problema de sobreposição de classes, o que é visível na validação cruzada, onde amostras não atribuídas no modelo SIMCA foram corretamente classificadas, indicando um modelo mais robusto.

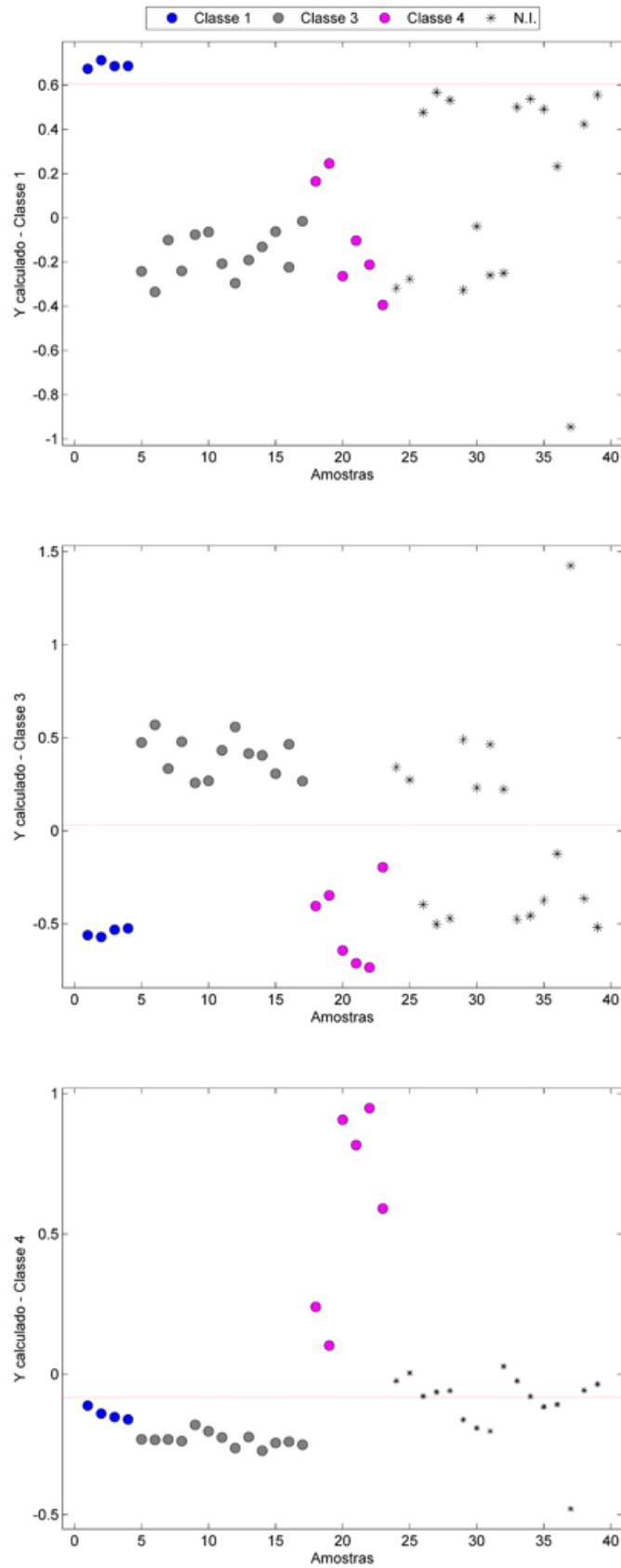
**Tabela 6.** Acurácia, especificidade e sensibilidade do modelo PLS-DA para amostras analisadas por ATR-FTIR utilizando a faixa 1.600-1.567 cm<sup>-1</sup>.

	AC <sup>1</sup>	Classe 1		Classe 3		Classe 4		Não atribuídas <sup>4</sup>
		ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	
<b>Treinamento</b>	100	100	100	100	100	100	100	0
<b>VC</b>	100	100	100	100	100	100	100	0

<sup>1</sup> Acurácia (%); <sup>2</sup> Especificidade (%); <sup>3</sup> Sensibilidade (%); <sup>4</sup> n<sup>o</sup> de amostras que não foram classificadas em nenhuma das classes

Os resultados do modelo PLS-DA para as amostras de treinamento (classes 1, 3 e 4) e predição (classe “5”) são apresentados na Figura 34. O limite (*threshold*) calculado para cada classe aparece na forma de uma linha tracejada vermelha. Das 16 amostras do conjunto de predição, 4 amostras foram classificadas como Classe 3, 7 como da Classe 4, e 5 não foram classificadas. Dentre estas 5 não atribuídas, 3 delas corresponderam com duas classes distintas, aparecendo tanto nos resultados da Classe 3, quanto da Classe 4, indicando que há semelhança entre as amostras, seja em relação ao ingrediente farmacêutico ativo ou seus excipientes. Esses resultados contribuiriam para o início de futura investigação por outras técnicas analíticas a fim de encontrar mais informações sobre essas amostras.





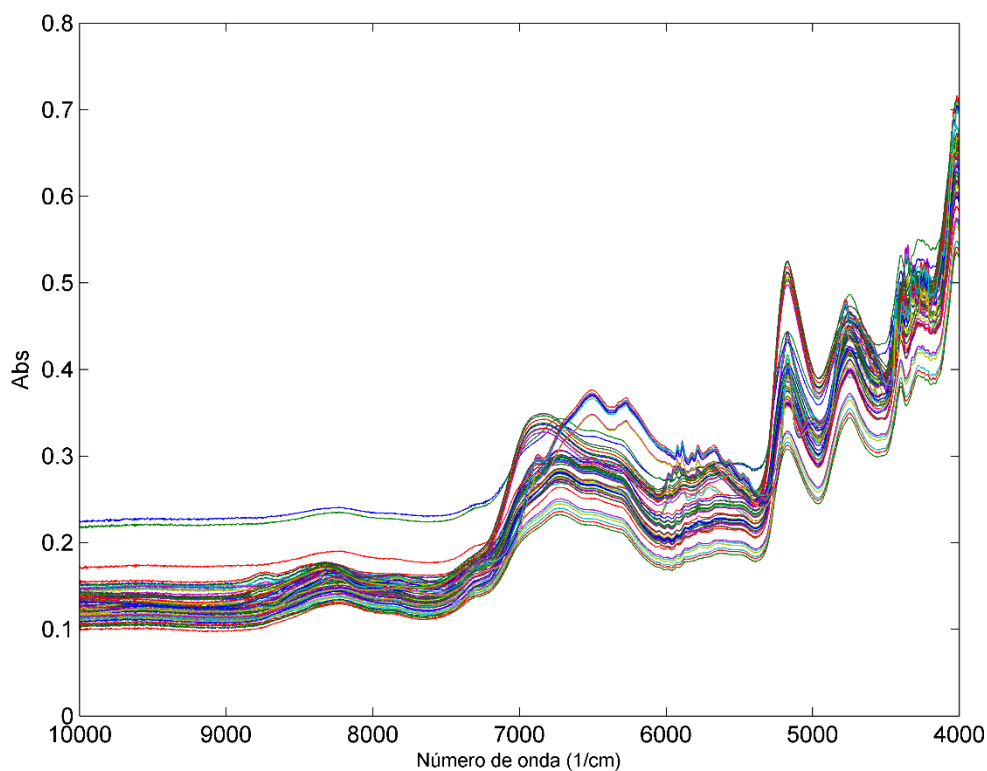
**Figura 34.** Classificação das amostras por PLS-DA utilizando a faixa espectral de 1.600 a 1.567  $\text{cm}^{-1}$ .

O método PLS-DA já é bem estabelecido na identificação de medicamentos e costuma gerar bons resultados, assim como os apresentados no presente trabalho. Outros estudos também abordam a técnica de PLS-DA com um provável substituto a técnicas tradicionais de cromatografia, alcançando uma taxa de acerto de 100% na classificação entre autênticos e falsificados de esteroides anabolizantes <sup>52</sup>.

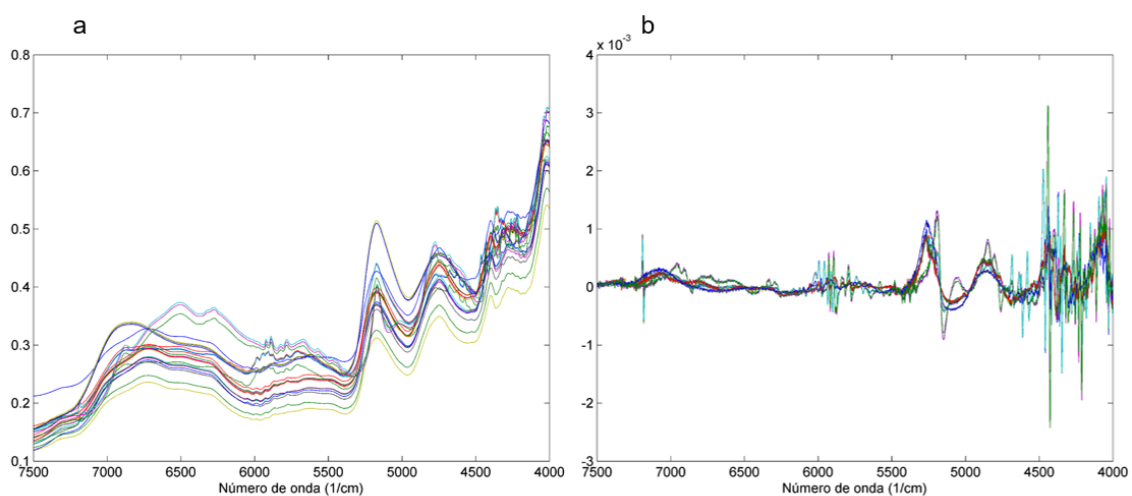
Assim como neste trabalho, que utilizou-se de uma técnica para reduzir o número de variáveis e selecionar aquelas que possuíam informação relevante, Anzanello e seus colaboradores (2014) <sup>92</sup> também apresentaram uma forma de seleção de variáveis na classificação de amostras autênticas e falsificadas de Cialis e Viagra por PLS-DA. Esta redução de variáveis resultou em uma melhora nos resultados, visto que o modelo ficou mais específico. A diminuição dos números de ondas utilizados possibilita também a criação de equipamentos portáteis, pois necessita-se de menos capacidade para o processamento dos dados, em especial em técnicas espectroscópicas que possuem milhares de variáveis.

#### **4.2 Infravermelho próximo**

As amostras das Classes 1,3 e 4 também foram analisadas em triplicata por NIR. Os espectros brutos se encontram na Figura 35. Foi selecionada então a faixa espectral entre 7.500 e 4.000  $\text{cm}^{-1}$  a fim de reduzir a interferência de ruído espectral nos modelos. Os dados foram corrigidos pela transformação MSC (*Multiplicative Scatter Correction*), utilizada para correção de efeitos de espalhamentos causados por fenômenos físicos. Então, foi aplicada a 1ª derivada de Savitzky-Golay com janela de 13 pontos e polinômio de 2ª ordem e centrados na média. Posteriormente, foi realizada a média da triplicata de cada amostra. Os espectros resultantes dos pré-processamentos podem ser visualizados na Figura 36.



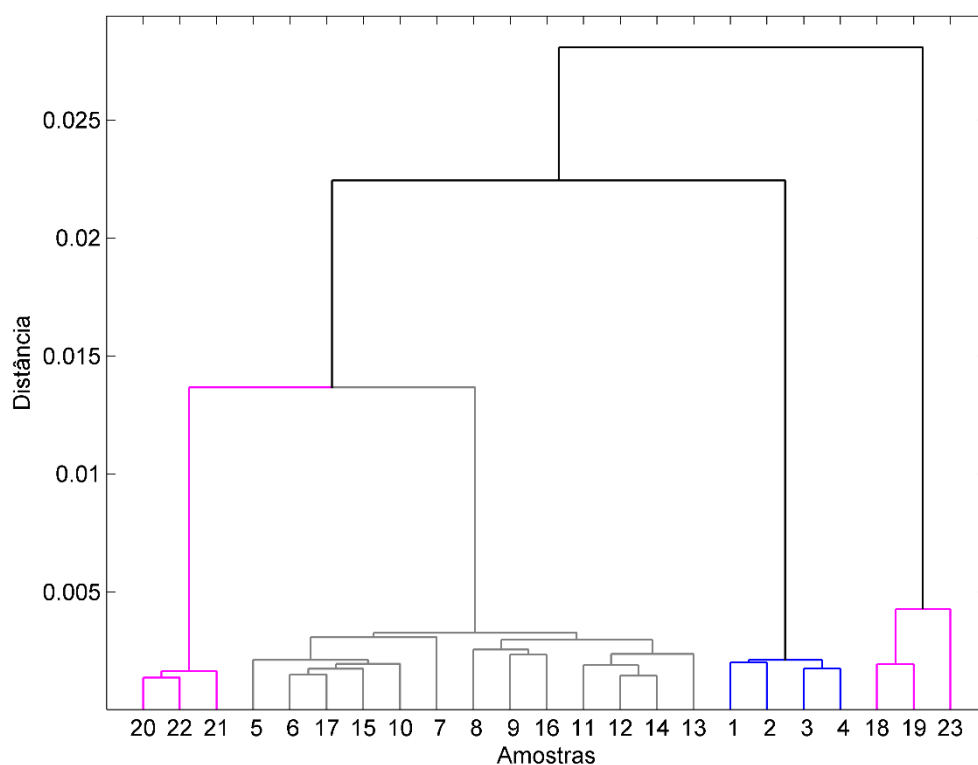
**Figura 35.** Espectros de infravermelho próximo das amostras sem pré-processamento.



**Figura 36.** Espectros de NIR para amostras de fármacos: a) espectros brutos da região espectral de 7.500 a 4.000  $\text{cm}^{-1}$  e b) média dos espectros pós-tratamentos MSC e 1ª derivada de Savitzky-Golay com janela de 13 pontos e polinômio de 2ª ordem.

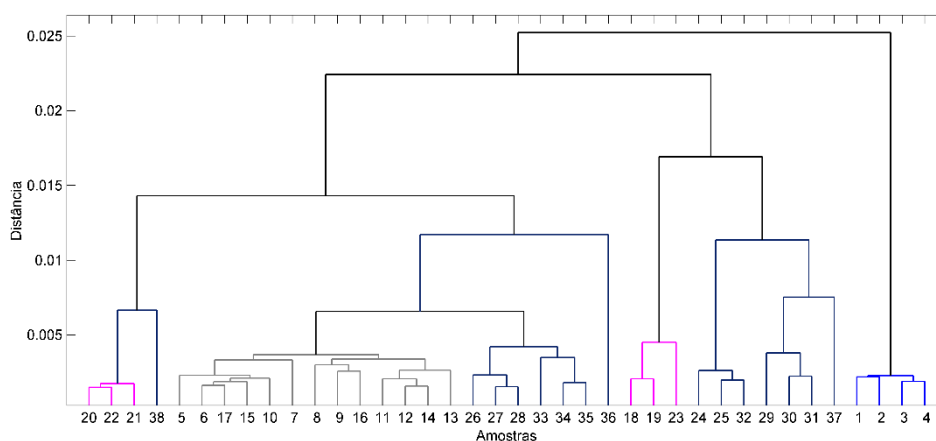
### 4.2.1 Análise exploratória

Os métodos de análise exploratória HCA e PCA foram utilizados como ferramentas para visualização de tendência de comportamento das amostras. A Figura 37 apresenta o dendrograma obtido na HCA para as amostras das três classes analisadas por NIR. Observa-se que ocorre formação de quatro grupos e que, assim como nas amostras analisadas por ATR-FTIR, a Classe 4 subdividiu-se em dois grupos. Esta diferenciação, que se deu devido à presença em grande quantidade de outra substância, não identificada. Já a Classe 3, em cinza, subdividiu-se em dois pequenos grupos, diferentes daqueles obtidos utilizando os espectros do infravermelho médio. Esta mudança de resultados em uma mesma classe se dá, possivelmente, devido a diferença de sensibilidade dos métodos analíticos utilizados em relação aos princípios ativos estudados.



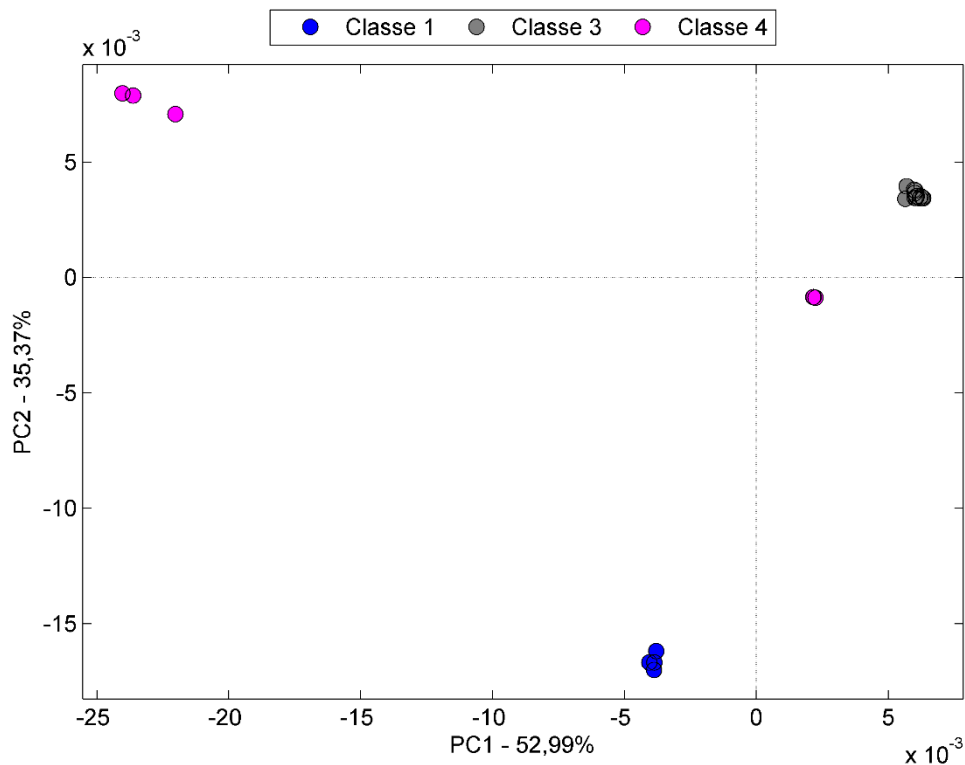
**Figura 37.** Análise de agrupamentos por métodos hierárquicos dos espectros obtidos por NIR.

A Figura 38 apresenta também as amostras da pseudo “Classe 5”. Nota-se que as amostras da Classe 5 (azul escuro), que não tiveram seus componentes identificados por cromatografia, agruparam-se com as amostras da Classe 3 (cinza), que possuem lidocaína e cafeína em sua composição e que foram comercializados ilegalmente como “rebite”. Esta associação pode indicar que os princípios ativos utilizados sejam semelhantes. Outras amostras da mesma classe se encontram junto as amostras da Classe 4 (rosa), que possuem em grande quantidade uma substância também não identificada.



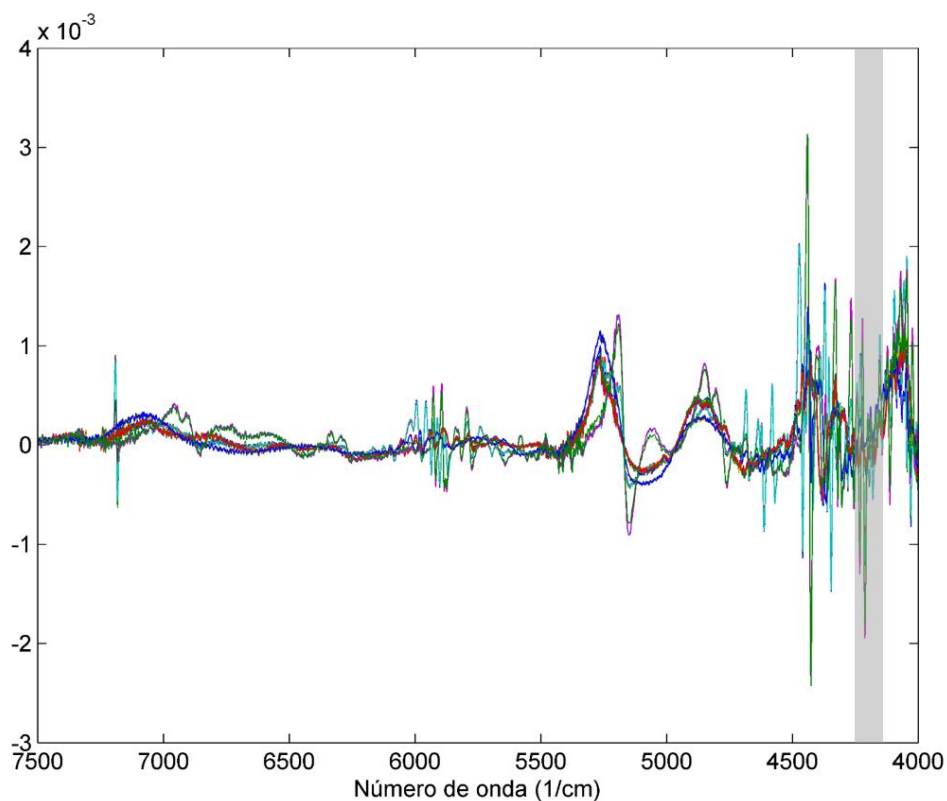
**Figura 38.** Análise de agrupamentos por métodos hierárquicos dos espectros obtidos por NIR, com a Classe 5 em azul escuro.

Na Figura 39 apresenta-se a PCA das amostras analisadas por NIR, sendo que as duas primeiras componentes principais representam 88,36% da variância explicada dos dados. Observa-se que amostras da Classe 1, que possuem femproporex como princípio ativo encontram-se nas partes negativas das PCs 1 e 2, bastante distantes dos demais grupos. Esta diferença se dá, possivelmente, devido a presença de apenas um princípio ativo em sua formulação. Já a Classe 3 apresenta-se de forma homogênea nas partes positivas das componentes principais 1 e 2, diferentemente da PCA dos dados de ATR-FTIR (Figura 21), indicando que no infravermelho próximo a diferenciação se dá, principalmente, pelo princípio ativo e não os excipientes

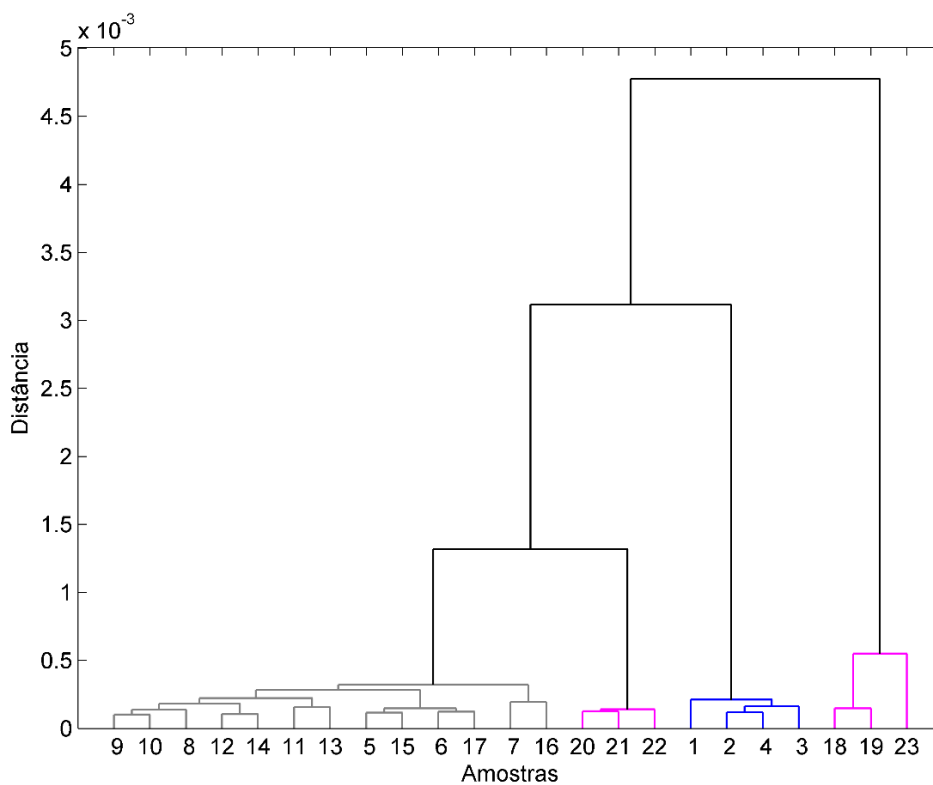


**Figura 39.** Gráfico de escores da análise de componentes principais da PC1xPC2.

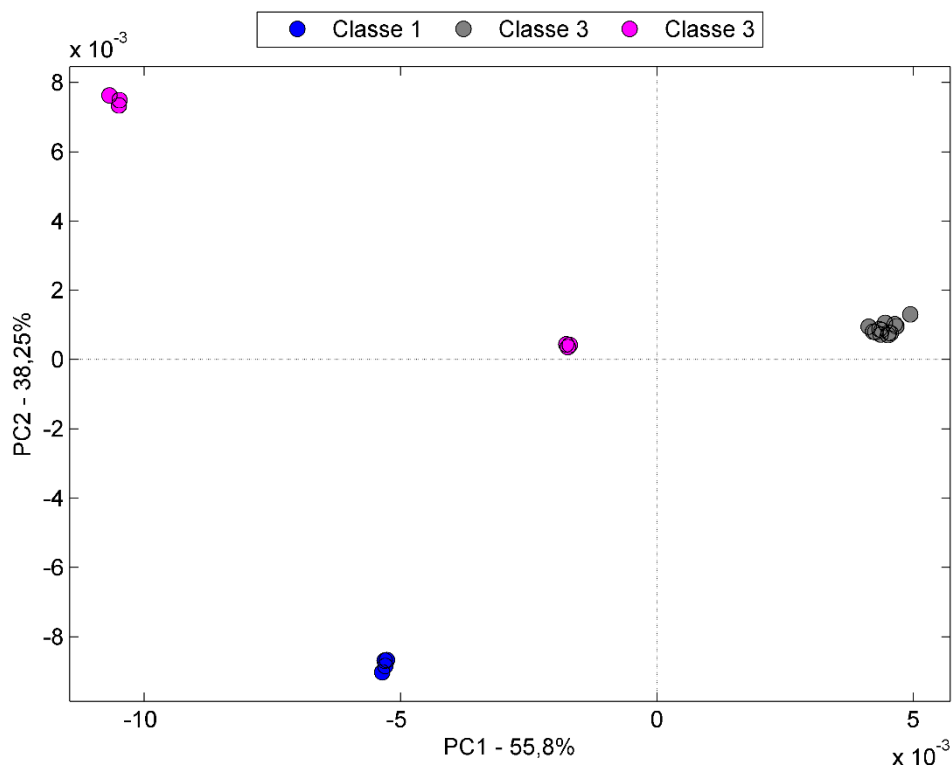
O conjunto de dados obtidos no infravermelho próximo também foi analisado por iPCA. Através do método iPCA foi selecionado o intervalo entre 4438 e 4330  $\text{cm}^{-1}$  (Figura 40) que apresentou melhor representação das classes farmacêuticas trabalhadas, constatada através da HCA e PCA (Figura 41 e Figura 42, respectivamente). As bandas apresentadas nessa região são, principalmente, combinações vibracionais dos grupos funcionais CH, CH<sub>2</sub> e CH<sub>3</sub>. Assim como nas análises de ATR-FTIR, as amostras da Classe 4 formam dois grupamentos distintos, tanto na HCA quanto na PCA. Com a seleção do intervalo mais representativo, também foi possível notar uma maior proximidade das amostras da Classe 5 com as da Classe 3 e 4, observado no dendrograma da Figura 43 e no gráfico de escores da PCA da Figura 44. Este intervalo de dados foi utilizado na construção dos modelos supervisionados kNN, SIMCA e PLS-DA.



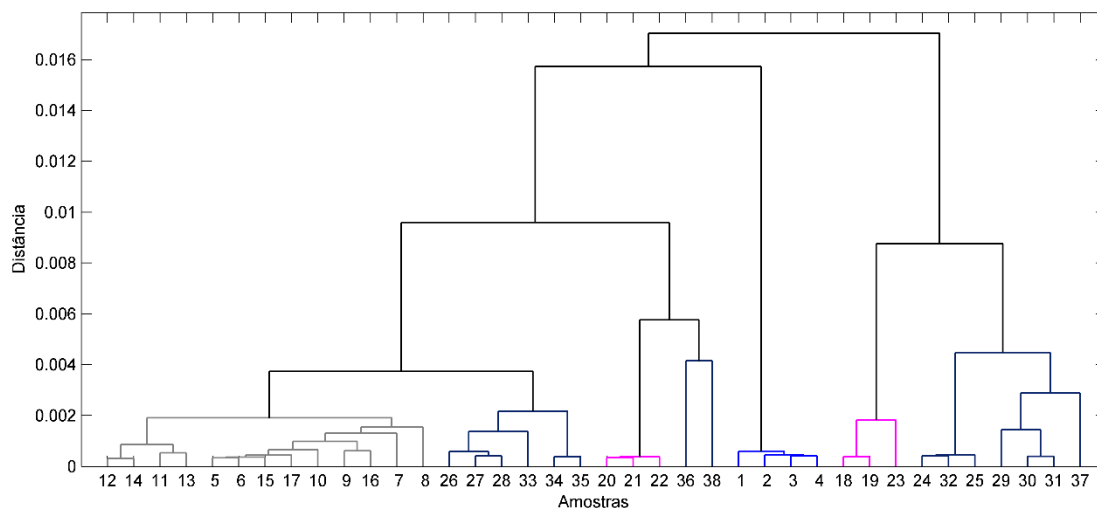
**Figura 40.** Espectros tratados com destaque para a região selecionada por iPCA ( $4.438 - 4.330 \text{ cm}^{-1}$ ).



**Figura 41.** Dendrograma resultante da HCA do intervalo espectral selecionado por iPCA ( $4.438 - 4.330 \text{ cm}^{-1}$ ).

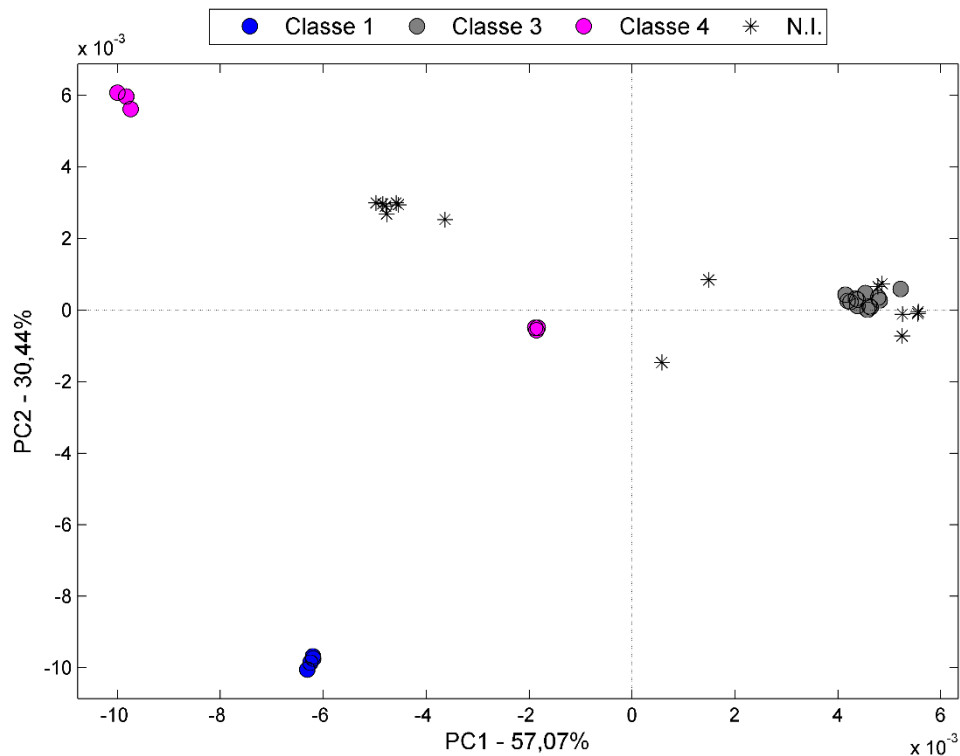


**Figura 42.** Gráfico de escores da PC1 e PC2 referente ao intervalo espectral selecionado por iPCA (4.438 – 4.330 cm<sup>-1</sup>)



**Figura 43.** Dendrograma resultante da HCA do intervalo espectral selecionado por iPCA (4.438 – 4.330 cm<sup>-1</sup>) com a adição das amostras da Classe 5.

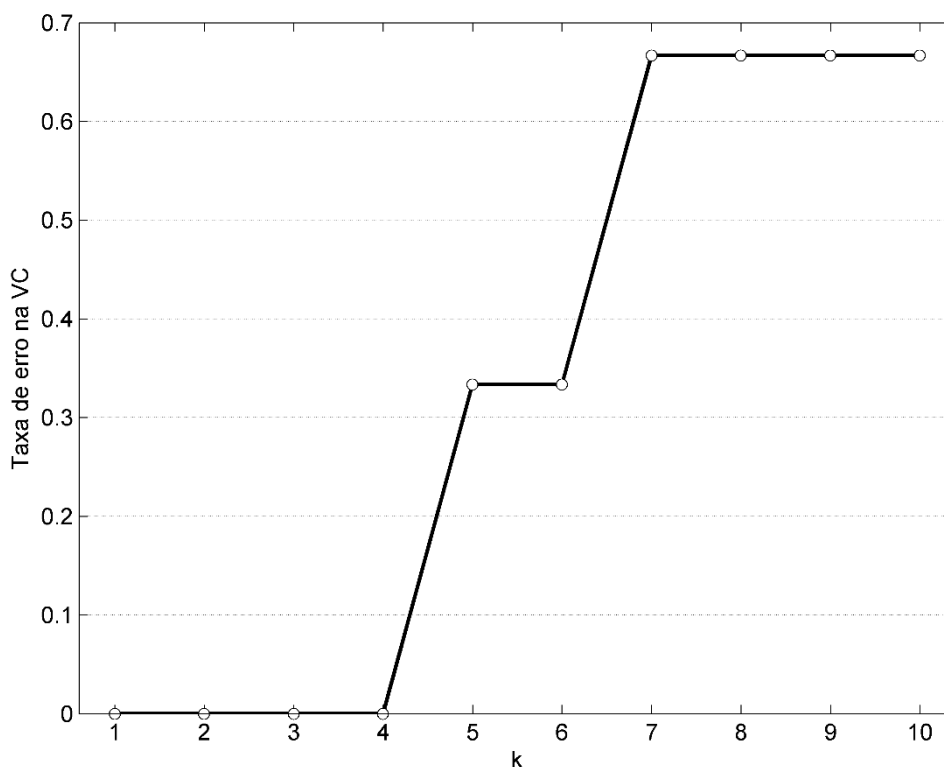




**Figura 44.** Gráfico de escores da PC1 e PC2 referente ao intervalo espectral selecionado por iPCA ( $4.438 - 4.330 \text{ cm}^{-1}$ ) com a adição das amostras da Classe 5.

#### 4.2.2 Métodos supervisionados

Para a elaboração do modelo kNN, inicialmente se otimizou o método buscando-se o número de vizinhos mais próximos cujo erro de validação cruzada fosse menor. O resultado dessa busca encontra-se na **Erro! Fonte de referência não encontrada.** Figura 45. Nota-se que não existe diferença na taxa de erro na escolha de k nos 5 primeiros vizinhos, portanto o modelo foi elaborado com  $k=3$  para que não ocorressem erros ou empates na previsão de novas amostras.



**Figura 45.** Taxa de erro na validação cruzada para a determinação do número de vizinhos mais próximos do modelo kNN.

A Tabela 7 apresenta os resultados de acurácia, sensibilidade e especificidade do modelo kNN no conjunto de treinamento e validação cruzada. Este método simples de classificação obteve 100% de acurácia na validação cruzada. Outros trabalhos utilizando NIR em conjunto com kNN encontraram resultados satisfatórios próximos a 100% de classificação correta de medicamentos falsificados <sup>57,93</sup>.

Na predição de amostras desconhecidas, o método kNN classificou 15 das 16 amostras como Classe 4, que possui fluoxetina em sua composição. Ao analisar resultados anteriores sobre essa classe, principalmente a PCA e HCA, observa-se que ela se apresenta de forma bem heterogênea devido a presença de um componente desconhecido em grande quantidade. Isso pode ter causado a classificação majoritária das amostras nesta classe. Apenas uma das amostras

foi classificada como Classe 3, indicando semelhança com a mesma, que possui lidocaína e cafeína.

**Tabela 7.** Acurácia, especificidade e sensibilidade do modelo kNN para amostras analisadas por NIR utilizando a faixa 4.438 a 4.330  $\text{cm}^{-1}$ .

	AC <sup>1</sup>	Classe 1		Classe 3		Classe 4	
		ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>
<b>Treinamento</b>	100	100	100	100	100	100	100
<b>VC</b>	100	100	100	100	100	100	100

<sup>1</sup> Acurácia (%); <sup>2</sup> Especificidade (%); <sup>3</sup> Sensibilidade (%)

No método de análise supervisionada SIMCA foi empregado o *leave-one-out* na validação cruzada para a seleção do número de componentes principais de cada classe. Para todas as classes, foram selecionadas duas componentes principais, uma vez que este modo apresentou o menor erro na validação cruzada. A Tabela 8 apresenta as figuras de mérito para este modelo. A validação cruzada do modelo apresentou 100% de acurácia.

**Tabela 8.** Acurácia, especificidade e sensibilidade do modelo SIMCA para amostras analisadas por NIR utilizando a faixa 4.438 a 4.330  $\text{cm}^{-1}$ .

	AC <sup>1</sup>	Classe 1		Classe 3		Classe 4	
		ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>
<b>Treinamento</b>	100	100	100	100	100	100	100
<b>VC</b>	100	100	100	100	100	100	100

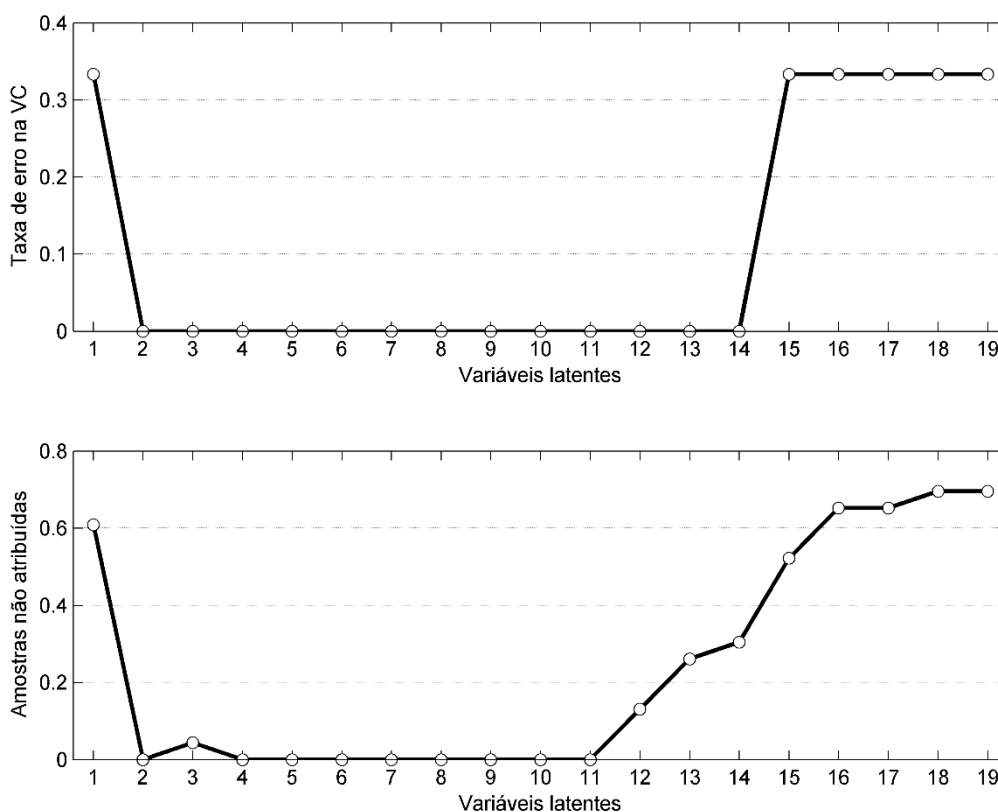
<sup>1</sup> Acurácia (%); <sup>2</sup> Especificidade (%); <sup>3</sup> Sensibilidade (%)

Na predição das amostras desconhecidas, sete amostras foram classificadas como Classe 3 e cinco como Classe 4, enquanto que duas não foram atribuídas. Analisando resultados anteriores de PCA, este comportamento já era esperado, visto que as amostras não identificadas ficam próximas a estas duas classes.

Hespanhol, Pasquini e Maldaner (2019) <sup>81</sup> avaliaram a aplicação de um equipamento de baixo custo e portátil de NIR com a finalidade de classificar diferentes formas de cocaína com o método SIMCA, alcançando resultados expressivos para distintas amostras. Outro trabalho <sup>91</sup> não encontrou diferença

entre os resultados de classificação utilizando NIR de bancada e portátil junto ao método SIMCA, ambos com 88% de classificação correta das amostras externas de antimicrobianos. Estudos como esses comprovam que é possível reduzir o número de dados coletados de forma que isso não prejudique os resultados finais.

Previamente à construção do modelo PLS-DA, utilizou-se do método *leave-one-out* na validação cruzada para a determinação do número de variáveis latentes utilizadas no modelo. Optou-se por empregar 2 VL, uma vez que a taxa de erro foi menor, como apresenta a Figura 46.



**Figura 46.** Taxa de erro e de amostras não atribuídas na validação cruzada para a determinação do número de variáveis latentes do modelo PLS-DA.

A Tabela 9 apresenta os resultados para esse modelo. A validação cruzada apresentou 100% de acurácia, além de atribuir classes para todas as amostras, indicando que o modelo é bastante robusto para amostras das classes trabalhadas.

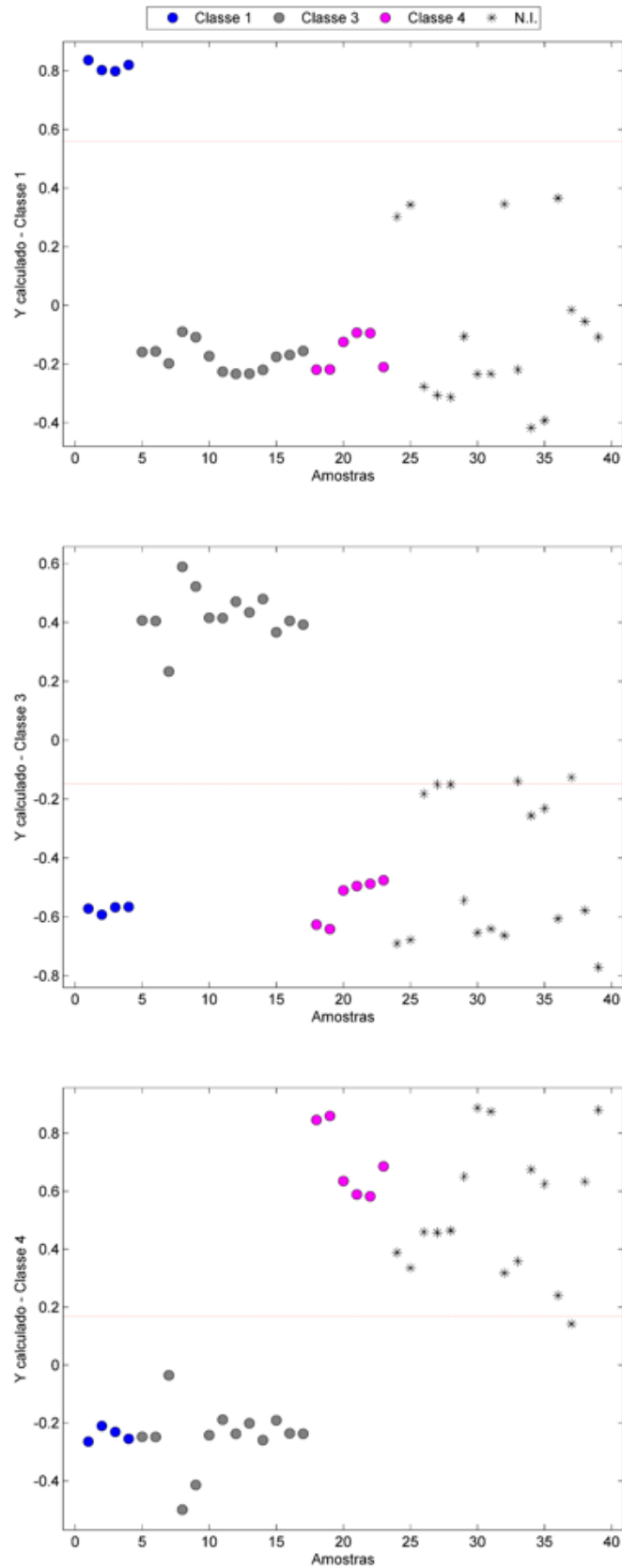
**Tabela 9.** Acurácia, especificidade e sensibilidade do modelo PLS-DA para amostras analisadas por NIR utilizando a faixa 4.438 a 4.330 cm<sup>-1</sup>.

	AC <sup>1</sup>	Classe 1		Classe 3		Classe 4		N.A. <sup>4</sup>
		ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	ESP <sup>2</sup>	SEN <sup>3</sup>	
<b>Treinamento</b>	100	100	100	100	100	100	100	0
<b>VC</b>	100	100	100	100	100	100	100	0

<sup>1</sup> Acurácia (%); <sup>2</sup> Especificidade (%); <sup>3</sup> Sensibilidade (%); <sup>4</sup> n<sup>o</sup> de amostras que não foram classificadas em nenhuma das classes

Na predição de novas amostras (classe “5”), apresentada na Figura 47, o modelo PLS-DA classificou-as majoritariamente como Classe 4. Como estabelecido anteriormente, a Classe 4 é a que possui menos homogeneidade intraclasses. Desse modo, as amostras são mais facilmente classificadas neste conjunto. Destaca-se também a diminuição de sensibilidade do NIR em comparação ao ATR-FTIR, o que poderia afetar estes resultados.

A combinação entre dados de NIR e PLS-DA é bastante comum e geralmente robusta em relação a classificação. Outros estudos, realizados com a combinação de NIR e PLS-DA, alcançaram 100% de classificação correta para amostras de antimicrobianos <sup>91</sup> e para antibióticos <sup>94</sup>.



**Figura 47.** Classificação das amostras por PLS-DA utilizando a faixa espectral de 4.438 a 4.330  $\text{cm}^{-1}$ .

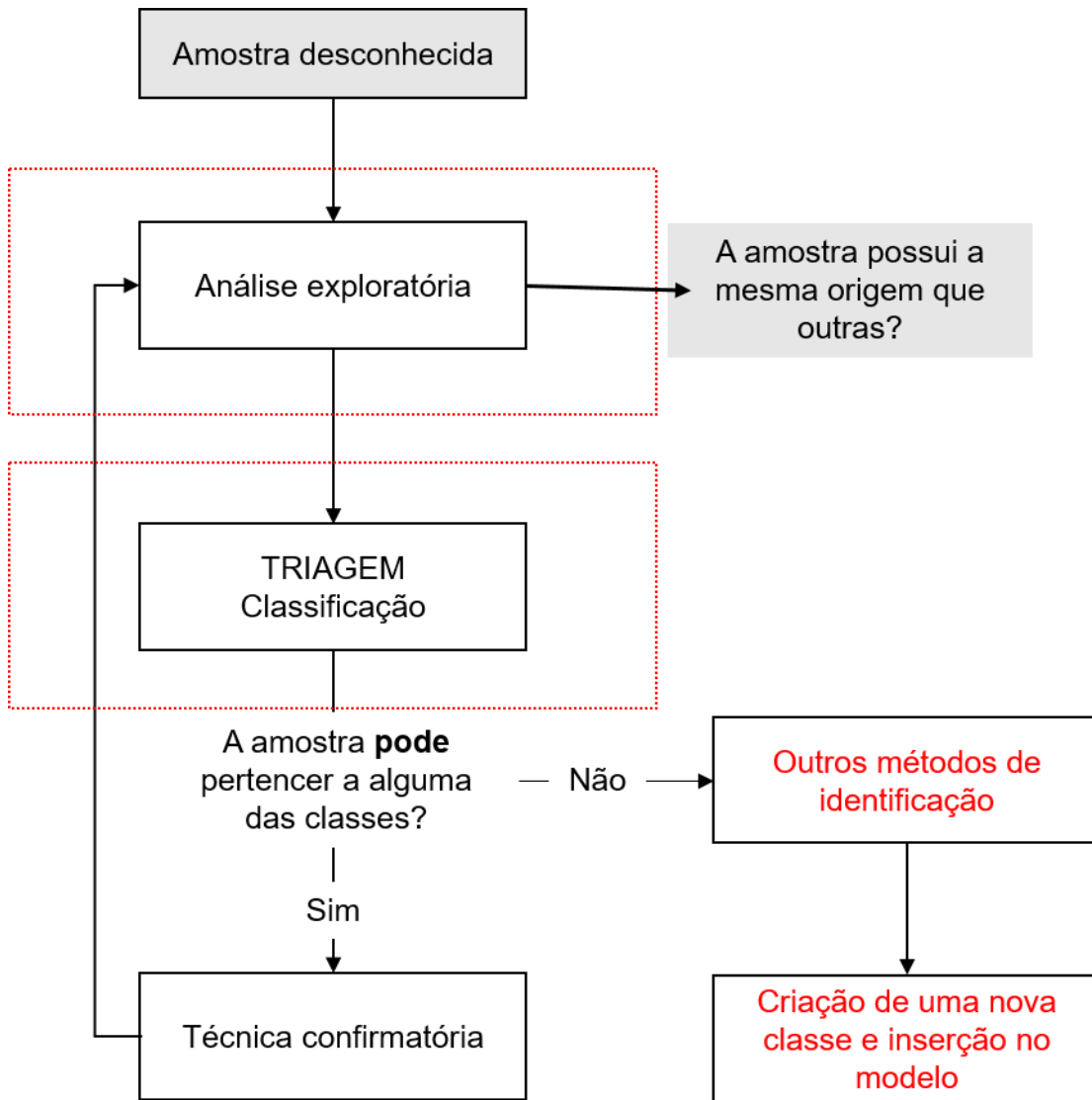
### 4.3 Aplicação na inteligência forense

O emprego dos modelos quimiométricos na ciência forense pode ser esquematizado conforme a Figura 48. A utilização inicial de técnicas espectroscópicas no infravermelho (FTIR e NIR) é sugerida, uma vez que são técnicas rápidas e não destrutivas, além de permitirem a utilização de equipamentos portáteis.

Inicialmente é realizada uma análise exploratória com os dados da amostra desconhecida inserida no conjunto de dados, onde pode-se observar o comportamento geral da mesma. Essa etapa pode ser realizada através de métodos não supervisionados como PCA e HCA.

Após, a classe da amostra pode ser predita, com o intuito realizar uma averiguação inicial com métodos supervisionados PLS-DA e SIMCA, respondendo a seguinte questão: “ela pode pertencer a alguma das classes previamente modeladas?”. Através dessa etapa de triagem, o analista é direcionado ao tipo de fármaco que deve procurar com técnicas mais avançadas, como GC-MS. Se a técnica confirmatória atestar a presença de um ingrediente farmacêutico ativo já modelado é possível inserir esta amostra nos modelos exploratórios, onde averigua a possibilidade da mesma ser da mesma origem de outras amostras previamente analisadas. Caso a amostra não pertença a nenhuma das classes já modeladas, outras técnicas de identificação são utilizadas, de forma que novas classes sejam criadas e inseridas no modelo.

Esse processo iterativo permite que o banco de dados seja constantemente alimentado por novas amostras, proporcionando informações atualizadas nos casos policiais.



**Figura 48.** Processo iterativo da aplicação de métodos quimiométricos na inteligência forense.



## 5 Conclusão

Com os resultados obtidos pelas diferentes técnicas espectroscópicas junto a quimiometria, pode-se definir regiões espectrais de interesse tanto no infravermelho próximo quanto no médio através da iPCA, o que possibilitou uma redução no número de variáveis e identificação de diferentes padrões nas amostras tanto na PCA e HCA, principalmente na Classe 3. Apesar de possuir a mesma mistura farmacêutica, lidocaína e cafeína, esse grupo de amostras apresentou um comportamento que indicou a presença de fabricantes e/ou lotes distintos nos métodos não supervisionados tanto ao utilizar o NIR quanto o ATR-FTIR. Uma possibilidade para esse comportamento da Classe 3 é a presença de distintos excipientes em cada um dos subgrupos formados nessa classe.

Além da seleção de um menor número de variáveis espectrais, os métodos supervisionados KNN, SIMCA e PLS-DA se mostraram muito eficientes para a classificação das amostras das classes pré-estabelecidas. Destacam-se os modelos PLS-DA com dados ATR-FTIR, que apresentaram resultados melhores na validação cruzada, classificando mais amostras que outros métodos supervisionados, sendo, portanto, mais robusto. A PLS-DA também apresentou melhores resultados utilizando dados de NIR. O método kNN, embora seja bastante simples e tenha apresentado excelentes resultados na validação cruzada em ambos os conjuntos de dados, não seria muito útil na identificação de amostras desconhecidas, pois seu algoritmo sempre classifica as amostras em alguma das classes previamente modeladas.

O trabalho também possibilitou a criação de um esquema para a aplicação na inteligência forense, de modo que seja um processo iterativo e constantemente atualizado. A praticidade de análise dos modelos quimiométricos, quando aplicada *in situ* através de equipamentos portáteis, vem a ser muito útil na detecção de amostras falsificadas em operações policiais.

## Referências bibliográficas

- (1) World Health Organization. WHO Global Surveillance and Monitoring System for Substandard and Falsified Medical Products. Genebra, Suíça 2017, p 55. <https://doi.org/978-92-4-151342-5>.
- (2) Dégardin, K.; Roggo, Y.; Margot, P. Understanding and Fighting the Medicine Counterfeit Market. *J. Pharm. Biomed. Anal.* **2014**, *87*, 167–175. <https://doi.org/10.1016/j.jpba.2013.01.009>.
- (3) Custers, D.; Cauwenbergh, T.; Bothy, J. L.; Courselle, P.; De Beer, J. O.; Apers, S.; Deconinck, E. ATR-FTIR Spectroscopy and Chemometrics: An Interesting Tool to Discriminate and Characterize Counterfeit Medicines. *J. Pharm. Biomed. Anal.* **2015**, *112*, 181–189. <https://doi.org/10.1016/j.jpba.2014.11.007>.
- (4) de Peinder, P.; Vredenburg, M. J.; Visser, T.; de Kaste, D. Detection of Lipitor® Counterfeits: A Comparison of NIR and Raman Spectroscopy in Combination with Chemometrics. *J. Pharm. Biomed. Anal.* **2008**, *47* (4–5), 688–694. <https://doi.org/10.1016/j.jpba.2008.02.016>.
- (5) Ferreira, M. M. C. *Quimiometria: Conceitos, Métodos e Aplicações*, 1ª ed.; Editora da Unicamp: Campinas, SP, Brasil, 2015. <https://doi.org/10.7476/9788526814714>.
- (6) Al-Saedi, D. B. A. Chemistry in Forensics and Crime Solving: Review. *J. Glob. Pharma Technol.* **2018**, *10* (3), 545–557.
- (7) Bovens, M.; Ahrens, B.; Alberink, I.; Nordgaard, A.; Salonen, T.; Huhtala, S. Chemometrics in Forensic Chemistry — Part I: Implications to the Forensic Workflow. *Forensic Sci. Int.* **2019**, *301*, 82–90. <https://doi.org/10.1016/j.forsciint.2019.05.030>.
- (8) Brasil. Código de Processo Penal (Decreto-Lei N° 3.689 de 03 de Outubro de 1941). Brasília - DF 1941.
- (9) Romão, W.; Schwab, N. V.; Bueno, M. I. M. S.; Sparrapan, R.; Eberlin, M. N.; Martiny, A.; Sabino, B. D.; Maldaner, A. O. Química Forense: Perspectivas Sobre Novos Métodos Analíticos Aplicados à Documentoscopia, Balística e Drogas de Abuso. *Quim. Nova* **2011**, *34* (10), 1717–1728. <https://doi.org/10.1590/S0100-40422011001000005>.
- (10) Koczwara, A.; Dressman, J. Poor-Quality and Counterfeit Drugs: A Systematic Assessment of Prevalence and Risks Based on Data Published From 2007 to 2016. *J. Pharm. Sci.* **2017**, *106* (10), 2921–2929. <https://doi.org/10.1016/j.xphs.2017.05.018>.
- (11) Koenraadt, R.; van de Ven, K. The Internet and Lifestyle Drugs: An Analysis of Demographic Characteristics, Methods, and Motives of Online Purchasers of Illicit Lifestyle Drugs in the Netherlands. *Drugs Educ. Prev. Policy* **2018**, *25* (4), 345–355. <https://doi.org/10.1080/09687637.2017.1369936>.
- (12) Morelato, M.; Medeiros Bozic, S.; Rhumorbarbe, D.; Broséus, J.; Staehli, L.; Esseiva, P.; Roux, C.; Rossy, Q. An Insight into the Sale of Prescription

- Drugs and Medicine on the AlphaBay Cryptomarket. *J. Drug Issues* **2020**, 50 (1), 15–34. <https://doi.org/10.1177/0022042619872955>.
- (13) Bastos, F. I. P. M.; Vasconcellos, M. T. L. de; De Boni, R. B.; Reis, N. B. dos; Coutinho, C. F. de S. III Levantamento Nacional Sobre o Uso de Drogas Pela População Brasileira. ICICT/FIOCRUZ: Rio de Janeiro 2017, p 528.
- (14) O’Dea, J.; Cinelli, R. L. Use of Drugs to Change Appearance in Girls and Female Adolescents. In *Chemically Modified Bodies: The Use of Diverse Substances for Appearance Enhancement*; Palgrave Macmillan UK: London, 2016; pp 51–76. [https://doi.org/10.1057/978-1-137-53535-1\\_4](https://doi.org/10.1057/978-1-137-53535-1_4).
- (15) Watkins, K.; Griffin, O. H.; Gardner, E. A. Investigation of Black Box Drugs Purchased from an Online Pharmacy. *J. Subst. Use* **2019**, 24 (4), 445–449. <https://doi.org/10.1080/14659891.2019.1595197>.
- (16) Arens, A. M.; van Wijk, X. M. R.; Vo, K. T.; Lynch, K. L.; Wu, A. H. B.; Smollin, C. G. Adverse Effects From Counterfeit Alprazolam Tablets. *JAMA Intern. Med.* **2016**, 176 (10), 1554. <https://doi.org/10.1001/jamainternmed.2016.4306>.
- (17) Armenian, P.; Vo, K. T.; Barr-Walker, J.; Lynch, K. L. Fentanyl, Fentanyl Analogs and Novel Synthetic Opioids: A Comprehensive Review. *Neuropharmacology*. Elsevier Ltd May 15, 2018, pp 121–132. <https://doi.org/10.1016/j.neuropharm.2017.10.016>.
- (18) Nielsen, S.; McAuley, A. Etizolam: A Rapid Review on Pharmacology, Non-medical Use and Harms. *Drug Alcohol Rev.* **2020**, 39 (4), 330–336. <https://doi.org/10.1111/dar.13052>.
- (19) Tivura, M.; Asante, I.; van Wyk, A.; Gyaase, S.; Malik, N.; Mahama, E.; Hostetler, D. M.; Fernandez, F. M.; Asante, K. P.; Kaur, H.; Owusu-Agyei, S. Quality of Artemisinin-Based Combination Therapy for Malaria Found in Ghanaian Markets and Public Health Implications of Their Use. *BMC Pharmacol. Toxicol.* **2016**, 17 (1), 48. <https://doi.org/10.1186/s40360-016-0089-2>.
- (20) Bolla, A. S.; Patel, A. R.; Priefer, R. The Silent Development of Counterfeit Medications in Developing Countries – A Systematic Review of Detection Technologies. *Int. J. Pharm.* **2020**, 587, 119702. <https://doi.org/10.1016/j.ijpharm.2020.119702>.
- (21) Venhuis, B. J.; Zwaagstra, M. E.; Keizers, P. H. J.; de Kaste, D. Dose-to-Dose Variations with Single Packages of Counterfeit Medicines and Adulterated Dietary Supplements as a Potential Source of False Negatives and Inaccurate Health Risk Assessments. *J. Pharm. Biomed. Anal.* **2014**, 89, 158–165. <https://doi.org/10.1016/j.jpba.2013.10.038>.
- (22) dos Santos, M. K.; de Cassia Mariotti, K.; Kahmann, A.; Anzanello, M. J.; Ferrão, M. F.; de Araújo Gomes, A.; Limberger, R. P.; Ortiz, R. S. Comparison between Counterfeit and Authentic Medicines: A Novel Approach Using Differential Scanning Calorimetry and Hierarchical Cluster Analysis. *J. Pharm. Biomed. Anal.* **2019**, 166, 304–309.

<https://doi.org/10.1016/j.jpba.2019.01.029>.

- (23) Blachford, S.; Krapp, K. M. *Drugs and Controlled Substances: Information for Students*, 1st ed.; Gale/Cengage Learning, 2003.
- (24) Childress, A. C.; Chow, H. Amphetamine Extended-Release Oral Suspension for Attention-Deficit/Hyperactivity Disorder. *Expert Rev. Clin. Pharmacol.* **2019**, *12* (10), 965–971. <https://doi.org/10.1080/17512433.2019.1659723>.
- (25) Brasil. Ministério Da Saúde - Portaria Nº 344, de 12 de Maio de 1998. Ministério da Saúde: Brasil 1998.
- (26) Tefft, B. C. Acute Sleep Deprivation and Culpable Motor Vehicle Crash Involvement. *Sleep* **2018**, *41* (10), 1–11. <https://doi.org/10.1093/sleep/zsy144>.
- (27) Bombana, H. S.; Gjerde, H.; dos Santos, M. F.; Jamt, R. E. G.; Yonamine, M.; Rohlf, W. J. C.; Muñoz, D. R.; Leyton, V. Prevalence of Drugs in Oral Fluid from Truck Drivers in Brazilian Highways. *Forensic Sci. Int.* **2017**, *273*, 140–143. <https://doi.org/10.1016/j.forsciint.2017.02.023>.
- (28) Leyton, V.; Sinagawa, D. M.; Oliveira, K. C. B. G.; Schmitz, W.; Andreuccetti, G.; De Martinis, B. S.; Yonamine, M.; Munoz, D. R. Amphetamine, Cocaine and Cannabinoids Use among Truck Drivers on the Roads in the State of Sao Paulo, Brazil. *Forensic Sci. Int.* **2012**. <https://doi.org/10.1016/j.forsciint.2011.03.032>.
- (29) Takitane, J.; Oliveira, L. G. de; Endo, L. G.; Oliveira, K. C. B. G. de; Muñoz, D. R.; Yonamine, M.; Leyton, V. Uso de Anfetaminas Por Motoristas de Caminhão Em Rodovias Do Estado de São Paulo: Um Risco à Ocorrência de Acidentes de Trânsito? *Cien. Saude Colet.* **2013**, *18* (5), 1247–1254. <https://doi.org/10.1590/S1413-81232013000500009>.
- (30) Cole, C.; Jones, L.; McVeigh, J.; Kicman, A.; Syed, Q.; Bellis, M. Adulterants in Illicit Drugs: A Review of Empirical Evidence. *Drug Test. Anal.* **2011**, *3* (2), 89–96. <https://doi.org/10.1002/dta.220>.
- (31) Who. Obesity and overweight <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> (accessed Apr 29, 2020).
- (32) ABESO. Mapa da obesidade <https://abeso.org.br/obesidade-e-sindrome-metabolica/mapa-da-obesidade/> (accessed Apr 29, 2020).
- (33) Souza, Á. P. S. Automedicação Com Anorexígenos No Tratamento Da Obesidade No Brasil. *Ref. em Saúde da Fac. Estácio Sá Goiás-RRS-FESGO* **2019**, *2* (1).
- (34) Lista de substancias sujeitas a controle especial - Anvisa <http://portal.anvisa.gov.br/lista-de-substancias-sujeitas-a-controle-especial> (accessed Apr 29, 2020).
- (35) Hocking, S.; Dear, A.; Cowley, M. A. Current and Emerging Pharmacotherapies for Obesity in Australia. *Obes. Res. Clin. Pract.* **2017**, *11* (5), 501–521. <https://doi.org/10.1016/j.orcp.2017.07.002>.

- (36) Sardela, V. F.; Carvalho, A. R.; da C. Nunes, I. K. The Black Market for Anorectic Agents: A Case Study of Amfepramone. *Toxicol. Anal. Clin.* **2018**, *30* (2), 149–153. <https://doi.org/10.1016/j.toxac.2018.03.001>.
- (37) Mariotti, K. de C.; Ortiz, R. S.; Souza, D. Z.; Mileski, T. C.; Fröhlich, P. E.; Limberger, R. P. Trends in Counterfeits Amphetamine-Type Stimulants after Its Prohibition in Brazil. *Forensic Sci. Int.* **2013**, *229* (1–3), 23–26. <https://doi.org/10.1016/j.forsciint.2013.03.026>.
- (38) Comiran, E.; Souza, D. Z.; Boehl, P. O.; Cássia Mariotti, K. de; Pechansky, F.; Duarte, P. D. C. A. V.; De Boni, R. B.; Fröhlich, P. E.; Limberger, R. P. Fenproporex and Amphetamine Pharmacokinetics in Oral Fluid After Controlled Oral Administration of Fenproporex. *Ther. Drug Monit.* **2012**, *34* (5), 545–553. <https://doi.org/10.1097/FTD.0b013e318263c6c5>.
- (39) Musshoff, F. Illegal Or Legitimate Use? Precursor Compounds To Amphetamine And Methamphetamine. *Drug Metab. Rev.* **2000**, *32* (1), 15–44. <https://doi.org/10.1081/DMR-100100562>.
- (40) Correia, R. M.; Domingos, E.; Tosato, F.; Santos, N. A. Dos; Leite, J. D. A.; Da Silva, M.; Marcelo, M. C. A.; Ortiz, R. S.; Filgueiras, P. R.; Romão, W. R. Portable near Infrared Spectroscopy Applied to Abuse Drugs and Medicine Analyses †. **2018**. <https://doi.org/10.1039/c7ay02998e>.
- (41) Greenway, F. L.; Bray, G. A. Combination Drugs for Treating Obesity. *Curr. Diab. Rep.* **2010**, *10* (2), 108–115. <https://doi.org/10.1007/s11892-010-0096-4>.
- (42) Sewell, C. R.; Rivey, M. P. A Case Report of Benzocaine-Induced Methemoglobinemia. *J. Pharm. Pract.* **2018**, *31* (5), 507–509. <https://doi.org/10.1177/0897190017723211>.
- (43) Ekar, T.; Kreft, S. Common Risks of Adulterated and Mislabeled Herbal Preparations. *Food Chem. Toxicol.* **2019**, *123*, 288–297. <https://doi.org/10.1016/j.fct.2018.10.043>.
- (44) Skånland, S. S.; Cieślak-Pobuda, A. Off-Label Uses of Drugs for Depression. *Eur. J. Pharmacol.* **2019**, *865*, 172732. <https://doi.org/10.1016/j.ejphar.2019.172732>.
- (45) Perez-Caballero, L.; Torres-Sanchez, S.; Bravo, L.; Mico, J. A.; Berrocoso, E. Fluoxetine: A Case History of Its Discovery and Preclinical Development. *Expert Opin. Drug Discov.* **2014**, *9* (5), 567–578. <https://doi.org/10.1517/17460441.2014.907790>.
- (46) de Freitas, S. C.; dos Santos, M. A. Z.; Berneira, L. M.; Ortiz, R. S.; de Pereira, C. M. P. Microwave-Assisted Extraction and Differential Scanning Calorimetry in the Chemical Identification of Sliming Agents Apprehended in the South Region of Brazil. *Sci. Justice* **2019**, *59* (2), 190–198. <https://doi.org/10.1016/j.scijus.2018.11.001>.
- (47) Neves, D. B. da J.; Caldas, E. D. Determination of Caffeine and Identification of Undeclared Substances in Dietary Supplements and Caffeine Dietary Exposure Assessment. *Food Chem. Toxicol.* **2017**, *105*, 194–202. <https://doi.org/10.1016/j.fct.2017.03.063>.

- (48) Dastjerdi, A. G.; Akhgari, M.; Kamali, A.; Mousavi, Z. Principal Component Analysis of Synthetic Adulterants in Herbal Supplements Advertised as Weight Loss Drugs. *Complement. Ther. Clin. Pract.* **2018**, *31*, 236–241. <https://doi.org/10.1016/j.ctcp.2018.03.007>.
- (49) van Amsterdam, J.; Opperhuizen, A.; Hartgens, F. Adverse Health Effects of Anabolic–Androgenic Steroids. *Regul. Toxicol. Pharmacol.* **2010**, *57* (1), 117–123. <https://doi.org/10.1016/j.yrtph.2010.02.001>.
- (50) Abrahin, O. S. C.; Souza, N. S. F.; Sousa, E. C. de; Moreira, J. K. R.; Nascimento, V. C. do. Prevalência Do Uso e Conhecimento de Esteroides Anabolizantes Androgênicos Por Estudantes e Professores de Educação Física Que Atuam Em Academias de Ginástica. *Rev. Bras. Med. do Esporte* **2013**, *19* (1), 27–30. <https://doi.org/10.1590/S1517-86922013000100005>.
- (51) Abrahin, O. S. C.; de Sousa, E. C.; Santos, A. M. Prevalence of the Use of Anabolic-Androgenic Steroids in Brazil: A Systematic Review. *Subst. Use Misuse* **2014**, *49* (9), 1156–1162. <https://doi.org/10.3109/10826084.2014.903750>.
- (52) Neves, D.; Talhavini, M.; Braga, J. W.; Zacca, J.; Caldas, E. Detection of Counterfeit Durateston® Using Fourier Transform Infrared Spectroscopy and Partial Least Squares - Discriminant Analysis. *J. Braz. Chem. Soc.* **2016**, *28* (7), 1288–1296. <https://doi.org/10.21577/0103-5053.20160293>.
- (53) Berneira, L. M.; de Freitas, S. C.; da Silva, C. C.; Machado, A. de M.; de Pereira, C. M. P.; dos Santos, M. A. Z. Application of Differential Scanning Calorimetry in the Analysis of Apprehended Formulations of Anabolic Androgenic Steroids. *Forensic Sci. Int.* **2019**, *296*, 15–21. <https://doi.org/10.1016/j.forsciint.2018.12.022>.
- (54) Ribaux, O.; Talbot Wright, B. Expanding Forensic Science through Forensic Intelligence. *Sci. Justice* **2014**, *54* (6), 494–501. <https://doi.org/10.1016/j.scijus.2014.05.001>.
- (55) Dégardin, K.; Roggo, Y.; Margot, P. Forensic Intelligence for Medicine Anti-Counterfeiting. *Forensic Sci. Int.* **2015**, *248* (December 2014), 15–32. <https://doi.org/10.1016/j.forsciint.2014.11.015>.
- (56) Dégardin, K.; Roggo, Y. Counterfeit Analysis Strategy Illustrated by a Case Study. *Drug Test. Anal.* **2016**, *8* (3–4), 388–397. <https://doi.org/10.1002/dta.1908>.
- (57) Been, F.; Roggo, Y.; Degardin, K.; Esseiva, P.; Margot, P. Profiling of Counterfeit Medicines by Vibrational Spectroscopy. *Forensic Sci. Int.* **2011**, *211* (1–3), 83–100. <https://doi.org/10.1016/j.forsciint.2011.04.023>.
- (58) Dégardin, K.; Roggo, Y.; Been, F.; Margot, P. Detection and Chemical Profiling of Medicine Counterfeits by Raman Spectroscopy and Chemometrics. *Anal. Chim. Acta* **2011**, *705* (1–2), 334–341. <https://doi.org/10.1016/j.aca.2011.07.043>.
- (59) Shinde, S. R.; Bhavsar, K.; Kimbahune, S.; Khandelwal, S.; Ghose, A.; Pal, A. Detection of Counterfeit Medicines Using Hyperspectral Sensing. In

- Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*; Institute of Electrical and Electronics Engineers Inc., 2020; Vol. 2020-July, pp 6155–6158. <https://doi.org/10.1109/EMBC44109.2020.9176419>.
- (60) SWGDRUG. Scientific Working Group for the Analysis of Seized Drugs Recommendations Edition 8.0 <http://www.swgdrug.org/approved.htm> (accessed Aug 24, 2020).
- (61) Pavia, D. L.; Lampman, G. M.; Kriz, G. S.; Vyvyan, J. R. *Introduction to Spectroscopy*, 4th ed.; Cengage Learning, 2004.
- (62) Skoog, D. A.; Holler, F. J.; Crouch, S. R. *Principles of Instrumental Analysis*; Cengage Learning, 2017.
- (63) Xiaobo, Z.; Jiewen, Z.; Povey, M. J. W. W.; Holmes, M.; Hanpin, M. Variables Selection Methods in Near-Infrared Spectroscopy. *Anal. Chim. Acta* **2010**, *667* (1–2), 14–32. <https://doi.org/10.1016/j.aca.2010.03.048>.
- (64) Smith, R. Chemometrics. In *Forensic Chemistry*; John Wiley & Sons, Ltd: Chichester, UK, UK, 2015; pp 469–503. <https://doi.org/10.1002/9781118897768.ch12>.
- (65) Yusoff, M. Z.; Chang, K. H.; Abdullah, A. F. L. Attenuated Total Reflectance – Fourier Transform Infra-Red Spectral Profiling of Illicit Heroin for Forensic Intelligence. *Aust. J. Forensic Sci.* **2018**, *50* (5), 543–551. <https://doi.org/10.1080/00450618.2016.1275794>.
- (66) Custers, D.; Canfyn, M.; Courselle, P.; De Beer, J. O. O.; Apers, S.; Deconinck, E. Headspace–Gas Chromatographic Fingerprints to Discriminate and Classify Counterfeit Medicines. *Talanta* **2014**, *123*, 78–88. <https://doi.org/10.1016/j.talanta.2014.01.020>.
- (67) Rodrigues, N. V. S. S.; Cardoso, E. M.; Andrade, M. V. O. O.; Donnici, C. L.; Sena, M. M. Analysis of Seized Cocaine Samples by Using Chemometric Methods and FTIR Spectroscopy. *J. Braz. Chem. Soc.* **2013**, *24* (3), 507–517. <https://doi.org/10.5935/0103-5053.20130066>.
- (68) Granato, D.; Santos, J. S.; Escher, G. B.; Ferreira, B. L.; Maggio, R. M. Use of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) for Multivariate Association between Bioactive Compounds and Functional Properties in Foods: A Critical Perspective. *Trends Food Sci. Technol.* **2018**, *72*, 83–90. <https://doi.org/10.1016/j.tifs.2017.12.006>.
- (69) Wold, S. Pattern Recognition by Means of Disjoint Principal Components Models. *Pattern Recognit.* **1976**, *8* (3), 127–139. [https://doi.org/10.1016/0031-3203\(76\)90014-5](https://doi.org/10.1016/0031-3203(76)90014-5).
- (70) Ballabio, D.; Todeschini, R. Multivariate Classification for Qualitative Analysis. In *Infrared Spectroscopy for Food Quality Analysis and Control*; Sun, D.-W., Ed.; Academic Press, 2009; pp 83–104.
- (71) Wold, S.; Sjöström, M.; Eriksson, L. PLS-Regression: A Basic Tool of Chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58* (2), 109–130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1).

- (72) Ballabio, D.; Consonni, V. Classification Tools in Chemistry. Part 1: Linear Models. PLS-DA. *Anal. Methods* **2013**, *5* (16), 3790. <https://doi.org/10.1039/c3ay40582f>.
- (73) Deconinck, E.; Van Campenhout, R.; Aouadi, C.; Canfyn, M.; Bothy, J. L.; Gremeaux, L.; Blanckaert, P.; Courselle, P. Combining Attenuated Total Reflectance- Infrared Spectroscopy and Chemometrics for the Identification and the Dosage Estimation of MDMA Tablets. *Talanta* **2019**, *195*, 142–151. <https://doi.org/10.1016/j.talanta.2018.11.027>.
- (74) Eid, S. M.; Soliman, S. S.; Elghobashy, M. R.; Abdalla, O. M. ATR-FTIR Coupled with Chemometrics for Quantification of Vildagliptin and Metformin in Pharmaceutical Combinations Having Diverged Concentration Ranges. *Vib. Spectrosc.* **2020**, *106*, 102995. <https://doi.org/10.1016/j.vibspec.2019.102995>.
- (75) Banas, A.; Banas, K.; Kalaiselvi, S. M. P.; Pawlicki, B.; Kwiatek, W. M.; Breese, M. B. H. Is It Possible to Find Presence of Lactose in Pharmaceuticals? — Preliminary Studies by ATR-FTIR Spectroscopy and Chemometrics. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2017**, *171*, 280–286. <https://doi.org/10.1016/j.saa.2016.08.003>.
- (76) Brito, J. B. G.; Bucco, G. B.; John, D. K.; Ferrão, M. F.; Ortiz, R. S.; Mariotti, K. C.; Anzanello, M. J. Wavenumber Selection Based on Singular Value Decomposition for Sample Classification. *Forensic Sci. Int.* **2020**, *309*, 110191. <https://doi.org/10.1016/j.forsciint.2020.110191>.
- (77) Deconinck, E.; Aouadi, C.; Bothy, J. L.; Courselle, P. Detection and Identification of Multiple Adulterants in Plant Food Supplements Using Attenuated Total Reflectance—Infrared Spectroscopy. *J. Pharm. Biomed. Anal.* **2018**, *152*, 111–119. <https://doi.org/10.1016/j.jpba.2018.01.047>.
- (78) Deconinck, E.; Djiogo, C. A. S.; Bothy, J. L.; Courselle, P. Detection of Regulated Herbs and Plants in Plant Food Supplements and Traditional Medicines Using Infrared Spectroscopy. *J. Pharm. Biomed. Anal.* **2017**. <https://doi.org/10.1016/j.jpba.2017.04.051>.
- (79) Rodionova, O. Y. Y.; Balyklova, K. S.; Titova, A. V.; Pomerantsev, A. L. Application of NIR Spectroscopy and Chemometrics for Revealing of the 'High Quality Fakes' among the Medicines. *Forensic Chem.* **2018**, *8* (2018), 82–89. <https://doi.org/https://doi.org/10.1016/j.forc.2018.02.004>.
- (80) Fuenffinger, N.; Arzhantsev, S.; Gryniowicz-Ruzicka, C. Classification of Ciprofloxacin Tablets Using Near-Infrared Spectroscopy and Chemometric Modeling. *Appl. Spectrosc.* **2017**, *71* (8), 1927–1937. <https://doi.org/10.1177/0003702817699624>.
- (81) Hespanhol, M. C.; Pasquini, C.; Maldaner, A. O. Evaluation of a Low-Cost Portable near-Infrared Spectrophotometer for in Situ Cocaine Profiling. *Talanta* **2019**, *200*, 553–561. <https://doi.org/10.1016/j.talanta.2019.03.091>.
- (82) Guillemain, A.; Dégardin, K.; Roggo, Y. Performance of NIR Handheld Spectrometers for the Detection of Counterfeit Tablets. *Talanta* **2017**, *165* (December 2016), 632–640. <https://doi.org/10.1016/j.talanta.2016.12.063>.



- (83) Zontov, Y. V.; Rodionova, O. Y.; Kucheryavskiy, S. V.; Pomerantsev, A. L. DD-SIMCA – A MATLAB GUI Tool for Data Driven SIMCA Approach. *Chemom. Intell. Lab. Syst.* **2017**, *167*, 23–28. <https://doi.org/10.1016/j.chemolab.2017.05.010>.
- (84) Ballabio, D. A MATLAB Toolbox for Principal Component Analysis and Unsupervised Exploration of Data Structure. *Chemom. Intell. Lab. Syst.* **2015**, *149*, 1–9. <https://doi.org/10.1016/j.chemolab.2015.10.003>.
- (85) Colthup, N. B.; Daly, L. H.; Wiberley, S. E. Introduction to Infrared and Raman Spectroscopy; Elsevier: San Diego, CA, 1990; p 542. <https://doi.org/10.1016/B978-0-08-091740-5.50002-8>.
- (86) Pereira, L. S. A.; Lisboa, F. L. C.; Coelho Neto, J.; Valladão, F. N.; Sena, M. M. Screening Method for Rapid Classification of Psychoactive Substances in Illicit Tablets Using Mid Infrared Spectroscopy and PLS-DA. *Forensic Sci. Int.* **2018**, *288*, 227–235. <https://doi.org/10.1016/j.forsciint.2018.05.001>.
- (87) Fuentes, A.; Pineda, M.; Venkata, K. Comprehension of Top 200 Prescribed Drugs in the US as a Resource for Pharmacy Teaching, Training and Practice. *Pharmacy* **2018**, *6* (2), 43. <https://doi.org/10.3390/pharmacy6020043>.
- (88) Nørgaard, L. IToolbox for MATLAB. KVL, Denmark 2004.
- (89) Ma, B.; Wang, L. An Application of Rapid Detection Technologies in a National Regulatory Laboratory Setting: Differentiating Imported and Domestic Drug Products of Oxcarbazepine Using Handheld Raman, near Infrared, and Portable FTIR Analyzers. *Am Pharm Rev* **2015**, *18*.
- (90) Kennard, R. W.; Stone, L. A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11* (1), 137–148. <https://doi.org/10.1080/00401706.1969.10490666>.
- (91) Tie, Y.; Duchateau, C.; Van de Steene, S.; Mees, C.; De Braekeleer, K.; De Beer, T.; Adams, E.; Deconinck, E. Spectroscopic Techniques Combined with Chemometrics for Fast On-Site Characterization of Suspected Illegal Antimicrobials. *Talanta* **2020**, *217*, 121026. <https://doi.org/10.1016/j.talanta.2020.121026>.
- (92) Anzanello, M. J.; Ortiz, R. S.; Limberger, R.; Mariotti, K. PLS-DA Wavenumber Selection for the Categorization of Medicine Samples Based on Multiple Criteria. *Forensic Sci. Int.* **2014**, *242*, 111–116. <https://doi.org/10.1016/j.forsciint.2014.06.034>.
- (93) Dégardin, D. K.; Guillemain, A.; Guerreiro, N. V.; Roggo, Y.; Dégardin, K.; Guillemain, A.; Guerreiro, N. V.; Roggo, Y. Near Infrared Spectroscopy for Counterfeit Detection Using a Large Database of Pharmaceutical Tablets. *J. Pharm. Biomed. Anal.* **2016**, *128* (April), 89–97. <https://doi.org/10.1016/j.jpba.2016.05.004>.
- (94) Lê, L. M. M.; Eveleigh, L.; Hasnaoui, I.; Prognon, P.; Baillet-Guffroy, A.; Caudron, E. Rapid Discrimination and Determination of Antibiotics Drugs in Plastic Syringes Using near Infrared Spectroscopy with Chemometric

Analysis: Application to Amoxicillin and Penicillin. *J. Pharm. Biomed. Anal.* **2017**, *138*, 249–255. <https://doi.org/10.1016/j.jpba.2017.02.019>.