

Evolução da resistência antimicrobiana no grupo ESKAPE

Luís Ferreira¹, Maria Pereira¹, Anália Lourenço²

¹ Centro de Engenharia Biológica, Universidade do Minho, 4710-057 Braga, Portugal

² Universidade de Vigo

Abstract. A resistência antimicrobiana a antibióticos é um fenômeno em que bactérias adquirem a capacidade de sobreviver e multiplicarem-se na presença de efeitos de antibióticos, que normalmente seriam eficazes no combate de infecções bacterianas. A lista da OMS destaca as bactérias ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* e *Enterobacter sp.*) como as mais importantes, sendo quatro delas consideradas críticas e duas altamente prioritárias. O estudo tem como objetivo investigar os processos evolutivos pelos quais os microrganismos adquiriram mecanismos de insensibilidade às concentrações mínimas dos antibióticos (MIC). A abordagem envolve a adoção de ferramentas de extração de resumos e anotações de artigos científicos, ferramentas de deteção de entidades nomeadas para identificar drogas e uso de abordagens baseadas em regras, para recolha de informações sobre organismos e medidas de quantificação da resistência antimicrobiana. Utilizando uma “query” específica, foram obtidos 7014 resultados, tendo sido recolhidas informações de cada parâmetro pretendido. No entanto, até ao momento, não é possível criar relações lógicas e direta entre as diferentes informações extraídas individualmente. A análise deste trabalho mostrou-se essencial para compreender as causas da resistência aos antibióticos e no combate às infecções bacterianas, fornecendo “insights” sobre as alterações nas concentrações necessárias para inibir os organismos.

Keywords: Resistência antimicrobiana · ESKAPE · MIC

1 Introdução

1.1 Contextualização e Motivação

Ao longo dos anos, desde o surgimento dos antibióticos, o problema da resistência antimicrobiana tem sido abordado e discutido devido às preocupações que esta têm provocado na saúde, bem como aos custos elevados no combate e na investigação desta problemática. Em 1935, Gerhard Domagk fez a descoberta da primeira sulfonamida utilizada para tratar infecções bacterianas [5] e na década de 1940, durante a Segunda Guerra Mundial, ocorreu a introdução da penicilina no combate de infecções graves, apesar de esta já ter sido descoberta em 1928 e ter demorado quase uma década a ser inserida no mercado dos antibióticos. Embora

o uso generalizado da penicilina só tenha ocorrido bastante tempo depois, nesta época já haviam sido descritas estirpes de *Staphylococcus* resistentes à penicilina. De modo a enfrentar esta problemática da resistência das penicilases e superar as complicações na indústria clínica na década de 1950, foram investigados novos antibióticos beta-lactâmicos, que proporcionaram uma melhor atuação destes medicamentos no combate das bactérias, tais como a introdução da meticilina (MRSA) em 1959, que demonstrou novamente a eficiência destes medicamentos. No entanto, esta foi uma solução temporária, pois os patógenos resistentes não pararam de surgir, sendo este fenômeno comprovado pelo aparecimento da *Staphylococcus aureus* resistente à meticilina (MRSA). Posto isto, embora tenham sido feitas muitas investigações com o intuito de produzir medicamentos que contrariem os mecanismos das bactérias, estas descobertas diminuíram abruptamente devido às dificuldades sentidas pelos especialistas para contrariar esta tendência (fig. 1). Como resultado desta preocupação, este assunto tornou-se num problema discutido e abordado em todo o mundo devido às suas consequências evidentes [7][21][32].

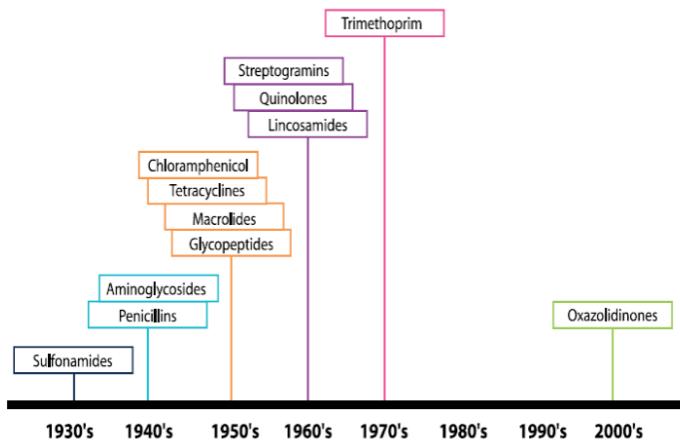


Fig. 1: Surgimento de novas categorias de medicamentos antimicrobianos (Fonte: artigo científico [21])

A resistência antimicrobiana diz respeito à capacidade dos microrganismos em sofrer mudanças que os tornam insensíveis aos efeitos dos medicamentos antimicrobianos, resultando numa infecção clínica. Essa resistência pode ter origem intrínseca, adquirida ou adaptativa [11]. Quando surge como característica intrínseca, significa que essa espécie de bactéria, possui a capacidade em resistir à ação de certos antibióticos com base nas suas características funcionais ou estruturais. É de origem adquirida quando surge através de mutações que ocorrem de forma espontânea durante a replicação celular ou por meio da aquisição de novo material genético a partir de uma fonte externa (transferência horizontal).

de genes). Essa fonte externa pode ser adquirida por meio de processos como a conjugação bacteriana, a transformação e a transdução. Outra forma de obter resistência é através da adaptação, incorporando material genético exógeno de microrganismos de outras espécies que contenham genes de resistência. A resistência aos antibióticos é um problema cada vez mais preocupante e é fundamental compreender os mecanismos que levam ao seu desenvolvimento, a fim de criar estratégias eficazes para combatê-la [25][20].

Existe também a possibilidade de classificar estes patógenos consoante o número de classes de antibióticos a quais estas bactérias são resistentes [17]. Temos os patógenos "Multidrug-resistant" (MDR), sendo definidos pela não suscetibilidade a pelo menos um antibiótico em três ou mais classes de antimicrobianos, os microrganismos "Extensively drug-resistant" (XDR), podem ser definidos como organismos não suscetíveis a pelo menos um agente em todas as classes de antimicrobianos, exceto uma ou duas categorias, por fim, temos os microrganismos "Pandrug-resistant" (PDR), sendo definidos como patógenos à qual não existe nenhum antibiótico em nenhuma das classes antimicrobianas, a qual estes sejam sensíveis [4]. Estes últimos, são considerados extremamente perigosos, pois mesmo os antibióticos da última linha de defesa são ineficazes.

Em 2017, a Organização Mundial da Saúde (OMS) divulgou uma relação global de bactérias resistentes a antibióticos que requerem atenção prioritária na pesquisa de novos medicamentos, pois representam uma grande ameaça à saúde pública em todo o mundo, limitando as opções de tratamento para infecções graves e aumentando o risco de morbidade e mortalidade. A lista destaca as bactérias ESKAPE como as mais importantes, sendo quatro delas consideradas críticas (*Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* e *Enterobacter spp.*) e duas altamente prioritárias (*S. aureus* e *Enterococcus faecium*). Diferentes mecanismos de resistência aos antibióticos foram desenvolvidos por estas bactérias, tais como a produção de enzimas que inativam os medicamentos, a modificação das proteínas alvo dos antibióticos, a redução da entrada dos medicamentos na célula bacteriana, o aumento da eficiência do bombeamento dos antibióticos para fora da célula bacteriana e a capacidade de formar biofilmes que dificultam a ação dos antimicrobianos [22]. Para enfrentar esta ameaça, é crucial adotar uma abordagem coordenada e global de modo a monitorar a resistência antimicrobiana e promover uma melhor administração de antibióticos por parte dos pacientes.

Os patógenos pertencentes ao grupo ESKAPE manifestam resistência a diversos tipos de antimicrobianos, apresentando variações nos valores de Concentração Inibitória Mínima (MIC) de acordo com o tipo de antibiótico utilizado. Uma análise global de algumas das bactérias deste grupo, sugere que um dos compostos mais efetivos no combate às bactérias *Pseudomonas aeruginosa* foi a ciprofloxacina, com uma frequência maior de MIC observada em 0,125 mg/L. Por outro lado, a fosfomicina e o florfenicol são exemplos de famílias de compostos menos efetivos, apresentando valores de MIC em torno de 64 mg/L. No que diz respeito às bactérias *Enterococcus faecium*, as famílias de compostos mais efetivas incluem a omadaciolina e tigeciclina, com valores de MIC predominan-

temente em torno de 0,06 mg/L. Já a nitrofurantoína é um exemplo de família menos efetiva no tratamento, apresentando valores de MIC em torno de 128 mg/L. Relativamente ao combate das bactérias *Staphylococcus aureus*, a família de compostos que se mostrou mais eficaz foi a delafloxacino, com um elevado número de observações de MICs em torno de 0,004 mg/L. Por outro lado, a cefazidima e o cloranfenicol se destacam como exemplos de famílias de compostos menos eficazes, apresentando valores de MICs em torno de 8 mg/L. De modo a validar os valores mencionados ou aprofundar o assunto, é possível consultar o Comité Europeu de Testes de Suscetibilidade Antimicrobiana, tendo os dados sido retirados do site de distribuição EUCAST MIC [1], no dia 06/04/2023.

1.2 Objetivos

O objetivo deste estudo consiste na compreensão das estratégias moleculares e genéticas envolvidas nos processos evolutivos pelos quais os microrganismos desenvolvem resistência aos antibióticos, visando investigar as MICs correspondentes aos antibióticos utilizados no tratamento de infecções. Para alcançar este objetivo, serão utilizadas técnicas de processamento de linguagem natural (NLP) para extrair automaticamente os resumos do PubMed e realizar o seu processamento. O processo será focado no reconhecimento automático de menções a drogas e parâmetros que quantificam a resistência antimicrobiana, utilizando dicionários numa primeira fase e com interesse em aplicar algoritmos de machine learning numa fase posterior.

2 Estado da arte

Nos últimos anos, tem havido um grande aumento na quantidade de literatura científica publicada na área biológica, o que tem criado desafios no processo de biocuração devido ao tempo e custo necessários para esta tarefa. Nesse contexto, a adoção de ferramentas modernas de mineração de texto e números, bem como a utilização de tecnologias de inteligência artificial, têm se mostrado úteis na evolução da pesquisa científica [8]. As técnicas de mineração de texto podem ser usadas para analisar dados de MIC, o tipo de droga usada ou o organismo afetado, de modo a extraír informações e padrões, como identificar tendências de resistência antimicrobiana em regiões específicas, entender o impacto de diferentes fatores na resistência antimicrobiana e descrever o processo evolutivo desses microrganismos. Este modelo de mineração também permite fazer um mapeamento dos agentes patogénicos de modo a relacionar os marcos históricos com os mecanismos dos organismos, permitindo analisar as estratégias adotadas pelas bactérias ao longo do tempo.

Um dos ramos da inteligência artificial (IA) e da ciência da computação que está em constante evolução é o Machine learning (ML), que se concentra no uso de dados e algoritmos para imitar a maneira como os humanos aprendem [2]. É uma abordagem computacional semiautomatizada na qual se introduz algoritmos capazes de analisar os dados de entrada, prever valores de saída com alta

precisão, identificar padrões e tendências em dados e aprender com experiências anteriores. Estes algoritmos de ML são a base da "inteligência" da máquina. Na medicina, por exemplo, o ML pode ser usado para melhorar a precisão dos algoritmos de diagnóstico e individualizar o tratamento do paciente. Em geral, o ML é classificado em quatro grupos: Aprendizado supervisionado que utiliza dados rotulados para treinar modelos a prever resultados para novas entradas; Aprendizado não supervisionado, que não usa dados rotulados e em vez disso identifica padrões e estruturas nos dados; Aprendizado semi-supervisionado que é uma combinação de aprendizado supervisionado e não supervisionado, onde o modelo usa dados rotulados e não rotulados para melhorar a precisão; Aprendizado por reforço, sendo um tipo de aprendizado em que um agente aprende a tomar decisões num ambiente interativo para maximizar uma recompensa [10][18].

As ferramentas de NLP podem ser aplicadas para extrair informações significativas de fontes de dados não estruturados, como artigos científicos e registos médicos [23]. Esta tecnologia é eficaz em auxiliar na triagem rápida de títulos e resumos ao atualizar uma revisão sistemática, permitindo também o reconhecimento de nomes de drogas e substâncias químicas [16][24]. Este reconhecimento consiste em identificar as menções aos compostos anteriores e classificá-los em categorias específicas, o que é crucial para sistemas de extração de informações médicas e extração de relações médicas, permitindo obter a relação entre as doenças e os respetivos medicamentos [27][26]. Além disso, esse reconhecimento é importante para associar essas entidades com terminologias disponíveis no domínio biomédico, como o SNOMED-CT, sendo um sistema de classificação de terminologias médicas e permite a partilha de informação em diferentes sistemas e organizações de saúde [16].

Existem diversas categorias de abordagens para o reconhecimento de entidades nomeadas (NER), tais como: abordagens baseadas em dicionário, abordagens baseadas em regras e abordagens baseadas em aprendizado de máquina [15]. Temos então diferentes técnicas que podem ser usadas para identificar nomes de medicamentos em textos médicos não estruturados e classificá-los em categorias específicas. As abordagens baseadas em dicionários comparam os nomes de drogas com dicionários de produtos químicos e medicamentos, utilizando recursos como DrugBank e ChEBI [12]. As abordagens baseadas em regras usam regras que descrevem os padrões de composição ou contexto dos nomes das drogas [27]. Já as abordagens baseadas em aprendizado de máquina, geralmente, formalizam a tarefa de identificação de entidades nomeadas como um problema de classificação ou rotulagem de sequência. [3].

O artigo de investigação de L. Haixia [14] teve como objetivo a investigação de maneira sistemática dos mecanismos moleculares da resistência à platina no cancro do ovário, bem como a identificação dos genes e vias regulatórias, sendo que para isso, foram utilizadas técnicas de mineração de texto e outras abordagens metodológicas. Como tal, diversas literaturas catalogadas no PubMed foram analisadas, e os genes e proteínas foram identificados por meio do AB-NER e verificados no banco de dados Entrez Gene. As análises de ontologia genética e de interação proteína-proteína (PPI) foram realizadas para exami-

nar os mecanismos subjacentes à resistência à platina. Foi destacado que genes específicos foram identificados em cada subtipo de cancro de ovário, e que os genes com frequência mais alta estavam intimamente relacionados à resistência à platina. Isso evidenciou a efetividade da avaliação quantitativa das funções dos genes e as suas relações, bem como da análise sistemática por meio da mineração de texto.

O diagnóstico preciso da resistência bacteriana a antibióticos é fundamental para o tratamento adequado do paciente. O sequenciamento do genoma completo surge como uma alternativa mais rápida aos testes fenotípicos, mas requer a identificação de genes específicos relacionados à resistência a antibióticos [30]. Nesse sentido, a extração automatizada de informações de literatura biomédica surgiu como uma solução promissora. O artigo de investigação de B. André [6] descreve uma pipeline de mineração de texto que utiliza modelos de aprendizado profundo para extrair relações de resistência gene-antibiótico de resumos biomédicos em inglês. Esta pipeline de processamento de linguagem natural, incluiu a aquisição de dados via PubMed e PubTator, reconhecimento de entidades nomeadas usando o PubTator complementado com dicionários personalizados e extração de relações (RE) por meio de supervisão distante usando modelos de aprendizado de máquina de última geração. Como resultado, foi possível identificar novos genes que causam resistência a antibióticos e que ainda não estão presentes em bases de conhecimento comuns de resistência a antibióticos.

3 Metodologia

Foram recolhidas algumas ferramentas de extração de resumos/títulos e de anotações de artigos científicos no PubMed, sendo que deste conjunto de ferramentas, vamos analisar qual ou quais se ajustam melhor às nossas necessidades no desenvolvimento do projeto. Temos a ferramenta “PubMed Retrieval Tool” [9] que suporta toda a sintaxe que a “interface” PubMed contém, retornando os campos que sejam pretendidos (títulos, resumos e datas de publicação), sendo que os campos devem estar completamente qualificados pelo caminho e se forem pretendidos vários campos, devem ser separados por vírgula. Outra forma de extrair informações do Pubmed é através do ”pubmed retrieval” [19], onde se utiliza uma implementação combinada do módulo Entrez e das NCBI E-utilities. Essa abordagem envolve a utilização das funcionalidades disponibilizadas pelas NCBI E-utilities, como esearch e efetch, que são parte do módulo Entrez. Por meio dessas ferramentas, é possível definir um período de tempo desejado e recuperar diversos pontos de dados do PubMed, como o PMID, título, resumo e data de publicação (mês e ano) de todos os registros arquivados pelo PubMed durante esse período. Essa integração entre o módulo Entrez e as NCBI E-utilities proporciona uma solução eficiente para extraer informações específicas do Pubmed de maneira automatizada. De modo a conseguirmos identificar drogas nos textos extraídos, foram analisadas três ferramentas. Temos a ferramenta “NER em documentos médicos” [31], que pode ser utilizada para detetar entidades nomeadas, como produtos químicos, doenças, genes e proteínas, em documentos médico,

temos a ferramenta “AutoNER” [13], que consiste também numa ferramenta de reconhecimento de entidade nomeada e que permite o treino de modelos NER através de dois dicionários, um contendo informações básicas (tipo, superfície tokenizada) e outro contendo frases tokenizadas de alta qualidade, [28][29] e por fim, temos SciSpacy, que é uma ferramenta projetada especificamente para facilitar a extração de informações e análise de texto em domínios como medicina, biologia e pesquisa científica. Ele se baseia em modelos pré-treinados que foram treinados em grandes volumes de texto biomédico, como artigos científicos e literatura médica.

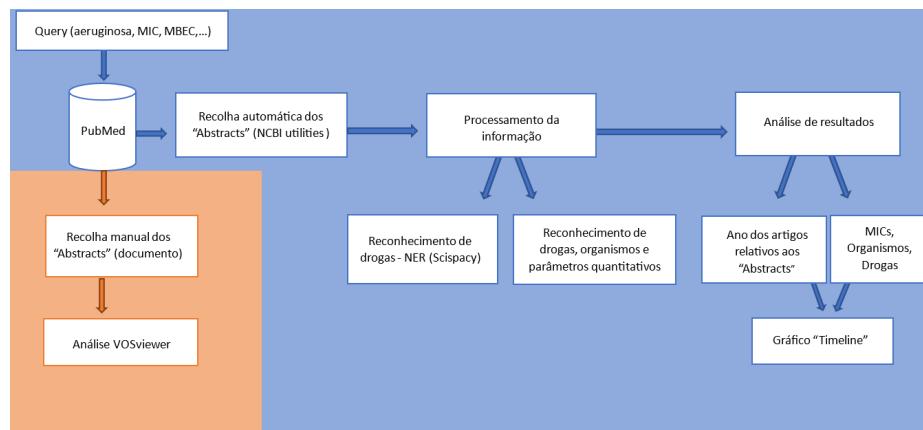


Fig. 2: Fluxograma completo do pipeline desenvolvido

Para facilitar a visualização das etapas necessárias que permitiram desenvolver o projeto, as informações referentes a este ponto foram organizadas em um fluxograma (fig. 2). Após as ideias terem sido estruturadas e a pesquisa das ferramentas necessárias estar concluída, inicia-se a etapa de recolha de dados. Nessa fase, automatizamos a extração de informações de fontes não estruturadas, como os resumos de artigos científicos de casos clínicos, disponíveis no banco de dados MEDLINE. Esta base de dados MEDLINE apresenta inúmeras referências bibliográficas de revistas de alta qualidade e consiste numa fonte confiável e amplamente utilizada na área de pesquisa científica, tornando os dados coletados robustos e precisos, tendo sido estes os motivos da sua escolha. Para obtermos apenas os artigos relevantes, foram especificadas “keywords”, tais como o nome do organismo, MIC e MBEC (Concentração mínima de erradicação do biofilme), de modo a ser possível identificar menções a compostos e drogas usadas para combater infecções associadas ao grupo ESKAPE, bem como identificar as principais medidas de quantificação de resistência antimicrobiana. A seleção específica e pormenorizada das “keywords” é necessária para obtermos uma busca mais completa e abrangente dos resultados pretendidos. A extração dos “abstracts” pode ser realizada de duas maneiras distintas. A primeira opção é utilizar a linguagem

Python e aproveitar as funcionalidades das NCBI E-utilities para extrair as informações dos "abstracts" de uma revisão sistemática de forma automatizada. A segunda opção é realizar o download manual dos "abstracts". Na primeira opção, os "abstract" e as informações relacionadas aos artigos científicos, título, pmid, data, autores e jornal, irão ser armazenadas numa base de dados MySQL.

Antes de prosseguir com o processamento automatizado dos abstracts, é crucial obter uma compreensão clara do tipo de dados com os quais estamos a trabalhar e estar ciente dos desafios que podem surgir ao recolher e agrupar essas informações. Para facilitar esse processo, será criado um arquivo Excel contendo dados extraídos de 100 "abstracts" selecionados aleatoriamente da base de dados no período de 2015 a 2023. Desta forma, iremos ter uma amostra representativa dos estudos realizados ao longo do tempo, abrangendo diferentes microrganismos, drogas e tipos de parâmetros, para além de fornecer uma base sólida para a implementação de modelos e scripts que nos auxiliarão na recolha automatizada de informações adicionais.

Após a extração do conteúdo de estudo e da análise manual do mesmo, são utilizadas bibliotecas de processamento de linguagem natural, como o spaCy, para o processamento do "abstract". Em particular, são empregados modelos pré-treinados, como o SciSpacy, que é uma extensão da biblioteca spaCy desenvolvida especificamente para o processamento de texto biomédico e científico. Esses modelos automatizam várias etapas do processamento, como tokenização, remoção de "stopwords", filtragem auxiliar e lematização ou redução de palavras à sua forma base. Dessa forma, o texto irá ser segmentado em frases ou palavras soltas, permitindo remover palavras curtas ou não pertinentes, e evitar repetição de palavras que possuam significado similar. Estes modelos do SciSpacy permitem o reconhecimento e classificação de entidades nos textos, tais como drogas e organismos.

Subsequentemente, será adotada uma abordagem baseada em dicionários, utilizando bases de dados como o DrugBank e o chEBI, para o reconhecimento de drogas, permitindo avaliar a eficácia do modelo NER em estudo, além da criação de uma lista abrangente de organismos ESKAPE que abranja suas diferentes variantes de nome. De modo a complementar as tarefas anteriores, será aplicada uma abordagem baseada em regras, utilizando expressões regulares, para identificar os organismos e extrair o valor numérico dos parâmetros MIC ou MBEC, juntamente com suas respectivas unidades de medida, presentes nos abstracts. Essas abordagens combinadas contribuirão para a precisão na identificação das diferentes entidades.

Por fim, utilizando a data de publicação de cada "abstract", em conjunto com as MICs, o organismo e a droga específica, será possível criar um gráfico com os resultados destes parâmetros quantitativos, permitindo visualizar uma "timeline" da evolução da resistência e identificar os padrões da mesma. De modo a complementar a análise dos nossos dados, também irá ser utilizada a ferramenta VOSviewer, que conjugada com todos os artigos extraídos, irá nos permitir observar as "keywords" mais relevantes na nossa pesquisa, de modo

a termos uma visão mais ampla das principais áreas temáticas relacionadas ao estudo.

4 Resultados e discussão

Para realizar a análise dos processos evolutivos e a aquisição de resistência antimicrobiana nos microrganismos do grupo ESKAPE, foram coletados, armazenados e analisados dados de diversos artigos científicos disponíveis no PubMed utilizando a "query" 'aeruginosa AND ("mic" OR "mbec" OR "mbic" OR "mbc" OR "ic50")'. Os 7014 resultados obtidos tiveram em foco a bactéria *Pseudomonas aeruginosa* e forneceram "insights" importantes sobre as características da bactéria, bem como os desafios que se tentam ultrapassar para a combater. No entanto, é importante destacar que até o momento, os resultados que foram obtidos, consistiram principalmente na recolha de informações de cada parâmetro individualmente, como ano, microrganismos, drogas (DrugBank e SciSpacy) e tipos de parâmetros (MIC, MIC90, MBEC, etc.), juntamente com seus respetivos valores, em relação a cada artigo. Na maioria dos casos, ainda não há uma relação lógica e direta entre os parâmetros.

Table 1: Resposta dos Organismos às Diversas Drogas: Análise de Abstracts (recolha automática)

Ano	Organismos	DrugBank	NER	Tipo	Valores
1998 [a]	<i>Pseudomonas aeruginosa, aeruginosa</i>	Meropenem	Beta-lactam, Meropenem	MIC90, MIC	4 mg/l
2006 [b]	<i>Pseudomonas aeruginosa, aeruginosa</i>	Ceftazidime	Ceftazidime	MIC	16 mg/l
2010 [c]	<i>P.aeruginosa, Pseudomonas aeruginosa, aeruginosa</i>	PolymyxinB, Doripenem, Colistin	Polymyxin B, Colistin, Doripenem, Penicillins/ Cephalosporins, etc	MIC, MIC	3mg/l, 2mg/l
2020 [d]	<i>Pseudomonas aeruginosa, aeruginosa</i>	Ciprofloxacin	Meropenem, Imipenem, Ceftazidime, etc	MIC	0.19 µg/ml

PMIDs: 10052554 [a], 16858611 [b], 20086165 [c], 26731974 [d].

A obtenção dos dados necessários para os parâmetros em questão nem sempre foram de fácil acesso (Table 1). Os organismos podiam ser escritos com várias nomenclaturas mas todos apresentavam a nomenclatura da espécie (*aeruginosa*, *aureus*,...), tendo sido esse o critério de recolha, quando não se encontrava o nome completo. As drogas nem sempre são mencionadas de forma concreta, por vezes temos extratos de uma determinada espécie de planta, o composto químico presente em um medicamento ou até mesmo a fórmula química por extenso. A estratégia que se adotou, foi usar o dicionário de drogas do DrugBank, para tentar cobrir a maior parte dos antimicrobianos e comparar com a informação recolhida dos modelos NER do SciSpacy, para testar a eficácia deste modelo no reconhecimento de drogas. A dificuldade encontrada reside no facto de que nem sempre são identificadas apenas as drogas para as quais temos informações sobre os parâmetros de sensibilidade, como as MICs, e em alguns casos, uma estirpe resistente a um determinado antibiótico pode ter exatamente o mesmo nome desse antibiótico, levando a uma possível confusão na identificação deste como droga. Relativamente ao tipo de parâmetro de sensibilidade, após analisar todos as formas como esta se pode encontrar, foi feita a recolha de forma apropriada, sem encontrar obstáculos significativos. No que diz respeito ao reconhecimento dos valores, também foram sentidas dificuldades devido às várias formas de escrita utilizadas. Essas formas incluíam a representação direta de um valor único, a indicação de intervalos tanto por meio de símbolos como por palavras, além da escrita separada dos valores e das suas unidades, quando existiam vários valores mencionados consecutivamente.

Table 2: Resposta das *Pseudomonas aeruginosa* à Tobramicina: Análise de Abstracts (recolha manual)

Ano	Organismo	Droga	Parâmetros	Valores
2015 [a]	<i>Pseudomonas aeruginosa</i>	Tobramycin	MIC90	0.002 mg/L
2016 [b]	<i>Pseudomonas aeruginosa</i>	Tobramycin	MIC90	0.032 mg/L
2018 [c]	<i>Pseudomonas aeruginosa</i>	Tobramycin	MIC	0.5 mg/L

PMIDs: 25980050 [a], 27928519 [b], 29437610 [c]

Com o intuito de estabelecer conexões entre os diversos parâmetros e exemplificar como a recolha das informações deve ser feita, foi realizada uma recolha manual dos dados restritamente presentes no arquivo Excel, tendo esta análise sido direcionada para obter informações que envolvessem um organismo e uma

droga específica ao longo de diferentes anos, de forma a ser possível compreender os níveis de resistência antimicrobiana ao longo do tempo. É importante ressaltar que diferentes parâmetros inibitórios originam diferentes interpretações, portanto é indicado fazer comparações apenas entre os mesmo critérios.

Com base nos dados fornecidos na Table 2, pode-se observar uma tendência de aumento gradual do valor da MIC da Tobramicina para o microrganismo *Pseudomonas aeruginosa* desde 2015 a 2018, indicando uma possível evolução e aquisição de resistência antimicrobiana por parte dessa bactéria. Esta informação é relevante, pois permite compreender que, ao longo desses anos, o uso contínuo do antibiótico resultou num aumento nos níveis de resistência desenvolvidos pela bactéria. Apesar do exemplo fornecido não seguir o mesmo critério de medição para todos os parâmetros, seja MIC ou MIC90 (concentração mínima necessária para inibir o crescimento de 90% das estirpes testadas), esta foi apenas uma demonstração do que era pretendido obter com os dados recolhidos, sendo que para cumprir o objetivo final, vai-se ter em conta as diferenças entre critérios e considerar essas variações ao interpretar e comparar os resultados.

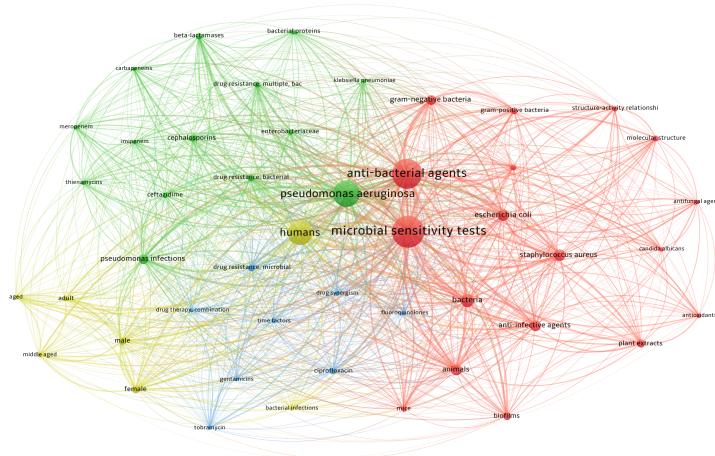


Fig. 3: Visualização no VOSviewer das co-ocorrências de "MeshKeywords" nos artigos recolhidos (Threshold: 200).

De forma a obtermos uma melhor compreensão dos resultados, foi realizada uma análise de coocorrência no VOSviewer (fig. 3), utilizando a unidade de análise "MeshKeywords" e um threshold de 200. Os resultados dessa análise revelaram um total de 47 correspondências, que foram agrupadas em 4 clusters distintos. Esses clusters indicam a existência de padrões de coocorrência entre os termos-chave relacionados aos artigos, o que pode sugerir a presença de áreas temáticas ou conceitos inter-relacionados. O primeiro cluster está associado às "*Pseudomonas aeruginosa*", uma espécie bacteriana do grupo ESKAPE. O se-

gundo cluster, é composto por termos relacionados aos estudos em "humanos". O terceiro cluster abrange os termos relacionados a "agentes antibacterianos" e "testes de sensibilidade microbiana", direcionando-nos para a identificação dos valores dos parâmetros de sensibilidade de diferentes drogas. Por fim, o quarto cluster está relacionado à "resistência microbiana a drogas", explorando o tema da resistência aos antimicrobianos.

5 Conclusão

A análise deste trabalho revelou-se de extrema importância no contexto do combate às infecções bacterianas. É essencial compreender as causas do aumento da insensibilidade às drogas ao longo dos anos e identificar quais medidas podem ser adotadas para combatê-lo. A conscientização sobre o uso prudente de antibióticos, a implementação de programas de vigilância e controle de infecções, bem como o desenvolvimento de novos antimicrobianos, devem ser considerados para enfrentar esse desafio em constante crescimento. No entanto, este trabalho ainda não atingiu seu objetivo final. É evidente que vários antibióticos possam ter apresentado um aumento na sua concentração inibitória mínima, porém é necessário realizar uma análise mais detalhada para cada um deles, afim de compreender se houve períodos em que a tendência foi diferente. Numa perspetiva futura, o estabelecimento e análise das associações entre os diferentes parâmetros é crucial para obter uma compreensão mais precisa da evolução dos microrganismos e da sua resistência aos antimicrobianos. Essas associações podem revelar padrões e tendências, permitindo-nos identificar fatores que influenciam o aumento ou diminuição da sensibilidade dos microrganismos ao longo do tempo. Ao explorar as relações entre as informações recolhidas, poderemos identificar se determinados eventos ou circunstâncias em certos anos, estão associados a alterações na sensibilidade dos microrganismos, isto porque, tanto para o aumento quanto para a diminuição desta concentração, há uma razão fundamental e é necessário esta ser compreendida. Como tal, tornasse essencial aprimorar esta análise, investigando e aplicando novas técnicas de recolhas de dados para obtermos efetivamente o resultado final pretendido.

References

1. MIC EUCAST (2021), https://mic.eucast.org/search/?search%5Bmethod%5D=mic&search%5Bantibiotic%5D=-1&search%5Bspecies%5D=411&search%5Bdisk_content%5D=-1&search%5Blimit%5D=50
2. Ahmed, Z., Mohamed, K., Zeeshan, S., Dong, X.Q.: Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database* **2020** (1 2020). <https://doi.org/10.1093/DATABASE/BAAA010>, <https://academic.oup.com/database/article/doi/10.1093/database/baaa010/5809229>
3. Armengol-Estabé, J., Soares, F., Marimon, M., Krallinger, M.: Pharamaconer tagger: a deep learning-based tool for automatically finding chemicals and drugs in spanish medical texts. *Genomics &*

- Informatics **17** (2019). <https://doi.org/10.5808/GI.2019.17.2.E15>, <http://genominfo.org/journal/view.php?doi=10.5808/GI.2019.17.2.e15>
4. Basak, S., Singh, P., Rajurkar, M.: Multidrug resistant and extensively drug resistant bacteria: A study. *Journal of Pathogens* **2016**, 1–5 (2016). <https://doi.org/10.1155/2016/4065603>
 5. Bickel, M.H.: Focal Point in the History of Chemotherapy. *Gesnerus: Swiss Journal for the History of Medicine and Science* **45**(1), 67–86 (1988)
 6. Brincat, A., Hofmann, M.: Automated extraction of genes associated with antibiotic resistance from the biomedical literature. *Database* **2022**, 1–14 (12 2022). <https://doi.org/10.1093/DATABASE/BAAB077>, <https://academic.oup.com/database/article/doi/10.1093/database/baab077/6520791>
 7. Christaki, E., Marcou, M., Tofarides, A.: Antimicrobial resistance in bacteria: Mechanisms, evolution, and persistence. *Journal of Molecular Evolution* **2019** 88:1 **88**, 26–40 (10 2019). <https://doi.org/10.1007/S00239-019-09914-3>, <https://link.springer.com/article/10.1007/s00239-019-09914-3>
 8. Costa, K.F., Araújo, F.A., Moraes, J., Frances, C.R.L., Ramos, R.T.: Text mining for identification of biological entities related to antibiotic resistant organisms. *PeerJ* **10**, e13351 (5 2022). <https://doi.org/10.7717/PEERJ.13351/SUPP-4>, <https://peerj.com/articles/13351>
 9. Dutta, S., Alberta, P., Lin: ReCiter-PubMed-Retrieval-Tool/README.md at master · wcmc-its/ReCiter-PubMed-Retrieval-Tool · GitHub (2022), <https://github.com/wcmc-its/ReCiter-PubMed-Retrieval-Tool/blob/master/README.md>
 10. Handelman, G.S., Kok, H.K., Chandra, R.V., Razavi, A.H., Lee, M.J., Asadi, .H., Gs, K., Hk, C., Rv, R., Ah, L., Mj, A.H., Victoria, R.: edoctor: machine learning and the future of medicine. *Journal of Internal Medicine* **284**, 603–619 (12 2018). <https://doi.org/10.1111/JOIM.12822>, <https://onlinelibrary.wiley.com/doi/full/10.1111/joim.12822>, <https://onlinelibrary.wiley.com/doi/abs/10.1111/joim.12822>
 11. Hanscheid, T., Hardisty, D.W.: How “resistant” is artemisinin resistant malaria? - the risks of ambiguity using the term “resistant” malaria. *Travel Medicine and Infectious Disease* **24**, 23–24 (7 2018). <https://doi.org/10.1016/J.TMAID.2018.04.008>
 12. Hettne, K.M., Stierum, R.H., Schuemie, M.J., Hendriksen, P.J., Schijvenaars, B.J., Mulligen, E.M.V., Kleinjans, J., Kors, J.A.: A dictionary to identify small molecules and drugs in free text. *Bioinformatics* **25**, 2983–2991 (11 2009). <https://doi.org/10.1093/BIOINFORMATICS/BTP535>, <https://academic.oup.com/bioinformatics/article/25/22/2983/180399>
 13. Jingbo Shang, Liyuan Liu, Xiaotao Gu, Xiang Ren, T.R., Han, J.: shangjingbo1226/AutoNER: Learning Named Entity Tagger from Domain-Specific Dictionary (2018), <https://github.com/shangjingbo1226/AutoNER>
 14. Li, H., Li, J., Gao, W., Zhen, C., Feng, L.: Systematic analysis of ovarian cancer platinum-resistance mechanisms via text mining. *Journal of Ovarian Research* **13**, 1–6 (3 2020). <https://doi.org/10.1186/S13048-020-00627-6/FIGURES/2>, <https://ovarianresearch.biomedcentral.com/articles/10.1186/s13048-020-00627-6>
 15. Liu, S., Tang, B., Chen, Q., Wang, X.: Drug name recognition: Approaches and resources. *Information* 2015, Vol. 6, Pages 790-810 **6**, 790–810 (11 2015). <https://doi.org/10.3390/INFO6040790>, <https://www.mdpi.com/2078-2489/6/4/790>

16. López-Úbeda, P., Díaz-Galiano, M.C., Ureña-López, L.A., Martín-Valdivia, M.T.: Combining word embeddings to extract chemical and drug entities in biomedical literature. *BMC Bioinformatics* **22**, 1–17 (2021). <https://doi.org/10.1186/s12859-021-04188-3>
17. Magiorakos, A.P., Srinivasan, A., Carey, R.B., Carmeli, Y., Falagas, M.E., Giske, C.G., Harbarth, S., Hindler, J.F., Kahlmeter, G., Olsson-Liljequist, B., Paterson, D.L., Rice, L.B., Stelling, J., Struelens, M.J., Vatopoulos, A., Weber, J.T., Monnet, D.L.: Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clinical Microbiology and Infection* **18**, 268–281 (2012). <https://doi.org/10.1111/j.1469-0691.2011.03570.x>
18. Mahesh, B.: Machine learning algorithms-a review. *International Journal of Science and Research* (2018). <https://doi.org/10.21275/ART20203995>, www.ijsr.net
19. McKittrick, A., Corvi, J.: GitHub - ProjectDebbie/pubmed_retrieval: python code to retrieve abstracts from pubmed (2020), https://github.com/ProjectDebbie/pubmed_retrieval
20. Munita, J.M., Arias, C.A.: Mechanisms of antibiotic resistance. *Microbiology spectrum* **4** (3 2016). <https://doi.org/10.1128/MICROBIOLSPEC.VMBF-0016-2015>, /pmc/articles/PMC4888801/ /pmc/articles/PMC4888801/?report=abstract <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4888801/>
21. Neto, D.M.: Infecções associadas aos microrganismos ESKAPE e respetiva resistência antimicrobiana (2015)
22. Oliveira, D.M.D., Forde, B.M., Kidd, T.J., Harris, P.N., Schembri, M.A., Beatson, S.A., Paterson, D.L., Walker, M.J.: Antimicrobial resistance in esophageal pathogens. *Clinical Microbiology Reviews* **33** (2020). <https://doi.org/10.1128/CMR.00181-19/ASSET/CBA1C1D1-CF90-43DE-A9DF-32D24A4334AC/ASSETS/GRAPHIC/CMR.00181-19-F0001.jpeg>, <https://journals.asm.org/doi/10.1128/CMR.00181-19>
23. Peng, J., Zhao, M., Havrilla, J., Liu, C., Weng, C., Guthrie, W., Schultz, R., Wang, K., Zhou, Y.: Natural language processing (nlp) tools in extracting biomedical concepts from research articles: a case study on autism spectrum disorder. *BMC Medical Informatics and Decision Making* **20**, 1–9 (12 2020). <https://doi.org/10.1186/S12911-020-01352-2/TABLES/3>, <https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-020-01352-2>
24. Qin, X., Liu, J., Wang, Y., Liu, Y., Deng, K., Ma, Y., Zou, K., Li, L., Sun, X.: Natural language processing was effective in assisting rapid title and abstract screening when updating systematic reviews. *Journal of Clinical Epidemiology* **133**, 121–129 (5 2021). <https://doi.org/10.1016/j.jclinepi.2021.01.010>, <http://www.jclinepi.com/article/S0895435621000147/fulltext> <http://www.jclinepi.com/article/S0895435621000147/abstract> [https://www.jclinepi.com/article/S0895-4356\(21\)00014-7/abstract](https://www.jclinepi.com/article/S0895-4356(21)00014-7/abstract)
25. Sefton, A.M.: Mechanisms of antimicrobial resistance: Their clinical relevance in the new millennium. *Drugs* **62**, 557–566 (9 2002). <https://doi.org/10.2165/00003495-200262040-00001/FIGURES/TAB4>, <https://link.springer.com/article/10.2165/00003495-200262040-00001>
26. Segura-Bedmar, I., Martínez, P., de Pablo-Sánchez, C.: Using a shallow linguistic kernel for drug–drug interaction extraction. *Journal of Biomedical Informatics* **44**(5), 789–804 (oct 2011). <https://doi.org/10.1016/J.JBI.2011.04.005>

27. Segura-Bedmar, I., Martínez, P., Segura-Bedmar, M.: Drug name recognition and classification in biomedical texts: A case study outlining approaches underpinning automated systems. *Drug Discovery Today* **13**(17-18), 816–823 (sep 2008). <https://doi.org/10.1016/J.DRUDIS.2008.06.001>
28. Shang, J., Liu, J., Jiang, M., Ren, X., Voss, C.R., Han, J.: Automated Phrase Mining from Massive Text Corpora. *IEEE Transactions on Knowledge and Data Engineering* **30**(10), 1825–1837 (feb 2017). <https://doi.org/10.1109/TKDE.2018.2812203>, <https://arxiv.org/abs/1702.04457v2>
29. Shang, J., Liu, L., Gu, X., Ren, X., Ren, T., Han, J.: Learning Named Entity Tagger using Domain-Specific Dictionary. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018* pp. 2054–2064 (sep 2018). <https://doi.org/10.18653/v1/d18-1230>, <https://arxiv.org/abs/1809.03599v1>
30. Su, M., Satola, S.W., Read, T.D.: Genome-based prediction of bacterial antibiotic resistance. *Journal of Clinical Microbiology* **57** (3 2019). <https://doi.org/10.1128/JCM.01405-18>, [/pmc/articles/PMC6425178/](https://pmc/articles/PMC6425178/), [/pmc/articles/PMC6425178/?report=abstract](https://pmc.ncbi.nlm.nih.gov/pmc/articles/PMC6425178/?report=abstract) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6425178/>
31. Wolf, M.: maxime7770/NER-Medical-Document: Named-entity recognition (drugs, diseases, ...) in medical documents. Visualization using streamlit., <https://github.com/maxime7770/NER-Medical-Document>
32. Zaremba, S., Ramos-Santacruz, M., Hampton, T., Shetty, P., Fedorko, J., Whitmore, J., Greene, J.M., Perna, N.T., Glasner, J.D., Plunkett, G., Shaker, M., Pot, D.: Text-mining of pubmed abstracts by natural language processing to create a public knowledge base on molecular mechanisms of bacterial enteropathogens. *BMC Bioinformatics* **10**, 1–11 (2009). <https://doi.org/10.1186/1471-2105-10-177>