

## A SYSTEMATIC REVIEW OF MACHINE LEARNING METHODS FOR FINDING MOLECULES WITH ANTIMICROBIAL ACTIVITY

---

**Lilyana Ivanova**

Training Sector "Assistant Pharmacist", Medical College,  
Medical University "Prof. Dr. Paraskev Stoyanov" - Varna, Bulgaria

**Ch. Asst. Prof. Momchil Lambev, PhD**

Training Sector "Assistant Pharmacist", Medical College,  
Medical University "Prof. Dr. Paraskev Stoyanov" - Varna, Bulgaria  
E-mail: momchil.lambev@mu-varna.bg

**Asst. Prof. Plamen Bekyarov**

Training Sector "Assistant Pharmacist", Medical College,  
Medical University "Prof. Dr. Paraskev Stoyanov" - Varna, Bulgaria  
E-mail: plamen.bekyarov@mu-varna.bg

**Assoc. Prof. Nadya Agova, PhD**

Department of Pharmaceutical Chemistry, Faculty of Pharmacy,  
Medical University "Prof. Dr. Paraskev Stoyanov" - Varna, Bulgaria  
E-mail: nadya.agova@mu-varna.bg

**Abstract:** According to the World Health Organization (WHO), antimicrobial resistance poses a significant threat to human health. Traditional antimicrobial treatments are becoming less effective, necessitating the exploration of alternative options such as the use of nanoparticles, bacteriophages, and the application of artificial intelligence in the development of new drugs and approaches to therapy.

Machine learning is a type of artificial intelligence in which computer algorithms are trained from known data to predict possible outcomes for various tasks. It is used in various healthcare fields, such as discovering new drug molecules, diagnostic methods, and clinical data analysis.

For the purposes of the present study, we conducted a literature search in the Scopus database using the PRISMA charting method. The inclusion and exclusion criteria in the present study are as follows: keywords (machine learning, antimicrobial activity), free access to publications, subject matter, and original articles. According to the set parameters, 13 articles published in the period 2018-2024 and only in English were selected.

**Keywords:** antimicrobial activity, antimicrobial resistance, machine learning, PRISMA diagram, artificial intelligence

### ВЪВЕДЕНИЕ

Справянето с антибиотичната резистентност е от решаващо значение, тъй като представлява значителна заплаха за глобалното здраве. Без ефективни антимикробни препарати възможността за лечение на инфекциозните заболявания ще бъде сериозно компрометирана. По тази причина трябва да се проучат алтернативни възможности като приложение на антимикробни пептиди, наночастици, терапия с бактериофаги и въвеждане на изкуствен интелект в практиката.

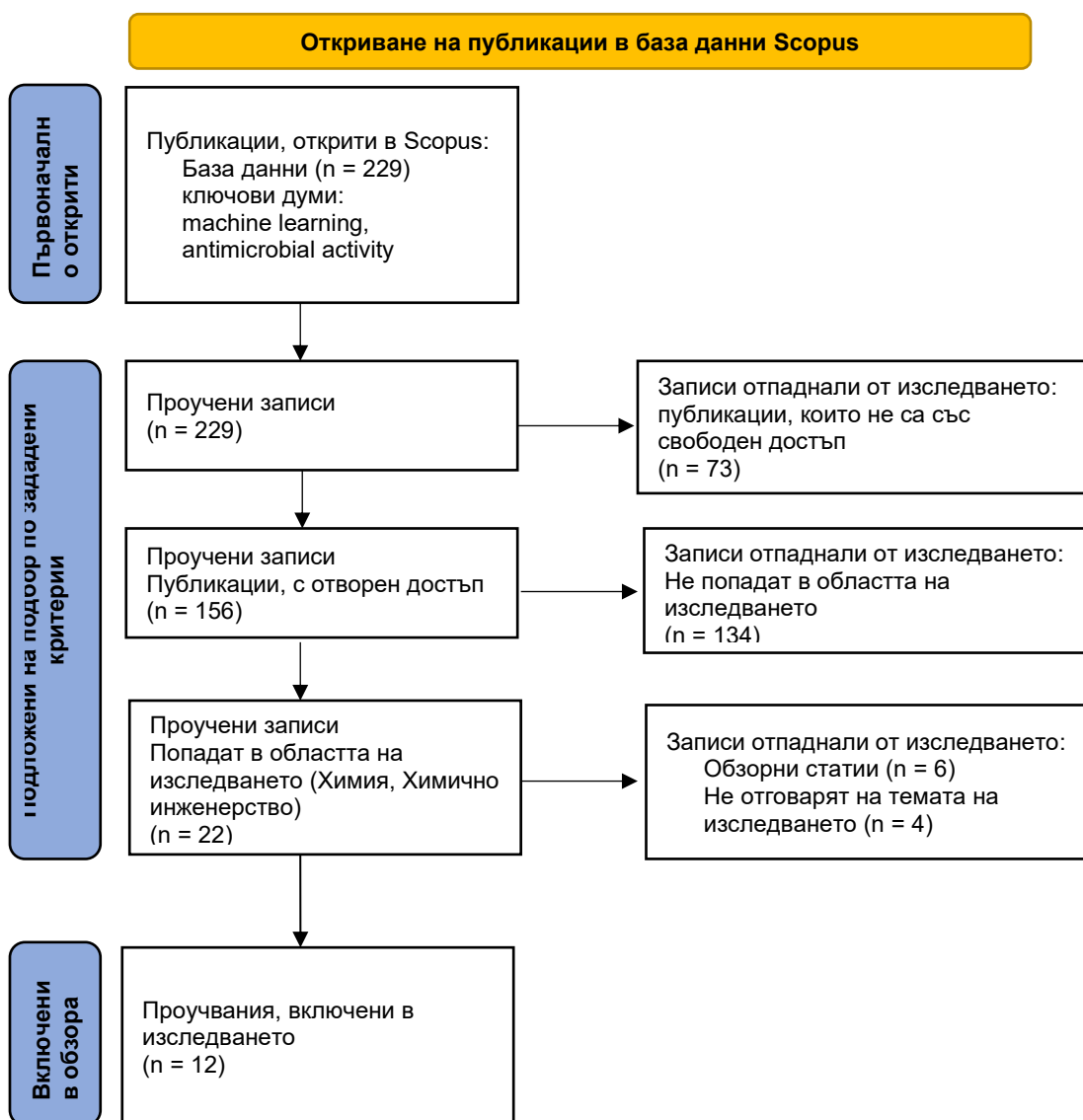
### ИЗЛОЖЕНИЕ

За изготвянето на настоящия литературен обзор беше използван статистически анализ с помощта на Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) диаграма (Фигура 1). Подобен тип изследвания съчетават констатациите от множество проучвания, сортирани на база различни критерии.

В областта на здравеопазването систематичните прегледи и мета-анализите стават все по-популярни като отправна точка за разработване на нови стратегии за терапия и синтез на

нови молекули. Целта на този вид анализ е да предостави систематизирано голям набор от данни, които са подбрани въз основа на редица параметри. По този начин се предоставя лесен достъп до информация както на медицинските специалисти, така и на фармацевтичните компании.

В тази статия разгледахме съществуващата литература за приложения за машинно обучение по отношение на прогнозирането на антимикробна резистентност. При разработване на анализа проведохме търсене в научна база данни (Scopus), въз основа на подбрани ключови думи. След това проведохме подбор на публикациите, според зададени от нашия екип критерии: свободен достъп на публикациите, да попадат в областта на изследването (химия и химично инженерство), както и да бъдат оригинални проучвания. Според зададените критерии за търсене резултата беше 16 публикации на английски език, всички публикувания в периода 2018 – 2024 година. В последствие, четири от статиите бяха премахнати от проучването, поради несъответствие с тематиката на обзора.



Фиг. 1. PRISM диаграма (Page et al., 2021)

### **Използване на machine learning (ML) алгоритми в прицелната терапия**

Оказва се, че компютърните модели се справят добре с предсказване на възможните мишени на single guide RNA (sgRNAs) в бактериалните геноми, но не успяват да прогнозираят активността им в бактериите, поради различни причини. Затова при прогнозиране не може да се направи ясно разделяне между терапевтична активност и токсичност. Недостатъците на съществуващите модели показват нуждата от създаване на подходящи набори от данни, които да предоставят нужната информация за валидиране и обучение.

Dalton и сътр., създават експериментален модел, който сравнява поведението на комбинацията от Cas9/sgRNA при различни състояния, за да отчете насочващата активност на sgRNAs. Тяхната работа води до създаването на набори от данни, които могат да се използват за обучението на *crispr machine transfer learning* (crispHAL) модел за предсказване. Резултатите от проведеният експеримент с *E. coli* показват, че набори от данни с малко количество информация, която е качествено подбрана за обучението на модела, водят и до повече правилни прогнози за антибактериална активност от страна на crispHAL модела (Nam et al., 2023).

### **Употреба на ML подходи при получаването на пептиди с антимикробни свойства**

Процесът на откриване на нови антимикробни пептиди (AMPs) е скъп, продължителен и най-важното не винаги е ефективен. Това довежда до търсенето на компютърни инструменти, които да провеждат скрининг на протеоми за AMPs с висока точност. Вече съществуват редица методи, използвани за предсказване на активността на AMPs. Въпреки това, все още съществуват предизвикателства, които трябва да се преодолеят, например откриването на AMPs с по-дълга пептидна верига (Burdukiewicz et al., 2020).

Компютърните проучвания подпомагат по-доброто разбиране на влиянието на физикохимичните характеристики на пептидите върху тяхната стабилност и активност. Това дава възможност да се преодолеят проблемите, свързани с дизайна на нови пептиди, с добра стабилност и активност срещу широк спектър от патогени.

В проучването на екипът на Söylemez, за първи път е създаден ML подход, който специфично определя активност само срещу Грам (+) или Грам (-) бактерии. За тази цел при започване на изследването са създадени два набора от данни, един за Грам (-) бактерии (*E. coli*, *P. aeruginosa*, *A. baumannii*) и друг за Грам (+) – *S. aureus*, *L. monocytogenes* и *B. cereus*. При създаването на наборите от данни са взети предвид: антибактериална активност, спрямо подбраните щамове, както и те да бъдат с механизъм на действие насочен срещу бактериалната мембрана. От данните са елиминирани пептиди с много сходна аминокиселинна последователност (80%).

Предсказването на антимикробната активност за двете групи става, като се вземат предвид различни физикохимични свойства (нетен заряд, изоелектрична точка, хидрофобност). При създаването на различните модели са използвани и различни класификационни алгоритми. Създаденият класификационен модел би послужил за по-прецизно предсказване и дизайн на AMPs с висока ефективност срещу специфични бактериални патогени (Söylemez et al., 2022).

Burdukiewicz и сътр. провеждат изследване с цел създаване на високо ефективен класификатор – AmpGram, който ефективно сканира протеоми не само от типични AMPs, но и на такива с по-дълга верига и т.нар. „прикрити“ AMPs. „Прикритите“ AMPs представляват пептидни последователности, съдържащи се в други протеини, които не проявяват антимикробни свойства. Типичен пример е тромбина, който няма антимикробни свойства, но част от неговата молекула проявява такива.

Екипът на Burdukiewicz провежда сравнителен анализ, в който сравняват AmpGram с AMPScanner, ADAM, iAMP-2L, CAMPR3 и iAMPpred, които са предикторите за откриване на AMPs с най-добро представяне. Те са сравнени спрямо стойностите на следните показатели: площ под кривата на концентрацията, прецизност, чувствителност и специфичност. От резултатите става ясно, че AmpGram демонстрира едни от най-добрите показатели, при всички изследвани параметри.

AmpGram показва висок потенциал за откриването на части от молекулите, които могат да проявят антимикробен ефект. Неговият процес на работа се основава на сканиране на 10 аминокиселини от молекулата, които след това се наслагват и се оценява дали могат да бъдат използвани като AMPs. Моделът дори открива вече доказани фрагменти от молекули, които имат активност срещу бактериални патогени, а именно лактоферицин и лактоферамин. Освен това AmpGram показва, че и други все още неизследвани фрагменти от лактоферин могат да се използват като AMPs (Burdukiewicz et al., 2020).

В своето проучване екипът на Sneideris използва ML алгоритъм за предсказване на фазовото разделяне на белтъци и комплекси от пептид с нуклеинова киселина. За целта на изследването учените подобряват вече съществуващ модел, който предсказва склонността за фазово разделяне само при белтъци и пептиди. Новополученият модел е дообучен и за взаимодействията между белтъци и нуклеинови киселини. По този начин се разширява неговият обхват за предсказване. Моделът предсказва голяма вероятност за фазово разделяне в присъствието на нуклеинови киселини, при AMPs включени в изследваните данни (Sneideris et al., 2023).

Амилоидните протеини са разнообразна група от молекули, които могат да образуват нишковидни агрегати с кръстосани  $\beta$ -структури. Те са свързани с разстройства като Алцхаймер, Паркинсон и Кройцфелд-Якоб. Амилоидите споделят редица прилики с AMPs, например те могат да убиват микроорганизми, като увреждат мембраните им. Амилоидните мономери се събират в олигомери и разрушават бактериалните мембрани, като създават канали или действат като детергенти. Антимикробните свойства на амилоидите, особено амилоид  $\beta$ , са добре проучени. Както амилоидите, така и AMPs имат деликатен баланс между антимикробни и цитотоксични свойства. Те също така споделят тенденцията да образуват нишковидни агрегати. Някои AMPs се отлагат като амилоиди при човешки амилоидопатии, докато други инхибират амилоидната агрегация. Друго сходство между AMPs и амилоидите е тяхната способност да модулират вродената имунна система. И двете могат да ангажират широк спектър от имунни рецептори, включително TLR2, TLR3, TLR4, TLR9, FPR2, FPR1 и NLRP3 (Bin Hafeez et al., 2021; Kumar et al., 2018).

Предвид тези прилики и факта, че антимикробните свойства на късоверижните амилоиди все още не са изследвани, Przemysław Gagat и екип от учени избира група от потенциално антимикробни късоверижните амилоиди, за да изследват тяхното въздействие върху бактериални и еукариотни клетки. Използват AmpGram, класификационен модел на AMPs, и избират десет амилоида. Според получените резултати четири от тестваните амилоиди образуват агрегати при условията, използвани за оценка на антимикробните свойства на пептидите. Въпреки това нито един от тестваните амилоиди не проявява антимикробни или цитотоксични свойства. Следователно, те трябва да бъдат включени в отрицателните набори от данни за обучение на моделите за прогнозиране на AMPs от следващо поколение, базирани на експериментално потвърдени AMP и не-AMP последователности (Gagat et al., 2023).

### **Приложение на ML за класификация на антимикробната активност на имидазолиевите производни**

Един от проблеми в областта на търсенето на нови активни съединения е множеството молекули, които могат да бъдат синтезирани. Не всички от тях обаче притежават желаните физико-химични качества и свойства, за да бъдат използвани в терапията. Процесът на търсене може да бъде рационализиран чрез използване на компютърни модели и ML в началните етапи на изследване. Подобен подход позволява намаляване на необходимостта от ресурси като време, пари и химикали. Крайната цел е да се идентифицират най-обещаващите съединения, които след това могат да бъдат синтезирани и подробно анализирани (Liu and Stokes, 2022).

В разгледаното проучване невронна мрежа Multilayer Perceptron (MLP), заедно с други техники за машинно обучение (Classification And Regression Decision Trees (CART), Principal Component Analysis (PCA)) е използвана за прогнозиране на антимикробната активност на 140 новоразработени имидазолони хлориди срещу *Klebsiella pneumoniae* (Lorenc et al., 2024).

В своето изследване Andželika Lorenc и колектив използват CART, за да се идентифицират най-добрите 200 предсказуеми молекулярни дескриптора. След това с помощта на PCA, намалят тези дескриптори до 5 компонента, запазвайки 99,57% от информацията за необработените данни.

Антимикробната активност се категоризира като висока или ниска въз основа на експериментално доказана минимална инхибиторна концентрация (MIC), с гранична точка при MIC=0,856 mol/L. 12-кратно кръстосано валидиране обучава MLP с архитектура от 5-12-2 и 5 основни компонента. MLP демонстрира добро представяне, постигайки почти 90% правилни класификации в процесите на обучение, валидиране и тестове, превъзхождайки моделите без намаляване на размерите на PCA. Ключови показатели, включително точност (0,907), чувствителност (0,905), специфичност (0,909) и прецизност (0,891), са високи. Тези резултати подчертават ефикасността на модела MLP с PCA като висококачествен класификатор за определяне на антимикробната активност.

В заключение, това проучване твърди, че MLP невронната мрежа, в комбинация с CART и PCA, е стабилен инструмент за прогнозиране на класа на антимикробна активност на имидазони хлориди срещу *Klebsiella pneumoniae*. Използването на CART и PCA в проучването позволява намаляване на входната променлива без значителна загуба на информация. Високата точност на класификацията и свързаните с нея показатели потвърждават потенциалната полезност на метода при предварителни оценки на синтеза, предоставяйки ценна информация за дизайна на антимикробни съединения.

### Дизайн на кандидат-анти tubеркулозни лекарства с ML модели

Туберкулозата (ТБ) продължава да представлява значителна глобална заплаха за здравето, въпреки постигнатия напредък в борбата с болестта. През 2014 г. Световната здравна асамблея одобри Стратегията за край на туберкулозата с амбициозната цел за намаляване на заболяемостта от туберкулоза, смъртността и разходите за лечение до 2030 г. Тази стратегия подчертава критичната необходимост от засилени изследвания и иновации за спиране на развитието на туберкуозна епидемия.

Лупанът е ключова структура в пентацикличните тритерпени, които са растителни метаболити. Съединения като бетулин, бетулинова киселина и лупеол са показали противоракови и противотуберкулозни свойства. От екипът на Kazakova, са синтезирани двадесет азепа-нотритерпеноиди с лупанов тип А-пръстен от бетулин и неговите сродни производни. Тяхната анти tubеркулозна активност е изследвана срещу *M. tuberculosis*, монорезистентни МТВ щамове и нетуберкулозни щамове *M. abscessus* и *M. Avium* (Kazakova et al., 2021).

От разгледаното проучване става ясно, че седемнадесет съединения показват активност, като шест съединения показват висока активност срещу дивия щам H37Rv. Едно от тях е идентифицирано като потенциален инхибитор на *M. tuberculosis* RNAP и показва адитивен ефект, когато се комбинира с рифампин, изониазид и етамбутол.

### Приложение на ML срещу антимикробна резистентност

Най-новите класификатори за ML са успешно използвани за анализиране на данни за секвениране на целия геном и идентифициране на генетични фактори, свързани с антимикробна резистентност (AMR). Тези класификатори обаче нямат способността да предоставят причинно-следствена интерпретация. В отговор на това ограничение е разработен класификатор за машинно обучение, базиран на метаболитен модел, наречен Класификатор на метаболитни алели (MAC). MAC използва анализ на баланса на потока, за да оцени биохимичните ефекти на генетичните вариации (Boolchandani et al., 2019).

В разгледаното проучване на Erol S. Kavvas и колектив е използван MAC върху набор от данни от с 1595 лекарства тествани върху *M. tuberculosis*. Те заключават, че MAC може да предвиди AMR фенотипове с точност, подобна на други модели на машинно обучение (например AUC на изониазид = 0,93). Важно е, че MAC също позволява биохимична интерпретация на връзката генотип-фенотип. Когато се прилагат към три специфични съединения (пиразинамид, пара-аминосалицилова киселина и изониазид). Резултатите от MAC се приравняват към известните механизми на AMR и дават представа за биохимичната основа за това как идентифицираните генетични вариации водят до AMR. Чрез разширяване на анализа на баланса на потока, за да се класифицират точно генетичните последователности, могат да се прогнозира промени в генома (GWAS) (Kavvas et al., 2020).

Baker и сътр. са използвали данни от изолати на *E. coli* и *S. enterica*, взети от пилета в различни ферми. Чрез геномно секвениране и анализ на антибиотичната чувствителност изследователите анализират еволюционните процеси, които тези микроорганизми претърпяват, за да се приспособят към антимикробната терапия и към конкурентната среда в птицефермите. Те използват ML, за да се предскажат еволюционните тенденции и да се открият ключови мутации, свързани с резистентността. Използвани са модели за предсказване на мутации, които водят до резистентност към антибиотици, както и методи за оценка на адаптивността на патогените към околната среда. Изследването показва, че при *E. coli* и *S. enterica* се наблюдават сходни еволюционни отговори по отношение на антибиотичната резистентност. Наблюдавано е сходство при резистентните механизми, при което и двата патогена развиват резистентност към едни и същи класове антибиотици. Това се дължи както на споделяне на резистентни гени чрез хоризонтален генен трансфер, така и на независими мутации (Baker et al., 2024).

### **Използване на ML алгоритми в разработването на антимикробни средства**

В проучване на Tiïhonen и сътр. се разглежда предсказването на антимикробната активност на молекули, наречени конюгирани олигоелектролити (COE), чрез машинно обучение. Този нов клас антибиотици привлича внимание поради възможността да се борят с резистентността към антибиотици, но техните свойства и структура все още не са добре проучени.

Изследователите създават модел за предсказване на MIC срещу *E. coli* K12, използвайки 21 избрани молекулни дескриптора от първоначалния набор от 5305 кандидата. Моделът постига голяма точност ( $R^2 = 0.65$ ), което е много добър резултат, предвид ограниченията в наличните данни за този клас молекули.

Използваният подход за машинно обучение е съставен от три стъпки и включва:

- Селектиране на молекулни дескриптори – От първоначалния набор от 5305 дескриптора са избрани 21, които оказват най-голямо влияние върху антимикробната активност;

- Представяне на молекулите – Ключова роля за точността на модела играе представянето на триизмерната форма на молекулите, което е от съществено значение за COE;

- Предсказване на MIC – Моделът предсказва минималната инхибиторна концентрация, като използва данни от експерименти с *E. coli* K12.

Въпреки че, моделът не изисква предварителни знания за механизма на действие на COE, резултатите показват, че 3D структурата на молекулите е основен фактор за точността на предсказването. Методологията може лесно да бъде адаптирана за други нови класове антибиотици и предоставя надежден инструмент за ускоряване на процеса на разработване на нови антимикробни агенти (Tiïhonen et al., 2021).

### **Използване на ML за предвиждане на резистентност към гъбички**

Машинното обучение може да бъде използвано за предсказване на важноста на гени и взаимодействията между химични вещества с цел намиране на нови мишени за противогъбична терапия.

В публикация на Fu и сътр. ML е използван, за да предскаже кои гени са жизненоважни за оцеляването на патогенните гъбички. Тези гени са потенциални цели за разработване на нови лекарства. Също така моделът анализира взаимодействията между химичните агенти и генома на гъбите. Целта е да се установи как различни химични съединения взаимодействат с генетични мишени, за да предложи нови противогъбични средства.

За реализирането на тези модели се използват големи набори от данни, съдържащи генетична информация, както и данни за хиляди химични взаимодействия. Те са получени от различни източници, включително експериментални данни и данни за геномите на гъбични патогени като *Candida*.

Създадените алгоритми предсказват кои гени са жизненоважни, като същевременно откриват нови химични съединения, които могат да взаимодействат с тези мишени. В резултат се генерират списъци с потенциални целеви гени и молекули. Този модел е обучен и валидиран с използване на вече известни данни за гъбични гени и техните взаимодействия с различни химични вещества. Приложени са техники за кръстосана валидация, за да се гарантира надеждността на предсказанията (Fu et al., 2021).

### Приложение на ML в drug repurposing

ML може да се използва и за откриване на антимикробна активност във вече разрешени за употреба лекарствени продукти. Изследване на екипът на Lara, публикувано през 2019 година цели да намери нови възможности за лечение на резистентни патогени, чрез откриване на нов ефект във вече одобрени лекарства. Чрез машинно обучение авторите анализират съществуващи данни за лекарства, които първоначално са били разработени за други заболявания, за да идентифицират такива, които имат потенциал да действат и срещу бактерии и други патогени.

За целта те използват хетероложно машинно обучение, което комбинира данни от различни източници, като генетични профили на бактерии и човешки таргети, с цел да се предскаже антимикробната активност на съществуващите лекарства. Използват се алгоритми като регресионни модели, support vector machine и невронни мрежи за създаване на модели, които могат да предскажат антимикробна активност въз основа на характеристиките на молекулите. Основното предимство на този подход е, че моделът се обучава на различни видове биологични данни (данни за химични структури, бактериалния геном, данни за бактериалните протеини) (Nava Lara et al., 2019).

### ИЗВОДИ

Алгоритмите за машинно обучение се прилагат все повече в здравеопазването. Сред множеството приложения на тези нови методи, използването им в борбата срещу антимикробната резистентност (AMR) е една от най-важните области на приложение, тъй като повишаването на резистентността към антибиотици и лечението на трудно лечими инфекции с мултирезистентност са значителни предизвикателства пред медицината и фармацията. Сред начините за справяне с този проблем са различните ML подходи разгледани в настоящата статия: синтез на нови антибактериални средства, разработване на класификатори за определяне на антимикробна активност и drug repurposing.

### REFERENCES

- Baker, M., Zhang, X., Maciel-Guerra, A., Babaarslan, K., Dong, Y., Wang, W., Hu, Y., Renney, D., Liu, L., Li, H., Hossain, M., Heeb, S., Tong, Z., Percy, N., Zhang, M., Geng, Y., Zhao, L., Hao, Z., Senin, N., Chen, J., Peng, Z., Li, F., Dottorini, T., 2024. Convergence of resistance and evolutionary responses in *Escherichia coli* and *Salmonella enterica* co-inhabiting chicken farms in China. *Nat Commun* 15, 206. <https://doi.org/10.1038/s41467-023-44272-1>
- Bin Hafeez, A., Jiang, X., Bergen, P.J., Zhu, Y., 2021. Antimicrobial Peptides: An Update on Classifications and Databases. *IJMS* 22, 11691. <https://doi.org/10.3390/ijms222111691>

- Boolchandani, M., D'Souza, A.W., Dantas, G., 2019. Sequencing-based methods and resources to study antimicrobial resistance. *Nat Rev Genet.* <https://doi.org/10.1038/s41576-019-0108-4>
- Burdukiewicz, M., Sidorchuk, K., Rafacz, D., Pietluch, F., Chilimoniuk, J., Rödiger, S., Gagat, P., 2020. Proteomic Screening for Prediction and Design of Antimicrobial Peptides with AmpGram. *IJMS* 21, 4310. <https://doi.org/10.3390/ijms21124310>
- Fu, C., Zhang, X., Veri, A.O., Iyer, K.R., Lash, E., Xue, A., Yan, H., Revie, N.M., Wong, C., Lin, Z.-Y., Polvi, E.J., Liston, S.D., VanderSluis, B., Hou, J., Yashiroda, Y., Gingras, A.-C., Boone, C., O'Meara, T.R., O'Meara, M.J., Noble, S., Robbins, N., Myers, C.L., Cowen, L.E., 2021. Leveraging machine learning essentiality predictions and chemogenomic interactions to identify antifungal targets. *Nat Commun* 12, 6497. <https://doi.org/10.1038/s41467-021-26850-3>
- Gagat, P., Duda-Madej, A., Ostrówka, M., Pietluch, F., Seniuk, A., Mackiewicz, P., Burdukiewicz, M., 2023. Testing Antimicrobial Properties of Selected Short Amyloids. *IJMS* 24, 804. <https://doi.org/10.3390/ijms24010804>
- Ham, D.T., Browne, T.S., Banglorewala, P.N., Wilson, T.L., Michael, R.K., Gloor, G.B., Edgell, D.R., 2023. A generalizable Cas9/sgRNA prediction model using machine transfer learning with small high-quality datasets. *Nat Commun* 14, 5514. <https://doi.org/10.1038/s41467-023-41143-7>
- Kavvas, E.S., Yang, L., Monk, J.M., Heckmann, D., Palsson, B.O., 2020. A biochemically-interpretable machine learning classifier for microbial GWAS. *Nat Commun* 11, 2580. <https://doi.org/10.1038/s41467-020-16310-9>
- Kazakova, O., Racoviceanu, R., Petrova, A., Mioc, M., Militaru, A., Udrescu, L., Udrescu, M., Voicu, A., Cummings, J., Robertson, G., Ordway, D.J., Slayden, R.A., Șoica, C., 2021. New Investigations with Lupane Type A-Ring Azepane Triterpenoids for Antimycobacterial Drug Candidate Design. *IJMS* 22, 12542. <https://doi.org/10.3390/ijms222212542>
- Kumar, P., Kizhakkedathu, J., Straus, S., 2018. Antimicrobial Peptides: Diversity, Mechanism of Action and Strategies to Improve the Activity and Biocompatibility In Vivo. *Biomolecules* 8, 4. <https://doi.org/10.3390/biom8010004>
- Liu, G., Stokes, J.M., 2022. A brief guide to machine learning for antibiotic discovery. *Current Opinion in Microbiology* 69, 102190. <https://doi.org/10.1016/j.mib.2022.102190>
- Lorenc, A., Badura, A., Karolak, M., Pałkowski, Ł., Kubik, Ł., Buciński, A., 2024. Antimicrobial Activity Classification of Imidazolium Derivatives Predicted by Artificial Neural Networks. *Pharm Res* 41, 891–898. <https://doi.org/10.1007/s11095-024-03699-x>
- Nava Lara, R.A., Aguilera-Mendoza, L., Brizuela, C.A., Peña, A., Del Rio, G., 2019. Heterologous Machine Learning for the Identification of Antimicrobial Activity in Human-Targeted Drugs. *Molecules* 24, 1258. <https://doi.org/10.3390/molecules24071258>
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* n71.
- Sneideris, T., Erkamp, N.A., Ausserwöger, H., Saar, K.L., Welsh, T.J., Qian, D., Katsuya-Gaviria, K., Johncock, M.L.L.Y., Krainer, G., Borodavka, A., Knowles, T.P.J., 2023. Targeting nucleic acid phase transitions as a mechanism of action for antimicrobial peptides. *Nat Commun* 14, 7170. <https://doi.org/10.1038/s41467-023-42374-4>
- Söylemez, Ü.G., Yousef, M., Kesmen, Z., Büyükkiraz, M.E., Bakir-Gungor, B., 2022. Prediction of Linear Cationic Antimicrobial Peptides Active against Gram-Negative and Gram-Positive Bacteria Based on Machine Learning Models. *Applied Sciences* 12, 3631. <https://doi.org/10.3390/app12073631>
- Tiihonen, A., Cox-Vazquez, S.J., Liang, Q., Ragab, M., Ren, Z., Hartono, N.T.P., Liu, Z., Sun, S., Zhou, C., Incandela, N.C., Limwongyut, J., Moreland, A.S., Jayavelu, S., Bazan, G.C., Buonassisi, T., 2021. Predicting antimicrobial activity of conjugated oligoelectrolyte molecules via machine learning. *J. Am. Chem. Soc.* 143, 18917–18931. <https://doi.org/10.1021/jacs.1c05055>