ECMC-P 2025 Conference

The 1st International Electronic Conference on Medicinal Chemistry and Pharmaceutics

01–30 November 2025 | Online

MDPI

Multitarget QSAR-based Virtual Screening of Natural Product and Drug Databases to Identify Novel Antifungals for Deep Mycoses

Henrique G. W. Martins[1]; Bernardo Mussa[1]; Raul E. L. Lazo[1]; Dile P. Stremel[1]; Roberto Pontarolo[1],*;
[1]Universidade Federal do Paraná, Curitiba - PR.

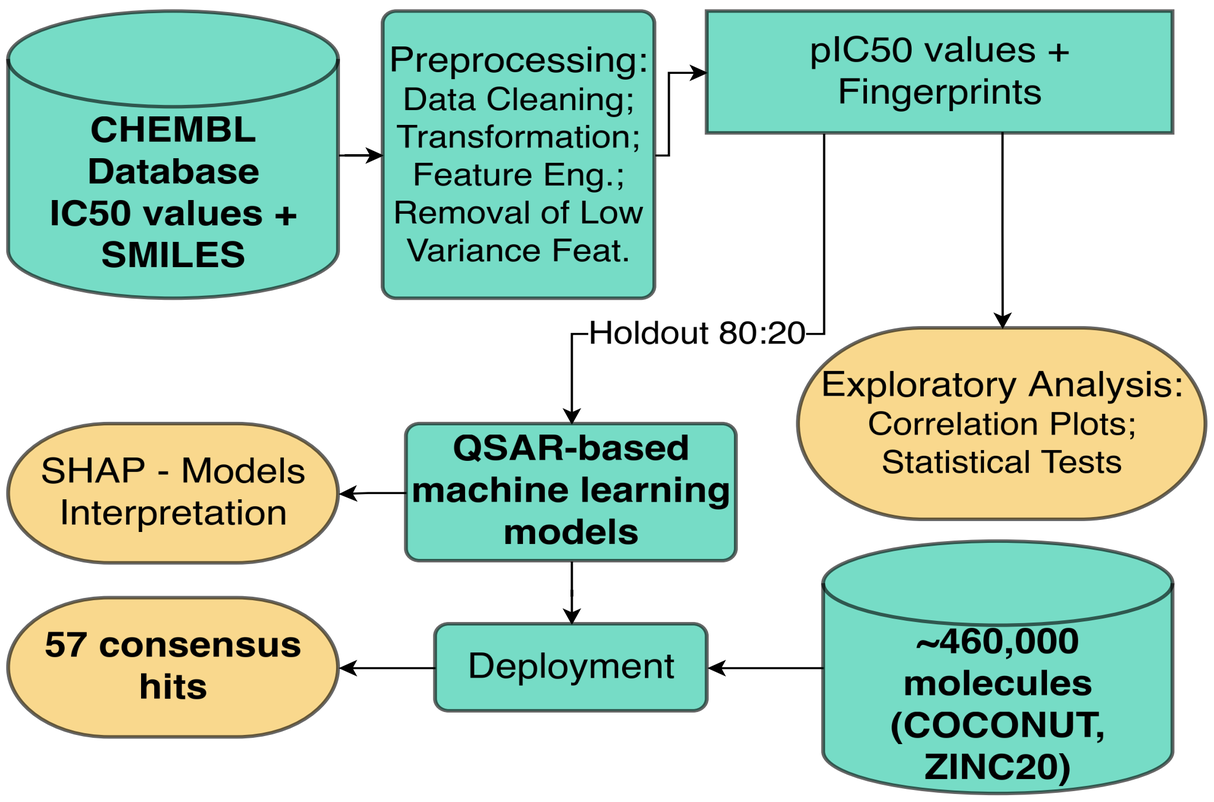
INTRODUCTION & AIM

Deep mycoses, according to the World Health Organization (WHO), are neglected tropical diseases, affecting hundreds of thousands of people around the world.[1] While there are treatments for fungal diseases, the rise of antifungal resistance urges the research of novel antifungal agents.[2] This study aims to screen natural products and approved drugs databases to discover new antifungal agent candidates against deep mycoses through machine learning models based on multitarget quantitative structure-activity relationships (QSAR).[3]

METHOD

Using IC50 data from 16 deep mycoses pathogens from ChEMBL, 5 machine learning regression algorithms (Bagging, Gradient Boosting, LightGBM, Random Forest, XGBoost) were selected from 40 candidates and used to create models. The models were assessed using 10-fold cross-validation (CV) and were then used in a consensus approach to predict the antifungal activity of nearly 460,000 compounds from natural products and approved drugs databases.

Figure 1 – Flowchart of the conducted study



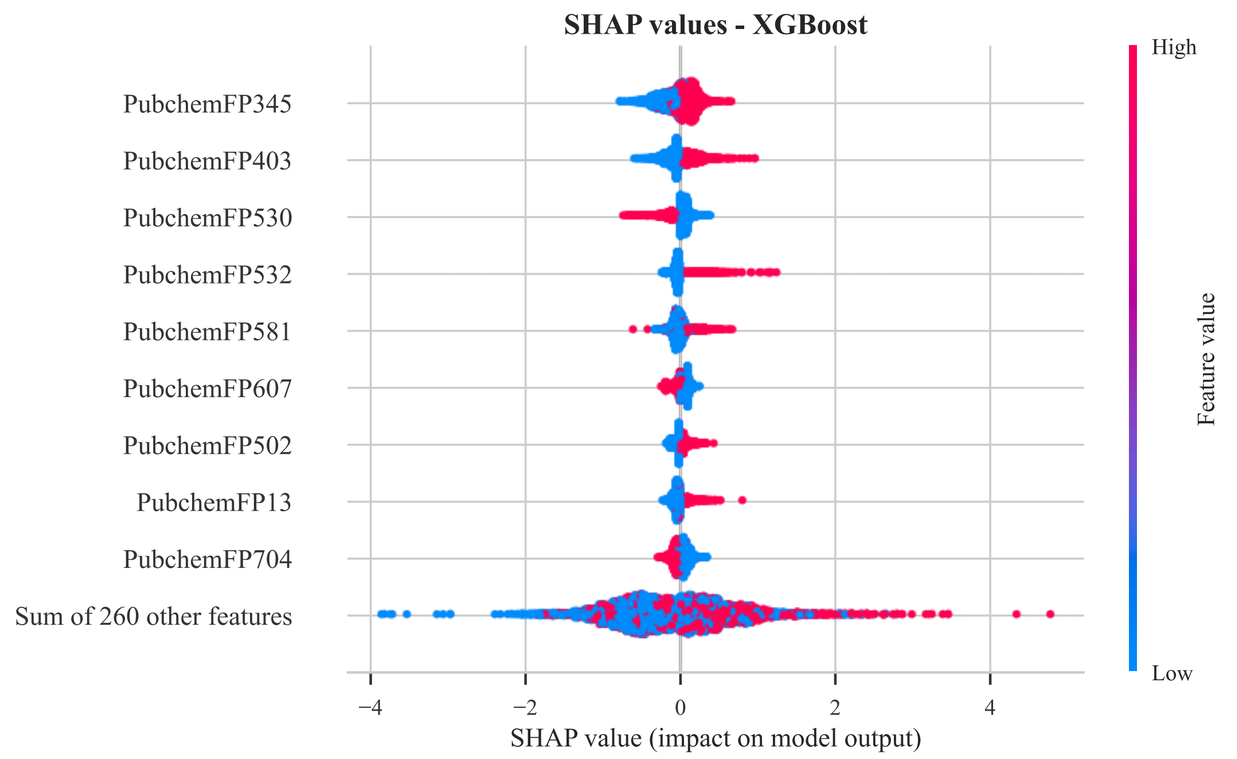
RESULTS & DISCUSSION

Exploratory data analysis revealed that there are statistically significant chemical differences between active and inactive substances.

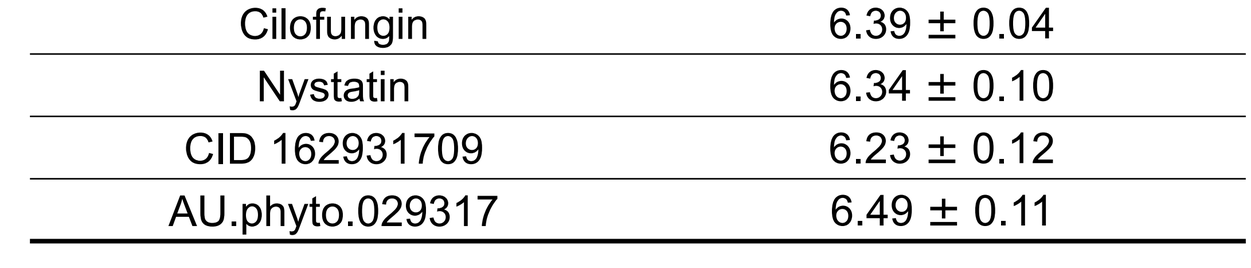
Table 1 - Average 10-fold CV score for the 5 models

R^2	MAE	RMSE
0.71 ± 0.03	0.42 ± 0.02	0.58 ±0.04

Figure 3 - SHAP values beeswarm plot and fingerprint structures



PubchemFP345 PubchemFP403 PubchemFP581



57 compounds were considered active (pIC50 > 6) by the consensus of the models (< 2% CV).

Table 2 – Some of the 57 compounds highlighted by the models.

Compound	Predicted pIC50
Cefoselis	6.66 ± 0.13
Cilofungin	6.39 ± 0.04
Nystatin	6.34 ± 0.10
CID 162931709	6.23 ± 0.12
AU.phyto.029317	6.49 ± 0.11

CONCLUSION

QSAR-based machine learning approaches proved to be a valuable tool for the virtual screening of large databases, though dependent on quality data. This method led to the identification of 57 novel and repurposed antifungal agent candidates, warranting further experimental studies.

FUTURE WORK / REFERENCES

Future work involves improving the machine learning pipeline through data and methodology, as well as conducting experimental analysis on hits. **References on QR code.**

This study was financed in part by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and partially supported by Consenho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (Proc. Nº 384022/2024-3).