enviromed

Next generation toolbox for greener pharmaceuticals design and manufacturing towards reduced environmental impact

# D6.1 – Report on the ecotoxicity studies

**CNR, Paola Italiani**

**January 2026**

Funded by the European Union

| Deliverable D6.1 – Report on the ecotoxicity studies | |
|---|---|
| Work Package(s) | WP6 – Better understanding of pharmaceuticals' environmental impact |
| Task(s) | T6.1 – Review of pharmaceuticals-related environmental issues<br>T6.2 – In-silico ecotoxicity models<br>T6.3 – In-vivo & in-vitro ecotoxicity studies & validation of in-silico models |
| Dissemination Level | Public |
| Due Date | 31-08-2025 |
| Actual Submission Date | 30-01-2026 |
| WP Leader | IRES |
| Task Leaders | IRES, CLOUD, CNR |
| Deliverable Leader | CNR |
| Contact Person | Paola Italiani |
| Email | paola.italiani@cnr.it |
| Authors | Paola Italiani |
| Contributors | Daniela Melillo (CNR), Eleni Chontzopoulou (CLOUD), Evangelos Tsoukas (CLOUD), Michail Papadourakis (CLOUD), Konstantina-Roxani Chatzipanagiotou (IRES), Adamantia Bon (IRES), George Antonaropoulos (IRES) |
| Reviewers | Zoe Zacharouli (MITERA)<br>Giorgos Katsouras (EYDAP) |

| Document History | | | | |
|---|---|---|---|---|
| **Version** | **Date** | **Description** | **Comments** | **Editor(s)** |
| v0.1 | 20-11-2025 | First Version | Table of Contents | P. Italiani (CNR) |
| v0.2 | 15-01-2026 | Second version | Addition of CLOUD contribution | P. Italiani (CNR) |
| v0.3 | 19-12-2025 | Third version | Addition of CNR and IRES contribution | P. Italiani (CNR) |
| v0.4 | 20-01-2026 | Fourth version | Submitted for internal review | P. Italiani (CNR) |
| v0.5 | 23-01-2026 | Reviewed document draft | First peer review | Z. Zacharouli (MITERA) |
| v0.6 | 23-01-2026 | Reviewed document draft | Second peer review | G. Katsouras (EYDAP) |
| v0.7 | 26-01-2026 | Version ready for quality review | Revised version based on reviewers' comments | P. Italiani (CNR) |
| v0.8 | 29-01-2026 | Reviewed document | Quality Manager review | M. Dimitriadou (RISA) |

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101057844

| | | | | |
|---|---|---|---|---|
| v1.0 | 30-01-2025 | Final version | Ready for submission | P. Italiani (CNR) |

# Executive Summary

The purpose of this deliverable is to document the activities carried out and the results obtained within the project in relation to ecotoxicity studies and the assessment of the environmental impact of pharmaceutical compounds. It provides a structured and integrated overview of the methodological approaches developed and applied to support ecotoxicological evaluation and environmental risk assessment. The scope of the deliverable includes a review of the current scientific knowledge and regulatory context concerning pharmaceuticals as emerging environmental contaminants, followed by the presentation of predictive computational approaches for ecotoxicity assessment. These approaches are designed to support the classification of chemical substances according to their ecotoxicological potential and to enable ML model interpretability for indicating the chemical fragments responsible for ecotoxicity. The models developed are integrated into a robust and automated, user-friendly web platform to facilitate wide access and use by researchers and regulators.

In addition, the deliverable reports on experimental investigations conducted using *in vitro* and *in vivo* ecotoxicological models. These activities are aimed at generating biological data for model evaluation and validation, with particular attention to physiological, immunological, and cellular-level responses in selected aquatic organisms. The extrapolation of toxicity predictions across species was further addressed by combining computational outputs with targeted experimental studies. The deliverable also presents a comparative analysis between experimentally derived toxicity endpoints and model-based predictions in order to assess the robustness, reliability, and applicability of the developed tools. Overall, it contributes to the advancement of integrated testing strategies, supports evidence-based decision-making, and aligns with European objectives promoting the development of alternative approaches for environmental safety assessment.

This deliverable directly contributes to the achievement of the project objectives related to the development and validation of integrated approaches for the environmental risk assessment of pharmaceutical compounds. In particular, it supports the project objective of advancing predictive ecotoxicology by combining computational ML modelling with experimental validation, thereby improving the scientific basis for the assessment of ecological hazards in aquatic environments. The activities and results documented in this deliverable are primarily linked to the work packages focused on modelling, experimental ecotoxicology, and validation. The development and application of *in silico* ecotoxicity models contribute to the objectives of the work package addressing data-driven and computational tools, while the experimental *in vitro* and *in vivo* studies support the work packages dedicated to biological testing and model evaluation. Furthermore, the comparison between experimental outcomes and model predictions contributes to cross-work package integration and validation activities. Overall, this deliverable plays a transversal role within the project by facilitating the integration of modelling and experimental work packages, ensuring coherence between methodological development and empirical validation, and supporting the delivery of project outcomes relevant to regulatory science, environmental protection, and the implementation of alternative testing strategies in line with European research and innovation priorities.

The ENVIROMED D6.1 deliverable summarizes efforts to assess pharmaceutical environmental impacts through literature review, *in silico* modeling, and experimental ecotoxicity studies. Key results include prioritization of high-risk drugs, accurate machine learning predictions of fish species toxicity via the G.AI.A platform, and sublethal effects observed in zebrafish validating the models. Computational-experimental integration effectively supports early ecotoxicity screening, exceeding the performance of some public tools and indicating the chemical structures driving toxicity. Furthermore, invertebrate

biomarkers reliably detect pharmaceutical stress in coastal ecosystems. These findings support the expansion of automated tools for regulatory use, the further development of chronic multi-species testing, the prioritization of metabolite monitoring, and the promotion of open-access platforms in support of EU green chemistry objectives.

The integrated computational-experimental approach demonstrates robust predictive power for pharmaceutical ecotoxicity, outperforming public tools like VEGA/T.E.S.T on external datasets (ADORE datasets), while highlighting specific chemical structural features driving toxicity. Invertebrate biomarkers effectively detect early stress from pharmaceuticals, supporting their use in monitoring Mediterranean coastal pollution hotspots. Recommendations include expanding G.AI.A for regulatory screening in green drug design, conducting chronic multi-species studies, prioritizing high-RQ metabolites for field monitoring, and disseminating models via open-access platforms to align with EU alternative testing strategies.

# Table of Contents

# List of Figures

## List of Tables

# List of Acronyms

*Table 1: Acronyms and abbreviations*

| Term | Definition |
|---|---|
| **1D** | One-Dimensional |
| **2D** | Two-Dimensional |
| **3D** | Three-Dimensional |
| **AD** | Applicability Domain |
| **ADORE** | Aquatic Database for Organism-level Response Evaluation |
| **ANN** | Artificial Neural Networks |
| **API** | Active Pharmaceutical Ingredient |
| **AUC** | Area Under the Curve |
| **CLP** | Classification, Labelling and Packaging |
| **CNR** | National Research Council of Italy |
| **CRED** | Criteria for Reporting and Evaluating Ecotoxicity Data |
| **CSS** | Chemicals Strategy for Sustainability |
| **DNN** | Deep Neural Networks |
| **ECOSAR** | Ecological Structure-Activity Relationship |
| **ENT** | Entropy Index |
| **FN** | False Negative |
| **FP** | False Positive |
| **GII** | Gini Index |
| **KNN** | K-Nearest Neighbors |
| **LC50** | Lethal Concentration |
| **LCA** | Life Cycle Assessment |
| **LCAI** | Life Cycle Impact Assessment |
| **LMS** | Lysosomal Membrane Stability |
| **logBCF** | logBioconcentration Factor |
| **MACCS** | Molecular ACCess System keys |
| **MCyt** | Micronucleus Cytome |
| **MEC** | Measured Environmental Concentration |
| **ML** | Machine Learning |
| **NIEHS** | National Institute of Environmental Health Sciences |
| **NSAIDs** | Nonsteroidal Anti-Inflammatory Drugs |

| | |
|---|---|
| **OECD** | Organization for Economic Co-operation and Development |
| **PCA** | Principal Component Analysis |
| **PI** | Phagocytic Index |
| **PNEC** | Predicted No-Effect Concentration |
| **PR** | Phagocytic Rate |
| **QSAR** | Quantitative Structure–Activity Relationship |
| **RASAR** | Read Across Structure Activity Relationships |
| **REACH** | European Union's Registration, Evaluation, Authorisation and Restriction of Chemicals |
| **RMSE** | Root Mean Square Error |
| **ROC** | Receiver Operating Characteristic |
| **RQ** | Risk Quotient |
| **SHAP** | SHapley Additive exPlanations |
| **SMARTS** | SMILES arbitrary target specification |
| **SMILES** | Simplified Molecular Input Line Entry System |
| **T.E.S.T** | Toxicity Estimation Software Tool |
| **TN** | true negative |
| **TP** | true positive |
| **t-SNE** | t-Distributed Stochastic Neighbor Embedding |
| **US EPA** | United States Environmental Protection Agency |
| **WWTPs** | Wastewater Treatment Plant |
| **XGBClassifier** | XGBoost Classifier |

# 1 Introduction

This deliverable address critical environmental challenges associated with pharmaceuticals by advancing the understanding of their occurrence, fate, and potential impacts across ecosystems. The activities carried out within WP6 adopt a holistic and multidisciplinary approach, integrating data-driven modelling, experimental validation, and life cycle–based environmental assessment to comprehensively evaluate pharmaceutical-related risks.

The work combines an extensive and systematic review of peer-reviewed scientific literature with the exploitation of open-source databases and monitoring data. These sources feed advanced computational efforts aimed at the development of in silico ecotoxicity models. In particular, machine learning techniques are employed to identify relevant molecular descriptors and structural features that are predictive of ecotoxicological effects, thereby enabling the screening and prioritisation of pharmaceutical compounds with potential environmental concern. The predictive capacity of these models is strengthened through validation against experimental evidence generated via in vivo and in vitro assays, conducted using selected marine organisms and environmental water samples collected from representative sites.

In parallel, WP6 evaluates the environmental performance of pharmaceutical products and manufacturing processes through the application of established Life Cycle Assessment (LCA) methodologies. This analysis focuses on commonly used industrial synthesis routes and production practices, allowing the identification of environmental hotspots and trade-offs across different stages of the pharmaceutical life cycle, from raw material extraction to end-of-life. A further key outcome of WP6 is the development of a comprehensive and structured registry of pharmaceutical micropollutants detected in wastewater and environmental matrices. This registry enhances data accessibility, transparency, and usability for researchers, regulators, industry, and other stakeholders, supporting evidence-based decision-making and risk assessment.

Through this integrated framework, WP6 provides a robust assessment of major environmental risks associated with pharmaceuticals throughout their entire life cycle and contributes actionable knowledge to support the transition towards greener pharmaceutical design, production, use, and disposal practices.

This document is structured as follows:

- *Chapter 1:* Introduction

This deliverable addresses the environmental challenges associated with pharmaceutical contamination by adopting an integrated and multidisciplinary approach within WP6. It combines data-driven modelling, experimental ecotoxicological validation, and life cycle assessment methodologies to evaluate the occurrence, fate, and potential impacts of pharmaceuticals across ecosystems. By integrating in silico predictions with in vivo and in vitro evidence and life cycle–based analyses, the work supports robust environmental risk assessment and contributes to the development of more sustainable pharmaceutical practices.

- *Chapter 2:* Review of Pharmaceuticals-Related Environmental Issues

In this task, a wide investigation regarding the presence of pharmaceuticals in the environment was performed, taking into account the entire lifecycle of APIs, i.e., considering the design, synthesis, the use phase, waste handling, and control of the emission sources using advanced technologies (e.g., for wastewater treatment). Special focus was given on providing input and insights towards updating the existing legislation on pharmaceutical processes and life cycle impact mitigation, the improvement of risk management options, the prioritization of the pharmaceuticals in the environment, and the improvement of the availability and collection of

pharmaceutical Environmental Risk Assessment data. The above considerations aim to contribute towards the development of a review – roadmap with specific actions, that will support the minimisation of releasing pharmaceuticals into the environment. The activities performed towards these aims, and the resulting insights and outcomes, are described in Section 2 of the present deliverable.

In addition, Task 6.1 results highlight the need of LCA following specific Product Category Rules (PCRs) for the pharmaceutical industry, which will help harmonise and facilitate the future use of LCA in this sector, and resolve the currently identified inconsistencies and inhomogeneity in the methodological choices among different LCA sources on the pharma sector. Finally, it aims to help overcome the lack of pharma-specific impacts in currently used Life Cycle Impact Assessment (LCIA) methods, proposing the development of pharma-PCRs and the inclusion of pharma-specific impact pathways into LCIA. These last two objective are not addressed in the present document, but are rather approached and discussed in the context of Deliverable 6.2, along with the outcomes of the LCA performed under Task 6.5.

- *Chapter 3:* In Silico Ecotoxicity Models

The increasing presence of pharmaceuticals and other organic chemicals in aquatic environments has raised concerns about their potential ecotoxic effects on fish and other aquatic organisms. Experimental assessment of these impacts is often time-consuming, costly, and relies on animal testing, highlighting the need for reliable computational approaches. Within Task 6.2, interpretable in silico ecotoxicity models were developed to predict the acute toxicity of organic chemicals in fish, based on their structural and physicochemical properties. The accuracy of these models was first evaluated in silico using an external dataset, and subsequently validated through in vivo fish experiments, with the latter described in detail in Chapter 4. These computational tools support regulatory compliance, environmental risk assessment, and the design of safer chemicals in line with the European Union's Chemicals Strategy for Sustainability (CSS) and related frameworks such as REACH and CLP. Building on earlier work in Task 3.6, which addressed bioconcentration and metabolite prediction, Task 6.2 integrates additional models for acute fish toxicity with structural analysis and computational workflows. These workflows serve as filtering tools to support early-stage green drug discovery. The resulting automated pipeline identifies structural moieties, functional groups, or fragments with ecotoxic potential, providing a foundation for safer chemical design while reducing reliance on experimental testing.

- *Chapter 4:* In Vivo & In Vitro Ecotoxicity Studies

This section presents a comprehensive experimental framework aimed at evaluating the ecotoxicological impacts of pharmaceutical contaminants by integrating in vivo, in vitro/ex vivo, and in silico methodologies. Given the increasing detection of pharmaceutical residues in aquatic environments and their potential to induce chronic and sublethal effects in non-target organisms, the activities described herein focus on generating robust, biologically relevant evidence to support environmental risk assessment and regulatory decision-making. Ecologically representative marine invertebrate models are employed to investigate immunotoxic, genotoxic, and physiological responses at the cellular and organismal levels, enabling the identification of sensitive early-warning biomarkers of adverse effects. These experimental data are further complemented by fish embryo toxicity assays and computational predictions, providing cross-species validation and strengthening the reliability of model-based assessments. Overall, the integrated approach adopted in this section contributes to a deeper understanding of pharmaceutical-induced stress responses across different levels of biological organization and supports the development and validation of predictive tools for assessing the environmental risks associated with emerging contaminants.

- *Chapter 5:* Conclusions

In this chapter, the main findings of the study are synthesized and the overall conclusions are presented.

# 2 Review of Pharmaceuticals-Related Environmental Issues

In the context of Task 6.1, a screening of scientific literature was performed, on the topic of different routes for pharmaceuticals to be released to the environment along their lifecycle. A complex pathway, involving multiple routes of environmental release, was developed, as is schematically represented in Figure 1. Among the identified routes, emissions have been primarily associated with the manufacturing phase, as well as the use phase (i.e., excreted API-related compounds in wastewater after human metabolism) (1) (2) (3) (4) (5) (6) (7) (5) (8) (9) (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20) (21) (22) (23) (24) (25).


*Figure 1: Routes for the emissions of pharmaceuticals (and their metabolites and transformation products) to the environment, along their lifecycle (modified, from (26)).*

To that end, these two major emission routes were considered within Task 6.1. For **emissions during the manufacture of pharmaceuticals**, considering the lack of publicly available datasets to quantify the impacts of released components, this exposure route was primarily discussed within the context of **Deliverable 6.2**. There, the opportunities for using the Registry of pharmaceutical micropollutants were identified and discussed, as the Registry can help contribute to more transparent and traceable quantification of the Ecotoxicity and Sustainability impacts during manufacture of pharmaceuticals, as well as enable more streamlined reporting and adherence to industry standards and environmental regulations.

The emissions during the use phase of pharmaceuticals were analysed through an extensive literature review. This involved compiling data from **monitoring studies** reporting the presence of APIs, metabolites and transformation products in Greek wastewater influent and effluent, surface waters and drinking water sources. These monitoring data were subsequently correlated with national **pharmaceutical consumption** statistics and **ecotoxicological data** to evaluate the potential risks posed by pharmaceuticals detected in the environment.

## 2.1 Data sources

The review integrated three complementary information data categories:

- **Consumption patterns**: National **pharmaceutical usage data** for Greece (2010–2021), expressed as daily dosage per 1000 inhabitants for several pharmaceutical classes, were retrieved from the Organization for Economic Co-operation and Development (OECD) database (27). These data provide an estimate of expected environmental loading but do not include over-the-counter medicines, hospital-dispensed pharmaceuticals, or non-reimbursed drugs, as retrieved from National Organisation for the Provision of Health Services (EOPY).
- **Environmental occurrence**: **Environmental monitoring data** were collected from 55 peer-reviewed scientific studies, through a review performed on Google scholar (duration: 1st – 25th October, 2022, keywords combination: "pharmaceuticals" or "drugs" and "effluent", "in wastewater", "in sewage", "in water", "in environment", "in drinking water") (28) (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41) (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54) (55) (56) (57) (58) (59) (60) (61) (62) (63) (64) (65) (66) (67) (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79). This research collectively reported **359 distinct substances**, including **APIs, metabolites and transformation products** across municipal and hospital wastewater, WWTP influent and effluent, sludge, seawater, surface waters (rivers, lakes), sediments and drinking water. Compounds were organised in pharmaceutical classes, to allow comparison with consumption data from OECD. For each class (and for each individual compound), the following data was collected: the number of studies it has been investigated in, the number of times detected in water bodies, and the highest measured concentration in different sample (effluent and environmental samples).
- **Ecotoxicological relevance:** The **environmental risk** of each identified compound, considering final effluents to the sea, rivers, lakes, as well as groundwater reservoirs, was evaluated based on **Risk Quotients (RQ)**, defined as the ratio between the maximum **Measured Environmental Concentration (MEC)** and the **lowest Predicted No-Effect Concentration (PNEC)**. MEC represents the maximum concentration observed in WWTP effluents, hospital discharges, or natural water bodies. PNEC values were retrieved from the **NORMAN Database System** (80), a harmonised European platform for contaminants of emerging concern that integrates **validated experimental data and expert-reviewed in-silico predictions**. NORMAN was selected because it provides consistent, comparable and quality-controlled ecotoxicity thresholds based on the CRED (Criteria for Reporting and Evaluating Ecotoxicity Data) framework, making it particularly suitable for conservative comparative assessments (81). Of the 359 compounds identified in Greece, 340 were found in NORMAN; the remaining 19 lacked sufficient ecotoxicity data and were excluded from the risk assessment. RQs could not be calculated for 77 substances that were analysed in effluent samples but not detected above analytical limits of detection. Additionally, 22 compounds were examined only in influent samples and were excluded from further evaluation, as the risk assessment focused on concentrations in wastewater effluents. Overall, RQs were successfully derived for 241 compounds, whereas 118 substances could not be assigned an RQ due to missing PNEC values or the absence of detectable MECs. **Three levels of ecotoxicity risk** were defined, i.e., $RQ \geq 1$ as **high**, $1 > RQ \geq 0.1$ as **medium** and $RQ \leq 0.1$ as **low**.

The **integration** of these three datasets, i.e., **consumption, occurrence and ecotoxicity**, allowed Task 6.1 to systematically identify where inconsistencies exist across the

pharmaceutical life cycle and to highlight substances that warrant regulatory attention for prioritisation.

## 2.2 Assessment of pharmaceutical classes in relation to consumption, monitoring and concentration

Figure 2 provides a comparative overview of pharmaceutical classes by integrating four indicators: dosage, count of searches, count of detection, and maximum concentrations measured in effluents. For each indicator, values are expressed as percentages representing the contribution of each therapeutic class relative to the total across all classes. This normalization allows direct comparison between indicators with different absolute scales.

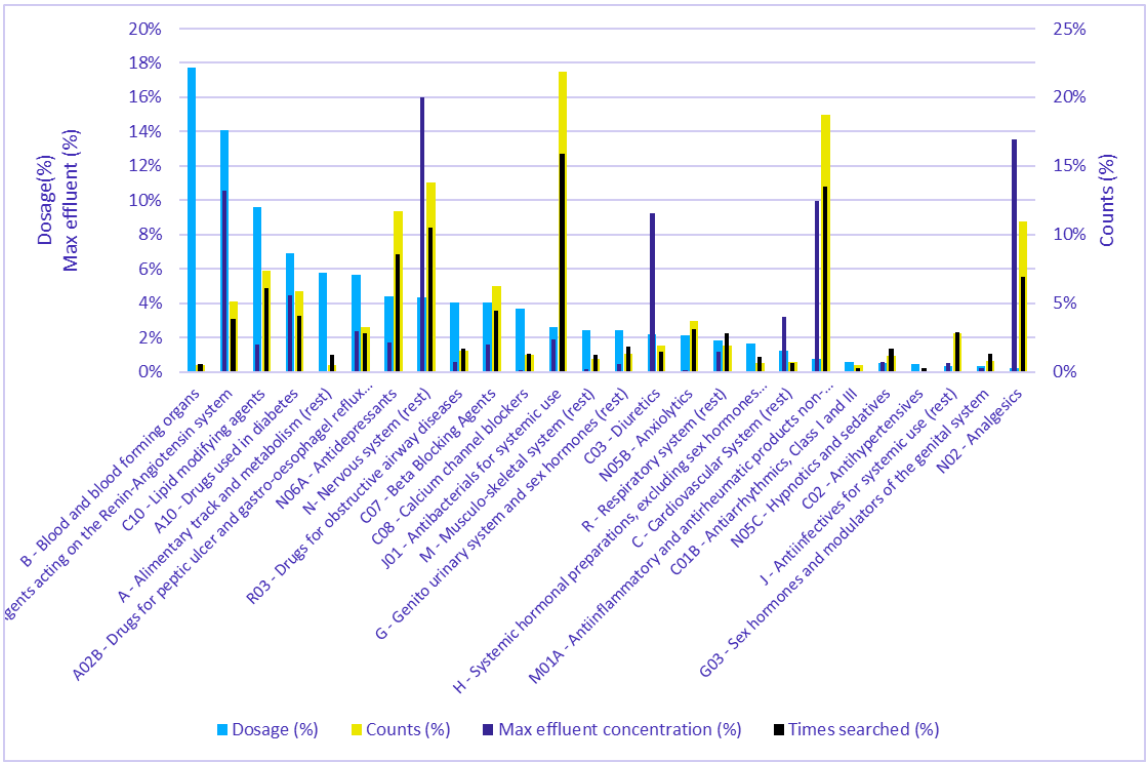



*Figure 2: Relative contribution (%) of each pharmaceutical class to (i) total pharmaceutical consumption (dosage, left axis), (ii) total maximum reported wastewater effluent concentrations (left axis), (iii) total number of times detected (counts, right axis), and (iv) total monitoring effort (times searched, right axis). Compounds are sorted in descending order, based on the pharmaceutical consumption data.*

National pharmaceutical **consumption** in Greece increased steadily between 2013 and 2020, rising from 1,096 to 1,408 defined daily doses per 1,000 inhabitants (82). Throughout this period, **Classes B** (Blood and blood forming organs), **C09** (Agents acting on the Renin-Angiotensin system), **C10** (Lipid modifying agents), **A10** (Drugs used in diabetes) and **A** (Alimentary track and metabolism) consistently representing the **largest share of use**, as presented in Figure 2. However, **environmental monitoring** does not reflect this pattern. Instead, most **analytical focus** in Greece (% times searched in Figure 2) has been placed on antibacterials (**class J01**), anti-inflammatories (**class M01A**), drugs used for the nervous system (**class N**), antidepressants (**class N06A**) and analgesics (**class N02**). This highlights an imbalance between substances most frequently used and those most frequently investigated.

Assuming that detection in natural waters is proportional to the total detection frequency reported in literature, additional trends emerge. Anti-inflammatory and antirheumatic drugs (**M01A**) and analgesics (**N02**) were detected **more frequently in natural waters** than expected, compared to the overall times detected in **any water body**, likely attributed to common and widely used over-the-counter medication in Greece (e.g., diclofenac, naproxen, ibuprofen, ketoprofen for M01A and paracetamol, salicylic acid, phenazone for N02). In contrast, antidepressants (class N06A) were detected less frequently in natural waters than anticipated.

Furthermore, assuming that detection in wastewater effluent samples is proportional to the maximum effluent concentration shows additional trends, as presented in Figure 2. In wastewater effluents, pharmaceuticals used for the nervous system (**Class N**), analgesics (**Class N02**), renin–angiotensin agents (**Class C09**) and diuretics (**Class C03**) appear at **higher concentrations** than would be expected **based on detection counts**. These patterns are driven **by a few individual compounds**, and particularly valproic acid (Class N), valsartan (Class C09), and furosemide (Class C03) that reach elevated effluent concentrations **despite relatively low detections** and rank among the highest concentrations reported in the literature. Conversely, **antibacterials (J01) and antidepressants (N06A)** appear at **lower concentrations despite** extensive **analytical coverage**. These deviations indicate that detection frequency does not necessarily predict environmental concentration.

Finally, the relationship between pharmaceutical occurrence in WWTP effluents (detection frequency and maximum concentrations) and dosage by therapeutic class was examined. Several classes deviated from the expected proportional relationship with dosage. Blood and blood-forming products (**Class B**) and alimentary tract and metabolism drugs (**Class A**) were consistently **under-represented in both detection and maximum concentration**. For these classes, relatively few compounds have been investigated in water bodies monitoring, compared to the high dosage and the total number of APIs included in each class, when further research in the ClinPGx database performed (83). Specifically, for Class B, only 4 of the 71 APIs listed in the ClinPGx database were reported in the literature, while for Class A only 6 of the 176 APIs were included.

In addition, **Class C09** was **under-represented** in terms of **detection frequency** despite exhibiting high maximum concentrations and high dosage, while **Class C10** was **under-represented** in terms of **maximum concentrations** relative to dosage. Similarly, analysis based on the ClinPGx database showed that only 10 of 23 APIs for Class C09 and 9 of 25 APIs for Class C10 were investigated, indicating moderate but incomplete representation in monitoring campaigns. The pattern of Class C10, with lower concentrations despite high dosage, may reflect either metabolic breakdown or transformation of APIs, or that the compounds currently monitored do not adequately represent overall consumption of this class in Greece.

In contrast, nervous system drugs (**Class N**), anti-inflammatory and antirheumatic (**M01A**), and analgesics (**N02**) were over-represented relative to their reported dosages. Both M01A and N02 classes include several widely used **over-the-counter** medications, suggesting that consumption may be underestimated in dosage datasets compared to prescription-only classes. Antidepressants (**N06A**) and antibacterials (**J01**) were over-represented in terms of detection counts, likely due to a combination of elevated consumption and increased concern regarding their ecotoxicological risks (84). Furthermore, the inclusion of multiple studies reporting measurements from psychiatric hospitals (40) (49) (52) (67) (72) (68) in the present analysis likely contributed to the higher detection counts observed for antidepressants. Diuretics (**Class C03**), although represented by relatively low dosage and detection, showed **elevated effluent**

**concentrations**, largely driven by furosemide, and therefore warrant further investigation in ecotoxicity assessment.

Based on the combined analysis of pharmaceutical consumption patterns and measured environmental concentrations, several therapeutic classes emerge as potential environmental hotspots. These hotspot classes are summarised in Table 2, together with the key factors supporting their prioritisation.

*Table 2. Therapeutic classes identified as potential hotspots based on consumption, concentration and environmental occurrence*

| Therapeutic class | Justification for hotspot consideration |
|---|---|
| **N – Nervous system** | High effluent concentrations driven by valproic acid combined with moderate consumption |
| **N02 – Analgesics** | High effluent concentrations, widespread use including over-the-counter medicines, and under-representation of class compounds in monitoring studies |
| **C03 – Diuretics** | High effluent concentrations, driven by a furosemide, despite limited detection frequency |
| **C09 – Renin–angiotensin system agents** | High daily dosage combined with high effluent concentrations |
| **A – Alimentary tract and metabolism** | High daily dosage coupled with limited environmental monitoring of class members |
| **B – Blood and blood-forming organs** | High daily dosage coupled with limited environmental monitoring of class members |
| **C10 – Lipid-modifying agents** | High daily dosage coupled with limited environmental monitoring of class members |
| **M01A – Anti-inflammatory and antirheumatic products non-steroids** | High detection counts and concentration coupled with low consumption, over-the-counter medications in this class |
| **J01 - Antibacterials for systemic use** | High detection counts with low consumption |
| **N06A - Antidepressants** | High detection counts, inclusion of multiple studies from psychiatric hospitals in this study |

## 2.3 Overview of ecotoxicological risk across pharmaceutical classes

The ecotoxicological assessment revealed that out of the 241 pharmaceutical compounds for which RQs could be calculated, 38 substances (16%) were classified as high risk, 60 substances (25%) as medium risk, and 143 substances (59%) as low risk. Antibacterials (**class J01**), anti-inflammatory and antirheumatic products (**class M01A**), antidepressants (**class N06A**) and a group of compounds **outside the defined OECD classes** ("NOT IN LIST") contribute the highest to the **medium- and high-risk compounds**. Table 3 presents only those therapeutic classes containing at least one high-risk compound, together with their associated medium-risk substances where applicable. Within these groups, **NSAIDs, antibiotics and sex hormones** contain compounds among the most high-risk substances assessed, as presented in Figure 3. Among NSAIDs, seven of fifteen compounds were classified as high risk, with ibuprofen and diclofenac ranking highest due to the combination of elevated effluent concentrations (up to 10 µg/L) and very low PNECs (0.01–0.04 µg/L). Antibacterials showed a similar pattern, with high RQs driven by low PNECs (0.005–0.13 µg/L) rather than high concentrations. Sex hormones (Class G03), although typically detected at low or sub-ng/L levels, exhibited some of the highest RQs due to their extremely low PNECs and potent risk effects.

*Table 3. High- and medium-risk compounds grouped by OECD therapeutic class*

| Therapeutic class | High-risk compounds | Medium-risk compounds |
|---|---|---|
| **J01 - antibacterials** | Dicloxacillin, ciprofloxacin, moxifloxacin, azithromycin, clarithromycin, amoxicillin, metronidazole, sulfamethoxazole, ampicillin, levofloxacin, clindamycin, trimethoprim | Ofloxacin, tetracycline, norfloxacin, erythromycin, lomefloxacin, chlortetracycline, thiamphenicol, sulfadiazine, cefaclor, oxytetracycline, doxycycline |
| **M01A - anti-inflammatory and antirheumatic drugs** | Ibuprofen, diclofenac, nimesulide, naproxen, niflumic acid, mefenamic acid, tolfenamic acid | Indomethacin, ketoprofen, meloxicam |
| **"NOT IN LIST"** | Diatrizoic acid, Triclosan, Iopamidol, Iomeprol, Tamoxifen | Methotrexate, Iohexol, Cyclophosphamide |
| **N06A - Antidepressants** | Serltraline, O-desmethyl venlafaxine | Doxepin, Venlafaxine, Amitriptiline, Fluoxetine, Clomipramine, Norsertraline, Nortriptyline |
| **N- Nervous system** | Valproic Acid, Cinnarizine, Carbamazepine | Clozapine, Norclozapine, Risperidone, Phenytoin |
| **G03 - Sex hormones and modulators of the genital system** | Etonogestrel, 17-a-ethinyl estradiol | Progesterone |
| **G - Genito urinary system and sex hormones** | Sildenafil | - |
| **C07 - Beta Blocking Agents** | Propranolol | Pindolol, Metoprolol |
| **C08 - Calcium channel blockers** | Nifedipine | Diltiazem |
| **C09 - Agents acting on the Renin-Angiotensin system** | Eprosartan | Losartan carboxylic acid |
| **C10 - Lipid modifying agents** | Gemfibrozil | Fenofibrate, Simvastatin, Rosuvastatin, Clofibrate, Bezafibrate |
| **N05B - Anxiolytics** | Lorazepam | Oxazepam, Bromazepam, Alprazolam |
| **R - Respiratory system** | Cetirizine (dihydrochloride) | - |

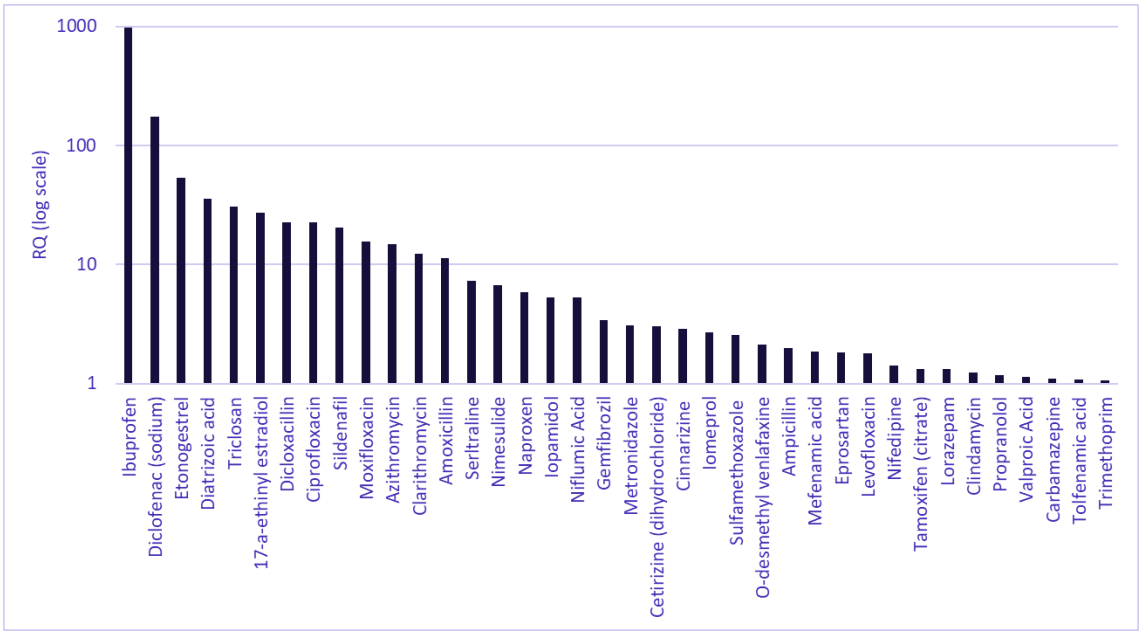



*Figure 3: High-risk pharmaceuticals in WWTP effluents based on RQ values (adapted from (26)).*

Figure 4 illustrates the distribution of pharmaceutical detections across therapeutic classes by ecotoxicological risk category. Detection frequencies are expressed as percentages of total detections and aggregated into high-, medium- and low-risk compounds, allowing comparison across classes. Table 4 complements this analysis by presenting the 50 most frequently detected compounds and their classification into medium- and high-risk categories.

Comparison of ecotoxicological risk with detection frequency **shows overlap** between **frequently detected** compounds in effluent samples and those classified as **high or medium** risk (Table 4), indicating both widespread occurrence and ecological relevance. Among the 50 most frequently detected pharmaceuticals, 20 compounds were classified as high risk and include several **NSAIDs, antibiotics, biocides** (triclosan), **psychiatric medication** (sertraline) and **lipid regulators** (gemfibrozil), while an additional 18 compounds were categorised as medium risk. In contrast, 14 of the 50 most frequently detected compounds fall into the low-risk category, including pharmaceuticals belonging to classes A10, C07, C09, N, N05, N06 and R03, demonstrating that frequent detection or high concentration does not necessarily correspond to elevated ecological risk.

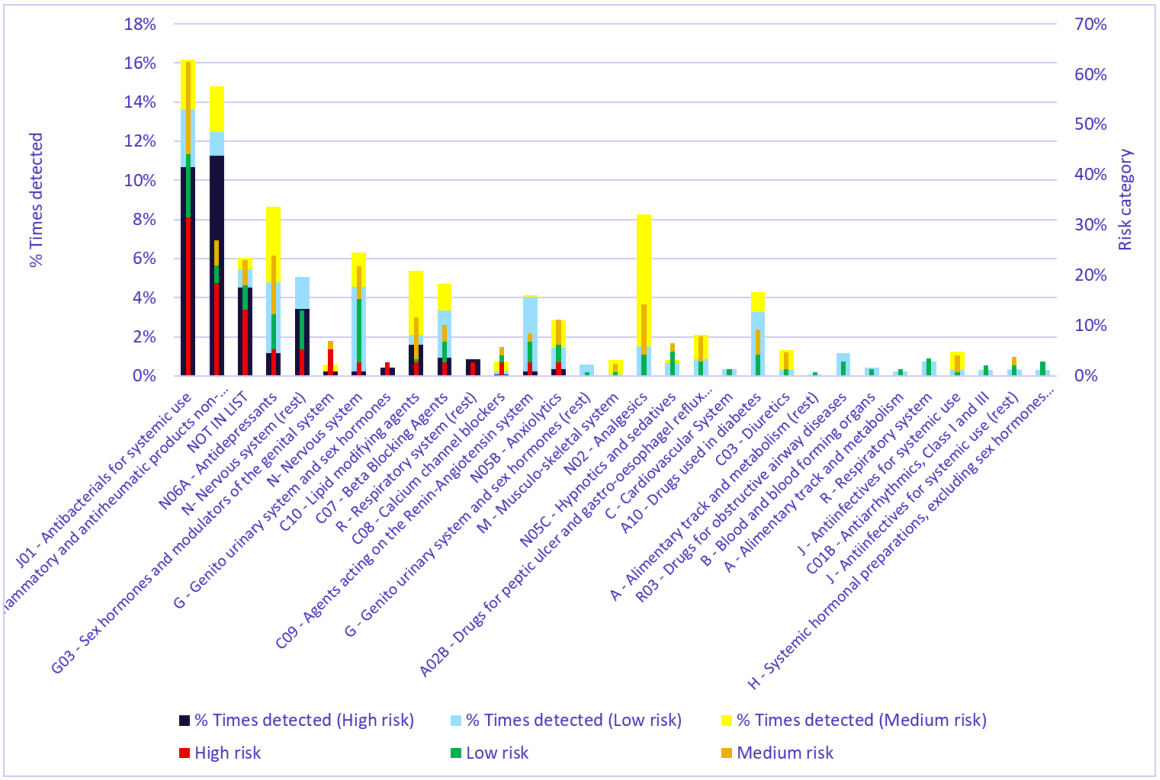



*Figure 4: Distribution of pharmaceutical detections across pharmaceutical classes by ecotoxicity risk category. Bars represent the percentage contribution of each class to the total number of detections, calculated separately for high-, medium-, and low-risk pharmaceuticals (left y-axis). Shaded bars indicate the relative proportion (%) of high-, medium-, and low-risk compounds within each ATC class (right y-axis), based on calculated RQ categories.*

*Table 4. High- and medium-risk pharmaceuticals among the 50 most frequently detected compounds, grouped by therapeutic class*

| Therapeutic class | High-risk compounds | Medium-risk compounds |
|---|---|---|
| **J01 - Antibacterials for systemic use** | Sulfamethoxazole, trimethoprim, clarithromycin, ciprofloxacin, azithromycin, moxifloxacin, metronidazole, levofloxacin | Erythromycin, sulfadiazine |
| **M01A – Anti-inflammatory and antirheumatic products non-steroids** | Diclofenac, naproxen, ibuprofen, mefenamic acid, nimesulide, tolfenamic acid | Ketoprofen |
| **A02B - Drugs for peptic ulcer and gastro-oesophagel reflux diseases** | - | Ranitidine |
| **C03 - Diuretics** | - | Furosemide |
| **C07 - Beta Blocking Agents** | Propranolol | Metoprolol |
| **C10 - Lipid modifying agents** | Gemfibrozil | Bezafibrate, clofibrate, fenofibrate |
| **J - Antiinfectives for systemic use** | - | Fluconazole |
| **M - Musculo-skeletal system** | - | Orphenadrine |

| Therapeutic class | High-risk compounds | Medium-risk compounds |
|---|---|---|
| **N- Nervous system** | Carbamazepine | Risperidone |
| **N02 - Analgesics** | - | Phenazone, tramadol, salicylic acid, acetaminophen |
| **N06A - Antidepressants** | Serltraline | Fluoxetine, venlafaxine |
| **Biocide ("NOT IN LIST")** | Triclosan | - |
| **R - Respiratory system** | Cetirizine | - |

To further assess the alignment between monitoring effort and environmental occurrence, a detection ratio was calculated for each compound, as the number of detections divided by the number of times the compound was analysed in wastewater effluents. This metric was subsequently compared with ecotoxicological risk to evaluate whether substances also represent priority environmental concerns, and the findings are presented in Table 5, focusing on high- and medium-risk categories. Furthermore, Figure 5 shows the relative number of compounds classified as balanced, inconclusive, over-analysed, or under-analysed across high-, medium-, and low-risk categories.

This assessment indicates that several ecotoxicologically relevant pharmaceuticals remain under-analysed, while others are repeatedly monitored despite low environmental risk. In particular, several APIs in the classes of **J01, M01A, N05C and compounds outside the OECD list** are frequently targeted despite low or inconsistent detections (**over-analysed**), likely due to historical monitoring priorities or analytical accessibility. Although such focus reflects established practices, it adds limited value to environmental risk assessment, as most of these **over-analysed** substances exhibit low environmental concentrations and **low risk quotients** (37 out of 62 compounds classified as low risk, as presented in Figure 5).

Conversely, among the **under-analysed** group, 7 out of 15 compounds exhibit **high or medium** ecotoxicological risk, indicating that priority should be given to substances combining elevated risk with low monitoring coverage, as highlighted in Table 5. In addition, **several medium-risk and a limited number of high-risk** substances remain insufficiently characterised due to sparse monitoring data. Targeted monitoring of these compounds, as identified in Table 5, is therefore recommended to improve understanding of their environmental occurrence and to refine their ecotoxicological risk characterisation.

At the same time, the majority of **under-analysed or inconclusive** compounds fall within the **low-risk category** (12 out of 19 under-analysed compounds and 57 out of 72 inconclusive compounds, as presented in : Distribution of pharmaceuticals across ecotoxicological risk categories according to monitoring coverage classification), indicating that their ecotoxicological relevance is likely limited despite sparse monitoring data. Consequently, these substances do not warrant immediate prioritisation, although targeted analyses may be considered to confirm this assessment. These compounds are predominantly associated with the nervous system (Class N) and antidepressant (Class N06A), which exhibit either high detection ratios but limited analytical coverage or inconclusive detection results, within the low-risk group.

Overall, the combined evaluation of monitoring coverage and ecotoxicological risk highlights the need for improved alignment between **monitoring effort and environmental relevance**, supporting a shift towards more targeted, risk-based monitoring strategies.

*Table 5. High- and medium-risk pharmaceuticals among the under-analysed and inconclusive category compounds, grouped by therapeutic class*

| Therapeutic class | High-risk compounds | Medium-risk compounds |
|---|---|---|
| **Under-analysed category** | | |
| **NOT IN LIST** | Iopamidol, iomeprol | - |
| **N06A - Antidepressants** | O-desmethyl venlafaxine | Nortriptyline norsertraline |
| **C09 - Agents acting on the Renin-Angiotensin system** | Eprosartan | - |
| **N- Nervous system** | Valproic acid | - |
| **Inconclusive category** | | |
| **G03 - Sex hormones and modulators of the genital system** | Etonogestrel | - |
| **NOT IN LIST** | Diatrizoic acid | - |
| **J01 - Antibacterials for systemic use** | Dicloxacillin | Doxycycline, cefaclor, lomefloxacin, thiamphenicol, chlortetracycline, oxytetracycline |
| **N05C - Hypnotics and sedatives** | - | Temazepam |
| **N02 - Analgesics** | - | Phenacetin |
| **N- Nervous system** | - | Phenytoin |
| **C03 - Diuretics** | - | Triamterene |
| **J - Antiinfectives for systemic use** | - | Atazanavir |
| **C09 - Agents acting on the Renin-Angiotensin system** | - | Losartan carboxylic acid |

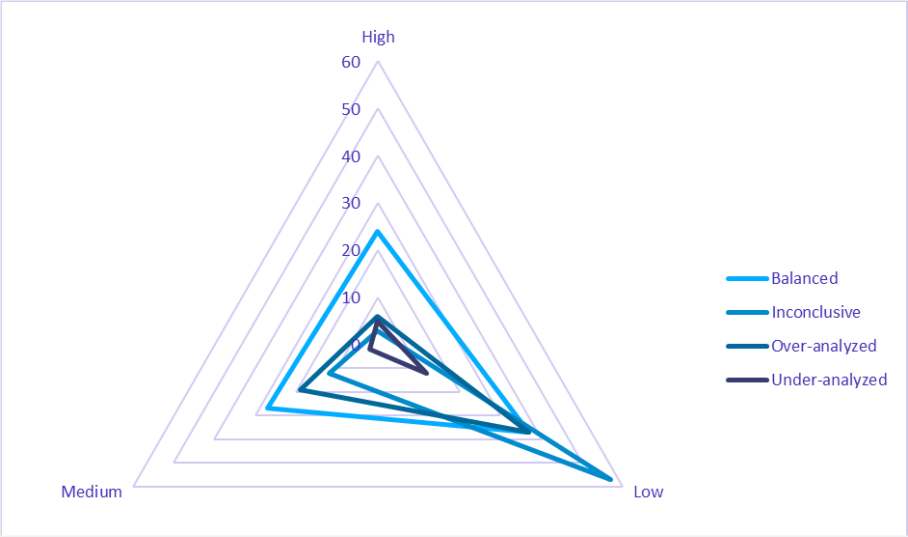



*Figure 5: Distribution of pharmaceuticals across ecotoxicological risk categories according to monitoring coverage classification*

Due to the fact that many pharmaceuticals undergo metabolic transformation in humans and wastewater systems, evaluation of environmental risk should consider both parent compounds and their metabolites. Although 49 metabolites and transformation products were included in the dataset from environmental monitoring literature, 44 of them were analysed only once or twice, resulting in largely inconclusive evidence regarding their environmental occurrence and risk. For 19 of those, no PNEC value was available, and they were excluded from the quantitative risk assessment. A comparative analysis for selected APIs and their metabolites was performed, focusing on compounds with more frequent monitoring in the dataset, and the results are presented in Figure 6.

Overall, the figure illustrates that parent APIs are generally monitored and detected more frequently than their corresponding metabolites. Among the compounds evaluated, **O-desmethylvenlafaxine**, metabolite of venlafaxine, combines limited monitoring coverage with a high RQ. Although the parent compound is detected relatively often, its metabolite remains under-analysed and warrants further investigation due to its elevated potential risk. In contrast, metabolites such as **norfluoxetine** (from fluoxetine) and **desloratadine** (from loratadine) exhibit both **low detection frequencies and low RQs**, suggesting limited environmental relevance under current conditions. For metformin, the parent compound is detected more frequently and shows a higher RQ than its transformation product **guanylurea**, consistent with previous findings indicating lower toxicity and environmental persistence of the metabolite, with both substances falling within the **low-risk category**. Similarly, within the flunitrazepam group, both the parent compound and its metabolite 7-aminoflunitrazepam display low detection frequencies and are classified as low risk. Overall, Figure 6 demonstrates that pharmaceutical metabolites can differ from their parent compounds in both environmental occurrence and toxicity. In several cases (e.g. O-desmethylvenlafaxine and norfluoxetine), metabolites exhibit higher RQs despite lower monitoring effort, suggesting that current monitoring practices may underestimate the contribution of transformation products to overall ecological risk.

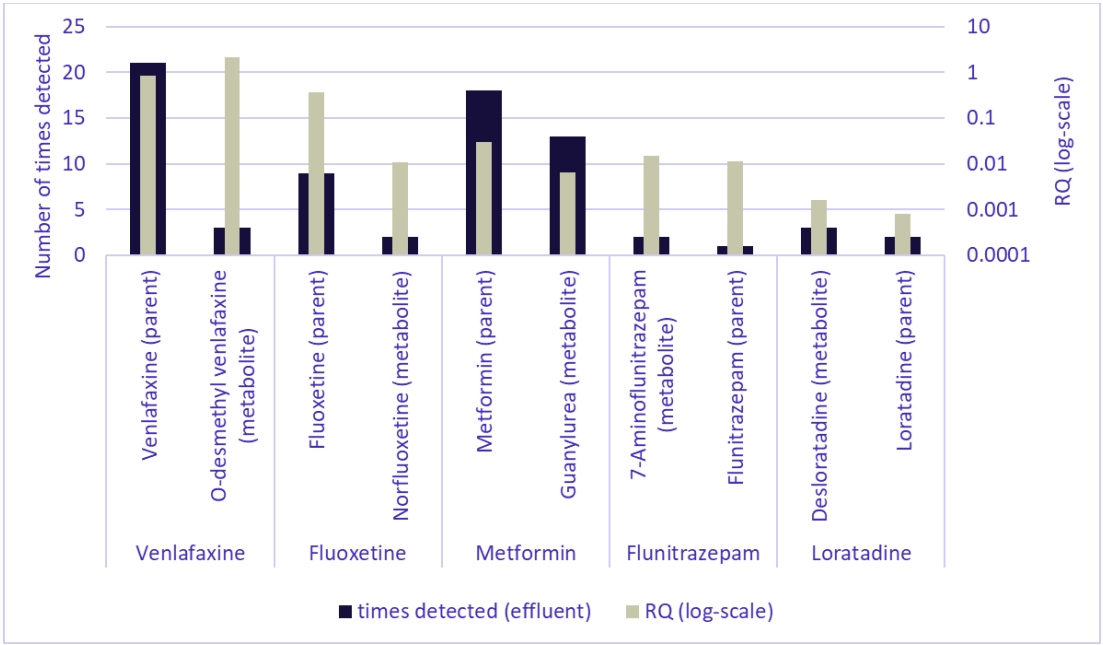



*Figure 6: Detection frequency and risk quotient (RQ) of selected API–metabolite pairs detected in WWTP effluents. Bars represent the number of detections per compound, while the line shows RQ values (log scale) – adapted from (26).*

## 2.4 Prioritisation of pharmaceutical hotspots and implications for monitoring

Overall, the scope of this task is to provide insights into the improvement of risk management options and to prioritize pharmaceuticals. By integrating pharmaceutical consumption, environmental monitoring frequency, detection patterns and ecotoxicological risk, several pharmaceutical classes and compounds were identified as priority hotspots for environmental monitoring (Table 6).

As it has been presented (Table 2), initial screening based on wastewater effluent concentrations highlighted Classes N02 (analgesics), C03 (diuretics) and N (nervous system medicines) as potential hotspots. However, when ecotoxicity was considered, only **Class N** clearly emerged as a **hotspot group**, as it contains both **high- and medium-risk compounds**, with **valproic acid** identified as a key driver of risk. In contrast, **Classes N02 and C03**, although associated with elevated concentrations, currently contain **no high-risk compounds**, and their hotspot status remains uncertain due to limited monitoring coverage and incomplete toxicological screening of all class members. It is important to note that classes with high percentage of inconclusively monitored compounds, i.e., **N02**, **further PNEC-based evaluation** of additional compounds within these classes is recommended, as these could be present but not yet identified in natural water and wastewaters.

Class C09 was also identified as a potential hotspot due to high consumption and elevated effluent concentrations. Nevertheless, most compounds in this class exhibit low ecotoxicological risk and are already well represented in monitoring studies; consequently, no additional monitoring effort is currently prioritized beyond selected medium- or high-risk APIs. In contrast, **Classes J01** (antibacterials), **N06A** (antidepressants) and **M01A** (anti-inflammatory drugs) appear over-represented in monitoring relative to consumption and concentration, yet contain a substantial proportion of **high-risk** as well as insufficiently characterized compounds. These classes should therefore **remain priority targets**, with monitoring efforts focused on ecotoxicologically relevant substances.

**Classes A** (alimentary tract and metabolism), **B** (blood and blood-forming organs) and **C10** (lipid-modifying agents) were also identified as priority groups due to high consumption combined with limited environmental monitoring. Although the compounds **currently** monitored from these classes show **low ecotoxicological risk**, screening of **additional APIs** using **PNEC values** indicates that several unmonitored substances **may pose greater ecological concern**. These findings indicate that **potential hotspots** may remain **undetected** due to incomplete class coverage, and targeted screening of the most toxicologically relevant compounds is recommended.

*Table 6. Identified priority pharmaceutical classes*

| Pharmaceutical class | Hotspot classification |
|---|---|
| **N- Nervous system** | Hotspot class |
| **N02 - Analgesics** | Emerging hotspot, further analysis needed |
| **J01 - Antibacterials for systemic use** | Hotspot class |
| **N06A - Antidepressants** | Hotspot class |
| **M01A – Anti-inflammatory and antirheumatic products** | Hotspot class |

| Pharmaceutical class | Hotspot classification |
|---|---|
| **A - Alimentary track and metabolism** | Emerging hotspot, further analysis needed |
| **B - Blood and blood forming organs** | Emerging hotspot, further analysis needed |
| **C10 - Lipid modifying agents** | Emerging hotspot, further analysis needed |

Beyond class-level prioritization, compounds classified as high risk should be considered priority. However, risk prioritization should not rely on concentration alone, as isolated high values may reflect site-specific or temporal factors. Instead, combining ecotoxicological risk with detection frequency and monitoring coverage allows identification of compounds that are both environmentally relevant and insufficiently monitored, providing a robust basis for prioritization. The overlapping compounds between the frequently detected and high- or medium-risk pharmaceuticals, as well as the overlapping compounds between the high- or medium-risk pharmaceuticals with the under-analysed and inconclusive groups are identified as hotspots (Table 4 and Table 5, respectively). Metabolites and transformation products represent an additional monitoring gap. Although generally under-represented in environmental studies, several metabolites exhibit higher ecotoxicological risk than their parent compounds. This indicates that current monitoring practices likely underestimate their contribution to environmental risk. Systematic identification and prioritization of relevant metabolites, supported by in silico prediction tools and PNEC screening, is therefore recommended.

Finally, several limitations should be acknowledged, including variability across monitoring studies, limited availability of ecotoxicity data for some compounds and metabolites, and aggregation of consumption data at class level, which underestimates non-prescription and over-the-counter use. These factors should be considered when interpreting results and designing future monitoring strategies.

Overall, in this study **17 priority substances** that require routine monitoring are proposed. The overview is presented Table 7. Routine quarterly monitoring of these compounds in WWTP influent, effluent and downstream surface waters is recommended, with monthly monitoring for NSAIDs, macrolide and fluoroquinolone antibiotics and synthetic estrogens due to consistently high risk. Broader analytical coverage is also needed for highly consumed pharmaceutical classes A and B and for metabolites that may exhibit greater ecotoxicological relevance than parent compounds.

*Table 7. Identified priority substances categorized by pharmaceutical class*

| Pharmaceutical class | Priority compounds |
|---|---|
| **M01A –anti-inflammatory and antirheumatic drugs** | Diclofenac, ibuprofen,naproxen |
| **J01 - Antibacterials for systemic use** | Ciprofloxacin, moxifloxacin, azithromycin, clarithromycin, metronidazole |
| **Biocide (“NOT IN LIST”)** | Triclosan |
| **Contrast agents (“NOT IN LIST”)** | Diatrizoic acid, iopamidol |
| **G03 - Sex hormones and modulators of the genital system** | 17α-ethinyl estradiol, etonogestrel |

| Pharmaceutical class | Priority compounds |
|---|---|
| **N – nervous system** | Valproic acid |
| **N06A - Antidepressants** | Sertraline, O-desmethylvenlafaxine |
| **C10 - Lipid modifying agents** | Gemfibrozil |

Overall, the activities carried out under Task 6.1 resulted in the following key outputs:

- Identification of mismatches between pharmaceutical consumption patterns and environmental monitoring practices.
- Prioritisation of pharmaceutical classes and individual compounds based on combined occurrence, concentration and ecotoxicological risk.
- Identification of under-monitored therapeutic classes with high consumption and potential environmental relevance.
- Evidence that metabolites and transformation products may contribute significantly to environmental risk and are currently under-represented in monitoring.
- Scientific input supporting a shift from frequency-based to risk-based prioritisation of pharmaceuticals in environmental monitoring programmes.

The outcomes of Task 6.1 directly support the development of a **review–roadmap** aimed at reducing pharmaceutical emissions into the environment. By identifying priority pharmaceutical classes and compounds based on integrated consumption, monitoring and ecotoxicity data, this task provides actionable input for **updating existing legislation** on pharmaceutical environmental risk assessment and monitoring requirements. The findings support more targeted allocation of monitoring resources, inclusion of under-monitored but ecotoxicologically relevant substances, and improved consideration of metabolites in regulatory frameworks. These insights are aligned with ongoing updates of EU water-related legislation and contribute to strengthening risk management options across the pharmaceutical life cycle.

To address the environmental risks posed by pharmaceuticals across their lifecycle, the overall work reported in D6.2 (Task 6.4 and Task 6.5) and the present document combines two complementary perspectives: production-related emissions and post-consumption environmental occurrence. Critical gaps in current methodologies for LCA assessment in pharmaceutical production were discussed in the context of D6.2, particularly the reliance on generic wastewater treatment datasets that fail to capture the ecotoxicological relevance of API-specific emissions, and the lack of pharma-specific impacts in current LCA impact assessment methods. To overcome this, the potential contribution to more robust LCAs for a dedicated registry is discussed within the context of Deliverable 6.2. The registry aims to track pharmaceuticals and intermediates in their corresponding wastewater streams during upstream manufacture, by collecting real wastewater analysis results and corresponding liquid volumes from different pharmaceutical manufacturers. Such a registry would enable more accurate LCAs, support compliance with evolving legislative requirements, and provide a foundation for pharma-specific impact pathways in Life Cycle Impact Assessment (LCIA). In parallel, the extensive review of APIs detected in wastewater and natural waters (presented here) highlights the need for targeted monitoring strategies, taking into consideration national consumption patterns and ecotoxicity profiles. Linking these findings with the proposed registry creates a robust framework for risk management: real-world manufacturing emission data can inform

LCA and regulatory decisions, while monitoring priorities ensure that environmental assessments capture the most relevant compounds. Together, these actions contribute to the EU's strategic objectives of minimizing pharmaceutical emissions, improving legislation, and advancing a roadmap for sustainable pharmaceutical lifecycle management.

# 3 In Silico Ecotoxicity Models

## 3.1 Introduction

Task 6.2 focuses on developing *in silico* tools to predict the ecotoxicity of chemicals in fish species. By applying computational methodologies, the aim is to enhance our understanding of the ecological impacts of various compounds and their potential adverse effects on aquatic ecosystems. This work is aligned with the European Union's Chemicals Strategy for Sustainability (CSS), a central initiative of the EU Green Deal. The CSS promotes safer chemical use and seeks to reduce environmental hazards in pursuit of climate neutrality by 2050 (85) (86) (87). As part of this effort, regulatory frameworks like the Classification, Labelling and Packaging (CLP) (88) regulation, European Union's European Union's Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) (89) , and the Restriction of Hazardous Substances Directive (90) are being revised to phase out harmful substances where safer alternatives exist (91). Among the regulated categories, pharmaceuticals represent a growing environmental concern due to their persistence and widespread presence in aquatic systems. These compounds enter the environment through wastewater effluents or agricultural runoff (92) (93), often remaining biologically active and affecting aquatic organisms. To mitigate these impacts (94), predicting the ecotoxicity of drug compounds is essential. Ecotoxicity is commonly assessed using the Lethal Concentration ($LC_{50}$) metric—i.e., the concentration of a substance that causes death in 50% of test organisms (such as fish) over a defined exposure period, typically 96 hours. Under current EU CLP regulations, substances with LC50 values below 1 mg/L are considered acutely toxic to aquatic life.

Numerous computational strategies have been developed to estimate LC50 values, ranging from traditional Quantitative Structure–Activity Relationship (QSAR) models to advanced machine learning (ML) and deep learning methods. These models aim to reduce reliance on animal testing by leveraging patterns in chemical structure and historical toxicity data (95) . Traditional QSAR models like ecological structure-activity relationship (ECOSAR (96)) and the Toxicity Estimation Software Tool (T.E.S.T. (97)) have long been used to estimate toxicity across species such as algae, daphnia, and fish using linear regression methods. Recent advancements in QSAR models have focused on improving the accuracy of predicting organic chemical toxicity in ecologically important fish species, utilizing both traditional approaches and ML technniques (98) (99) (100) (101). However, the focus has increasingly shifted toward ML approaches in recent developments. Li et al. utilized ML models to predict the acute toxicity of pesticides to fish, achieving a balanced accuracy of 0.83 on a small dataset (102) . Tuulaikhuu et al. used EcoTox data with taxonomy features and random forests, reaching $R^2$ = 0.85 for acute fish toxicity (103). Wu et al. combined ML and Read Across Structure Activity Relationships (RASAR) models, predicting fish toxicity with over 93% accuracy (104). Similar approaches have used both conventional ML algorithms and graph convolutional networks, achieving classification accuracies exceeding 83% in predicting fish toxicity (105). Viljanen et al. tested several ML methods on ECOTOX data (2,431 chemicals, 1,506 species) to predict $LC_{50}$ values (106). Finally, Gasser et al. evaluated ML models in the "t-F2F" challenge (107) using six molecular representations, with the best model reaching Root Mean Square Error (RMSE) equal to 0.90 (108).

Despite these advances, many existing models face limitations in usability and interpretability. Some tools restrict the number of compounds that can be tested simultaneously (109) or present results in an unclear format, making it hard to link inputs to results (96). Others lack transparency offering no access to underlying code or web services (110) (111) (112) (113)

(114), which makes it difficult to evaluate or reproduce the outcomes—an important drawback in regulatory contexts. Moreover, many models are trained on small or narrow datasets (111) (114) (115), which can hinder their generalizability to real-world pharmaceutical compounds. In this context, we present the development of interpretable machine learning models for predicting the ecotoxicity of organic chemicals in fish. Our analysis explored structural features and key physicochemical descriptors linked to compounds with high toxicity potential. The best-performing ecotoxicity model was benchmarked against other publicly available tools and external datasets, as well as validated against experimental results obtained under standardized conditions on a carefully selected set of 19 compounds. This work is closely aligned with the objectives of Task 3.6, which focused on developing *in silico* tools to predict the bioconcentration of chemicals in fish, as well as conducting a literature review and assessment of metabolite prediction software. The latter aimed to identify potential metabolites and transformation products (e.g., those generated during wastewater treatment) of parent compounds and assess their bioconcentration and ecotoxicity potential. Tasks 3.6 and 6.2 culminated in the creation of an integrated, automated pipeline capable of processing chemical information, identifying structural moieties, functional groups, or fragments with ecotoxic potential, and providing computational workflows and filtering tools to support early green drug discovery. An overview of the workflow employed in Tasks 3.6 and 6.2 is presented in Figure 7.

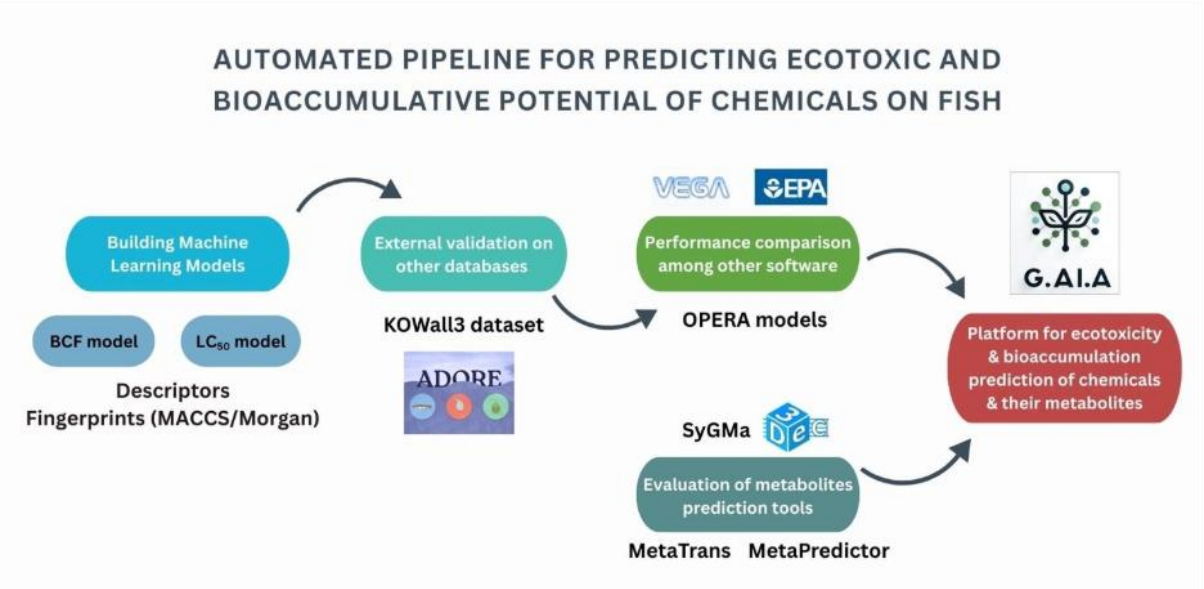



*Figure 7: Overview of the workflow employed in Tasks 3.6 and 6.2.*

Finally, it is important to highlight that this deliverable presents snapshots of the automated pipeline in its current state, acknowledging that it will be continuously enhanced with additional features and functionalities beyond the scope of the ENVIROMED project.

## 3.2 Classification of a given chemical as ecotoxic or non-ecotoxic

To develop and validate machine learning models for predicting aquatic toxicity, extensive experimental LC50 data were gathered from multiple reputable sources. The core dataset comprised 96-hour LC50 values for a wide range of organic compounds tested on

various aquatic species, emphasizing fish as the primary test organisms. At the start of the project, our initial plan was to train the machine learning models using LC50 data from organic compounds tested on aquatic species that National Research Council of Italy (CNR) could later validate through in vitro experiments. However, following an extensive literature review, we found that there was insufficient data available for these species to effectively train the models. As a result, we shifted our focus to LC50 data from fish species. Two main databases were used to compile the initial ecotoxicity dataset: the EnviroTox platform (116) and the OPERA database[28]. These sources include high-quality, peer-reviewed toxicity studies, regulatory submissions, and curated data from environmental risk assessments. From these, we extracted LC50 values across diverse endpoints and aquatic species, ultimately assembling a dataset of 3,005 unique organic compounds. Each compound was categorized according to its minimum reported 96-hour LC50 value. Classification followed the threshold defined by the CLP regulation, designating substances with LC50 ≤ 1 mg/L as "ecotoxic" and those above the threshold as "non-ecotoxic".

For external validation and model benchmarking, we used data from the ADORE (107) (Aquatic Database for Organism-level Response Evaluation) benchmark dataset, which aggregates ecotoxicity records – primarily acute mortality outcomes – from numerous regulatory and scientific sources. ADORE is rooted in a harmonized and pre-filtered version of the United States Environmental Protection Agency (US EPA)'s ECOTOX database and was enriched with curated chemical and taxonomic metadata. The final ADORE dataset includes over 33,000 entries, with approximately 26,000 involving fish species. Two key subsets of ADORE were selected for model benchmarking. The t_F2F dataset contains 1,118 compounds with experimentally measured LC50 values across multiple fish species, divided into 614 non-ecotoxic and 504 ecotoxic compounds. A second subset, the s-F2F-2 dataset, focuses exclusively on *Pimephales promelas* and includes 497 compounds (348 non-ecotoxic and 149 ecotoxic). Both datasets were specifically chosen for their consistency, coverage of structurally diverse chemicals, and relevance to acute aquatic toxicity modelling. Figure 8 presents the distribution of logLC50 values (ranging from –6.00 to 4.76) alongside three key chemical properties: molecular weight, water solubility, and log P highlighting the broad applicability domain of the dataset used for model training and validation.

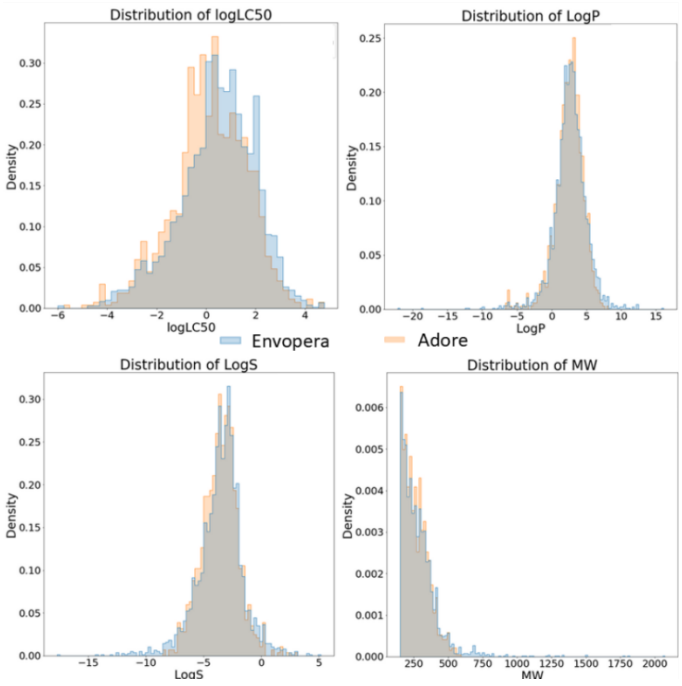



*Figure 8: Distribution of LC50 values in our dataset and the external validation set, along with three key chemical properties: molecular weight, logS and logP.*

To surpass the limitations of the ADORE dataset—namely its integration of data from heterogeneous experimental conditions, including variations in fish species, laboratory protocols, and testing parameters—we took a more controlled approach to external validation. Specifically, we assembled and purchased an independent set of 19 organic compounds and conducted experimental testing under standardized conditions. Given that CNR is limited to validating the model using data from fish aquatic species, we opted to first carry out direct testing on a single fish species to ensure accurate model evaluation. To support this effort, we commissioned an academic subcontractor to perform the necessary experiments using zebrafish larvae up to 96 hours post-fertilization. The subcontractor selected is Dr. Dimitris Beis, Associate Professor of Biological Chemistry at the Faculty of Medicine, University of Ioannina. His laboratory specializes in the use of zebrafish models to study human diseases and screen for novel bioactive compounds. Dr. Beis was selected following a thorough market assessment to ensure the best value for money.

On our initial efforts for the development of an ecotoxicity model, we utilized the PaDEL-descriptor software v 2.2.1 (117) to generate 1D and 2D physicochemical descriptors, resulting in an initial dataset comprising 1,167 descriptors. To address missing values (NaNs) in descriptors that PaDEL could not compute, we implemented a data cleaning strategy based on a tunable threshold. This threshold, ranging from 1% to 100%, was used to determine whether to retain or discard specific columns (descriptors) or rows (compounds). Specifically, if the proportion of NaN values in a column exceeded the threshold, the entire column was removed; otherwise, only the rows containing NaNs in that column were discarded. We evaluated the impact of various threshold levels on the resulting datasets and identified 1%, 10%, and 70% as optimal cutoffs, providing a favorable trade-off between the number of retained descriptors and the number of usable compounds for downstream modelling.

During model training, we employed stratified 10-fold cross-validation to ensure robust and reliable evaluation across all models (118). We explored a range of supervised learning algorithms, including AdaBoost (119), K-Nearest Neighbors (KNN) (120), Random Forest (121), Artificial Neural Networks (ANN), and Deep Neural Networks (DNN) (122). Model

performance was evaluated using several metrics—accuracy, precision, recall, F1 score, and specificity (123)—with particular emphasis placed on accuracy and F1 score to capture both overall correctness and class balance in predictions.

Validating predictive models is critical to ensuring their reliability, reproducibility, and real-world applicability. While internal validation—such as k-fold cross-validation—provides insight into a model's performance within a given dataset, external validation is essential for evaluating its scalability and generalizability to new, unseen data. K-fold cross-validation partitions the dataset into k subsets, sequentially using one subset for validation and the remaining for training. This iterative process ensures that each data point is used for validation exactly once, yielding a robust estimate of model consistency. However, external validation using independent datasets, such as the eChemPortal database (124), is vital for assessing how well a model performs beyond its training domain.

To investigate the external validity of our models, we adopted a systematic approach centered around our top-performing model: a Random Forest algorithm configured with 300 estimators. We deployed this model using the eChemPortal dataset as an external test set. However, upon evaluation, we observed a significant bias—our model predominantly classified all compounds as non-ecotoxic. To understand the source of this bias, we analyzed the class distribution within our training data. We found a strong class imbalance, with a majority of compounds labeled as non-toxic to fish. This imbalance likely contributed to the model's tendency to overpredict the non-ecotoxic class, compromising its predictive reliability. To enhance the validity of our external evaluations and mitigate extrapolative errors, we incorporated applicability domain (AD) analysis using the Python-based pyADA framework (125). This approach involved generating molecular fingerprints for all compounds in the training dataset. Compounds from the eChemPortal validation set were then compared based on Tanimoto structural similarity (126), and only those falling within the applicability domain of the training data were selected for evaluation. This strategy ensured that model predictions were made within a chemically relevant space, thereby improving accuracy and reducing the risk of extrapolation—unlike earlier validation methods that often relied on synthetic or poorly matched test compounds.

Following extensive experimentation and parameter tuning, we established specific thresholds for our applicability domain analysis: a maximum Tanimoto similarity of 0.7, a standard deviation threshold of 0.3, and an average similarity of 0.08. Using these criteria, we filtered the eChemPortal validation set to identify compounds with structural characteristics sufficiently similar to those in our training dataset. Out of the 1,393 compounds available in eChemPortal, only 115 met these stringent similarity thresholds, indicating they fell within the defined applicability domain. These 115 compounds were subsequently used for external validation of our predictive models, which had been trained on a class-balanced dataset. We processed this refined validation set by aligning it with the descriptor columns used during model training. These preprocessed validation sets were then used to evaluate the robustness and generalizability of six machine learning models—Random Forest, KNN, AdaBoost, and others—whose performance results are presented in Table 17.

Upon observing poor performance across all six models in Table 17, we revised our approach. We retrained the models using a feature selection strategy based on the feature_importances_ attribute from tree-based classifiers to retain only descriptors with higher statistical relevance. The retraining phase employed several algorithms, including Gradient Boosting (127), Extra Trees Classifier (128), Decision Tree Classifier (129), AdaBoost, and Random Forest. This step aimed to enhance model performance by focusing on the most informative features while reducing noise and overfitting. At each stage of model development, we explored various

hyperparameters tailored to the specific machine learning algorithm. For Gradient Boosting, models were trained with 100 and 200 estimators. The Extra Trees Classifier used 100 and 150 estimators, while the Decision Tree Classifier was tested with maximum depths of 3 and 6. For AdaBoost, both the SAMME.R and SAMME algorithms were employed, each with 500 estimators. Random Forest was evaluated using both Gini Index (GII) and Entropy Index (ENT) as split criteria, along with varying numbers of estimators (50, 100, 150, and 300). A central component of our approach was feature selection—identifying the most relevant molecular descriptors to train the models. We employed the feature_importances_ attribute from tree-based models to assign importance scores to each descriptor, enabling the selection of subsets containing 8–12 features with the highest significance.

Model training was conducted over 30 epochs, and we employed 50-fold cross-validation to ensure robust performance assessment. This strategy not only improved evaluation reliability but also aided in identifying optimal descriptor combinations and the most effective model architecture. Following this extensive series of training experiments, we identified the most promising model–descriptor combination. The optimal descriptor set consisted of 11 key features: ATS0p, ATS3m, ATS2m, VP-3, ATS0m, MLFER_L, ZMIC0, ZMIC1, MW, CrippenLogP, and XLogP. To further refine our feature selection process, we analyzed the frequency of descriptor occurrence in 20 most informative features across models. Thirteen descriptors stood out based on their appearance frequency (Table 18).

Using this refined 13-descriptor set, we retrained all machine learning models to evaluate how frequently occurring features affect predictive performance. This additional step enabled a comparative analysis between high-importance descriptors and those frequently selected across models. Based on these two training strategies, we identified two top-performing model–descriptor combinations. Their performance metrics are summarized in Table 19, in which a special column, "Confusion Matrix" (130), allows us to evaluate model performance comprehensively by showing true negative (TN), false positive (FP), false negative (FN) and true positive (TP) values. We then validated these models using the eChemPortal dataset. However, the external validation results were less favourable, as shown in Table 20, suggesting limited generalizability despite strong performance on the internal dataset.

To further explore potential performance improvements, we partitioned our original dataset into 80% training and 20% external validation subsets. This internal validation approach leveraged known data while simulating a real-world testing scenario. Table 21 presents the training results, while Table 22 shows validation performance on this held-out 20%, offering insight into model robustness and generalization when faced with more representative external data. In earlier stages, we opted for oversampling rather than the more commonly used under-sampling approach to address class imbalance. Using Python's RandomOverSampler, we expanded our training set from 2,200 to 6,434 compounds. This strategy allowed us to preserve minority-class data while increasing model exposure to diverse examples. As before, we prioritized the identification of meaningful descriptors using two complementary feature selection approaches, single-model feature importance and cross-model feature occurrence frequency. Table 23 23 presents the training results using the oversampled dataset, highlighting both approaches. For completeness, Table 24 reports model performance on the corresponding 20% validation set, offering a side-by-side comparison of feature importance strategies and their effect on generalizability and classification accuracy.

Following the initial model application, a comprehensive analysis of its limitations and areas for improvement was conducted. The analysis revealed two major issues: overfitting, where the model performed well on the training data but poorly on unseen data, and data leakage, which compromised the validity of the evaluation by inadvertently allowing information from

the validation set to influence the training process. These findings guided the development of a second, more robust modeling procedure, which was applied to the initial dataset of 3,005 compounds to improve generalizability and ensure methodological integrity.

In the new round of model development, three types of molecular representations were tested: one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D) descriptors. These descriptors were computed using PaDEL software (v0.1.16) (117), based on canonical Simplified Molecular Input Line Entry System (SMILES) strings. Additionally, two fingerprint-based encodings were generated using the RDKit package (131) : Molecular ACCess System keys (MACCS (132)) keys and Morgan (133) fingerprints. For the Morgan fingerprints, various configurations were explored, including radii of 2, 3, and 4, each with lengths of 1024 and 2048 bits.

Prior to model training, all datasets underwent rigorous preprocessing to ensure data quality and consistency. Features exhibiting zero variance, NaN values, high correlation (Pearson correlation coefficient > 0.80), or low variance (standard deviation < 0.1) were removed. For continuous descriptors, feature scaling was performed using Scikit-learn's StandardScaler, which normalized each feature to have zero mean and unit variance. Scaling was not applied to MACCS or Morgan fingerprints due to their inherently binary nature. Importantly, normalization was applied separately to each training set, and the corresponding transformation parameters were then used to normalize the associated test set, preserving data integrity during validation.

A total of five machine learning classifiers were evaluated for binary classification of compounds as non-ecotoxic (class 0) or ecotoxic (class 1): Random Forest Classifier, Extra Trees Classifier, XGBoost Classifier (134) (XGBClassifier), AdaBoost Classifier, Gradient Boosting Classifier. Random Forest and Extra Trees utilize bagging techniques by constructing multiple decision trees and aggregating their predictions. In contrast, XGBoost, AdaBoost, and Gradient Boosting follow boosting strategies that iteratively focus on misclassified samples to improve performance.

To refine input features for each molecular encoding (PaDEL descriptors, MACCS keys, and Morgan fingerprints), we implemented a feature selection pipeline. Across 200 training epochs, each classifier was trained on a randomly sampled and scaled 70% subset of the dataset. At every epoch, the top 20 most important features – as determined by the classifier's built-in feature importance algorithm – were recorded for each encoding. The features that appeared most frequently across all epochs were retained for further evaluation.

To determine the optimal subset of features for each classifier and encoding type, we employed an incremental feature selection strategy: starting with the top 5 features, one feature was added at each step until all 20 were included. Model performance was assessed using mean ROC AUC scores from two perspectives: internal validation through repeated cross-validation on the training set and external validation on 10 held-out test sets to simulate real-world performance. This comprehensive and systematic approach enabled us to identify the most predictive and computationally efficient model–feature combinations, ensuring robust generalizability and improved classification performance across different molecular representations.

The final dataset was partitioned into 70% training and 30% testing subsets using the train_test_split function from the Scikit-learn library. To address class imbalance and enrich the training data, various oversampling techniques were applied exclusively to the training set. Specifically, the SMOTE (135) algorithm (from the imblearn.over_sampling library) was used to generate synthetic minority class examples,

ensuring a balanced class distribution. This augmentation step aimed to reduce overfitting and improve the model's robustness, particularly when working with limited sample sizes. For datasets using binary molecular encodings (e.g., MACCS and Morgan fingerprints), we employed SMOTENC (136) instead of SMOTE. Unlike SMOTE, which can generate non-binary (float) values that are unsuitable for categorical or binary data, SMOTENC is explicitly designed to handle categorical features, ensuring that binary patterns in fingerprints are preserved during oversampling. To further reduce overfitting and improve model generalizability, we implemented 7-fold cross-validation, repeated 50 times with randomized fold assignments in each iteration. This was performed using the RepeatedStratifiedKFold function from Scikit-learn. In each repetition, six folds were used for training and one for internal validation. This repeated cross-validation strategy provided a robust estimate of model stability across different data partitions. After final model training on the full 70% training set, performance was evaluated on the independent 30% test set, which was never used during training or cross-validation. This approach provided an unbiased assessment of each model's ability to generalize to unseen data, reducing the risk of performance inflation. Model performance was assessed using several key classification metrics, including accuracy, sensitivity (recall), and specificity. Additionally, Receiver Operating Characteristic (ROC) curves were generated, and Area Under the Curve (AUC) values were calculated to further quantify predictive performance. To ensure statistical reliability, the entire evaluation process was repeated across 10 independent epochs, and final results were reported as the mean ± standard deviation of each performance metric.

To enhance the interpretability of our machine learning models, we employed SHapley Additive exPlanations (SHAP) (137), a model-agnostic interpretability framework grounded in game theory. SHAP assigns each input feature a SHAP value, representing its contribution to a given model prediction. These values quantify how much each molecular descriptor or fingerprint influences the model's output for a specific compound, allowing us to understand which features drive predictions related to acute ecotoxicity. SHAP treats each feature as a "player" in a cooperative game, where the final prediction represents the game's payout. By computing the marginal contribution of each feature across all possible feature subsets, SHAP ensures a fair and consistent allocation of importance—even when features interact. A SHAP value threshold of 0.1 was established: features with SHAP values above this threshold were considered influential, while those below were deemed less impactful. For the model built on PaDEL descriptors, SHAP values were combined with insights from a 2D Principal Component Analysis (PCA) (138) plot. PCA was applied to the scaled descriptor dataset, and the first two principal components captured the majority of variance among compounds. The coefficients (loadings) of each descriptor in these components reflected their directional influence on the chemical space. SHAP was then used to identify which key features contributed most to the separation of class 1 compounds (ecotoxic) in this space. To further interpret these features, we examined their distribution across the two classes using boxplots generated from the original PaDEL descriptor values. Confidence intervals for each descriptor within class 0 and class 1 provided insight into trends differentiating ecotoxic and non-ecotoxic compounds. Additionally, a Mann-Whitney U test (139)—a non-parametric statistical method—was applied to rank features by their ability to distinguish between the two classes.

For models trained on MACCS keys and Morgan fingerprints, we quantified the frequency of each binary feature (bit = 1) across classes. SHAP was again applied to evaluate the influence of class 1-enriched substructures on model predictions. This approach allowed us to pinpoint specific chemical substructures associated with increased bioconcentration or

ecotoxic potential. These findings offer interpretable, structure-based insights that may inform environmental risk assessment and guide safer chemical or drug design.

The optimal LC50 classification model was generated by classifiers trained on PaDEL descriptors. As summarized in Table 8, the Random Forest, XGBoost, and ExtraTrees classifiers consistently outperformed other algorithms, each achieving accuracy scores above 85%. In contrast, AdaBoost and Gradient Boosting delivered lower performance, with accuracies of 76.84% ± 1.13 and 75.36% ± 0.68, respectively. Among the top performers, Random Forest and ExtraTrees stood out by delivering specificity and ROC AUC values exceeding 90%, indicating strong capabilities in correctly identifying non-ecotoxic compounds. Moreover, both models demonstrated reliable sensitivity in detecting ecotoxic compounds, with Random Forest achieving 74.98% ± 3.73 and ExtraTrees achieving 73.38% ± 3.16. The robustness of the ExtraTrees model was further validated through external testing. Figure 9 illustrates its predictive performance over ten independent validation runs, using ROC AUC as the primary metric for evaluating specificity across randomly generated test sets. Owing to its superior accuracy, generalization capability, and consistent performance across multiple fish species, the ExtraTrees model was ultimately selected for integration into the G.AI.A platform.

*Table 8. Cross-model validation metrics for LC50 models trained on PaDEL descriptors after 10 epochs of classification on random test sets.*

| Classifier | Accuracy mean (± std) % | Specificity mean (± std) % | Sensitivity mean (± std) % | ROC AUC mean (± std) % |
|---|---|---|---|---|
| **Random Forest** | 86.76 (± 1.04) | 90.16 (± 1.30) | 74.98 (± 3.73) | 90.73 (± 1.57) |
| **XGBoost** | 85.59 (± 1.09) | 89.70 (± 0.95) | 71.31 (± 3.60) | 88.96 (± 1.51) |
| **ExtraTrees** | 87.99 (± 1.13) | 92.21 (± 0.97) | 73.38 (± 3.16) | 90.96 (± 1.35) |
| **ADABoost** | 76.84 (± 1.13) | 78.13 (± 1.88) | 72.39 (± 4.08) | 83.29 (± 1.70) |
| **Gradient Boosting** | 75.36 (± 0.68) | 77.24 (± 1.06) | 70.00 (± 3.17) | 80.53 (± 1.28) |

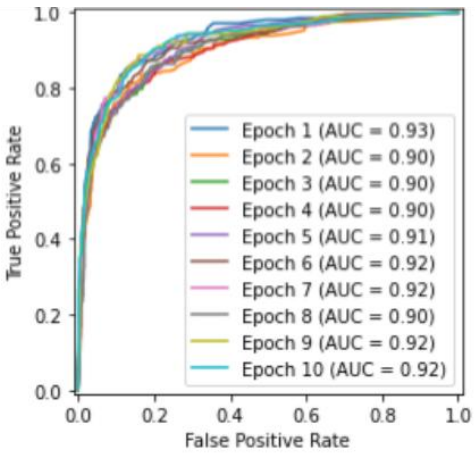



*Figure 9: Receiver Operating Characteristic (ROC) curves from 10-epoch validation runs using the ExtraTrees classifier, the best-performing model for predicting chemical ecotoxicity potential.*

The contribution of each feature to the performance of the ExtraTrees model was evaluated using SHAP analysis. The complete list of the 14 descriptors used in model training is detailed in Table 9, with their SHAP value distributions visualized in Figure 10. A focused summary of the eight most influential features for distinguishing ecotoxic compounds in fish is provided in Table 10. Notably, "CrippenLogP" and "ALogP", both indicators of molecular lipophilicity, were among the top contributors, exhibiting the highest mean absolute SHAP values. Their strong influence reinforces the ecotoxicological principle that compounds with increased lipophilicity are more likely to accumulate in fatty tissues, thereby enhancing their toxic potential. "AATS5p", a descriptor reflecting molecular polarizability, also emerged as a shared key predictor across both bioconcentration and ecotoxicity models, further highlighting the relevance of electronic properties in toxicity assessment. Additional descriptors such as "SpMax5_Bhm" and "ATS3m", which integrate topological and mass-related features, proved essential to model accuracy. These findings are consistent with the notion that larger and more structurally complex molecules often exhibit higher toxicity, potentially due to enhanced lipophilic behavior (140). Boxplot visualizations of these features across ecotoxic (class 1) and non-ecotoxic (class 0) compounds, presented in Figure 11, clearly illustrate the trends and distributions underpinning their predictive power. Finally, "VP-3", which describes valence electron pathways through three connected heavy atoms, also showed high SHAP values, emphasizing the role of electronic structure in distinguishing toxic compounds within the multi-species fish dataset.

*Table 9. The full set of descriptors used in generating our best ecotoxic predictive model, along with a brief description of their type, their P-values derived from the Mann-Whitney test, their weights in PC1 and PC2 from the PCA analysis, and their mean SHAP values.*

| Descriptors | Descriptor type | P-value | Weight in PC1 (36.81% of variance) | Weight in PC2 (19.66% of variance) | SHAP value (mean) |
|---|---|---|---|---|---|
| **ALogP** | ALOGP | 2.08e-76 | - 0.001 | 0.686 | 0.078 |
| **AATS5p** | Autocorrelation | 2.60e-106 | 0.015 | 0.103 | 0.074 |
| **SpMax5_Bhm** | Burden modified eigenvalues | 9.90e-125 | 0.401 | - 0.225 | 0.069 |
| **VP-3** | Chi path | 1.20e-123 | 0.167 | 0.120 | 0.065 |
| **CrippenLogP** | Crippen logP | 1.30e-136 | 0.225 | 0.314 | 0.064 |
| **XLogP** | XLogP | 1.25e-92 | -0.059 | 0.323 | 0.064 |
| **ZMIC1** | Information content | 9.80e-124 | 0.011 | -0.128 | 0.057 |
| **ATS3m** | Autocorrelation | 4.96e-127 | 0.422 | 0.074 | 0.055 |
| **SpMax4_Bhm** | Burden modified eigenvalues | 1.58e-118 | -0.424 | 0.106 | 0.053 |
| **ATS2m** | Autocorrelation | 1.09e-124 | -0.209 | 0.204 | 0.048 |
| **SpMax1_Bhp** | Burden modified eigenvalues | 1.92e-79 | 0.028 | -0.112 | 0.047 |
| **ATS4m** | Autocorrelation | 1.21e-122 | 0.395 | 0.107 | 0.045 |
| **ATS1m** | Autocorrelation | 7.16e-117 | 0.128 | -0.341 | 0.045 |
| **SpMax3_Bhm** | Burden modified eigenvalues | 3.68e-119 | -0.427 | -0.188 | 0.044 |

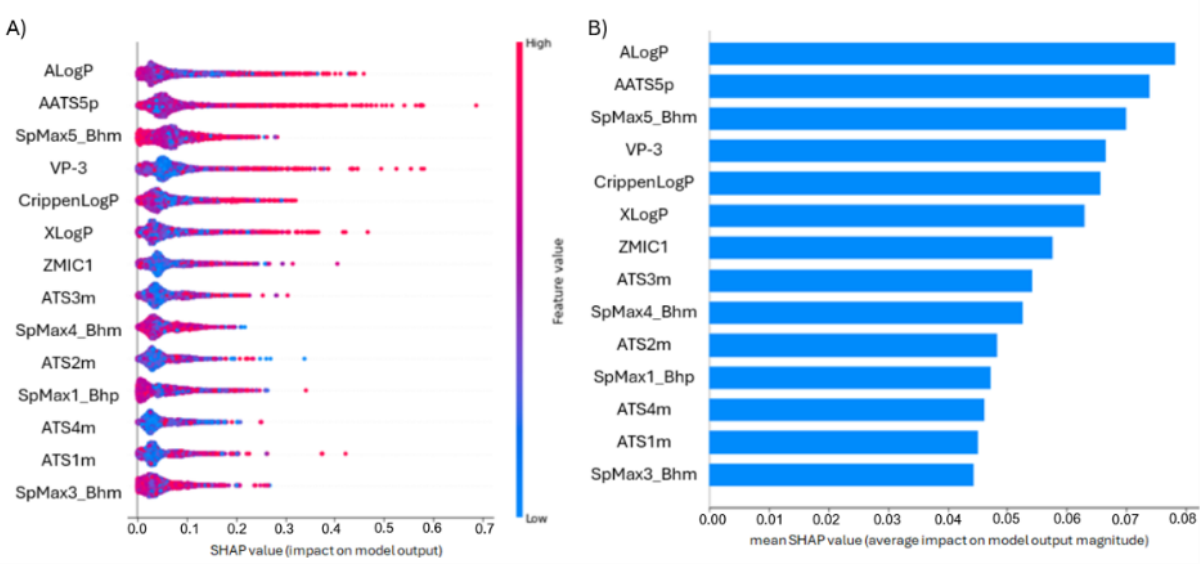



*Figure 10: A) Histogram displaying the mean absolute SHAP values of all features employed in training the ExtraTrees classifier. B) Distribution of local SHAP values for each feature used in the ExtraTrees model.*

*Table 10. The eight most important descriptors used in generating our best ecotoxic predictive model, along with a brief description of their type, their P-values derived from the Mann-Whitney test, their weights in PC1 and PC2 from the PCA analysis, and their mean SHAP values.*

| Descriptors | Descriptor type | P-value | Weight in PC1 (36.81% of variance) | Weight in PC2 (19.66% of variance) | SHAP value (mean) |
|---|---|---|---|---|---|
| **ALogP** | ALOGP | 2.08e-76 | - 0.001 | 0.686 | 0.078 |
| **AATS5p** | Autocorrelation | 2.60e-106 | 0.015 | 0.103 | 0.074 |
| **SpMax5_Bhm** | Burden modified eigenvalues | 9.90e-125 | 0.401 | - 0.225 | 0.069 |
| **VP-3** | Chi path | 1.20e-123 | 0.167 | 0.120 | 0.065 |
| **CrippenLogP** | Crippen logP | 1.30e-136 | 0.225 | 0.314 | 0.064 |
| **ZMIC1** | Information content | 9.80e-124 | 0.011 | -0.128 | 0.057 |
| **ATS3m** | Autocorrelation | 4.96e-127 | 0.422 | 0.074 | 0.055 |
| **SpMax1_Bhp** | Burden modified eigenvalues | 1.92e-79 | 0.028 | -0.112 | 0.047 |

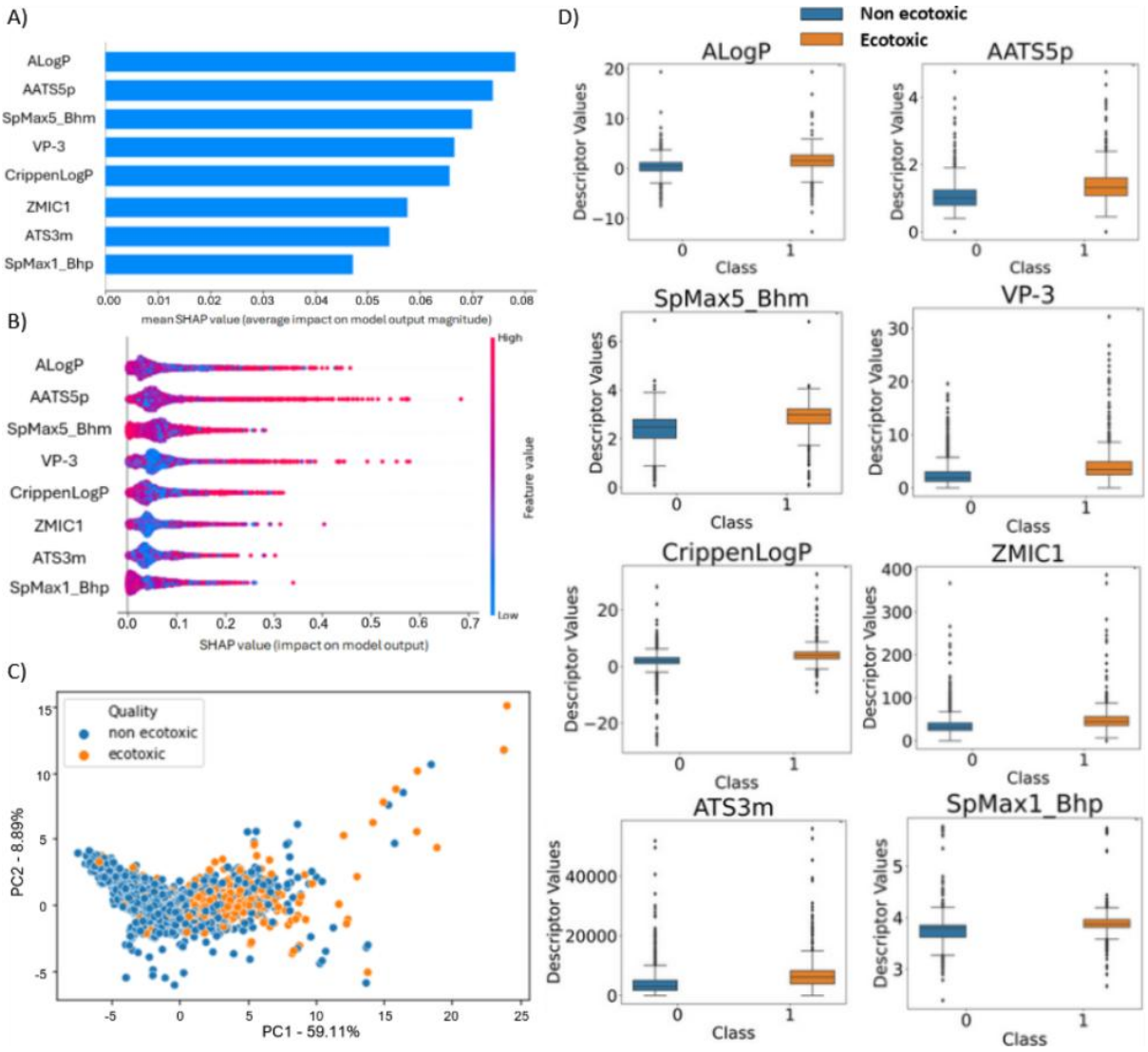



*Figure 11: A) Histogram of mean absolute SHAP values for the most influential features in differentiating ecotoxic compounds. B) Distribution of local SHAP values for the most influential features in differentiating ecotoxic compounds. C) 2D PCA plot depicting the training features of the ecotoxic ML model. D) Boxplots representing the distribution values of the 8 most influential features in distinguishing ecotoxic compounds across the two ecotoxic classes.*

While the descriptor-based models offered valuable insights into the physicochemical drivers of ecotoxicity, we extended our analysis to include molecular fingerprint-based features to uncover additional structural and chemical patterns. To this end, the multi-species fish dataset was further analyzed using Morgan fingerprints – generated with varying radii and bit lengths – and MACCS keys as input features. This strategy was designed to capture substructural motifs associated with ecotoxic compounds and potentially enhance the predictive capacity of our machine learning models. The outcomes of this analysis, summarized in Table 25, reflect the performance of the top fingerprint-based models across ten independent external validation runs. Compared to the best-performing descriptor-based model—which achieved a mean ROC AUC of 90.96% ± 1.35%—models trained on fingerprint features generally exhibited lower predictive performance. Among the Morgan fingerprint configurations, the Random Forest model trained with radius 3 and 1024-bit fingerprints delivered the highest performance, reaching a ROC AUC of 79.31% ± 1.10%. Likewise, the best MACCS-based model attained a ROC AUC of 80.14% ± 1.70%, also falling short of the descriptor-based benchmark. Despite

the modest performance gap, the fingerprint-based models offered critical interpretability benefits.

To uncover structural features of small molecules that drive ecotoxicity in fish species, we conducted SHAP analysis on the top-performing machine learning models trained with Morgan and MACCS fingerprints. This analysis focused on the top 20 most influential fingerprint bits for distinguishing ecotoxic (class 1) from non-ecotoxic (class 0) compounds, as shown in Figure 12. The distribution of each fingerprint bit across both classes is illustrated in Figure 13. To refine our identification of key ecotoxicity-related substructures, we further filtered the most impactful features by selecting those with both high mean absolute SHAP values and a minimum twofold prevalence in the ecotoxic class compared to the non-ecotoxic class. This dual criterion was applied to both Morgan (Figure 14) and MACCS-based models (Figure 15) to highlight the most ecologically relevant structural alerts. In the highest-performing MACCS-based model – which surpassed all other fingerprint-based approaches – several features stood out. Bit 145 and bit 125, which encode for the presence of single or fused aromatic rings, ranked first and third in mean SHAP importance (Figure 15B), respectively. These aromatic features appeared in nearly 40% of ecotoxic compounds, but in less than 20% of non-ecotoxic ones, reinforcing findings from prior studies that associate aromatic systems (e.g., carbonyl-benzene structures) with elevated aquatic toxicity (141) (142) (143). Additionally, bit 134, indicative of halogen atoms, was the most frequent feature in the ecotoxic class (present in 50% of compounds) and held the second-highest SHAP value. Bit 87, also linked to halogenated fragments and previously highlighted in our bioconcentration models, showed a similar trend – occurring in about 40% of ecotoxic compounds versus 20% in non-ecotoxic ones – despite having a lower SHAP value. These findings are consistent with toxicological evidence that halogen substitution patterns increase the toxicity potential of organic pollutants in aquatic systems (144) (145). These structural alerts serve as practical and interpretable indicators, supporting both risk assessment and regulatory decision-making efforts in environmental toxicology.

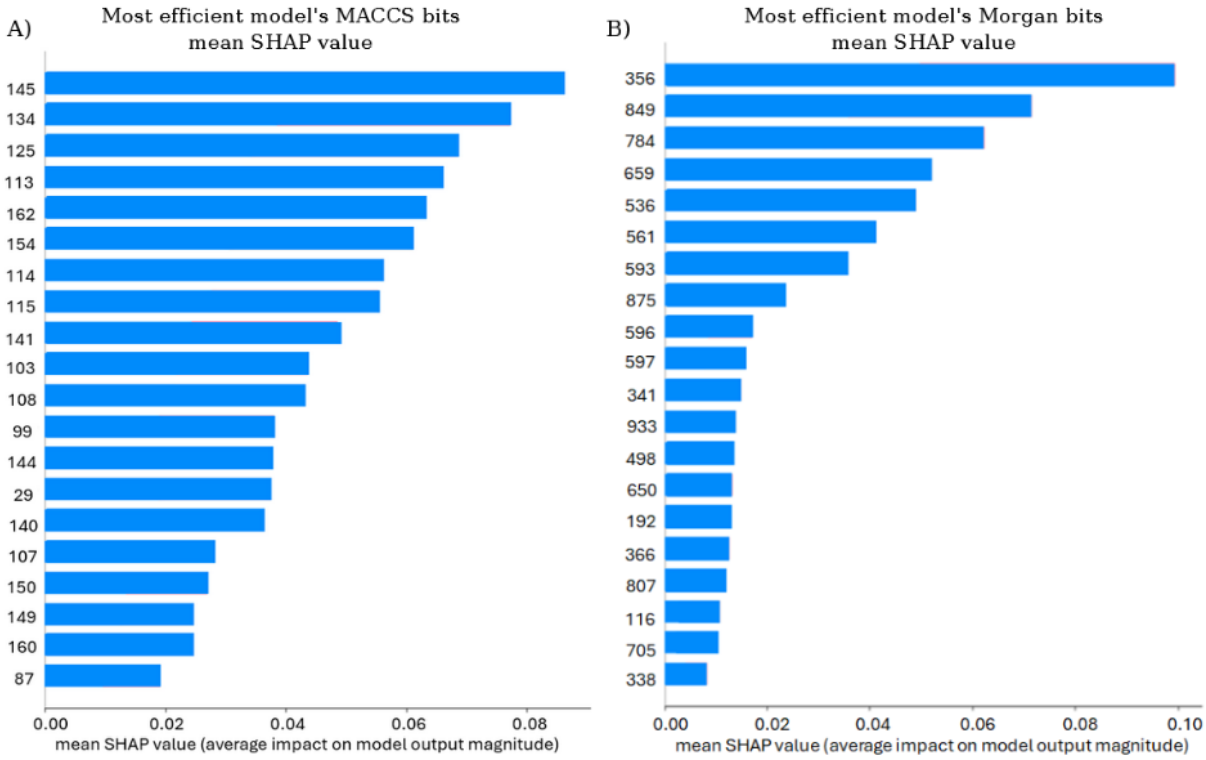



*Figure 12: Mean absolute SHAP values of the top 20 structural features contributing to the prediction of ecotoxicity potential in the best-performing ML models based on A) MACCS fingerprints and B) Morgan fingerprints.*

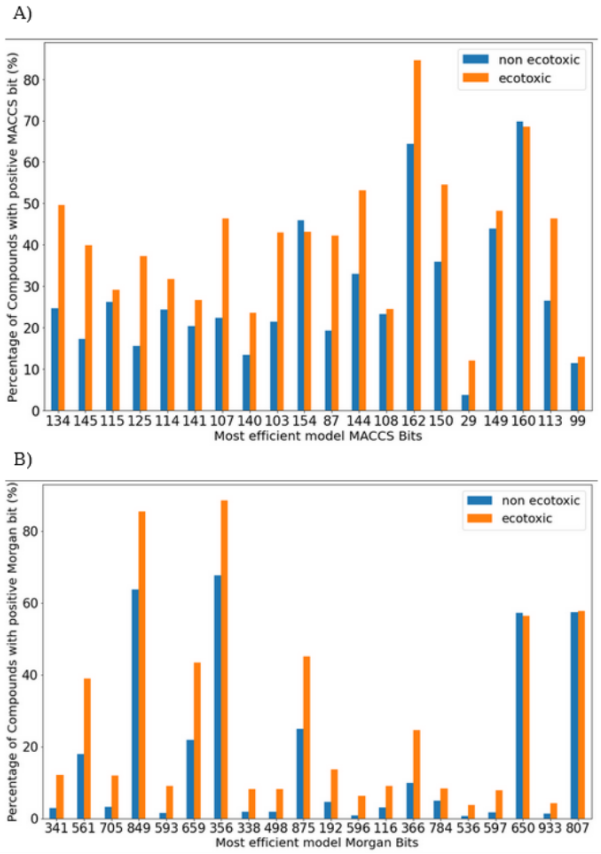



*Figure 13: Class-wise distribution of the top 20 structural features in the best-performing ML models based on A) MACCS fingerprints and B) Morgan fingerprints for predicting the ecotoxicity potential of chemical compounds.*

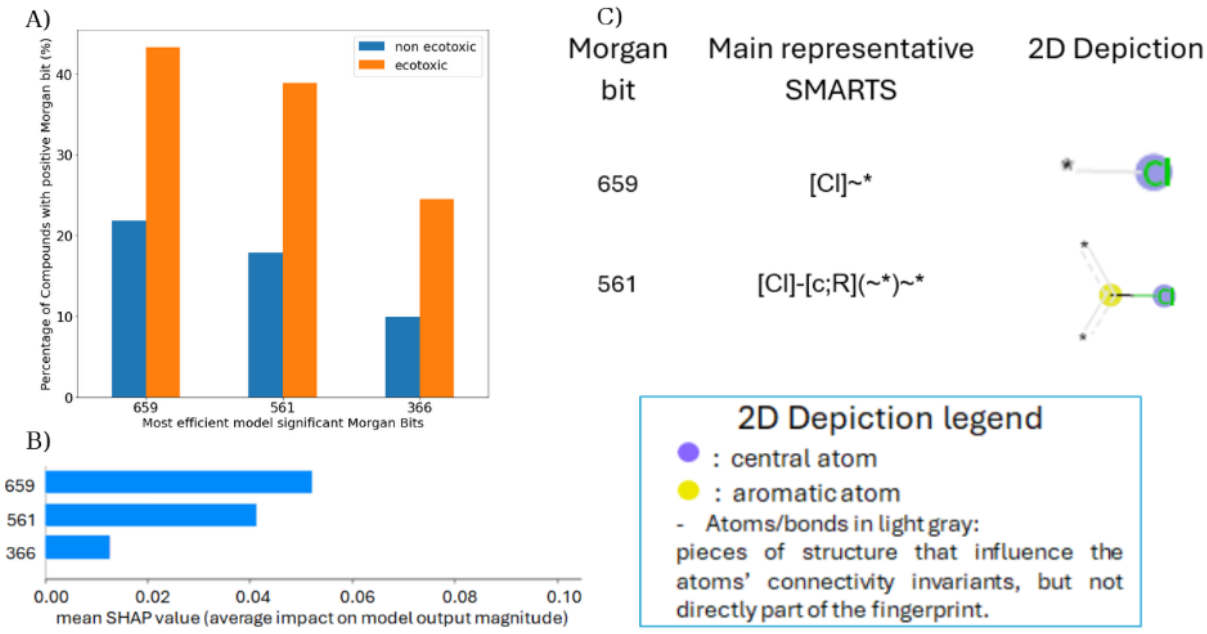



*Figure 14: A) Class-wise distribution of the selected structural features in the best-performing Morgan-based machine learning model (Random Forest model trained with Morgan fingerprints, radius 3, 1024 bits) for predicting the ecotoxicity potential of chemical compounds. B) Mean absolute SHAP values for the selected structural features in the best-performing Morgan-based ML model C) The most significant Morgan bits along with their SMARTS representations and 2D chemical depictions.*

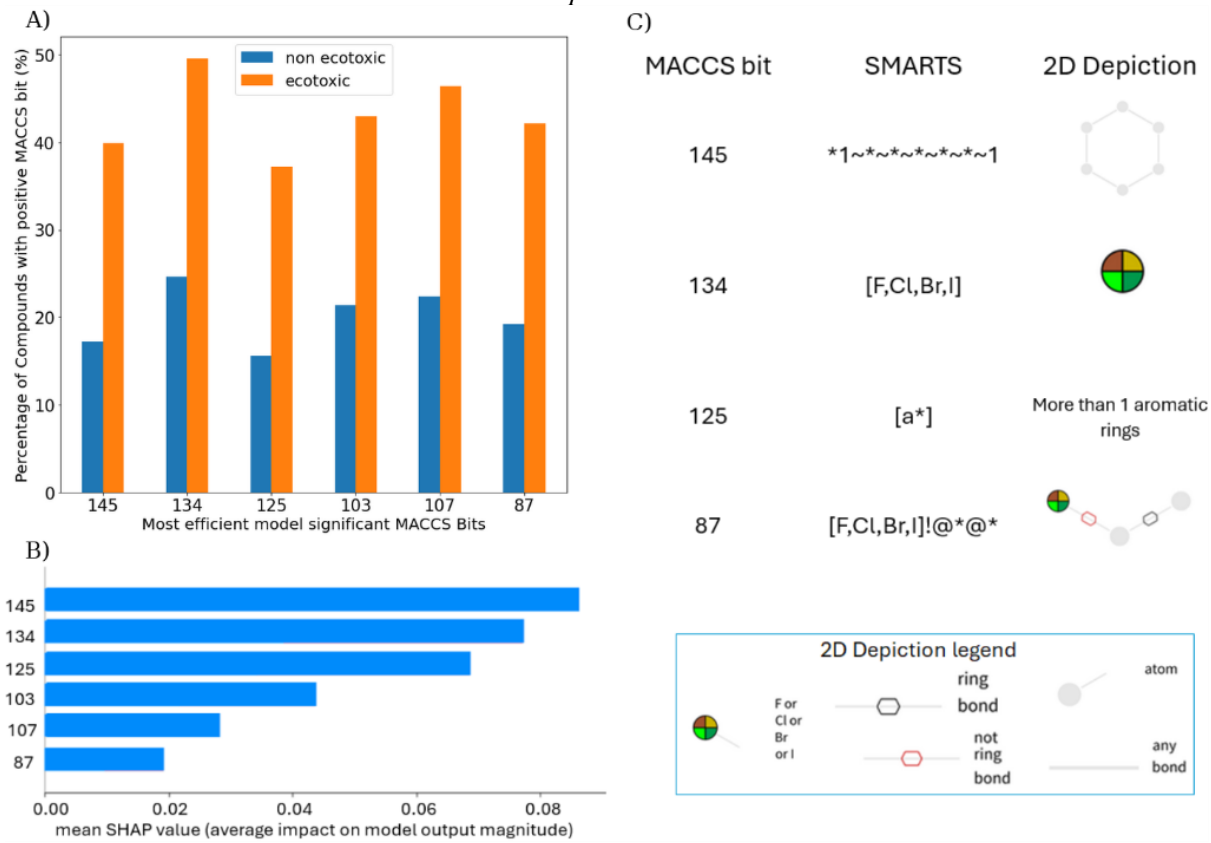



*Figure 15: A) Distribution of the selected structural features across both classes in the best-performing MACCS-based ML model. B) Mean absolute SHAP values for the selected structural features in the best-performing MACCS-based ML model C) The most significant MACCS bits along with their SMARTS representations and 2D chemical depictions. Bits 103 and 107 are omitted, as their representative substructures closely resemble bit 134.*

To benchmark the performance of our best LC50 predictive model, we conducted a comparative evaluation against several widely used, publicly available tools: VEGA (146), OPERA (147), and T.E.S.T.. This evaluation was based on both our curated LC50 dataset and an external validation set, enabling an objective assessment of model generalizability. For the VEGA platform, we focused on two established models for acute aquatic toxicity prediction: IRFMN-Combase, designed to estimate LC50 values for fish species using curated toxicity databases, and KNN-IRFMN, a k-nearest neighbours model optimized specifically for predicting LC50 in Fathead Minnow (Pimephales promelas). Additionally, the T.E.S.T. (v5.1.2) software developed by the U.S. Environmental Protection Agency was employed. Among its various modeling approaches, we selected the Consensus method, which combines multiple algorithms and has demonstrated superior predictive performance in external validation studies. The endpoint selected for comparison was Fathead Minnow LC50 (96-hour exposure). We also utilized OPERA (version 2.9), an open-source QSAR modeling suite developed by the U.S. National Institute of Environmental Health Sciences (NIEHS). OPERA is designed to predict a range of physicochemical and toxicological endpoints relevant to environmental risk assessment. LC50 predictions were generated using OPERA's standard configurations, as specified in its official documentation and GitHub repository (https://github.com/kmansouri/OPERA). This comparative analysis allowed us to rigorously evaluate the predictive accuracy of our LC50 model in relation to other publicly available tools, providing insights into its potential applicability for regulatory and environmental decision-making.

All models—our LC50 prediction model, along with other publicly available tools, including VEGA and T.E.S.T.—were first benchmarked using the same external LC50 dataset compiled for training and validating our machine learning approach (Table 11). This consistent evaluation framework allowed for a direct and fair comparison of predictive performance across tools. Among the models assessed, the G.AI.A LC50 model significantly outperformed the publicly available alternatives. It achieved a ROC AUC of 90.96%, notably higher than T.E.S.T (67.00%) and VEGA (55.00%). In terms of overall accuracy and specificity, G.AI.A also led with 87.99% accuracy and 92.21% specificity, compared to VEGA (85.04% and 88.11%) and T.E.S.T (78.49% and 80.97%). Importantly, the sensitivity of each model—its ability to correctly identify ecotoxic compounds—varied significantly. VEGA demonstrated poor sensitivity at just 33.33%, indicating a substantial limitation in detecting toxic compounds, which are underrepresented in the dataset. In contrast, G.AI.A (73.38%) and T.E.S.T (66.90%) showed a much stronger ability to detect these minority-class compounds, offering greater practical relevance in ecological risk screening. While VEGA and T.E.S.T. provide limited interpretability, the G.AI.A platform offers not only high predictive accuracy but also interpretable chemical insights through model explainability techniques. This enables the identification of key structural features driving ecotoxicity, providing valuable guidance for chemical risk assessment and regulatory applications.

*Table 11. Validation metrics obtained from testing publicly available online software using the LC50 dataset.*

| Online Software/Model | Accuracy (%) | Specificity (%) | Sensitivity (%) | ROC AUC (%) |
|---|---|---|---|---|
| **VEGA** | 85.04 | 88.11 | 33.33 | 55.00 |
| **T.E.S.T** | 78.49 | 80.97 | 66.40 | 67.00 |
| **(our) G.AI.A MODEL** | 87.99 | 92.21 | 73.38 | 90.96 |

The ExtraTrees model exhibits a strong overall predictive performance, achieving an accuracy of 87.99%. However, its sensitivity – reflecting the model's ability to correctly identify ecotoxic compounds – stands at a moderate 73.38%. Despite this, the model demonstrates solid generalization capacity, as evidenced by its performance on independent validation using the benchmark dataset ADORE. To evaluate external applicability, two curated datasets were utilized: t_F2F (comprising 1,118 compounds from various fish species) and s-F2F-2 (containing 497 compounds specific to Pimephales promelas). These dataset were preprocessed by removing non-96-hour LC50/EC50 values within the non-ecotoxic range (LC50/EC50 >1 mg/L), selecting the minimum LC50/EC50 value per fish, filtering out compounds with missing values (NANs) in the training features of our LC50 model, and eliminating compounds with molecular weights below 150 Da. Importantly, all compounds used for model training were excluded from these external sets, which were standardized to match the modeling criteria. As illustrated in Figure 16, there is both overlap and diversity in chemical space between the training set and these external datasets, highlighting the broad applicability domain of our LC50 dataset in comparison with a widely accepted external reference. Table 12 presents a summary of the benchmarking results across three models. On the t_F2F dataset, the ExtraTrees model delivered superior performance, with an accuracy of 90.44% and a specificity of 96.42%. In comparison, VEGA and T.E.S.T achieved significantly lower accuracies (44.31% and 67.18%, respectively) and specificities (77.88% and 82.05%, respectively). Notably, our G.AI.A model achieved a substantial improvement in sensitivity (83.88%), outperforming both its internal validation score and the sensitivities reported for VEGA (8.09%) and T.E.S.T (51.14%). Furthermore, G.AI.A demonstrated exceptional robustness, reflected in its ROC AUC of 96.06%, significantly surpassing VEGA (35.57%) and T.E.S.T (65.15%).

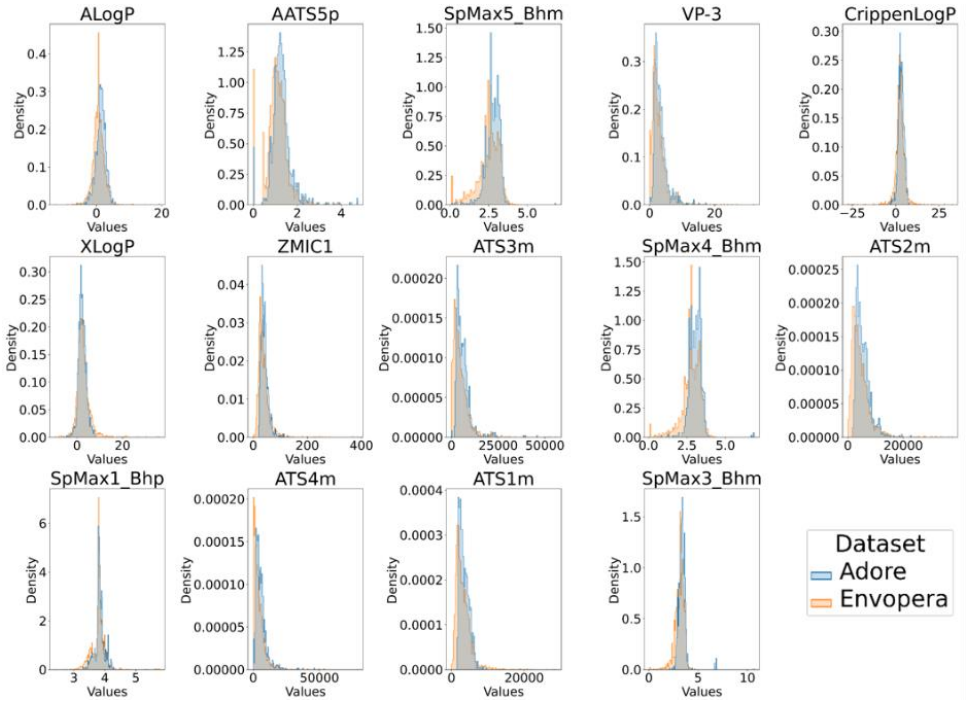



*Figure 16: Comparison of chemical space covered by the compounds in the ADORE dataset of multiple fish species (t_F2F) and our acute ecotoxicity dataset (Envopera). Each plot delineates the chemical space as defined by the values of training features used in the ExtraTrees model. The density values along the y-axis represent the count of compounds per value normalized by the total number of compounds within each dataset. This metric is employed to facilitate a comparative analysis between the two datasets. The x-axis displays the values of each training feature.*

It is worth noting that, during external validation, only compounds used to train our model were excluded, while those used to train VEGA and T.E.S.T were retained—potentially placing our models at a comparative disadvantage. To address a known limitation of T.E.S.T (which is trained solely on Pimephales promelas LC50 data), we conducted an additional evaluation using the s-F2F-2 dataset, which consists exclusively of compounds tested on this species. On this subset, G.AI.A again demonstrated superior performance, with the highest ROC AUC (90.01%) and sensitivity (89.12%), while maintaining accuracy comparable to T.E.S.T (84.62% vs. 84.22%). Although G.AI.A exhibited slightly lower specificity (82.59%) compared to T.E.S.T (91.00%), it still delivered a balanced and reliable performance. In summary, the G.AI.A model shows strong accuracy, robustness, and generalizability across both diverse and species-specific external datasets, underscoring its potential for reliable ecotoxicity prediction in a variety of environmental contexts.

*Table 12. Validation metrics obtained from testing publicly available online software using the t-F2F (multiple fish species) and s-F2F-2 (Pimephales promelas) external datasets of ADORE database.*

| Online Software/Model | Fish species | Accuracy (%) | Specificity (%) | Sensitivity (%) | ROC AUC (%) |
|---|---|---|---|---|---|
| **VEGA** | multiple fish species | 44.31 | 77.88 | 8.09 | 35.57 |
| **VEGA** | *Pimephales promelas* | 41.18 | 56.23 | 2.73 | 39.94 |
| **T.E.S.T** | multiple fish species | 67.18 | 82.05 | 51.14 | 65.15 |
| **T.E.S.T** | *Pimephales promelas* | 84.22 | 91.00 | 69.06 | 79.78 |
| **(our) G.AI.A MODEL** | multiple fish species | 90.44 | 96.42 | 83.88 | 96.06 |
| **(our) G.AI.A MODEL** | *Pimephales promelas* | 84.62 | 82.59 | 89.12 | 90.01 |

## 3.3 Integration of the developed *in silico* models into an automated pipeline

This deliverable focused on the development and integration of novel computational tools into a streamlined framework designed to enhance early-stage drug discovery, particularly through environmental safety screening. To this end, we created an online software that integrates the most accurate machine learning models we established for predicting bioconcentration, aquatic toxicity, and drug metabolite behaviour. This automated system interprets chemical information, detects substructures or chemical motifs associated with ecological risk, and provides predictive analytics that support green-by-design principles in pharmaceutical research. It enables systematic filtering and ranking of compounds based on potential environmental impact. The pipeline structure and process are depicted in Figure 17. Technically, the platform is built using the Django web framework, with the Python programming language responsible for handling the back-end logic. The user interface is constructed using HTML and CSS, ensuring accessibility and usability for end-users.

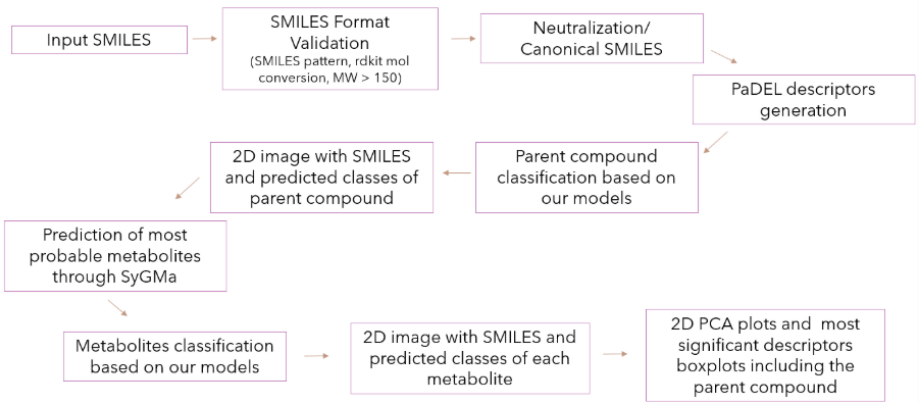



*Figure 17: Automated pipeline workflow.*

Users can submit SMILES strings of organic compounds through two modes (Figure 18 offered by the webserver:

1. "Analyse up to 3 SMILES" mode - users can input up to three SMILES strings for analysis. The tool outputs bioconcentration and ecotoxicity classifications, visualizes molecular structures, provides classification confidence scores, and displays predicted metabolites along with their associated predicted classes. Metabolites with molecular

weights below 150 Da are excluded from classification. Metabolite relevance is ranked using SyGMa scores, and the likelihood of class 0 (non-toxic, non-bioaccumulative) outcomes is presented as a weighted percentage based on predicted probabilities. Interactive t-SNE plots, generated using Plotly, show the positioning of input compounds relative to the training data: class 0 and class 1 compounds are marked with green and red circles, respectively, while user-submitted compounds are shown as blue stars. Separate visualizations are provided for bioconcentration and ecotoxicity predictions. For compounds classified as class 1, the parent molecule's structural features contributing to this classification are highlighted. Additional insights include boxplots of key PaDEL descriptor distributions from the training dataset, showing where the input compounds fall in relation to each class.

2. "Bulk SMILES Analysis" mode - users can upload structure files in ".txt", ".csv", or ".xlsx" format for high-throughput screening. Specific example templates for each accepted format are provided to guide users in structuring their input files correctly. The tool generates a CSV output that includes the input SMILES strings, predicted bioconcentration and ecotoxicity classes along with their associated classification confidence scores, up to 24 most probable metabolites per compound (where applicable), and a summary of any entries that could not be processed. For deeper interpretation and visualization, users are encouraged to utilize the "Analyse up to 3 SMILES" mode.

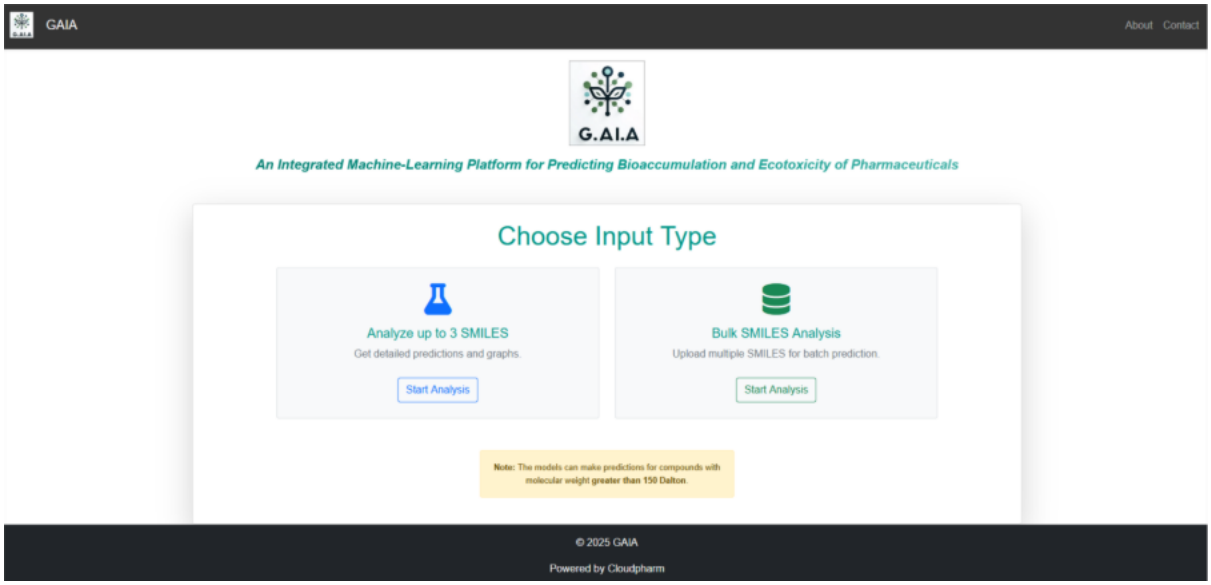



*Figure 18:Input types accepted by G.AI.A server.*

Below, we elaborate on the "Analyse up to 3 SMILES" mode, which enables users to submit up to three chemical structures simultaneously for quick, small-scale screening with interactive plots. This mode serves as a streamlined summary of the broader analysis pipeline and is designed to showcase its core functionalities, particularly for exploratory or demonstrative purposes.

The user interface allows users to input chemical compounds in the form of SMILES notation via a dedicated entry field (Figure 19). Submitted SMILES strings must be valid for processing with RDKit and must describe molecules with a molecular mass above 150 Daltons. To demonstrate the pipeline's functionality, we applied it to a test compound – zileuton, an antiasthmatic medication – entered using its SMILES:

CC(C1=CC2=CC=CC=C2S1)N(C(=O)N)O. The system analysed zileuton in terms of both bioaccumulation potential and acute fish toxicity.

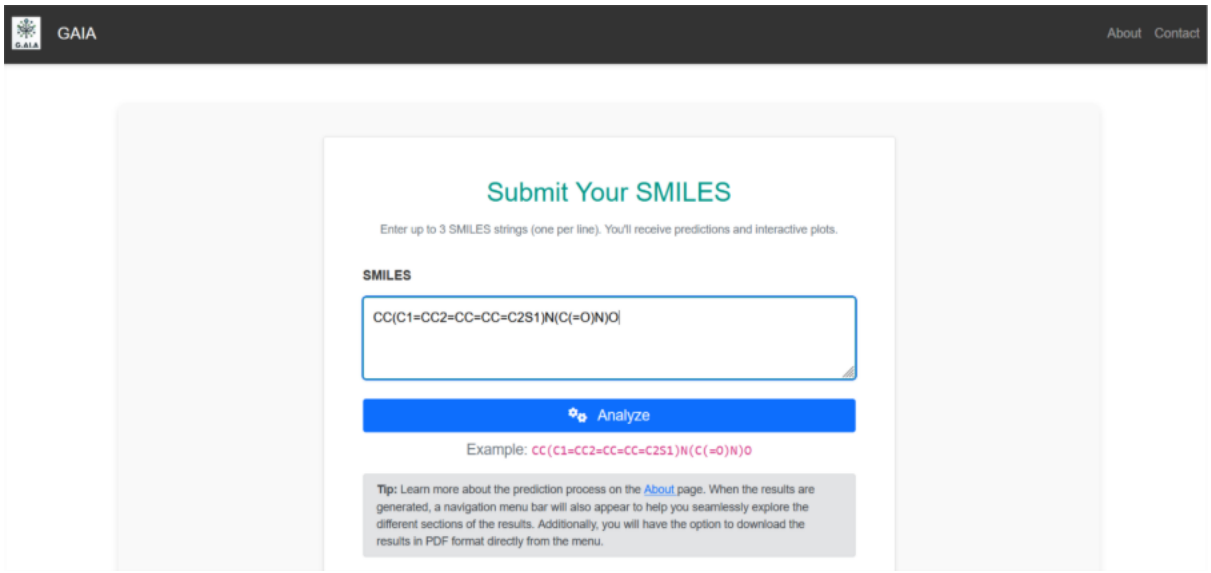



*Figure 19: Input dialog box of our online server.*

Upon submission, the input undergoes standardization, involving both neutralization and canonicalization, to ensure consistency across computational steps. Molecular descriptors are then computed using PaDEL, but only those features used during the training phase of each model are retained. These selected descriptors are normalized using StandardScaler from the sklearn.preprocessing module, based on the statistical distribution of the original training data. These processed features are then passed through trained classifiers to generate predictions for bioconcentration (bioaccumulative or not) and acute toxicity (ecotoxic or not) in fish.

Once entered, the SMILES input undergoes neutralization and canonicalization to ensure consistency before generating the appropriate PaDEL descriptors. Only the descriptors utilized as training features in each model are retained for further processing. Using the StandardScaler function from the sklearn.preprocessing library, these descriptors are standardized based on the mean and standard deviation of the corresponding training dataset. The trained models then classify the parent compound in terms of bioconcentration (bioaccumulative or non-bioaccumulative) and acute toxicity (ecotoxic or non-ecotoxic) in fish (Figure 20). In the case of zileuton, the platform classified the compound as non-bioaccumulative and non-ecotoxic, with confidence levels of 94.66% and 77.74% respectively.

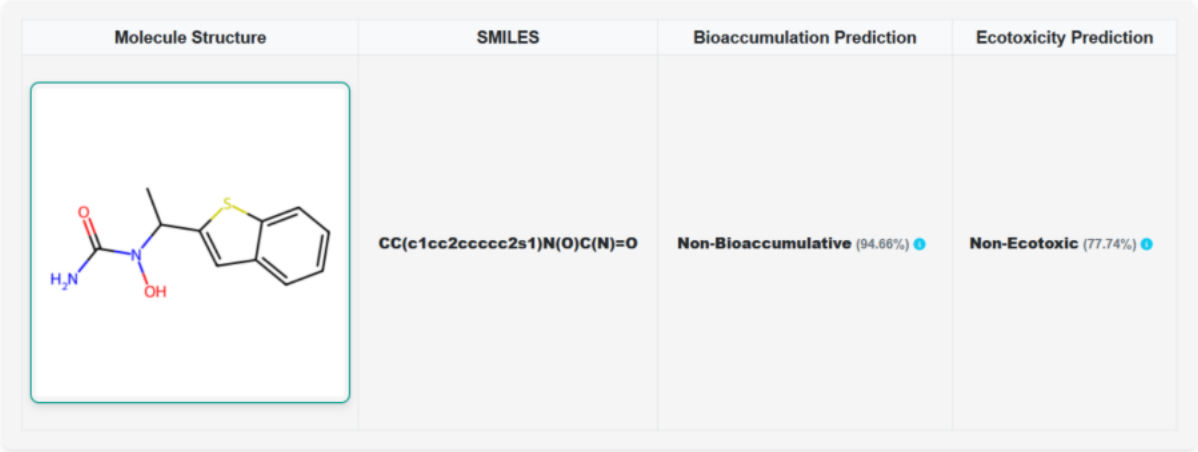



*Figure 20: Predicted bioconcentration and ecotoxicity classifications of zileuton along with associated prediction confidence scores, following the submission of its SMILES string in the automated pipeline.*

For visualization and further interpretation, the pipeline uses RDKit to generate a canonical version of the SMILES input, which is then encoded into a binary format for display within the web interface. This encoding is performed via the b64encode function from Python's base64 library. To better contextualize the model predictions, the system generates two-dimensional t-SNE visualizations using the Plotly library and the TSNE algorithm from sklearn.manifold. These plots project the chemical descriptors of the training dataset and the input molecule into reduced-dimensional space.

Each t-SNE plot distinguishes the compound classes visually. Compounds labelled as non-bioaccumulative or non-ecotoxic (Class 0) are shown as green dots and compounds identified as bioaccumulative or ecotoxic (Class 1) are represented by red dots. The test compound is displayed as a blue star. Two separate t-SNE charts are generated—one for bioconcentration and the other for acute toxicity (Figure 21). In the zileuton case, the blue star is placed within the green cluster for bioconcentration, reinforcing the classification. For ecotoxicity, the star appears near the boundary between green and red clusters, suggesting an intermediate position and potential borderline properties.

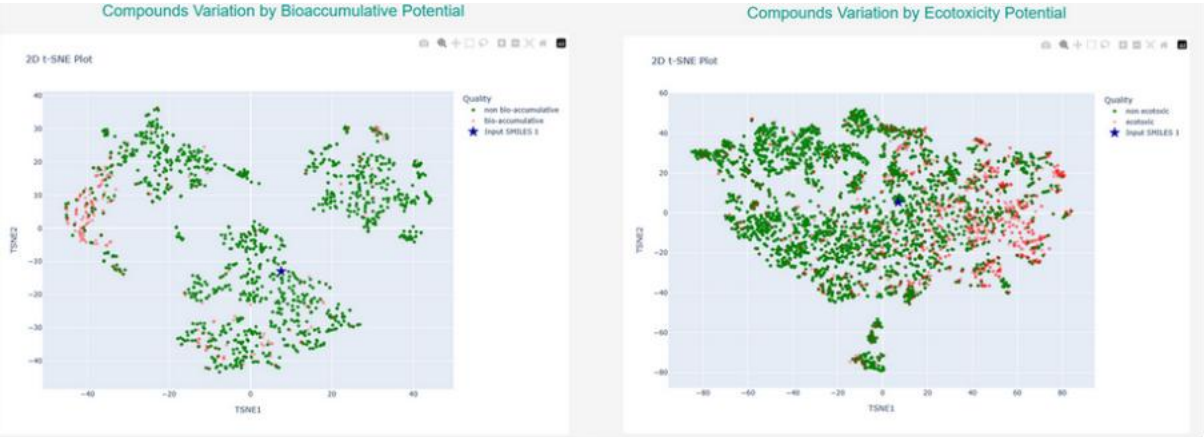



*Figure 21: 2D t-SNE plots of training data used for bioconcentration and acute ecotoxicity models. The "star" symbol represents the input compound (zileuton).*

In addition to t-SNE, the pipeline offers detailed insight through descriptor value distributions. It identifies five significant molecular descriptors from each model and presents their distributions across the training data. These are visualized using boxplots that allow for direct

comparison between the training data and the input compound's values. Each boxplot displays green distributions for non-bioaccumulative or non-ecotoxic compounds (Class 0), orange distributions for bioaccumulative or ecotoxic compounds (Class 1) and a star marker showing the test compound's descriptor value. The following descriptors were used across the models: TopoPSA, AATS5p, gmin, ALogP, MAXDN, CrippenLogP, SpMax5_Bhm, VP-3, ATS3m, and AATS5p (Figure 22). The analysis helps clarify which physicochemical traits may require modification to reduce environmental hazards.

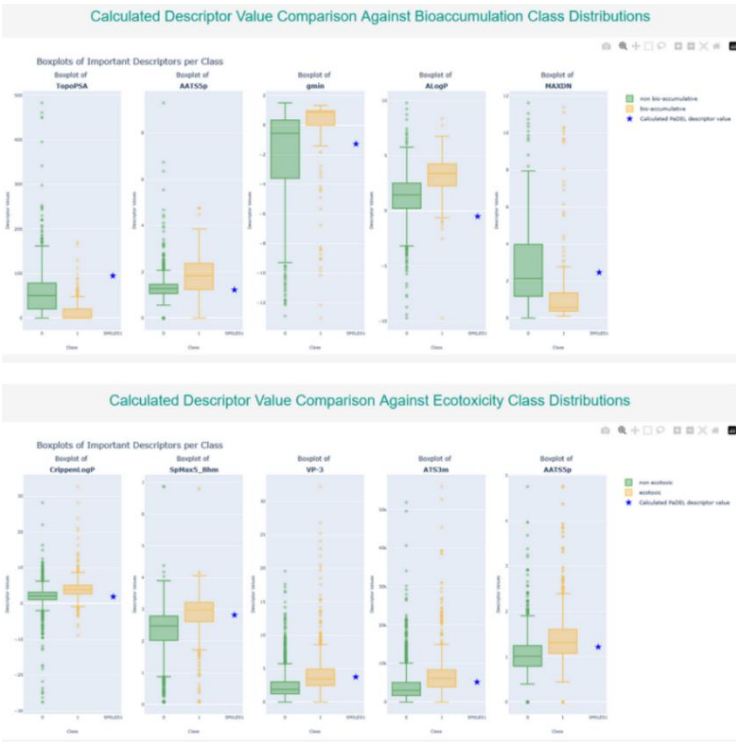
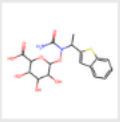
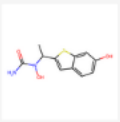
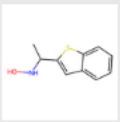
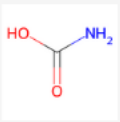
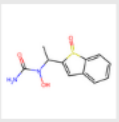
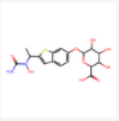
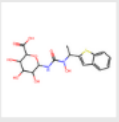
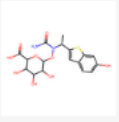



*Figure 22: Boxplots of the five most significant descriptors for the bioconcentration and acute ecotoxicity models. The "star" symbol represents the input compound (zileuton).*

To extend environmental risk assessment to potential drug metabolites, the pipeline incorporates SyGMA for metabolic transformation prediction. The twenty-four most likely metabolites are generated for the input molecule, and their PaDEL descriptors – consistent with those used during model training – are calculated. These are classified using the same models as for the parent compound. Each metabolite is visualized using RDKit and rendered within the interface using base64 encoding. Metabolites with a molecular weight below 150 Daltons are flagged as "MW < 150 Dalton" and excluded from classification. For zileuton, fourteen of the fifteen predicted metabolites shared the same classification as the parent (i.e., non-bioaccumulative and non-ecotoxic). One metabolite, due to insufficient molecular weight, was excluded from prediction (Figure 23).

| Structure | SMILES | Bioaccumulation | Ecotoxicity | Sygma Score |
|---|---|---|---|---|
|  | CC(c1cc2ccccc2s1)N(OC1OC(C(=O)O)C(O)C(O)C1O)C(N)=O | Non-bioaccumulative | Non-ecotoxic | 0.213 |
|  | CC(c1cc2ccc(O)cc2s1)N(O)C(N)=O | Non-bioaccumulative | Non-ecotoxic | 0.061 |
|  | CC(NO)c1cc2ccccc2s1 | Non-bioaccumulative | Non-ecotoxic | 0.053 |
|  | NC(=O)O | MW < 150 | | 0.053 |
|  | CC(C1=Cc2ccccc2S1=O)N(O)C(N)=O | Non-bioaccumulative | Non-ecotoxic | 0.025 |
|  | CC(c1cc2ccc(OC3OC(C(=O)O)C(O)C(O)C3O)cc2s1)N(O)C(N)=O | Non-bioaccumulative | Non-ecotoxic | 0.0152 |
|  | CC(c1cc2ccccc2s1)N(O)C(=O)NC1OC(C(=O)O)C(O)C(O)C1O | Non-bioaccumulative | Non-ecotoxic | 0.013 |
|  | CC(c1cc2ccc(O)cc2s1)N(OC1OC(C(=O)O)C(O)C(O)C1O)C(N)=O | Non-bioaccumulative | Non-ecotoxic | 0.013 |

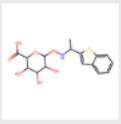
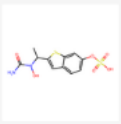
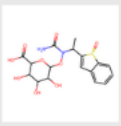
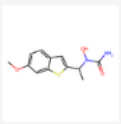
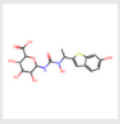
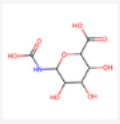
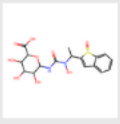
| | | | | |
|---|---|---|---|---|
|  | CC(NOC1OC(C(=O)O)C(O)C(O)C1O)c1cc2ccccc2s1 | Non-bioaccumulative | Non-ecotoxic | 0.0113 |
|  | CC(c1cc2ccc(OS(=O)(=O)O)cc2s1)N(O)C(N)=O | Non-bioaccumulative | Non-ecotoxic | 0.0073 |
|  | CC(C1=Cc2ccccc2S1=O)N(OC1OC(C(=O)O)C(O)C(O)C1O)C(N)=O | Non-bioaccumulative | Non-ecotoxic | 0.0053 |
|  | COc1ccc2cc(C(C)N(O)C(N)=O)sc2c1 | Non-bioaccumulative | Non-ecotoxic | 0.0033 |
|  | CC(c1cc2ccc(O)cc2s1)N(O)C(=O)NC1OC(C(=O)O)C(O)C(O)C1O | Non-bioaccumulative | Non-ecotoxic | 7.93e-04 |
|  | O=C(O)NC1OC(C(=O)O)C(O)C(O)C1O | Non-bioaccumulative | Non-ecotoxic | 6.89e-04 |
|  | CC(C1=Cc2ccccc2S1=O)N(O)C(=O)NC1OC(C(=O)O)C(O)C(O)C1O | Non-bioaccumulative | Non-ecotoxic | 3.25e-04 |

*Figure 23: The 15 main predicted metabolites of zileuton, along with their predicted bioconcentration and acute ecotoxicity classifications.*

The objectives of Task 6.2, in conjunction with Task 3.6, centered on developing *in silico* tools to support the design of environmentally friendly Active Pharmaceutical Ingredients (APIs). These tools provide green toxicology metrics, thereby enhancing chemical design strategies used by medicinal chemists and supporting computational approaches such as virtual screening and binding free energy calculations. Together, they contribute to a more sustainable and comprehensive framework for drug discovery. Within this effort, we developed state-of-the-art ML classification models to predict the bioaccumulation potential of diverse organic chemicals in fish, using experimental logBioconcentration Factor (logBCF) values collected through an extensive literature review. A thorough benchmarking study was conducted, comparing our models to publicly available tools on both our curated dataset and an independent external dataset. Results highlighted superior robustness and stability, with ROC AUC values exceeding 90%. Importantly, our model also provided interpretable insights into the physicochemical properties and structural features driving bioconcentration. Building on the accuracy and interpretability of the classification model, we further developed a

regression model that leveraged physicochemical properties as training features, showing strong promise for future applications. Recognizing the importance of assessing drug metabolites, we also investigated their potential ecotoxicity. To this end, we conducted a comprehensive review of metabolite prediction software and assembled a dataset of 221 drugs and their metabolites from the CHEMBL database. Among the algorithms tested, SyGMA emerged as the most effective, correctly predicting 52% of drug metabolites, though improvements remain necessary to achieve higher accuracy. Finally, we integrated our top-performing ML models for bioconcentration, ecotoxicity, and metabolite prediction into a fully automated pipeline. This advanced cheminformatics-based framework enables predictive assessment of the environmental impact of pharmaceuticals. With further optimization, it holds strong potential as a novel screening methodology in early-stage drug discovery and could ultimately serve as a regulatory checkpoint within the pharmaceutical industry. An overview of the G.AI.A platform is shown in Figure 24. The complete *in silico* workflow, encompassing model and platform development, is reported in a scientific article (DOI: 10.1021/acs.jcim.5c02286) accepted in the *Journal of Chemical Information and Modeling*.

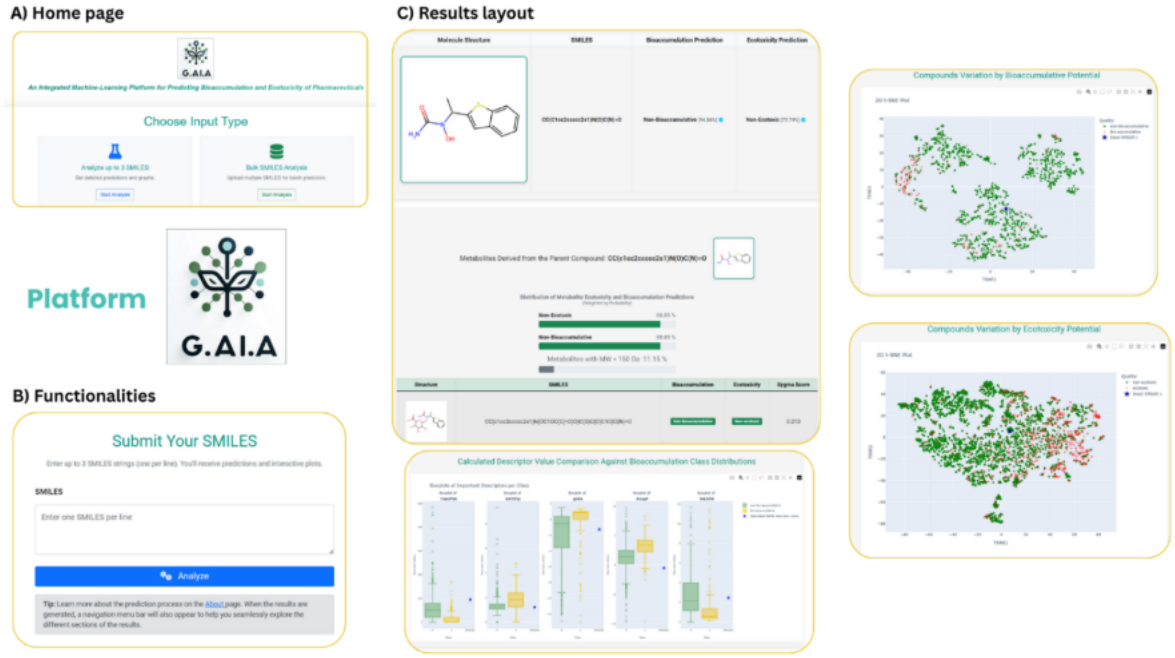



*Figure 24: Screenshots from the G.AI.A platform.*

# 4 In Vivo & In Vitro Ecotoxicity Studies

The increasing release of pharmaceutical residues into aquatic environments represents a growing environmental challenge, requiring integrated and preventive assessment strategies. Pharmaceuticals are designed to be biologically active at low concentrations, and their continuous input into marine ecosystems may lead to chronic, sublethal effects on non-target organisms, potentially altering immune competence, cellular homeostasis, and ecosystem functioning (148). Marine invertebrates are particularly suitable model organisms for investigating these effects, as they are widely distributed, ecologically relevant, and rely primarily on innate immune mechanisms that are highly sensitive to chemical stressors.

On this basis, the main objective of this task is to investigate the immunotoxic and genotoxic effects of selected pharmaceutical compounds using marine invertebrate models, with the aim of identifying early-warning biomarkers of adverse biological responses. By characterizing compound-specific effects and distinguishing them from background environmental stress, the project seeks to improve the understanding of the interactions between pharmaceutical contamination and biological systems.

In vivo and in vitro ecotoxicity studies provide experimental evidence to support and validate the modelling and assessment activities carried out within Task 6.3. These studies assess the effects of selected pharmaceutical compounds and environmental seawater samples (see D7.1) on representative marine organisms, generating biological data essential for confirming toxicity predictions and strengthening the overall environmental risk assessment framework.

## 4.1 Selection of Model Marine Invertebrates

Marine invertebrates constitute a valuable biological model for immuno-toxicological investigations due to their ecological relevance, the predictability of their physiological responses to environmental stressors, and the high sensitivity of their innate immune system (149). Their continuous exposure to the surrounding marine environment enables a direct linkage between waterborne contaminants - such as pharmaceutical residues - and quantifiable biological outcomes. Moreover, these organisms possess conserved immune and stress-response pathways that facilitate the identification of early immunological alterations, which can be used as robust indicators of environmental disturbance. Consequently, marine invertebrates provide an effective framework for evaluating the ecological risks associated with emerging pollutants, including pharmaceutical residues in marine ecosystems.

### *4.1.1 Mytilus galloprovincialis*

*Mytilus galloprovincialis* is a filter-feeding bivalve capable of processing substantial volumes of seawater, resulting in the efficient bioaccumulation of pharmaceuticals and other micropollutants even at low environmental concentrations. This characteristic makes the species particularly relevant for assessing chronic, low-dose exposure scenarios. Its hemocytes, which play a key role in the organism's innate immune system, display quantifiable changes in phagocytic activity, DNA integrity and lysosomal stability following exposure to contaminants. The availability of these well-established immunological and cellular biomarkers, together with the species' wide geographic distribution and ecological relevance, positions *M. galloprovincialis* as a robust and widely accepted model for evaluating immunotoxic effects associated with marine pollution.

### 4.1.2 *Paracentrotus lividus*

*Paracentrotus lividus* represents a complementary and highly informative model for immunotoxicological assessment, owing to the pronounced sensitivity of its coelomocytes to chemical stressors. These immune cells display rapid and well-defined responses – such as morphological changes, variations in lysosomal membrane stability, and alterations in phagocytic activity – following exposure to pharmaceuticals and other environmental contaminants. The species' developmental and physiological features are extensively documented, allowing for accurate interpretation of immunological and cellular endpoints. Moreover, its ecological relevance within Mediterranean benthic communities further strengthens its role as a biological indicator, contributing to comprehensive evaluations of the impacts of pharmaceutical pollution on marine ecosystem health.

## 4.2 Experimental Design and Exposure Protocols

The experimental activities carried out were designed to generate a comprehensive evaluation of the immunotoxic and physiological effects induced by exposure to a list of pharmaceuticals in the selected marine invertebrates, mussels *M. galloprovincialis* and sea urchins *P. lividus*. Seven different compounds were selected as frequently detected in coastal environments and known for their potential to interfere with cellular homeostasis:

- Carbamazepine (anti-convulsant)
- Clopidogrel COOH (antiplatelet)
- Diclofenac (anti-inflammatory)
- Hydrochlorothiazide (anti-hypertensive)
- Ibuprofen (anti-inflammatory)
- Metformin hydrochloride (anti hyperglycemic)
- Metoprolol tartrate (anti-hypertensive)

To what concern the experimental procedure on mussels, controlled laboratory assays were implemented using both Neapolitan/local and Greek populations of *M. galloprovincialis*, allowing comparisons between organisms with different environmental histories. Greek mussels collected in the surrounding of WWTP of Psyttalia island should represent an example of *life-long exposure conditions*, as they could have been continuously subjected to a complex mixture of pharmaceutical compounds and other anthropogenic stressors. In contrast, control mussels collected from the Gulf of Naples are assumed to have experienced comparatively lower levels of anthropogenic pressure. The comparison between these two populations provides a useful framework to evaluate how long-term environmental contamination may influence physiological status, stress responses, and immune competence in mussels.

All animals were held in holding tanks for 24/48 h to allow the recovery from sampling/transport-related stress before being used in the exposure experiments.

Mussels (3/treatment) were directly exposed (*in vivo* experiments) to predefined concentrations of pharmaceuticals namely 10 ng/L (environmental realistic/expected concentration in marine environment) and 1 mg/L (therapeutic concentration, able to sort an effect in marine organisms) for a short, yet biologically relevant, 2-hour interval to simulate acute contact with contaminants. Control experiments (animals exposed to Neapolitan/local seawater) were run in parallel. After that, haemocytes (immune cells) were collected in physiological artificial seawater and subjected to selected immunotoxicity assays, able to assess stressful level, immunological ability and DNA integrity. This integrated approach provides a detailed characterization of the cellular pathways possible disrupted by pharmaceutical contaminants

and supports the identification of early warning signals of environmental impact in marine organisms.

The experimental protocol for the sea urchins *P. lividus* involved collecting coelomocytes in physiological artificial seawater, followed by their exposure to 1 mg/L of the selected pharmaceuticals for 30 minutes or seawater as control (*in vitro*/*ex vivo* experiments). The chosen concentration corresponds to the lowest dose capable of eliciting measurable biological effects in sea urchins, as determined through preliminary dose–response analyses. Following exposure, coelomocytes were subsequently employed to carry out the immunotoxicity assays for assessing stress level of the cells and their overall immunological competence.

## 4.3 Assessment of Physiological and Immunological Parameters

The selected model organisms, *M. galloprovincialis* and *P. lividus*, together provide a comprehensive and complementary framework for assessing the effects of emerging contaminants—particularly pharmaceuticals—on marine ecosystem health. In both species, key biomarkers such as cellular viability, lysosomal membrane stability, and immune cell functional activity (i.e., phagocytosis) act as sensitive indicators for detecting early biological disturbances. Hemocytes in mussels and coelomocytes in sea urchins offer valuable insight into innate immune performance, exhibiting rapid and quantifiable modifications in phagocytic capacity and stress-related pathways following exposure to xenobiotics. The combined analysis of these physiological and immunological responses enables a detailed evaluation of organismal health status and contaminant-induced dysfunction, ultimately supporting more robust and accurate risk assessments of pharmaceutical pollution in coastal environments.

### 4.3.1 Functional Cellular Assays (LMS, Phagocytosis, Genotoxicity)

To investigate the immunotoxic effects of pharmaceuticals, *in vivo* experiments were performed on *M. galloprovincialis* mussels from both local and Greek populations, while *ex vivo* analyses were conducted on coelomocytes isolated from local specimens of *P. lividus*. The aim of these complementary approaches was to determine how defined concentrations of selected pharmaceuticals influence the immune cell performance of the two species. To achieve this, three standardized assays, the Lysosomal Membrane Stability (LMS) assay, the phagocytic assay and the Micronucleus Cytome (MCyt) assay were applied to assess key aspects respectively of cellular stress, immune competence, and genomic integrity.

The LMS assay served as a sensitive early marker of cellular injury, detecting alterations in lysosomal integrity associated with toxic exposure. The lysosomal alterations observed through the LMS assay can be categorized into six distinct scores (Figure 25) that reflect progressive stages of cellular stress and lysosomal dysfunction. **Score 0** indicates no effect, with dye fully retained in intact lysosomes. **Score 1** shows lysosomal enlargement without leakage, while **Score 2** marks the first signs of membrane destabilization through dye release into the cytosol. **Score 3** involves both enlargement and leakage, indicating more advanced impairment. **Score 4** describes enlarged, colorless lysosomes unable to retain the dye. **Score 5** represents severe damage, with widespread cytosolic dye accumulation and cell rounding, signaling pronounced cytotoxicity (150).

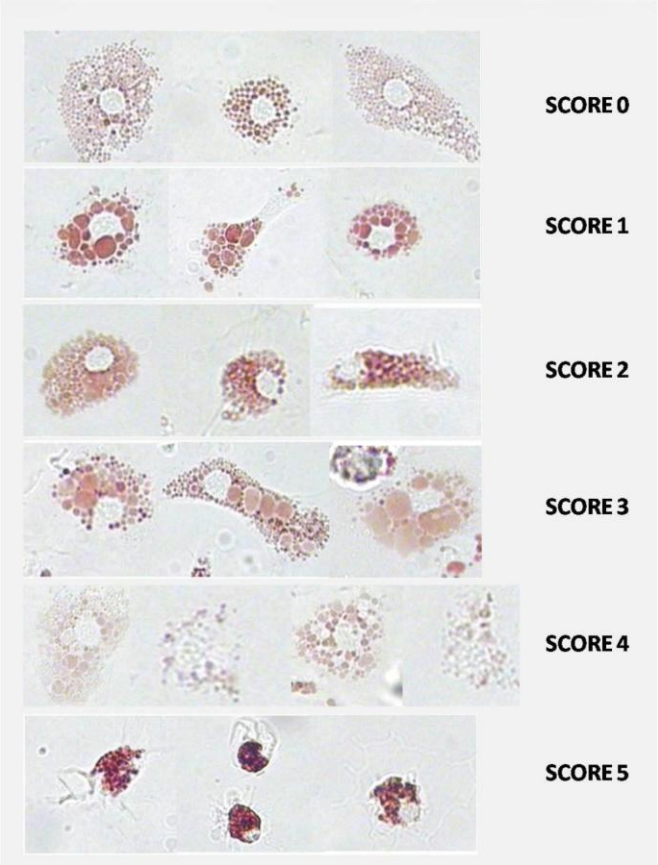



*Figure 25: Illustrations of M. galloprovincialis granulocytes exhibiting different lysosomal anomalies and associated scores. (from ICES TECHNIQUES IN MARINE ENVIRONMENTAL SCIENCES, 2015 – ref.150)*

The phagocytic assay measures the functional activity of innate immune cells by evaluating their capacity to recognize, engulf, and internalize foreign particles. In this method, isolated immune cells are exposed to fluorescent heat-inactivated bacteria for mussels or fluorescent latex beads for sea urchins, under controlled laboratory conditions. The percentage of phagocytic cells (PR – Phagocytic Rate) and, when possible, the number of particles ingested per cell (PI – Phagocytic Index) were determined to understand the efficiency of key immune processes.

The third assay performed only on mussels, the MCyt assay assesses genomic stability by detecting micronuclei within exposed immune cells. By evaluating the frequency of these structures, along with a limited set of related nuclear abnormalities, the assay provides a clear indication of chromosomal damage and genotoxic stress.

## 4.4 *In Vivo* and *In vitro/Ex Vivo* Outcomes

Mussels *M. galloprovincialis* from the Gulf of Naples (control/local animals) and from Greece, Saronikos Gulf, collected in different periods covering seasonality (see D7.1) were used to perform *in vivo* experiments with local seawater (from Gulf of Naples) as control and with the selected pharmaceuticals such as Carbamazepine, Clopidogrel COOH, Diclofenac, Hydrochlorothiazide, Ibuprofen, Metformin and Metoprolol.

*In vitro/Ex vivo* studies were performed on coelomocytes from local sea urchins *P. lividus* exposed to local seawater and to the selected pharmaceuticals. After dose-response trials, the

selected concentration of drugs was 1mg/L, a "therapeutic dose", the only dose able to induce an immunotoxic effect.

The main purpose of these studies is to investigate the effect of known concentrations of pharmaceuticals on the immune performance of the selected marine invertebrates and to compare the stressful level and immunological capabilities of Greek mussels with Neapolitan's.

To this aim, local mussels were exposed *in vivo* to 1mg/L of drugs for 2h and then immunotoxicity assays were performed on live haemocytes. Control experiments (animals exposed to filtered local seawater) were run in parallel. The LMS assay demonstrated that almost all the selected pharmaceuticals induced a certain level of stress (Figure 26) particularly evident after treatment with Ibuprofen, Metformin and Metoprolol as evidenced by the abundance of cells showing lysosomal enlargement and leakage phenotype.

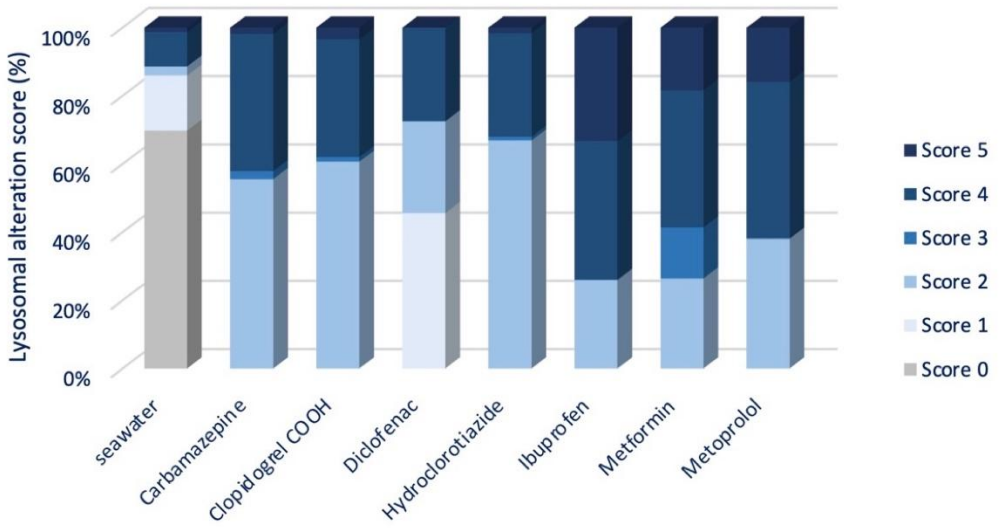



*Figure 26: Lysosomal membrane stability measured on Neapolitan mussels after pharmaceutical exposure. Lysosomal alterations are classified into five scores representing increasing levels of cellular stress and lysosomal dysfunction: Score 0, no effect with dye retained in intact lysosomes; Score 1, lysosomal enlargement without leakage; Score 2, initial membrane destabilization with dye release into the cytosol; Score 3, enlargement with leakage; Score 4, enlarged lysosomes unable to retain the dye; Score 5, severe damage with widespread cytosolic dye accumulation and cell rounding.*

The immunological ability of mussels' haemocytes (Figure 27) measured in terms of phagocytic rate (PR) was decreased only after treatment with Carbamazepine, Metformin and Metoprolol thus reflecting an impaired cellular responsiveness and a potential suppression of innate immune function. On the contrary, Diclofenac induced in increased PR indicating a physiological stress and a potential disruption of immunological homeostasis. The phagocytic index (PI), a measure of effective phagocytosed bacteria, was not affected by pharmaceutical exposure meaning that basic phagocytic function is retained and that the drug does not interfere with the cellular mechanisms required to engulf bacteria.

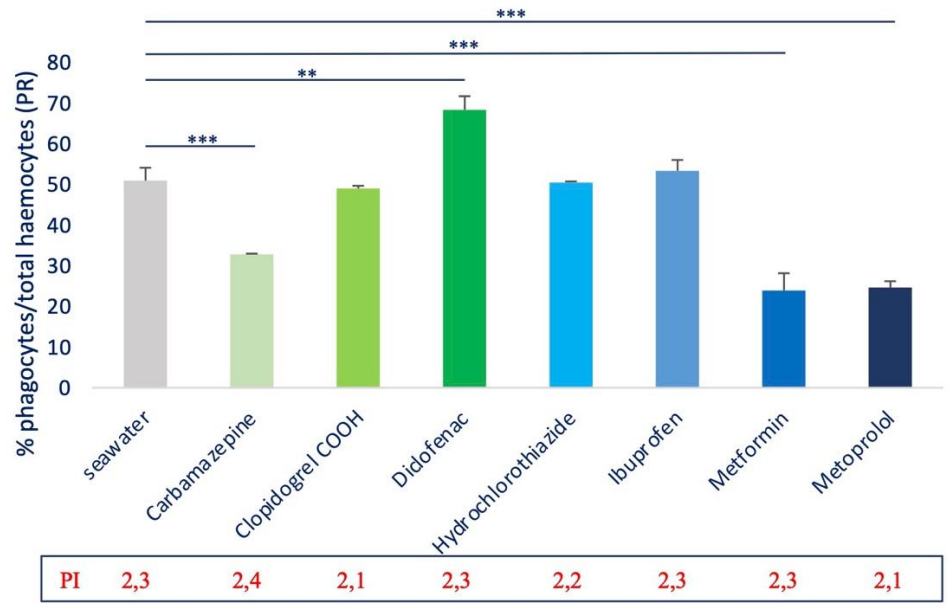



*Figure 27: Effect of pharmaceuticals on phagocytic capacity of Neapolitan mussels. Statistical significance was determined using the Mann–Whitney U test; $p < 0.05$.*

The genotoxic effect (Figure 28) was particularly evident after exposure to Clopidogrel COOH, Metformin and Metoprolol as for the abundance of cells with micronuclei. The effect of Diclofenac was even worst since the huge number of cells with nuclear abnormalities.

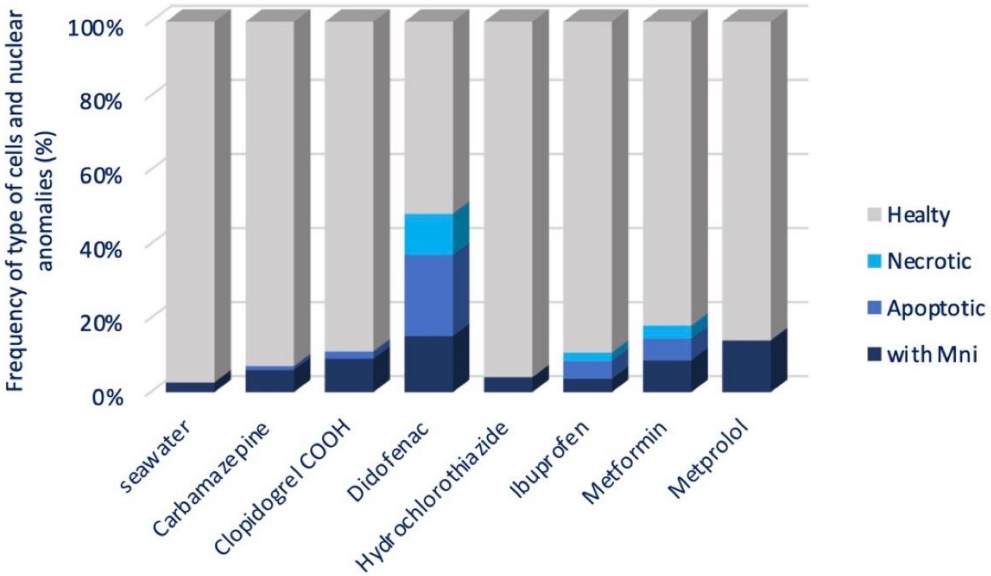



*Figure 28: Genotoxic effect of pharmaceuticals on Neapolitan mussels.*

The same experimental procedure was then performed on Greek mussels in order to verify their general status and to assess the effect of pharmaceutical exposure. The results of immonocitotoxic assays revealed that Greek mussels were *di per se* more stressed than

Neapolitan's as evidenced by the lysosomal instability degree measured after exposure to local seawater (Figure 26 and Figure 29). Their status became even worst after exposure to drugs, above all Hydrochlorothiazide, Metformin and Metoprolol.

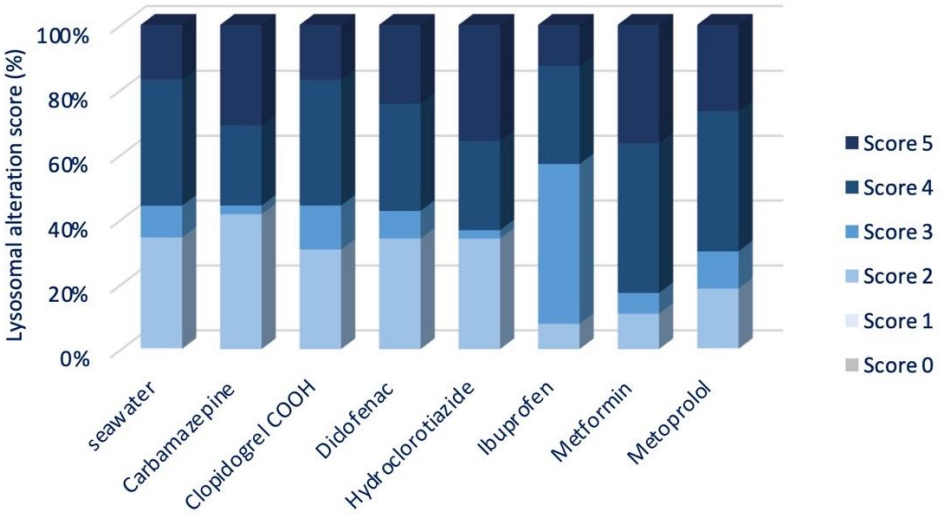



*Figure 29: Lysosomal membrane stability measured on Greek mussels after pharmaceutical exposure. Classification of Lysosomal alteration score was described in Figure 26 legend.*

Regarding the immunological abilities (Figure 30) and genotoxicity level (Figure 31), no significative differences were observed in Greek mussels either after exposure to local seawater or after exposure to drugs compared to results from Neapolitan mussels (Figure 27 and Figure 28).

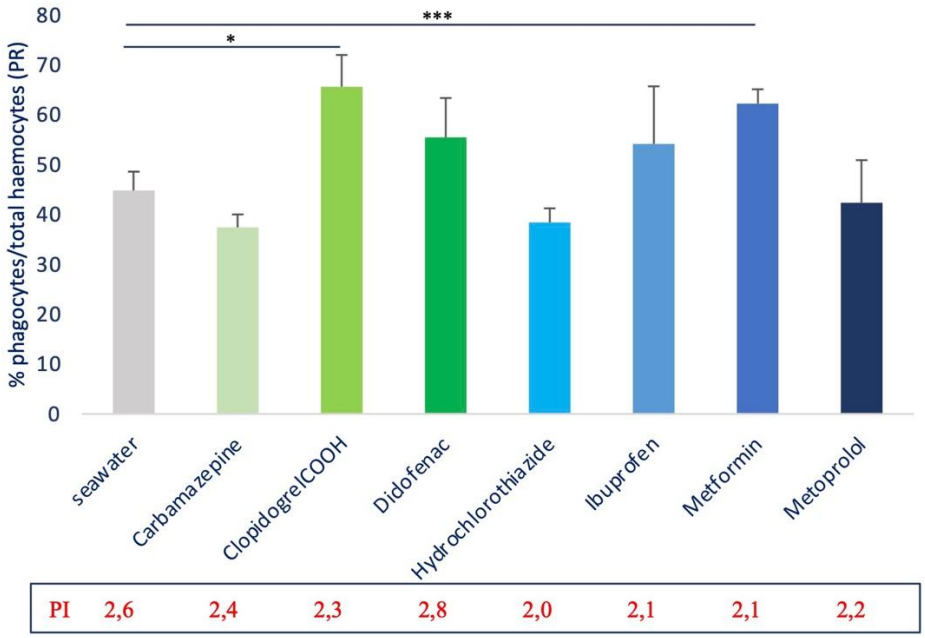



*Figure 30: Effect of pharmaceuticals on phagocytic capacity of Greek mussels. Statistical significance was determined using the Mann–Whitney U test; $p < 0.05$.*

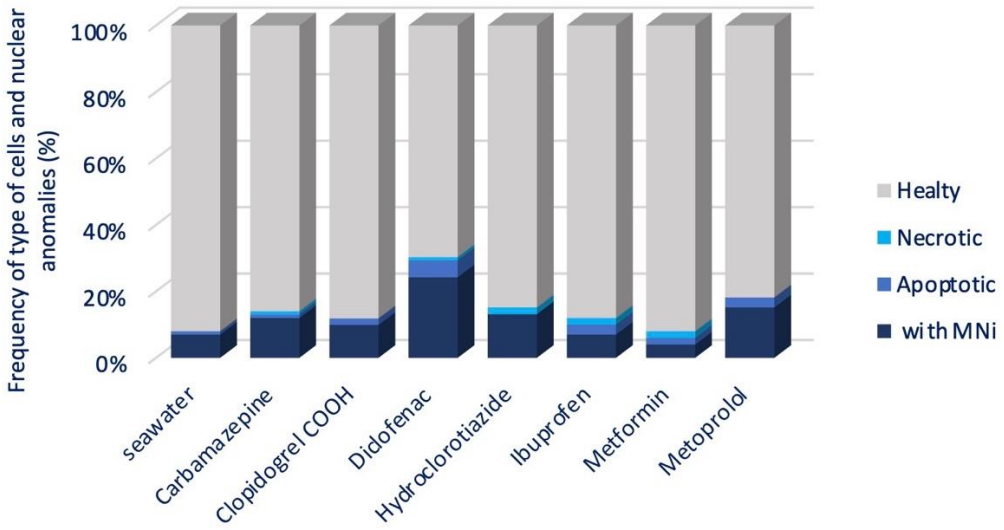



*Figure 31: Genotoxic effect of pharmaceuticals on Greek mussels.*

To extend this analysis to another ecologically relevant invertebrate species, coelomocytes from Neapolitan Sea urchins *P. lividus,* were exposed to 1mg/L of selected drugs for 30 min and immunotoxicity studies were performed immediately afterwards. Results from the LMS assay (Figure 32) demonstrated that the lysosomal integrity was compromised above all after exposure to Clopidogrel COOH, Diclofenac, Ibuprofen and Metoprolol compared to control experiments.

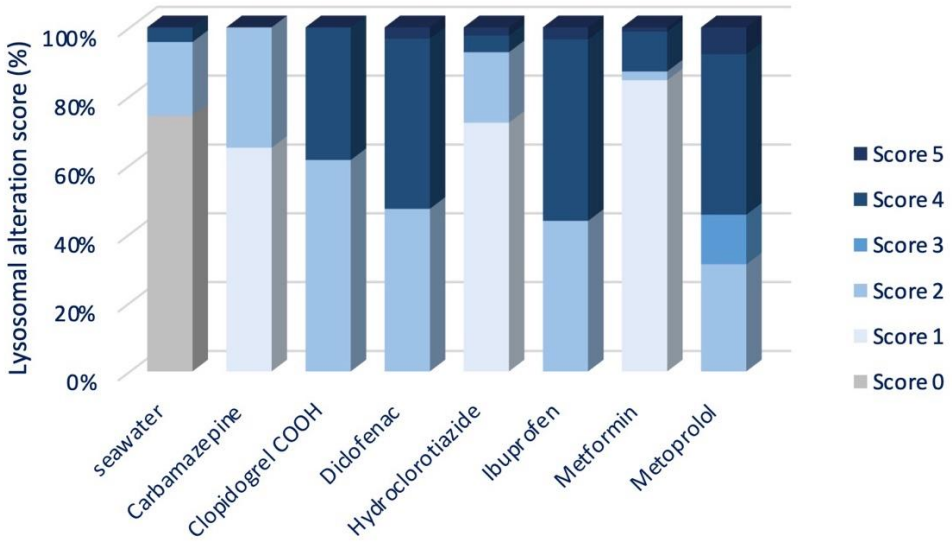



*Figure 32: Lysosomal membrane stability measured on Neapolitan Sea urchins after pharmaceutical exposure. Classification of Lysosomal alteration score was described in Figure 21 legend.*

As concern the immunological ability of coelomocyte, PR (Figure 33) was increased above all after exposure with Clopidogrel COOH, Hydrochlorothiazide and Metoprolol confirming a physiological stress of exposed coelomocytes. The PI value could not be calculated due to

technical issues related to the excessively small size of these cells. Moreover, for the same reason, the genotoxic effect on Neapolitan Sea Urchin could not be calculated.

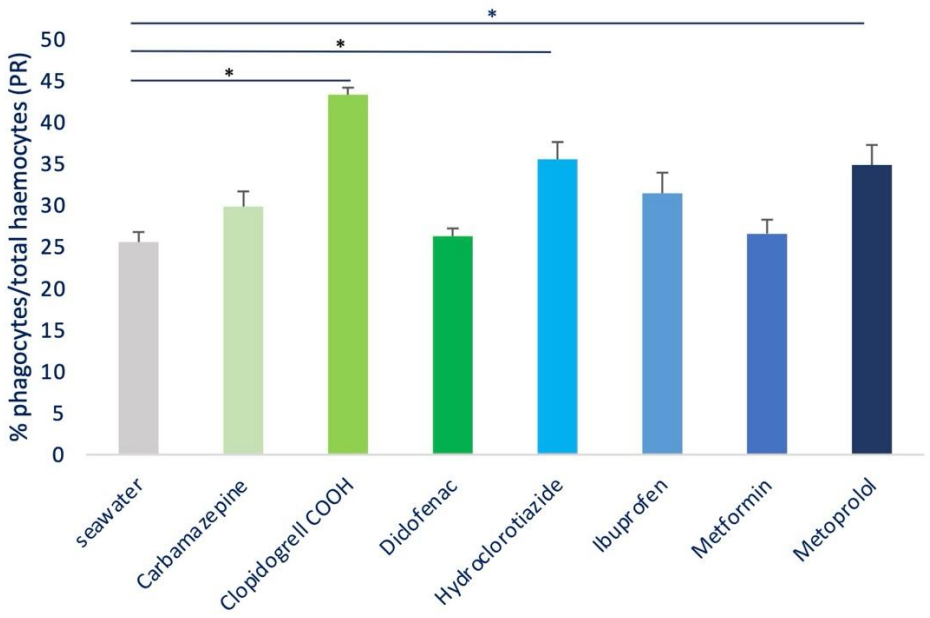



*Figure 33: Effect of pharmaceuticals on phagocytic capacity of Neapolitan Sea urchins. Statistical significance was determined using the Mann–Whitney U test; $p < 0.05$.*

### 4.4.1 Conclusions and considerations on Ecotoxicological studies

Overall, these results demonstrate that pharmaceutical exposure induces comparable immunotoxic and genotoxic effects in *M. galloprovincialis* from both the Gulf of Naples and Saronikos Gulf, while highlighting clear differences in their basal physiological status. Neapolitan mussels, considered as reference organisms, exhibited drug-specific stress responses, including lysosomal destabilization, altered phagocytic activity, and increased genotoxic damage, particularly after exposure to Metformin and Metoprolol.

In contrast, Greek mussels, collected from an area potentially influenced by pharmaceutical contamination, were *per se* more stressed, as indicated by higher basal lysosomal instability under control conditions, and this condition was further exacerbated by drug exposure. However, chemical analyses of seawater revealed no substantial differences in the relative concentrations of pharmaceuticals between Greek and Neapolitan sites. Therefore, the higher basal stress observed in Greek mussels cannot be directly ascribed to a greater presence of pharmaceuticals in their environment, suggesting the contribution of other site-related stressors, “long-term environmental pressures” and/or “transportation” stress.

Consistently, sea urchin (*P. lividus*) coelomocytes showed rapid lysosomal destabilization and increased phagocytic activity following short-term pharmaceutical exposure, confirming that these compounds elicit conserved stress-related cellular responses across marine invertebrates.

Overall, for the ecotoxicological studies only short-term experiments have been conducted because we consider that the wild animals collected from Greece have already experienced long-term exposure to pharmaceuticals in their environment. Their basal immunological reactivity and stress levels (without direct exposure to pharmaceuticals in vitro) when

compared to animals from another environment such as those collected in the Gulf of Naples, provide direct evidence of the immuno-ecotoxicological effects of the seawater contaminants. Moreover, the aspects investigated upon short-term exposure, such as stress, immune response, and genotoxicity, can contribute to risk assessment, but only to a limited and carefully interpreted extent.

Stress in terms of lysosomal membrane integrity, immune response in terms of phagocytosis, and genotoxicity in terms of micronuclei formation are early-warning indicators. They are sensitive and can reveal that an organism is responding to a contaminant before overt adverse effects occur. For this reason, they are useful for identifying potential hazards and for supporting mechanistic understanding within an Adverse Outcome Pathway.

In any case, it is necessary to clarify that the effects measured after short exposures do not necessarily translate into long-term adverse outcomes at the organism or population level. Stress and immune parameters are often transient, adaptive, and non-specific, and may recover once exposure ceases. Similarly, genotoxic markers, such as micronuclei, indicate DNA damage, but their ecological relevance depends on persistence, repair capacity, and whether they ultimately lead to impaired fitness, reproduction, or survival.

Therefore, while these short-term markers are valuable as screening tools and mechanistic endpoints, they are insufficient on their own for robust ecological risk assessment.

In summary, short-exposure responses can inform risk assessment by indicating early biological perturbations, but they should be interpreted as supportive evidence rather than definitive proof of ecological risk.

## 4.5 *In silico* predictions of fish toxicity for compounds evaluated in invertebrates

As illustrated in Table 13, the compounds experimentally evaluated in invertebrates were also assessed *in silico* using our G.AI.A platform to predict their potential ecotoxicity and bioaccumulation in fish, both for the parent compounds and their metabolites. Table 13 presents the predicted LC50 values for each compound along with the corresponding confidence scores for these predictions. Metformin was not included in the predictions, as the G.AI.A platform is trained exclusively on compounds with molecular weights exceeding 150 Da for bioaccumulation and ecotoxicity assessment. Additionally, a summary of the most probable metabolites and their predicted ecotoxicity potential is provided. Corresponding predictions for the BCF of both parent compounds and their metabolites are also included.

*Table 13. Ecotoxicity and bioaccumulation predictions in fish for compounds already evaluated in invertebrates.*

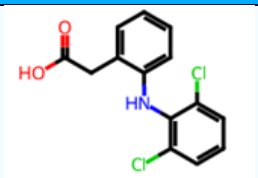
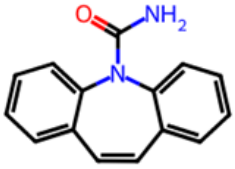
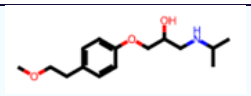
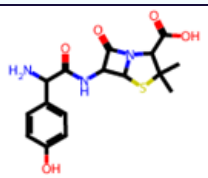
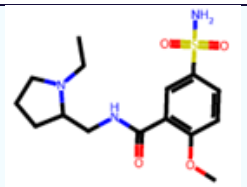
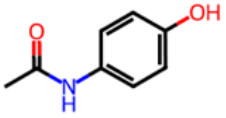
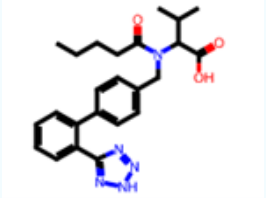
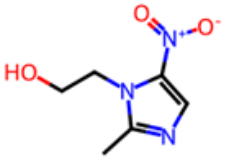
| Compound | LC50 prediction (confidence score) | Metabolite Ecotoxicity Summary (Ecotoxicity) | BCF prediction (confidence score) | Metabolite Bioaccumulation Summary (Bioconcentration) |
|---|---|---|---|---|
| **Metoprolol** | Non-ecotoxic (65.67%) | 94.38 % Non-ecotoxic | non-bioaccumulative (95.50%) | 94.38 % non-bioaccumulative |
| **Diclofenac** | Non-ecotoxic (89.21%) | 51.40 % Non-ecotoxic | Bioaccumulative (61.47%) | 68.98 % non-bioaccumulative |
| **Ibuprofen** | Non-ecotoxic (100.00 %) | 92.63 % Non-ecotoxic | non-bioaccumulative (90.31%) | 96.61 % |

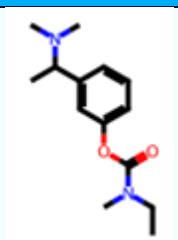
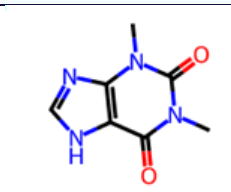
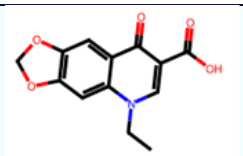
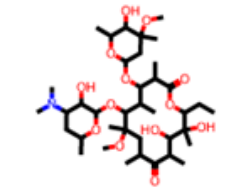
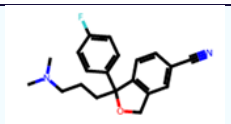
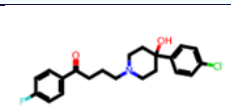
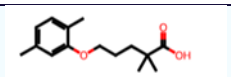
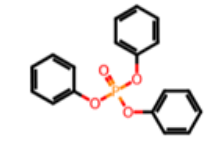
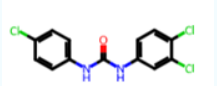
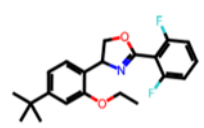
| Compound | LC50 prediction (confidence score) | Metabolite Ecotoxicity Summary (Ecotoxicity) | BCF prediction (confidence score) | Metabolite Bioaccumulation Summary (Bioconcentration) |
|---|---|---|---|---|
| | | | | non-bioaccumulative |
| **Metformin** | - | - | - | - |
| **Hydrochlorothiazide** | Non-ecotoxic (84.59%) | 100.00 % Non-ecotoxic | non-bioaccumulative (98.60%) | 100.00 % non-bioaccumulative |
| **Carbamazepine** | Non-ecotoxic (100.00 %) | 76.18 % Non-ecotoxic | non-bioaccumulative (84.92%) | 88.48 % non-bioaccumulative |
| **Clopidogrel-COOH** | Non-ecotoxic (51.54%) | 100.00 % Non-ecotoxic | non-bioaccumulative (92.25%) | 100.00 % non-bioaccumulative |

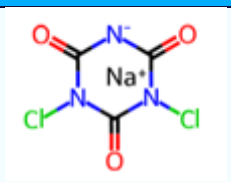
## 4.6 In vivo ecotoxicity studies for model evaluation in fish

Besides externally validating our ExtraTrees model using two subsets from the ADORE benchmarking dataset—t_F2F (1,118 compounds across various fish species) and s-F2F-2 (497 compounds specific to *Pimephales promelas*)—we aimed to further assess model performance on an external dataset composed of organic compounds tested for LC50 on the same fish species under identical experimental conditions. While ADORE is a comprehensive and well-documented resource for acute aquatic toxicity, it includes data from varying experimental setups, such as differences in laboratories, fish species, and specific testing parameters. To address this limitation, we assembled and purchased an external dataset of 19 organic compounds for testing under consistent conditions. Selection criteria focused on compounds with widespread pharmaceutical use and structural diversity. However, based on minimum LC50 values reported in the literature (Table 14 14), all selected compounds were classified as non-ecotoxic. Only three – diclofenac (5.22 mg/L), citalopram (3.18 mg/L), and haloperidol (2.40 mg/L) – had LC50 values approaching the 1 mg/L threshold, indicating borderline toxicity. To ensure representation from both ecotoxicity classes, we also included five compounds from our training dataset with LC50 values below 1 mg/L, classifying them as ecotoxic (X1-X5). Given that CNR is limited to validating the model using data from fish aquatic species, we opted to first carry out direct testing on a single fish species to ensure accurate model evaluation. To support this effort, we commissioned an academic subcontractor to perform the necessary experiments using zebrafish larvae up to 96 hours post-fertilization. The subcontractor selected is Dr. Dimitris Beis, Associate Professor of Biological Chemistry at the Faculty of Medicine, University of Ioannina. His laboratory specializes in the use of zebrafish models to study human diseases and screen for novel bioactive compounds. Dr. Beis was selected following a thorough market assessment to ensure the best value for money.

*Table 14.The 19 chemical compounds assembled and purchased for the external validation set are listed in the table, which includes each compound's CAS number, 2D structure, primary clinical or industrial application, the minimum LC50 value reported in the literature.*

| ID (Chemicals name) | CAS Number | 2D Structure | Clinical/ Industrial Use | Literature's Minimum Reported LC 50 (mg/L) | Ecotoxicity Class |
|---|---|---|---|---|---|
| **diclofenac** | 15307-86-5 |  | NSAID (nonsteroidal anti-inflammatory drugs) | 5.22 (151) | non-ecotoxic |
| **carbamazepine** | 298-46-4 |  | anticonvulsant | 19.90 (152) | non-ecotoxic |
| **metoprolol** | 51384-51-1 |  | Antihypertension (beta-blocker) | 130 (153) | non-ecotoxic |
| **amoxicillin** | 26787-78-0 |  | Antibiotic (penicillin derivative) | 1000 (154) | non-ecotoxic |
| **sulpiride** | 15676-16-1 |  | antipsychotic | 72.12 (155) | non-ecotoxic |
| **paracetamol** | 103-90-2 |  | NSAID (nonsteroidal anti-inflammatory drugs) | 814 (156) | non-ecotoxic |
| **valsartan** | 137862-53-4 |  | antihypertensive (angiotensin-receptor I blocker) | 100 (157) | non-ecotoxic |
| **metronidazole** | 443-48-1 |  | antibiotic (nitroimidazole ) | 100 (158) | non-ecotoxic |

| ID (Chemicals name) | CAS Number | 2D Structure | Clinical/ Industrial Use | Literature's Minimum Reported LC 50 (mg/L) | Ecotoxicity Class |
|---|---|---|---|---|---|
| **rivastigmine** | 123441-03-2 |  | anti-dementia drug (cholinesterase inhibitor) | 31.80 (159) | non-ecotoxic |
| **Theophylline** | 58-55-9 |  | bronchodilator (xanthine) | 100 (160) | non-ecotoxic |
| **Oxolinic acid** | 14698-29-4 |  | antibiotic (quinolone) | - | - |
| **Clarithromycin** | 81103-11-9 |  | antibiotic (macrolide) | 100 (161) | non-ecotoxic |
| **citalopram** | 59729-33-8 |  | antidepressant (Selective Serotonin Reuptake Inhibitor) | 3.18 (162) | non-ecotoxic |
| **haloperidol** | 52-86-8 |  | antipsychotic | 2.40 (163) | non-ecotoxic |
| **X2 (Gemfibrozil)** | 25812-30-0 |  | lipid regulator | 0.85 (164) | ecotoxic |
| **X3 (Triphenyl Ester Phosphate)** | 115-86-6 |  | plasticizer, fire-retardant | 0.30 (165) | ecotoxic |
| **X5 (Triclocarban)** | 101-20-2 |  | antiseptic in cleaning products | 0.04 (166) | ecotoxic |
| **X1 (Etoxazole)** | 153233-91-1 |  | pesticide | 0.16 (167) | ecotoxic |

| ID (Chemicals name) | CAS Number | 2D Structure | Clinical/ Industrial Use | Literature's Minimum Reported LC 50 (mg/L) | Ecotoxicity Class |
|---|---|---|---|---|---|
| **X4 (Sodium dichlorocyanuric acid)** | 2893-78-9 |  | bleaching agent | 0.22 (167) | ecotoxic |

## 4.7 Experimental procedure and conditions

The nineteen compounds were dissolved in DMSO (Sigma D2438) to a concentration of 200mg/ml, except for X4 that was dissolved directly to $dH_2O$ at 200mg/ml. Stocks were kept at 4°C. They were thawed, vortexed and diluted to working concentrations in embryo water (60 µg/ml Ocean Salts), before adding the fertilized embryos.

Fertilized wild-type embryos (AB strain) were placed individually in 96-well plates and the compounds at different concentrations were added. The maximum concentration tested was 2mg/ml, as 1% DMSO is the maximum tolerated concentration for proper zebrafish embryo development. At least, 3 replicates with 6 embryos per concentration were used and 1% DMSO in embryo water was the positive control (no malformations observed). To verify the results and narrow the confidence limits, additional experiments were conducted at the estimated LC50 and at half that concentration (LC50/2), using 18 embryos per concentration for each compound. Embryos were monitored daily up to 96hours post fertilization and malformations, deaths or any other morphological observations were documented in a lab book. Dead embryos were considered when the heart stopped or no circulation was observed. Representative images (and videos) were taken using a Zeiss Stereoscope and an Axiocam camera (ZEN software).

The experimental protocols described in this study were carried out with zebrafish larvae up to 96 hours post-fertilization (hpf), and therefore these experiments are not considered animal experiments and do not fall under the protection guidelines of the directive 2010/63/EU revising directive 86/609/EEC on the protection of animals used for scientific purposes as adopted on 22nd September 2010.

LD/LC50 (Table 15) was calculated using an excel file available on: https://probitanalysis.wordpress.com/. Mekapogu, A.R. (2021) Finney's probit analysis spreadsheet calculator (Version 2021). Based on the experimental observations, all five toxic compounds (X1-X5) show LC50 <25mg/L, together with Diclofenac and Haloperidol that showed a pronounced heart defect in developing embryos (low heartbeat, heart edemas).

A limitation of the study is that the concentration available due to solubility in the DMSO stock solution and water might be overestimated [concentration calculations are made based on 100% solubility in DMSO (stock solutions) and water]. Clarithromycin, carbamazepine, haloperidol, X1, citalopram, metronidazole and theophylline showed precipitation in the stock solution of DMSO at 200mg/ml. They were thawed and vortexed before making working concentration embryo water, where only in clarithromycin, X1 remained a precipitate but the others showed good solubility in embryo water even at the highest concentration. Diclofenac, valsartan, and to a lesser extent sulpiride, along with compounds X2 and X3, although fully soluble in DMSO at 200 mg/mL, exhibited visible precipitation in embryo water at the highest tested concentration (2 mg/mL). For the lower concentrations, no visible precipitation was noticed. X4 apparently reacted with DMSO and gave a highly volatile product that could not

be tested. However, when stock solution was made in water, it was soluble up to 200mg/ml and showed high toxicity.

All things considered, a clear separation of the 5 putative toxic compounds (X1-X5) in this set up, along with haloperidol and diclofenac, can be observed using the 25mg/L as a threshold. Some toxic compounds are still above the 1mg/L threshold, but this could also be due to the bioavailability and the limited time-window of exposure the OECD236 protocol adheres to (up to 96 hours post fertilization), mainly due to bioethical concerns (embryos up to 5 days do not fall under the protection guidelines of the directive 2010/63/EU).

Other observed phenotypes (Figure 34) include truncated body axis (X1, X3), inhibition of melanin production (X2, X3, paracetamol, amoxicillin).

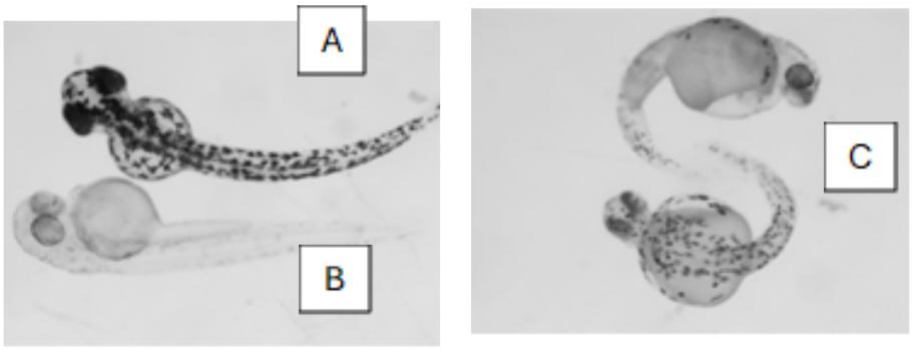



*Figure 34: A wild-type embryo at 72hours post fertilization (A), a paracetamol treated one showing inhibition of melanin synthesis (B) and an X3 treated one showing inhibition of melanin synthesis as well as truncated/shorter A-P axis.*

*Table 15. LC50 values for the 19-compound external validation dataset, determined using zebrafish larvae up to 96 hours post-fertilization. Each value is presented with its corresponding 95% fiducial confidence interval.*

| ID | CAS number | LC50 (mg/L) | 95%Fiducial Confidence lower Limits (mg/L) | 95%Fiducial Confidence upper Limits (mg/L) |
|---|---|---|---|---|
| **diclofenac** | 15307-86-5 | 1.5 | 0,7 | 3,1 |
| **carbamazepine** | 298-46-4 | 100.3 | 20,7 | 486,1 |
| **metoprolol** | 51384-51-1 | 459.7 | 189,7 | 1113,7 |
| **amoxicillin** | 26787-78-0 | 1617.3 | 1003,3 | 2606,9 |
| **sulpiride** | 15676-16-1 | 261.2 | 140,3 | 486,2 |
| **paracetamol** | 103-90-2 | 947.9 | 661,6 | 1358,2 |
| **valsartan** | 137862-53-4 | 465.6 | 383,2 | 565,8 |
| **metronidazole** | 443-48-1 | 5459.3 | 1552,5 | 19201,4 |
| **rivastigmine** | 123441-03-2 | 808.6 | 529,5 | 1234,8 |
| **theophylline** | 58-55-9 | 746.6 | 567,4 | 982,4 |
| **oxolinic acid** | 14698-29-4 | 151 | 85,9 | 265,5 |
| **clarithromycin** | 81103-11-9 | 648.7 | 461,5 | 911,6 |
| **citalopram** | 59729-33-8 | 1066.5 | 663,7 | 1713,3 |
| **haloperidol** | 52-86-8 | 7.1 | 3,2 | 15,7 |
| **X2** | 25812-30-0 | 23.5 | 9,2 | 59,6 |
| **X3** | 115-86-6 | 2.64 | 0,88 | 7,9 |

| ID | CAS number | LC50 (mg/L) | 95%Fiducial Confidence lower Limits (mg/L) | 95%Fiducial Confidence upper Limits (mg/L) |
|---|---|---|---|---|
| **X5** | 101-20-2 | 10.6 | 5,2 | 21,5 |
| **X1** | 153233-91-1 | 20.7 | 7,3 | 59,1 |
| **X4** | 2893-78-9 | 3.15 | 1,4 | 6,9 |

## 4.8 Comparison of experimental LC50 results with model predictions

To assess the predictive performance of our acute ecotoxicity classification model, experimental LC50 assays were conducted by Dr. Beis following OECD Guideline 236, using Danio rerio embryos. A total of 19 compounds were tested and compared against model predictions. The model classifies compounds as ecotoxic when LC50 < 1 mg/L, a threshold based on established regulatory standards.

Based on the experimental results, seven compounds (diclofenac, haloperidol, and five chemicals denoted as X1–X5) exhibited LC50 values below 25 mg/L. These compounds form a distinctive toxic subset under the experimental conditions, demonstrating clear adverse effects in zebrafish embryos. It should be noted again that due to solubility limitations in DMSO and water, the experimentally available concentrations may be overestimated, potentially affecting the bioavailability of tested compounds and, by extension, the accuracy of the LC50 values. Given these constraints, a pragmatic cutoff of 25mg/L was adopted for defining experimentally "toxic" compounds within this study. Among the seven toxic compounds identified experimentally, only Diclofenac was misclassified as non-ecotoxic (Table 16). However, additional evidence from the G.AI.A platform supports Diclofenac's ecotoxic potential. In the t-SNE plot (Figure 35.A), Diclofenac (represented by Input SMILES 1, star symbol) is embedded within the ecotoxicity cluster (red points), strongly suggesting alignment with known toxic compounds. Distribution boxplots (Figure 35.B) for key PaDEL descriptors (CrippenLogP, SpMax5_Bhm, VP-3, ATS3m, AATS5p) further support this. Diclofenac's values for these descriptors consistently fall within the upper quartile (Q3+) of the ecotoxic compound distributions, while overlapping with the non-ecotoxic range but remaining outside its upper quartile. The remaining five ecotoxic compounds (X1–X5) were previously reported to have LC50 values below 1 mg/L. The references for the corresponding LC50 experiments are provided in Table 14. These compounds were included in the experimental set to re-evaluate their LC50 values under consistent testing conditions. The results of the conducted assay confirmed their ecotoxic nature.

Conversely, regarding compounds identified as non-ecotoxic experimentally, two compounds—citalopram and rivastigmine—were predicted to be ecotoxic by the model, despite having LC50 values well above 25 mg/L. However, both predictions had low model confidence scores (54.79% for citalopram, 55.99% for rivastigmine), indicating classification uncertainty. This uncertainty is corroborated by insights from the G.AI.A platform. In the t-Distributed Stochastic Neighbor Embedding (t-SNE) plot (Figure 35.A), both compounds (Input SMILES 2 and 3, star symbols) are located near the interface between ecotoxic and non-ecotoxic clusters, highlighting their borderline nature. Descriptor-level analysis (Figure 35.B) reveals that both compounds' CrippenLogP values and the AATS5p descriptor for rivastigmine align more closely with the non-ecotoxic distribution, suggesting the model's classifications for these compounds may be false positives due to structural or physicochemical ambiguity. In terms of structural features identified by the platform, both compounds contain chemical groups that are more prevalent among ecotoxic compounds than in non-ecotoxic ones (Figure 36). This

structural similarity may have contributed to their misclassification by the model. Finally, according to literature data (Table 14), both compounds exhibit borderline LC50 values (citalopram at 3.18 mg/L, rivastigmine at 31.80 mg/L) which explain their ambiguous ecotoxicity profiles.

Overall, our LC50 model correctly predicted the ecotoxicity class for sixteen out of nineteen compounds evaluated experimentally by Professor Beis, resulting in an accuracy of 84.21%. This performance is comparable to the model's accuracy on the external validation set from ADORE's t_F2F dataset, where it achieved 90.44% accuracy across 1,118 compounds.

*Table 16. Comparison of experimental results with the predicted classes produced by the ecotoxicity model. The experimental class is based on the LC50 cutoff of 50mg/L, where compounds with LC50 < 50 mg/L are considered ecotoxic under the tested conditions.*

| ID | CAS number | Predicted Acute Ecotoxicity Label | Model Confidence (%) | Prof. Beis' LC50 (mg/L) | Experimental Acute Ecotoxicity Label |
|---|---|---|---|---|---|
| **diclofenac** | 15307-86-5 | non-ecotoxic | 89.21 | 1.5 | ecotoxic |
| **carbamazepine** | 298-46-4 | non-ecotoxic | 100 | 100.3 | non-ecotoxic |
| **metoprolol** | 51384-51-1 | non-ecotoxic | 65.67 | 459.7 | non-ecotoxic |
| **amoxicillin** | 26787-78-0 | non-ecotoxic | 96.06 | 1617.3 | non-ecotoxic |
| **sulpiride** | 15676-16-1 | non-ecotoxic | 85.45 | 261.2 | non-ecotoxic |
| **paracetamol** | 103-90-2 | non-ecotoxic | 100 | 947.9 | non-ecotoxic |
| **valsartan** | 137862-53-4 | non-ecotoxic | 63.87 | 465.6 | non-ecotoxic |
| **metronidazole** | 443-48-1 | non-ecotoxic | 100 | 5459.3 | non-ecotoxic |
| **rivastigmine** | 123441-03-2 | ecotoxic | 55.99 | 808.6 | non-ecotoxic |
| **Theophylline** | 58-55-9 | non-ecotoxic | 67.64 | 746.6 | non-ecotoxic |
| **Oxolinic acid** | 14698-29-4 | non-ecotoxic | 73.46 | 151 | non-ecotoxic |
| **Clarithromycin** | 81103-11-9 | non-ecotoxic | 100 | 648.7 | non-ecotoxic |
| **citalopram** | 59729-33-8 | ecotoxic | 54.79 | 1066.5 | non-ecotoxic |
| **haloperidol** | 52-86-8 | ecotoxic | 53.08 | 7.1 | ecotoxic |
| **X2** | 25812-30-0 | ecotoxic | 100 | 23.5 | ecotoxic |
| **X3** | 115-86-6 | ecotoxic | 100 | 2.64 | ecotoxic |
| **X5** | 101-20-2 | ecotoxic | 95.72 | 10.6 | ecotoxic |
| **X1** | 153233-91-1 | ecotoxic | 100 | 20.7 | ecotoxic |
| **X4** | 2893-78-9 | ecotoxic | 100 | 3.15 | ecotoxic |


*Figure 35: A) 2D t-SNE plot of training data used for acute ecotoxicity model with the potisions of diclofenac (Input SMILES 1), rivastigmine (Input SMILES 2) and citalopram (Input SMILES 3) indicated by map-coloured star symbols. B) Boxplots of the five most influential descriptors used by the acute ecotoxicity model to distinguish between ecotoxic and non-ecotoxic compounds, with the calculated values for diclofenac (SMILES 1), rivastigmine (SMILES 2), and citalopram (SMILES 3) overlaid to show their position relative to the ecotoxic and non-ecotoxic distributions.*

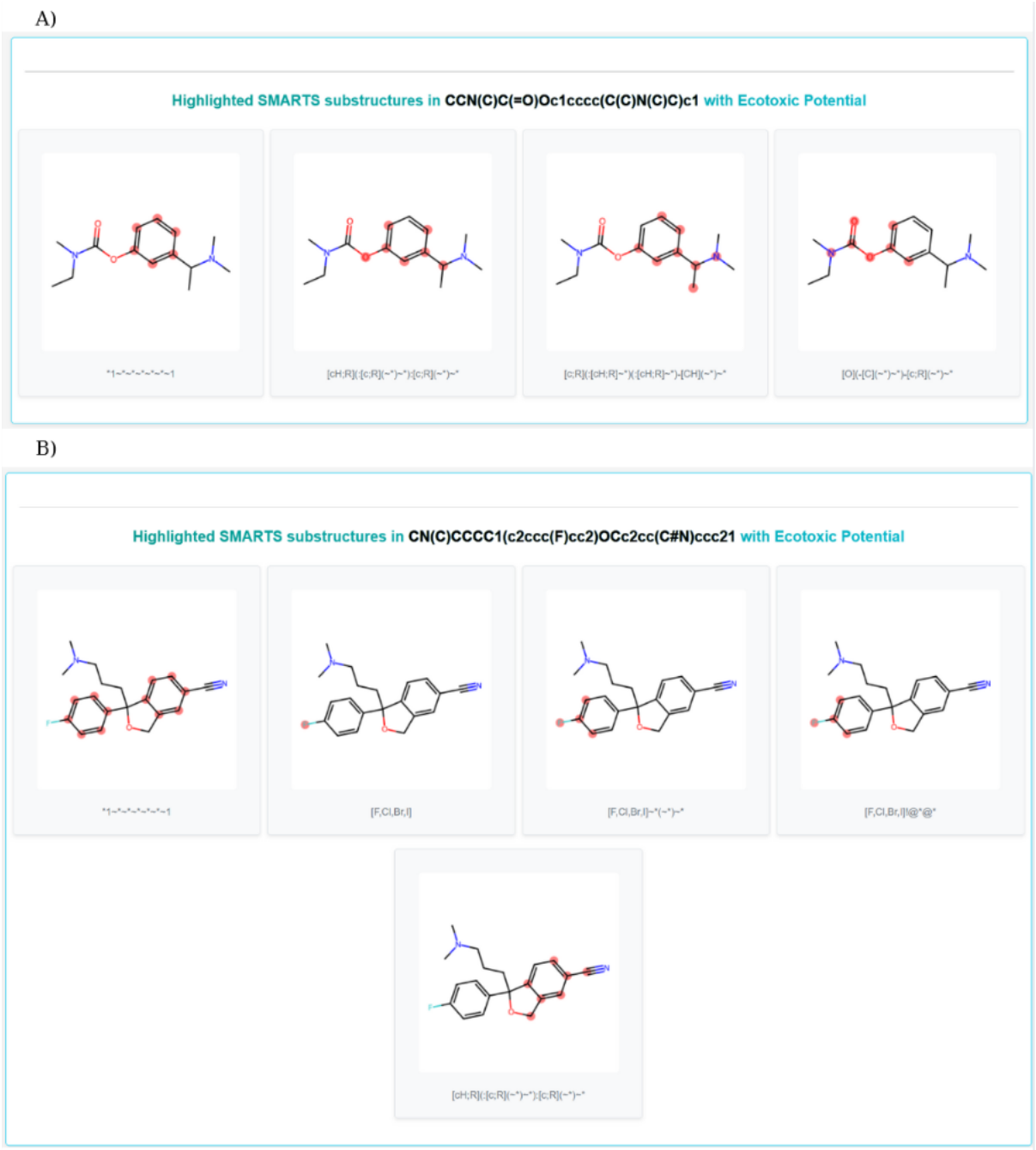



*Figure 36: SMILES arbitrary target specification (SMARTS) substructures linked with ecotoxicity potential highlighted in the chemical structure of A) rivastigmine and B) citalopram.*

# 5 Conclusions

The review of pharmaceutical-related environmental issues shows that **risk-based prioritization** is important for effective environmental monitoring of pharmaceuticals, providing input for updating existing legislation and enabling better allocation of resources and improved risk assessment. Integrating pharmaceutical consumption, environmental occurrence, monitoring coverage and ecotoxicological risk provides a more robust basis for prioritization. The main findings of this review highlight:

- The primary **hotspot classes** identified: **class N** (nervous system), **class J01** (antibacterials), **class N06A** (antidepressants) and **class M01A** (anti-inflammatory drugs), regularly containing both high- and medium-risk compounds.

- Additional classes identified as **potential hotspots requiring further screening,** to avoid overlooking high-risk compounds: **class N02** (analgesics), **class A** (alimentary track and metabolism), **class B** (blood and blood forming organs) and **class C10** (lipid modifying agents).

**- Monitoring gaps** related to **metabolites and transformation products,** some of which may pose greater ecological risk than parent compounds, remain important.

**- Seventeen priority substances** are proposed for routine monitoring, with **increased frequency** recommended for **NSAIDs, selected antibiotics and synthetic hormones** due to their consistently high risk.

Building directly on this prioritization framework, through Task 6.2, CLOUD has delivered a comprehensive computational framework to mitigate the environmental impact of pharmaceuticals by embedding ecotoxicity considerations early in the drug design process. Machine learning models capable of reliably classifying compounds as ecotoxic or non-ecotoxic to aquatic fish organisms were developed and extensively benchmarked against state-of-the-art tools and external datasets, demonstrating the robustness and competitiveness of the proposed approach. Importantly, the best-performing ecotoxicity model was experimentally validated through LC50 assays in *Danio rerio* embryos under controlled and homogeneous conditions, enhancing the validity of our tools. The best-performing ML model, together with the ML tools for predicting bioconcentration and metabolic products previously developed and documented under Deliverable 3.2, was integrated into a unified, automated analysis pipeline, ensuring usability and scalability for real-world applications in early-stage drug design. Overall, the development of these tools significantly advances the project's objective of enabling environmentally sustainable pharmaceutical design and supports regulatory-relevant decision-making aligned with EU environmental protection goals.

Complementing both the review and the computational prioritization framework, the **eco-immunotoxicological studies provide experimental, effect-based evidence of biological vulnerability in real-world marine organisms**. Greek mussels collected from the Saronikos Gulf, close to the Psyttalia wastewater treatment plant (WWTP), an area influenced by multiple anthropogenic pressures (potentially at risk of pharmaceutical contamination) exhibited higher basal stress levels, as indicated by increased lysosomal instability under control conditions, with this stress further exacerbated by drug exposure. However, chemical analyses of seawater did not reveal substantial differences in pharmaceutical concentrations between the Greek and Neapolitan sites. This suggests that the elevated basal stress observed in Greek mussels cannot be directly attributed to higher pharmaceutical exposure alone, but rather reflects the influence

of **site-specific stressors, long-term environmental pressures and/or transportation-related stress**, which may increase organismal susceptibility to chemical stress.

Overall, these studies show that the identified biomarkers are valuable as **screening tools and mechanistic endpoints**, but are insufficient on their own for robust ecological risk assessment. Short-term exposure responses can reveal early biological perturbations and support hazard identification, yet they should be interpreted as complementary evidence rather than definitive proof of ecological risk. Taken together, the findings strongly support the **integration of effect-based assays and marine invertebrate models – particularly mussels – into regulatory risk assessment frameworks**. This integrated approach reinforces the need to account for sublethal immunotoxic endpoints, site-specific vulnerability, and mixture effects within European marine monitoring and management strategies, thereby bridging chemical prioritization, predictive modeling, and biological effects assessment into a coherent environmental risk paradigm.

# 6 References

1. *Hospice Nurses and Drug Disposal.* Ortner, Pamela and McCullagh, Marjorie. 1, 1 2010, Journal of Hospice & Palliative Nursing, Vol. 12, pp. 15-26. 1522-2179.

2. *Promoting resistance by the emission of antibiotics from hospitals and households into effluent.* Kümmerer, K. and Henninger, A. 12, 12 2003, Clinical Microbiology and Infection, Vol. 9, pp. 1203-1214. 1198743X.

3. *Transformation products of pharmaceuticals in surface waters and wastewater formed during photolysis and advanced oxidation processes – Degradation, elucidation of byproducts and assessment of their biological potency.* Fatta-Kassinos, D., Vasquez, M.I. and Kümmerer, K. 5, 10 2011, Chemosphere, Vol. 85, pp. 693-709. 00456535.

4. *Environmental toxicology and risk assessment of pharmaceuticals from hospital wastewater.* Escher, Beate I., et al. 1, 1 2011, Water Research, Vol. 45, pp. 75-92. 00431354.

5. *Calibration and field evaluation of polar organic chemical integrative sampler (POCIS) for monitoring pharmaceuticals in hospital wastewater.* Bailly, Emilie, Levi, Yves and Karolak, Sara. 3 2013, Environmental Pollution, Vol. 174, pp. 100-105. 02697491.

6. *Micropollutant Point Sources in the Built Environment: Identification and Monitoring of Priority Pharmaceutical Substances in Hospital Effluents.* K, Helwig. 04, 2013, Journal of Environmental & Analytical Toxicology, Vol. 03. 21610525.

7. *A framework for the assessment of the environmental risk posed by pharmaceuticals originating from hospital effluents.* Al Aukidy, Mustafa, Verlicchi, Paola and Voulvoulis, Nikolaos. 9 2014, Science of The Total Environment, Vol. 493, pp. 54-64. 00489697.

8. *Potential Upstream Strategies for the Mitigation of Pharmaceuticals in the Aquatic Environment: a Brief Review.* Blair, Benjamin D. 2, 6 2016, Current Environmental Health Reports, Vol. 3, pp. 153-160. 2196-5412.

9. *Distribution and fate of pharmaceuticals and their metabolite conjugates in a municipal wastewater treatment plant.* Brown, Alistair K. and Wong, Charles S. 11 2018, Water Research, Vol. 144, pp. 774-783. 00431354.

10. *Sources of Pharmaceutical Residues in the Environment and their Control.* Caldwell, Daniel J. s.l. : The Royal Society of Chemistry, 9 2015, Pharmaceuticals in the Environment, pp. 92-119.

11. *The SIPIBEL project: treatment of hospital and urban wastewater in a conventional urban wastewater treatment plant.* Chonova, Teofana, et al. 10, 4 2018, Environmental Science and Pollution Research, Vol. 25, pp. 9197-9206. 0944-1344.

12. *Dynamics of active pharmaceutical ingredients loads in a Swiss university hospital wastewaters and prediction of the related environmental risk for the aquatic ecosystems.* Daouk, Silwan, et al. 3 2016, Science of The Total Environment, Vol. 547, pp. 244-253. 00489697.

13. *Modeling pharmaceutical emissions and their toxicity-related effects in life cycle assessment (LCA): A review.* Emara, Yasmine, et al. 1, 1 2019, Integrated Environmental Assessment and Management, Vol. 15, pp. 6-18. 1551-3777.

14. *Pharmaceuticals in hospital wastewater: Their ecotoxicity and contribution to the environmental hazard of the effluent.* Frédéric, Orias and Yves, Perrodin. 11 2014, Chemosphere, Vol. 115, pp. 31-39. 00456535.

15. *Endocrine Disrupting Compounds Removal Methods from Wastewater in the United Kingdom: A Review.* Gadupudi, China K., et al. 1, 2 2021, Sci, Vol. 3, p. 11. 2413-4155.

16. *Ecotoxicity screening evaluation of selected pharmaceuticals and their transformation products towards various organisms.* Grabarczyk, Łukasz, et al. 21, 7 2020, Environmental Science and Pollution Research, Vol. 27, pp. 26103-26114. 0944-1344.

17. *Significance of metabolites in the environmental risk assessment of pharmaceuticals consumed by human.* Han, Eun Jeong and Lee, Dong Soo. 8 2017, Science of The Total Environment, Vol. 592, pp. 600-607. 00489697.

18. *N-nitrosamines, emerging disinfection by-products of health concern: an overview of occurrence, mechanisms of formation, control and analysis in water.* Kadmi, Y., Favier, L. and Wolbert, D. 1, 2 2015, Water Supply, Vol. 15, pp. 11-25. 1606-9749.

19. *Advanced oxidation process-mediated removal of pharmaceuticals from water: A review.* Kanakaraju, Devagi, Glass, Beverley D. and Oelgemöller, Michael. 8 2018, Journal of Environmental Management, Vol. 219, pp. 189-207. 03014797.

20. *Diclofenac and its transformation products: Environmental occurrence and toxicity - A review.* Lonappan, Linson, et al. 11 2016, Environment International, Vol. 96, pp. 127-138. 01604120.

21. *Investigating the formation and toxicity of nitrogen transformation products of diclofenac and sulfamethoxazole in wastewater treatment plants.* Osorio, Victoria, et al. 5 2016, Journal of Hazardous Materials, Vol. 309, pp. 157-164. 03043894.

22. *The impact of on-site hospital wastewater treatment on the downstream communal wastewater system in terms of antibiotics and antibiotic resistance genes.* Paulus, Gabriela K., et al. 4, 5 2019, International Journal of Hygiene and Environmental Health, Vol. 222, pp. 635-644. 14384639.

23. *A review of combined sewer overflows as a source of wastewater-derived emerging contaminants in the environment and their management.* Petrie, Bruce. 25, 7 2021, Environmental Science and Pollution Research, Vol. 28, pp. 32095-32110. 0944-1344.

24. *Determination of metals and pharmaceutical compounds released in hospital wastewater from Toluca, Mexico, and evaluation of their toxic impact.* Pérez-Alvarez, Itzayana, et al. 9 2018, Environmental Pollution, Vol. 240, pp. 330-341. 02697491.

25. *The fate and risk assessment of psychiatric pharmaceuticals from psychiatric hospital effluent.* Xiang, Jiajia, et al. 4 2018, Ecotoxicology and Environmental Safety, Vol. 150, pp. 289-296. 01476513.

26. *Prioritizing Pharmaceuticals for Environmental Monitoring in Greece: A Comprehensive Review of Consumption, Occurrence, and Ecological Risk.* Chatzipanagiotou, K. R., et al. 2025, Submitted, under review, Toxics.

27. Pharmaceutical Market Dataset. *OECD website.* [Online] 19 10 2022. https://stats.oecd.org/viewhtml.aspx?datasetcode=HEALTH_PHMC&lang=en.

28. *Determination of polar pharmaceuticals in sewage water of Greece by gas chromatography–mass spectrometry.* Koutsouba, V, et al. 2, 4 2003, Chemosphere, Vol. 51, pp. 69-75. 00456535.

29. *Pharmaceuticals in STP effluents and their solar photodegradation in aquatic environment.* Andreozzi, Roberto, Raffaele, Marotta and Nicklas, Paxéus. 10, 3 2003, Chemosphere, Vol. 50, pp. 1319-1330. 00456535.

30. *Determination of pharmaceuticals from different therapeutic classes in wastewaters by liquid chromatography–electrospray ionization–tandem mass spectrometry.* Botitsi, Eleni, Frosyni, Charalampia and Tsipi, Despina. 4, 2 2007, Analytical and Bioanalytical Chemistry, Vol. 387, pp. 1317-1327. 1618-2642.

31. *Simultaneous determination of the endocrine disrupting compounds nonylphenol, nonylphenol ethoxylates, triclosan and bisphenol A in wastewater and sewage sludge by gas chromatography–mass spectrometry.* Gatidou, Georgia, et al. 1-2, 1 2007, Journal of Chromatography A, Vol. 1138, pp. 32-41. 00219673.

32. *Determination of phenolic and steroid endocrine disrupting compounds in environmental matrices.* Arditsoglou, Anastasia and Voutsa, Dimitra. 3, 5 2008, Environmental Science and Pollution Research, Vol. 15, pp. 228-236. 0944-1344.

33. *Endocrine disrupting compounds in municipal and industrial wastewater treatment plants in Northern Greece.* Pothitou, Paraskevi and Voutsa, Dimitra. 11, 12 2008, Chemosphere, Vol. 73, pp. 1716-1723. 00456535.

34. *Occurrence and fate of endocrine disrupters in Greek sewage treatment plants.* Stasinakis, Athanasios S., et al. 6-7, 3 2008, Water Research, Vol. 42, pp. 1796-1804. 00431354.

35. *Passive sampling of selected endocrine disrupting compounds using polar organic chemical integrative samplers.* Arditsoglou, Anastasia and Voutsa, Dimitra. 2, 11 2008, Environmental Pollution, Vol. 156, pp. 316-324. 02697491.

36. *Analysis of Selected Pharmaceutical Compounds and Endocrine Disruptors in Municipal Wastewater Using Solid-Phase Microextraction and Gas Chromatography.* Antoniou, Chrysoula V., Koukouraki, Elisavet E. and Diamadopoulos, Evan. 7, 7 2009, Water Environment Research, Vol. 81, pp. 664-669. 1061-4303.

37. *Determination of selected non-steroidal anti-inflammatory drugs in wastewater by gas chromatography-mass spectrometry.* Samaras, Vasilios G., et al. 3-6, 3 2010, International Journal of Environmental Analytical Chemistry, Vol. 90, pp. 219-229. 0306-7319.

38. *Partitioning of endocrine disrupting compounds in inland waters and wastewaters discharged into the coastal area of Thessaloniki, Northern Greece.* Arditsoglou, Anastasia and Voutsa, Dimitra. 3, 3 2010, Environmental Science and Pollution Research, Vol. 17, pp. 529-538. 0944-1344.

39. *Occurrence of endocrine disrupters and selected pharmaceuticals in Aisonas River (Greece) and environmental risk assessment using hazard indexes.* Stasinakis, Athanasios S., et al. 5, 6 2012, Environmental Science and Pollution Research, Vol. 19, pp. 1574-1583. 0944-1344.

40. *Occurrence and removal of PPCPs in municipal and hospital wastewaters in Greece.* Kosma, Christina I., Lambropoulou, Dimitra A. and Albanis, Triantafyllos A. 1-3, 7 2010, Journal of Hazardous Materials, Vol. 179, pp. 804-817. 03043894.

41. *Occurrence and removal of fungicides in municipal sewage treatment plant.* Stamatis, N., Hela, D. and Konstantinou, I. 1-3, 3 2010, Journal of Hazardous Materials, Vol. 175, pp. 829-835. 03043894.

42. *Occurrence and partitioning of endocrine-disrupting compounds in the marine environment of Thermaikos Gulf, Northern Aegean Sea, Greece.* Arditsoglou, Anastasia and Voutsa, Dimitra. 11, 11 2012, Marine Pollution Bulletin, Vol. 64, pp. 2443-2452. 0025326X.

43. *Contribution of primary and secondary treatment on the removal of benzothiazoles, benzotriazoles, endocrine disruptors, pharmaceuticals and perfluorinated compounds in a sewage treatment plant.* Stasinakis, Athanasios S., et al. 10 2013, Science of The Total Environment, Vols. 463-464, pp. 1067-1075. 00489697.

44. *Fate of selected pharmaceuticals and synthetic endocrine disrupting compounds during wastewater treatment and sludge anaerobic digestion.* Samaras, Vasilios G., et al. 1 2013, Journal of Hazardous Materials, Vols. 244-245, pp. 259-267. 03043894.

45. *EU-wide monitoring survey on emerging polar organic contaminants in wastewater treatment plant effluents.* Loos, Robert, et al. 17, 11 2013, Water Research, Vol. 47, pp. 6475-6487. 00431354.

46. *An analytical method for the simultaneous trace determination of acidic pharmaceuticals and phenolic endocrine disrupting chemicals in wastewater and sewage sludge by gas chromatography-mass spectrometry.* Samaras, Vasilios G., et al. 7, 3 2011, Analytical and Bioanalytical Chemistry, Vol. 399, pp. 2549-2561. 1618-2642.

47. *Occurrence and distribution of selected pharmaceutical compounds on sewage-impacted section of River Acheloos, Western Greece.* Stamatis, Nikolaos, et al. 15, 12 2013, International Journal of Environmental Analytical Chemistry, Vol. 93, pp. 1602-1619. 0306-7319.

48. *Highly sensitive determination of 68 psychoactive pharmaceuticals, illicit drugs, and related human metabolites in wastewater by liquid chromatography–tandem mass spectrometry.* Borova, Viola L., et al. 17, 7 2014, Analytical and Bioanalytical Chemistry, Vol. 406, pp. 4273-4285. 1618-2642.

49. *Investigation of PPCPs in wastewater treatment plants in Greece: Occurrence, removal and environmental risk assessment.* Kosma, Christina I., Lambropoulou, Dimitra A. and Albanis, Triantafyllos A. 1 2014, Science of The Total Environment, Vols. 466-467, pp. 421-438. 00489697.

50. *Occurrence and removal of emerging pharmaceutical, personal care compounds and caffeine tracer in municipal sewage treatment plant in Western Greece.* Stamatis, Nikolaos K. and Konstantinou, Ioannis K. 9, 9 2013, Journal of Environmental Science and Health, Part B, Vol. 48, pp. 800-813. 0360-1234.

51. *Polar organic micropollutants in the coastal environment of different marine systems.* Nödler, Karsten, Voutsa, Dimitra and Licha, Tobias. 1, 8 2014, Marine Pollution Bulletin, Vol. 85, pp. 50-59. 0025326X.

52. *Comprehensive study of the antidiabetic drug metformin and its transformation product guanylurea in Greek wastewaters.* Kosma, Christina I., Lambropoulou, Dimitra A. and Albanis, Triantafyllos A. 3 2015, Water Research, Vol. 70, pp. 436-448. 00431354.

53. *Is there a risk for the aquatic environment due to the existence of emerging organic contaminants in treated domestic wastewater? Greece as a case-study.* Thomaidi, Vasiliki S., et al. 2 2015, Journal of Hazardous Materials, Vol. 283, pp. 740-747. 03043894.

54. *A multi-residue method for determination of 70 organic micropollutants in surface waters by solid-phase extraction followed by gas chromatography coupled to tandem mass spectrometry.* Terzopoulou, Evangelia, Voutsa, Dimitra and Kaklamanos, George. 2, 1 2015, Environmental Science and Pollution Research, Vol. 22, pp. 1095-1112. 0944-1344.

55. *Multianalyte method for the determination of pharmaceuticals in wastewater samples using solid-phase extraction and liquid chromatography–tandem mass spectrometry.* Dasenaki, Marilena E. and Thomaidis, Nikolaos S. 15, 6 2015, Analytical and Bioanalytical Chemistry, Vol. 407, pp. 4229-4245. 1618-2642.

56. *Occurrence of pharmaceuticals in surface waters: analytical method development and environmental risk assessment.* Nannou, Christina I., Kosma, Christina I. and Albanis, Triantafyllos A. 13, 10 2015, International Journal of Environmental Analytical Chemistry, Vol. 95, pp. 1242-1262. 0306-7319.

57. *Simultaneous determination of 148 pharmaceuticals and illicit drugs in sewage sludge based on ultrasound-assisted extraction and liquid chromatography–tandem mass spectrometry.* Gago-Ferrero, Pablo, et al. 15, 6 2015, Analytical and Bioanalytical Chemistry, Vol. 407, pp. 4287-4297. 1618-2642.

58. *Occurrence and spatial distribution of 158 pharmaceuticals, drugs of abuse and related metabolites in offshore seawater.* Alygizakis, Nikiforos A., et al. 1 2016, Science of The Total Environment, Vol. 541, pp. 1097-1105. 00489697.

59. *Assessing the risk associated with the presence of emerging organic contaminants in sludge-amended soil: A country-level analysis.* Thomaidi, Vasiliki S., et al. 4 2016, Science of The Total Environment, Vols. 548-549, pp. 280-288. 00489697.

60. *Evaluation of polar organic micropollutants as indicators for wastewater-related coastal water quality impairment.* Nödler, Karsten, et al. 4 2016, Environmental Pollution, Vol. 211, pp. 282-290. 02697491.

61. *Seasonal occurrence, removal, mass loading and environmental risk assessment of 55 pharmaceuticals and personal care products in a municipal wastewater treatment plant in Central Greece.* Papageorgiou, Myrsini, Kosma, Christina and Lambropoulou, Dimitra. 2 2016, Science of The Total Environment, Vol. 543, pp. 547-569. 00489697.

62. *UHPLC-QTOF MS screening of pharmaceuticals and their metabolites in treated wastewater samples from Athens.* Ibáñez, M., et al. 2 2017, Journal of Hazardous Materials, Vol. 323, pp. 26-35. 03043894.

63. *Environmental fate of non-steroidal anti-inflammatory drugs in river water/sediment systems.* Koumaki, Elena, Mamais, Daniel and Noutsopoulos, Constantinos. 2 2017, Journal of Hazardous Materials, Vol. 323, pp. 233-241. 03043894.

64. *Reflection of Socioeconomic Changes in Wastewater: Licit and Illicit Drug Use Patterns.* Thomaidis, Nikolaos S., et al. 18, 9 2016, Environmental Science & Technology, Vol. 50, pp. 10065-10072. 0013-936X.

65. *Assessment of the environmental fate of endocrine disrupting chemicals in rivers.* Koumaki, Elena, Mamais, Daniel and Noutsopoulos, Constantinos. 7 2018, Science of The Total Environment, Vols. 628-629, pp. 947-958. 00489697.

66. *Comprehensive investigation of a wide range of pharmaceuticals and personal care products in urban and hospital wastewaters in Greece.* Papageorgiou, Myrsini, et al. 12 2019, Science of The Total Environment, Vol. 694, p. 133565. 00489697.

67. *Photocatalytic Treatment of Pharmaceuticals in Real Hospital Wastewaters for Effluent Quality Amelioration.* Konstas, Panagiotis-Spyridon, et al. 10, 10 2019, Water, Vol. 11, p. 2165. 2073-4441.

68. *Psychiatrics and selected metabolites in hospital and urban wastewaters: Occurrence, removal, mass loading, seasonal influence and risk assessment.* Kosma, Christina I., et al. 4 2019, Science of The Total Environment, Vol. 659, pp. 1473-1483. 00489697.

69. *A modified QuEChERS approach for the analysis of pharmaceuticals in sediments by LC-Orbitrap HRMS.* Nannou, Christina I., Boti, Vasiliki I. and Albanis, Triantafyllos A. 7, 3 2019, Analytical and Bioanalytical Chemistry, Vol. 411, pp. 1383-1396. 1618-2642.

70. *Analytical and mathematical assessment of emerging pollutants fate in a river system.* Noutsopoulos, Constantinos, et al. 2 2019, Journal of Hazardous Materials, Vol. 364, pp. 48-58. 03043894.

71. *Contamination patterns and attenuation of pharmaceuticals in a temporary Mediterranean river.* Mandaric, Ladislav, et al. 1 2019, Science of The Total Environment, Vol. 647, pp. 561-569. 00489697.

72. *Exploring the Efficiency of UHPLC-Orbitrap MS for the Determination of 20 Pharmaceuticals and Acesulfame K in Hospital and Urban Wastewaters with the Aid of FPSE.* Kalaboka, Maria, et al. 3, 9 2020, Separations, Vol. 7, p. 46. 2297-8739.

73. *Determination of illicit drugs and psychoactive pharmaceuticals in wastewater from the area of Thessaloniki (Greece) using LC–MS/MS: estimation of drug consumption.* Christophoridis, Christophoros, et al. 5, 5 2021, Environmental Monitoring and Assessment, Vol. 193, p. 249. 0167-6369.

74. *Occurrence of pharmaceuticals and personal care products in bottled water and assessment of the associated risks.* Wang, Chengfei, et al. 10 2021, Environment International, Vol. 155, p. 106651. 01604120.

75. *Patterns of pharmaceuticals use during the first wave of COVID-19 pandemic in Athens, Greece as revealed by wastewater-based epidemiology.* Galani, Aikaterini, et al. 12 2021, Science of The Total Environment, Vol. 798, p. 149014. 00489697.

76. *Monitoring of a Broad Set of Pharmaceuticals in Wastewaters by High-Resolution Mass Spectrometry and Evaluation of Heterogenous Catalytic Ozonation for Their Removal in a Pre-Industrial Level Unit.* Nannou, Christina, et al. 2, 4 2022, Analytica, Vol. 3, pp. 195-212. 2673-4532.

77. *Simultaneous determination of selected pharmaceuticals and plasticisers in urban stormwater in Chania (Greece).* Kotti, Melina, Papafilippaki, Androniki and Stavroulakis, George. 16, 12 2023, International Journal of Environmental Analytical Chemistry, Vol. 103, pp. 3790-3800. 0306-7319.

78. *Assessment of a wide array of organic micropollutants of emerging concern in wastewater treatment plants in Greece: Occurrence, removals, mass loading and potential risks.* Ofrydopoulou, Anna, et al. 1 2022, Science of The Total Environment, Vol. 802, p. 149860. 00489697.

79. *Removal of organic pollutants (pharmaceuticals and pesticides) from sewage sludge by hydrothermal carbonization using response surface methodology ( <scp>RSM</scp> ).* Miserli, Kleopatra, Nastopoulou, Amalia and Konstantinou, Ioannis. 11, 11 2022, Journal of Chemical Technology & Biotechnology, Vol. 97, pp. 3111-3120. 0268-2575.

80. *The NORMAN Association and the European Partnership for Chemicals Risk Assessment (PARC): let's cooperate!* Dulio, Valeria, et al. 1, 12 2020, Environmental Sciences Europe, Vol. 32, p. 100. 2190-4707.

81. NORMAN Database System website. [Online] [Cited: 10 10 2025.] https://www.norman-network.com/nds/.

82. https://stats.oecd.org/viewhtml.aspx?datasetcode=HEALTH_PHMC&lang=en. [Online] [Cited: 19 10 2022.] https://stats.oecd.org/viewhtml.aspx?datasetcode=HEALTH_PHMC&lang=en.

83. ClinPGx database website. [Online] 17 10 2025. https://www.clinpgx.org.

84. *Antibiotics: An overview on the environmental occurrence, toxicity, degradation, and removal methods.* Yang, Q., et al. 2021, Bioengineered, Vol. 12, pp. 7376-7416.

85. *The EU Green Deal: The challenge of greening medical technologies.* Rampi V, Bisazza O. 4, 2023, Clin Chem Lab Med., Vol. 61, pp. 651-653.

86. The European Green Deal - European Commission. . [Online] [Cited: ] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en .

87. The Paris Agreement. [Online] [Cited: ] https://www.un.org/en/climatechange/paris-agreement .

88. Regulation - 1272/2008 - EN - clp regulation - EUR-Lex. . [Online] [Cited: ] https://eur-lex.europa.eu/eli/reg/2008/1272/oj/eng .

89. REACH Regulation - European Commission. . [Online] [Cited: ] https://environment.ec.europa.eu/topics/chemicals/reach-regulation_en .

90. RoHS Directive - European Commission. [Online] [Cited: ] https://environment.ec.europa.eu/topics/waste-and-recycling/rohs-directive_en#timeline .

91. *Chemical strategies for sustainable medical laboratories.* Ozben T, Fragão-Marques M. 4, 2023, Clin Chem Lab Med., Vol. 61, pp. 642-650.

92. *A critical evaluation of different parameters for estimating pharmaceutical exposure seeking an improved environmental risk assessment.* Pereira AMPT, Silva LJG, Lino CM, Meisel LM, Pena A. 2017, Science of The Total Environment. , pp. 603-604:226-236.

93. *Ecotoxicity evaluation of selected sulfonamides.* Białk-Bielińska A, Stolte S, Arning J, et al. 6, 2011, Chemosphere, Vol. 85, pp. 928-933.

94. *Ecotoxicity screening evaluation of selected pharmaceuticals and their transformation products towards various organisms.* Grabarczyk Ł, Mulkiewicz E, Stolte S, et al. 21, 2020, Environ Sci Pollut Res Int., Vol. 27, pp. 26103-26114.

95. *ToxAIcology-The evolving role of artificial intelligence in advancing toxicology and modernizing regulatory science.* T., Hartung. 4, 2023, ALTEX - Alternatives to animal experimentation., Vol. 40, pp. 559-570.

96. *ECOSAR model performance with a large test set of industrial chemicals.* Reuschenbach P, Silvani M, Dammann M, Warnecke D, Knacker T. 10, 2008, Chemosphere, Vol. 71, pp. 1986-1995.

97. Toxicity Estimation Software Tool (TEST) | US EPA. [Online] [Cited: ] https://www.epa.gov/comptox-tools/toxicity-estimation-software-tool-test .

98. *Ecotoxicological QSAR modeling of organic compounds against fish: Application of fragment based descriptors in feature analysis.* Khan K, Baderna D, Cappelli C, et al. 2019, Aquatic Toxicology, Vol. 212, pp. 162-174.

99. *The Bigger Fish: A Comparison of Meta-Learning QSAR Models on Low-Resourced Aquatic Toxicity Regression Tasks.* Schlender T, Viljanen M, van Rijn JN, et al. 46, 2023, Environ Sci Technol., Vol. 57, pp. 17818-17830.

100. *Physiological variables in machine learning QSARs allow for both cross-chemical and cross-species predictions.* Zubrod JP, Galic N, Vaugeois M, Dreier DA. 2023, Ecotoxicol Environ Saf., Vol. 263, p. 115250.

101. *Implementing comprehensive machine learning models of multispecies toxicity assessment to improve regulation of organic compounds.* He Y, Liu G, Hu S, et al. 2023, J Hazard Mater., Vol. 458, p. 131942.

102. *In silico prediction of pesticide aquatic toxicity with chemical category approaches.* Li F, Fan D, Wang H, et al. 6, 2017, Toxicol Res (Camb), Vol. 6, pp. 831-842.

103. *Examining predictors of chemical toxicity in freshwater fish using the random forest technique.* Tuulaikhuu BA, Guasch H, García-Berthou E. 11, 2017, Environmental Science and Pollution Research, Vol. 24, pp. 10172-10181.

104. *Predicting chemical hazard across taxa through machine learning.* Wu J, D'Ambrosi S, Ammann L, Stadnicka-Michalak J, Schirmer K, Baity-Jesi M. 2022, Environ Int., Vol. 163, p. 107184.

105. *In silico prediction of chemical aquatic toxicity by multiple machine learning and deep learning approaches.* Xu M, Yang H, Liu G, Tang Y, Li W. 11, 2022, Journal of Applied Toxicology, Vol. 42, pp. 1766-1776.

106. *What is the ecotoxicity of a given chemical for a given aquatic species? Predicting interactions between species and chemicals using recommender system techniques.* Viljanen M, Minnema J, Wassenaar PNH, Rorije E, Peijnenburg W. 10, 2023, SAR QSAR Environ Res., Vol. 34, pp. 765-788.

107. *A benchmark dataset for machine learning in ecotoxicology.* . Schür C, Gasser L, Perez-Cruz F, Schirmer K, Baity-Jesi M. 1, 2023 , Scientific Data, Vol. 10, pp. 1-20.

108. *Machine learning-based prediction of fish acute mortality: Implementation, interpretation, and regulatory relevance.* Gasser L, Schür C, Fernando Perez-Cruz ‡, Schirmer K, Baity-Jesi M. 2024, bioRxiv.

109. *Development of an ecotoxicity QSAR model for the KAshinhou Tool for Ecotoxicity (KATE) system.* Furuhama A, Toida T, Nishikawa N, Aoki Y, Yoshioka Y, Shiraishi H. 5-6, 2010, SAR QSAR Environ Res., Vol. 21, pp. 403-413.

110. *Aquatic ecotoxicity of personal care products: QSAR models and ranking for prioritization and safer alternatives' design.* Gramatica P, Cassani S, Sangion A. 16, 2016, Green Chemistry, Vol. 18, pp. 4393-4406.

111. *Multispecies QSAR modeling for predicting the aquatic toxicity of diverse organic chemicals for regulatory toxicology.* . Singh KP, Gupta S, Kumar A, Mohan D. 5, 2014, Chem Res Toxicol., Vol. 27, pp. 741-753.

112. *Ensemble QSAR Modeling to Predict Multispecies Fish Toxicity Lethal Concentrations and Points of Departure.* Sheffield TY, Judson RS. 21, 2019, Environ Sci Technol., Vol. 53, pp. 12793-12802.

113. *Comparing LD50/LC50Machine Learning Models for Multiple Species.* Lane TR, Harris J, Urbina F, Ekins S. 2, 2023, ACS Chemical Health and Safety, Vol. 30, pp. 83-97.

114. *QSAR modelling study of the bioconcentration factor and toxicity of organic compounds to aquatic organisms using machine learning and ensemble methods.* Ai H, Wu X, Zhang L, et al. 2019, Ecotoxicol Environ Saf., Vol. 179, pp. 71-78.

115. *Improved environmental chemistry property prediction of molecules with graph machine learning.* Zhu S, Nguyen BH, Xia Y, et al. 17, 2023, Green Chemistry, Vol. 25, pp. 6612-6617.

116. *Creation of a Curated Aquatic Toxicology Database: EnviroTox.* . Connors KA, Beasley A, Barron MG, et al. 5, 2019, Environ Toxicol Chem., Vol. 38, pp. 1062-1073.

117. *PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints.* CW., Yap. 7, 2011, J Comput Chem. , Vol. 32, pp. 1466-1474.

118. *Evaluation of Classifier Models Using Stratified Tenfold Cross Validation Techniques.* Purushotham S, Tripathy BK. 2012, Communications in Computer and Information Science, Vol. 270 CCIS(PART II), pp. 680-690.

119. *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting.* . Freund Y, Schapire RE. 1, 1997, J Comput Syst Sci., Vol. 55, pp. 119-139.

120. *k-Nearest Neighbor Classification.* Mucherino A, Papajorgji PJ, Pardalos PM. 2009, Springer Optimization and Its Applications., pp. 83-106.

121. *Random forests.* L., Breiman. 1, 2001, Mach Learn. , Vol. 45, pp. 5-32.

122. Haykin, Simon. *Neural networks: A comprehensive foundation.* [ed.] Tadeusiewicz R. s.l. : Macmillan College Publishing, 1995. pp. 746-747.

123. *On evaluation metrics for medical applications of artificial intelligence.* Hicks SA, Strümke I, Thambawita V, et al. 1, 2022, Sci Rep. , Vol. 12, p. 5979.

124. eChemPortal provides free public access to information on properties of chemicals. [Online] [Cited: ] https://www.echemportal.org/echemportal/ .

125. GitHub - jeffrichardchemistry/pyADA: A cheminformatics package to perform Applicability Domain of molecular fingerprints based in similarity calculation. [Online] [Cited: ] https://github.com/jeffrichardchemistry/pyADA .

126. *Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?* . Bajusz D, Rácz A, Héberger K. 1, 2015, J Cheminform, Vol. 7, pp. 1-13.

127. *Greedy function approximation: A gradient boosting machine.* JH., Friedman. 5, 2001, Vol. 29, pp. 1189-1232.

128. *Extremely randomized trees.* Geurts P, Ernst D, Wehenkel L. 1, 2006, Mach Learn. , Vol. 63, pp. 3-42.

129. *Decision tree methods: applications for classification and prediction.* Song YY, Lu Y. 2, 2015, Shanghai Arch Psychiatry, Vol. 27, p. 130.

130. KM., Ting. Confusion Matrix. . [book auth.] C., Webb, G.I. Sammut. *Encyclopedia of Machine Learning.* s.l. : Springer, 2011, p. 209.

131. Landrum G, Tosco P, Kelley B, et al. rdkit/rdkit: 2024_09_5 (Q3 2024) Release. [Online]

132. *Reoptimization of MDL keys for use in drug discovery.* Durant JL, Leland BA, Henry DR, Nourse JG. 6, 2002, J Chem Inf Comput Sci., Vol. 42, pp. 1273-1280.

133. *The Generation of a Unique Machine Description for Chemical Structures—A Technique Developed at Chemical Abstracts Service.* HL., Morgan. 2, 1965, J Chem Doc., Vol. 5, pp. 107-113.

134. *XGBoost: A scalable tree boosting system.* Chen T, Guestrin C. 2016. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. . pp. 785-794.

135. *SMOTE: Synthetic Minority Over-sampling Technique.* Chawla N V., Bowyer KW, Hall LO, Kegelmeyer WP. 2002, Journal of Artificial Intelligence Research, Vol. 16, pp. 321-357.

136. *SMOTE-ENC: A Novel SMOTE-Based Method to Generate Synthetic Data for Nominal and Continuous Features.* Mukherjee M, Khushi M. 1, 2021, Applied System Innovation , Vol. 4, p. 18.

137. *Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach.* Mangalathu S, Hwang SH, Jeon JS. 2020, Eng Struct., Vol. 219, p. 110927.

138. *Principal component analysis: a review and recent developments.* Jollife IT, Cadima J. 2016, Philos Trans A Math Phys Eng Sci., Vol. 374, p. 2065.

139. *Mann-Whitney test is not just a test of medians: differences in spread can be important.* A., Hart. 7309, 2001, BMJ : British Medical Journal, Vol. 323, p. 391.

140. Michael C. Newman, Michael C. Newman. *Fundamentals of Ecotoxicology, The Science of Pollution, Fourth Edition.* 2014.

141. *Safer and greener chemicals for the aquatic ecosystem: Chemometric modeling of the prolonged and chronic aquatic toxicity of chemicals on Oryzias latipes.* Kumar A, Ojha PK, Roy K. 2024, Aquatic Toxicology, Vol. 273, p. 106985.

142. *QSAR modeling of Daphnia magna and fish toxicities of biocides using 2D descriptors.* Khan K, Khan PM, Lavado G, et al. 2019, Chemosphere, Vol. 229, pp. 8-17.

143. *Modeling and insights into the structural basis of chemical acute aquatic toxicity.* Zhang R, Guo H, Hua Y, Cui X, Shi Y, Li X. 2022, Ecotoxicol Environ Saf., Vol. 242, p. 113940.

144. *Metabolic biotransformation half-lives in fish: QSAR modeling and consensus analysis. .* Papa E, van der Wal L, Arnot JA, Gramatica P. 2014, Science of The Total Environment. , Vols. 470-471, pp. 1040-1046.

145. *Ecotoxicological QSAR modelling of the acute toxicity of fused and non-fused polycyclic aromatic hydrocarbons (FNFPAHs) against two aquatic organisms: Consensus modelling and comparison with ECOSAR.* Li F, Sun G, Fan T, et al. 2023, Aquatic Toxicology, Vol. 255, p. 106393.

146. VEGA-QSAR: AI Inside a Platform for Predictive Toxicology. | BibSonomy. . [Online] [Cited: ] https://www.bibsonomy.org/bibtex/af59ab1ab58794cafcc7ad48d1bb3b4f .

147. *OPERA models for predicting physicochemical properties and environmental fate endpoints.* Mansouri K, Grulke CM, Judson RS, Williams AJ. 1, 2018, J Cheminform. , Vol. 10.

148. *Occurrence and ecotoxicological assessment of pharmaceuticals: Is there a risk for the Mediterranean aquatic environment?* Desbiolles, Fanny, et al. s.l. : Science of The Total Environment, 2018, Vol. 639.

149. *Methodological Approaches To Assess Innate Immunity and Innate Memory in Marine Invertebrates and Humans.* Auguste, Manon, et al. s.l. : Toxicol., 2022, Vol. 4.

150. *Lysosomal membrane stability.* Martinez-Gomez C., Bignell J., Lowe D. s.l. : ICES Techniques in Marine Environmental Science, 2015, Vol. 56.

151. *Effects of anti-inflammatory diclofenac assessed by toxicity tests and biomarkers in adults and larvae of Danio rerio.* Penha LC de C, Rola RC, Martinez CB dos R, Martins C de MG. 2021, Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology., Vol. 242, p. 108955.

152. *Acute toxicity of carbamazepine to juvenile rainbow trout (Oncorhynchus mykiss): Effects on antioxidant responses, hematological parameters and hepatic EROD.* Li ZH, Zlabek V, Velisek J, et al. 3, 2011, Ecotoxicol Environ Saf. , Vol. 74, pp. 319-327.

153. Environmental Risk Assessment Data Metoprolol. . [Online] [Cited: ] http://ec.europa.eu/eurostat/web/population-demography-migration-projections/population-data/main-tables .

154. *Amoxicillin Occurrence and Toxicity in the Aquatic Environment and Mechanisms of Antibiotic Resistance: A Review.* Saldaña CM, Woodley CM. 2024, Vol. 28.

155. Enhanced Reader. [Online] [Cited: ] https://assets.lgcstandards.com/sys-master%2Fpdfs%2Fh8e%2Fhce%2F10675058868254%2FSDS_LGCAMP0051.00-01_ST-WB-MSDS-6055745-1-1-1.PDF?utm_source=chatgpt.com .

156. Acetaminophen SDS, 103-90-2 Safety Data Sheets - ECHEMI. [Online] [Cited: ] https://www.echemi.com/sds/paracetamol-pd20160328140412996.html .

157. CENTER FOR DRUG EVALUATION AND RESEARCH ENVIRONMENTAL ASSESSMENT. . [Online]

158. SECTION 1: IDENTIFICATION OF THE SUBSTANCE/MIXTURE AND OF THE COMPANY/UNDERTAKING. . [Online]

159. SAFETY DATA SHEET. . [Online] [Cited: ] www.usp.org .

160. Safety Data Sheet: Theophylline | Enhanced Reader. . [Online] [Cited: ] https://www.carlroth.com/medias/SDB-1EX3-IE-EN.pdf?context=bWFzdGVyfHNlY3VyaXR5RGF0YXNoZWV0c3wzNDkzMzZ8YXBwbGljYXRpb24vcGRmfGFEUXdMMmczTXk4NU1UWTVPVGN4T1RNM016RXdMMU5FUWw4eFJWZ3pYMGxGWDBWT0xuQmtaZ3w4ODM0YzM0YjU2NTRhNzUyNTRlMzViYzRhYTE2MDExZDBlMzNmYjc5ZGQ3.

161. assessment, reliable data for the no observed effect concentration (NOEC) for reproductive inhibition in the crustacean. . [Online]

162. *Endocrine Effects of Selective Endocrine Effects of Selective Serotonin Reuptake Inhibitors Serotonin Reuptake Inhibitors ( ( SSRIs SSRIs ) on Aquatic Organisms ) on Aquatic Organisms.* . Black MC, Rogers ED, Henry TB.

163. *Haloperidol alters the behavioral, hematological and biochemical parameters of freshwater African catfish, Clarias gariepinus (Burchell 1822).* Chiejina CO, Anih L, Okoye C, et al. 2022, Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology, Vol. 254, p. 109292.

164. *Identification of environmental chemicals that induce yolk malabsorption in zebrafish using automated image segmentation.* Kalasekar SM, Zacharia E, Kessler N, et al. 2015, Reprod Toxicol. , Vol. 55, pp. 20-29.

165. *Some toxicological effects of phosphate esters on rainbow trout and bluegill.* S., Sitthichaikasem. 1978.

166. *Monitoring Strategies for Chemicals of Emerging Concern (CECs) in California's Aquatic Ecosystems Recommendations of a Science Advisory Panel. .* Anderson PD, Denslow ND, Drewes JE, et al. 2012.

167. *Complexation with ATP as a Cause of Pollutants Toxicity to Aquatic Life.* Saratovskikh EA, Saratovskikh EA. 6, 2013, J Environ Prot (Irvine, Calif)., Vol. 4, pp. 585-594.

# ANNEX

*Table 17. Accuracy and F1 score validation metrics for initially trained models evaluated on an external validation set (EchemPortal database).*

| Algorithm | Parameters | Accuracy | F1 score |
|---|---|---|---|
| **Random forest** | 300 estimators (entropy) | 47% | 20% |
| **KNN** | 6 neighbors | 55% | 17% |
| **Random forest** | 300 estimators (entropy) | 55% | 17% |
| **Adaboost** | samme | 58% | 15% |
| **KNN** | 8 neighbors | 62% | 17% |
| **Random forest** | 150 estimators (entropy) | 74% | 12% |

*Table 18. Summary of the 13 Descriptors with Highest Occurrence Frequency*

| PaDEL descriptor | Appearance Frequency |
|---|---|
| **CrippenLogP** | 95 |
| **MW** | 85 |
| **XLogP** | 80 |
| **ZMIC1** | 74 |
| **ATS2m** | 67 |
| **ZMIC0** | 64 |
| **ATS3m** | 59 |
| **ATS1m** | 58 |
| **MLFER_L** | 54 |
| **ATS0p** | 54 |
| **VP-3** | 39 |
| **ATS0m** | 34 |
| **ZMIC2** | 34 |

*Table 19. Internal validation metrics and confusion matrices for top-performing model-descriptors combinations.*

| Model | Accuracy | Precision | Recall | F1 Score | Specificity | Confusion Matrix (TN, FP, FN, TP) | Combination of Columns |
|---|---|---|---|---|---|---|---|
| **RandomForestClassifier** | 0.84 | 0.84 | 0.86 | 0.85 | 0.82 | 912, 188, 144, 956 | 'ATS0p', 'ATS3m', 'ATS2m', 'VP-3', 'ATS0m', 'MLFER_L', 'ZMIC0', 'ZMIC1', 'MW', 'CrippenLogP', 'XLogP' |
| **ExtraTreesClassifier** | 0.85 | 0.83 | 0.88 | 0.86 | 0.82 | 904, 196, 123, 977 | 'ATS0p', 'ATS3m', |

| Model | Accuracy | Precision | Recall | F1 Score | Specificity | Confusion Matrix (TN, FP, FN, TP) | Combination of Columns |
|---|---|---|---|---|---|---|---|
| | | | | | | | 'ATS2m', 'VP-3', 'ATS0m', 'MLFER_L', 'ZMIC0', 'ZMIC1', 'MW', 'CrippenLogP', 'XLogP' |

*Table 20. External validation metrics and confusion matrices for top-performing model-descriptors combinations. External validation was conducted on eChemPortal database.*

| Model | Accuracy | Precision | Recall | F1 Score | Specificity | Confusion Matrix (TN, FP, FN, TP) | Combination of Columns |
|---|---|---|---|---|---|---|---|
| **RandomForestClassifier** | 0.40 | 0.09 | 0.54 | 0.16 | 0.38 | 36, 57, 5, 6 | 'ATS0p', 'ATS3m', 'ATS2m', 'VP-3', 'ATS0m', 'MLFER_L', 'ZMIC0', 'ZMIC1', 'MW', 'CrippenLogP', 'XLogP' |
| **ExtraTreesClassifier** | 0.44 | 0.05 | 0.27 | 0.09 | 0.46 | 43, 50, 8, 3 | 'ATS0p', 'ATS3m', 'ATS2m', 'VP-3', 'ATS0m', 'MLFER_L', 'ZMIC0', 'ZMIC1', 'MW', 'CrippenLogP', 'XLogP' |

*Table 21. Internal validation metrics and confusion matrices for top-performing model-descriptors combinations evaluated on the 80% training subset of the initial dataset.*

| Model | Accuracy | Precision | Recall | F1 Score | Specificity | Confusion Matrix (TN, FP, FN, TP) | Combination of Columns |
|---|---|---|---|---|---|---|---|
| **RandomForestClassifier** | 0.82 | 0.81 | 0.86 | 0.83 | 0.79 | 694,179, 124, 763 | 'ATS0p', 'ATS3m', 'ATS2m', 'VP-3', 'ATS0m', |

| Model | Accuracy | Precision | Recall | F1 Score | Specificity | Confusion Matrix (TN, FP, FN, TP) | Combination of Columns |
|---|---|---|---|---|---|---|---|
| | | | | | | | 'MLFER_L', 'ZMIC0', 'ZMIC1', 'MW', 'CrippenLogP', 'XLogP' |
| **ExtraTreesClassifier** | 0.83 | 0.82 | 0.86 | 0.84 | 0.80 | 703,170, 119, 768 | 'ATS0p', 'ATS3m', 'ATS2m', 'VP-3', 'ATS0m', 'MLFER_L', 'ZMIC0', 'ZMIC1', 'MW', 'CrippenLogP', 'XLogP' |

*Table 22. Internal validation metrics and confusion matrices for top-performing model-descriptors combinations evaluated on the 20% training subset of the initial dataset (held-out set).*

| Model | Accuracy | Precision | Recall | F1 Score | Specificity | Confusion Matrix (TN, FP, FN, TP) | Combination of Columns |
|---|---|---|---|---|---|---|---|
| **RandomForestClassifier** | 0.82 | 0.80 | 0.83 | 0.82 | 0.81 | 184, 43, 35, 178 | 'ATS0p', 'ATS3m', 'ATS2m', 'VP-3', 'ATS0m', 'MLFER_L', 'ZMIC0', 'ZMIC1', 'MW', 'CrippenLogP', 'XLogP' |
| **ExtraTreesClassifier** | 0.81 | 0.79 | 0.82 | 0.81 | 0.80 | 183, 44, 38, 175 | 'ATS0p', 'ATS3m', 'ATS2m', 'VP-3', 'ATS0m', 'MLFER_L', 'ZMIC0', 'ZMIC1', 'MW', 'CrippenLogP', 'XLogP' |

*Table 23. Internal validation metrics and confusion matrices for top-performing model-descriptors combinations evaluated on the 80% training subset of the initial dataset augmented via random oversampling.*

| Algorithm | Parameters | Most important features | Cross model features | Accuracy(%) | F1 score(%) | Confusion matrix | |
|---|---|---|---|---|---|---|---|
| **RandomForest** | criterion='entropy' n_estimators=150 | 'ATS2m', 'ATS0p', 'VP-4', 'ZMIC1', 'VP-5', 'ATS1m', 'MLFER_L', 'VP-3', 'MW', 'XLogP', 'CrippenLogP' | No | 96 | 96 | 3020 | 197 |
| | | | | | | 35 | 3182 |
| **Extratree** | N_estimators=100 | 'VP-3', 'CrippenLogP', 'MLFER_L', 'MW', 'XLogP', 'ATS1m', 'ATS0p', 'VP-5', 'ZMIC1', 'ZMIC0', 'VP-4', 'ATS2m', 'nHeteroRing', 'SpDiam_DzZ' | Yes | 97 | 97 | 3086 | 131 |
| | | | | | | 38 | 3179 |

*Table 24. Internal validation metrics and confusion matrices for top-performing model-descriptors combinations evaluated on the 20% training subset of the initial dataset (held-out set).*

| Algorithm | Parameters | Most important features | Cross model features | Accuracy(%) | F1 score(%) | Confusion matrix | |
|---|---|---|---|---|---|---|---|
| **RandomForest** | criterion='entropy' n_estimators=150 | 'ATS2m', 'ATS0p', 'VP-4', 'ZMIC1', 'VP-5', 'ATS1m', 'MLFER_L', 'VP-3', 'MW', 'XLogP', 'CrippenLogP' | No | 95 | 95 | 588 | 50 |
| | | | | | | 12 | 637 |
| **Extratree** | N_estimators=100 | 'VP-3', 'CrippenLogP', 'MLFER_L', 'MW', 'XLogP', 'ATS1m', 'ATS0p', 'VP-5', 'ZMIC1', 'ZMIC0', 'VP-4', 'ATS2m', 'nHeteroRing', 'SpDiam_DzZ' | Yes | 96 | 96 | 602 | 36 |
| | | | | | | 11 | 638 |

*Table 25. Cross-model validation metrics for LC50 models trained on MACCS and Morgan fingerprints after 10 epochs of classification on random test sets. The test sets were derived from a dataset with LC50 data from multiple fish species.*

| Fingerprints (radius, n_bits) | Classifier | Accuracy mean (± std) % | Specificity mean (± std) % | Sensitivity mean (± std) % | ROC AUC mean (± std) % |
|---|---|---|---|---|---|
| **Morgan (2, 1024)** | Random Forest | 75.81 (± 1.77) | 80.05 (± 1.77) | 63.50 (± 3.04) | 76.29 (± 1.47) |
| **Morgan (2, 2048)** | XGBoost | 75.64 (± 1.15) | 83.42 (± 1.95) | 53.09 (± 2.13) | 73.67 (± 1.28) |
| **Morgan (3, 1024)** | Random Forest | 76.36 (± 2.38) | 78.99 (± 3.98) | 67.51 (± 3.83) | 79.31 (± 1.10) |
| **Morgan (3, 2048)** | XGBoost | 75.07 (± 2.34) | 78.94 (± 3.04) | 62.09 (± 3.99) | 76.58 (± 2.46) |
| **Morgan (4, 1024)** | ExtraTrees | 76.98 (± 1.07) | 81.06 (± 1.43) | 63.28 (± 2.72) | 77.21 (± 1.77) |
| **Morgan (4, 2048)** | Gradient Boosting | 74.24 (± 1.84) | 78.89 (± 2.21) | 58.64 (± 1.33) | 74.06 (± 2.18) |
| **MACCS (-, 166)** | ExtraTrees | 75.23 (± 1.61) | 78.10 (± 2.94) | 65.59 (± 4.10) | 80.14 (± 1.70) |