ToxTwin V2.3 — User Guide

Twingital Institute · April 2026 · Version 2.3

> **Regulatory disclaimer.** ToxTwin is a computational decision-support tool. Its predictions do not replace the in vitro and in vivo tests required under ICH S2(R1), S7A and S7B guidelines. They do not constitute a regulatory opinion.

Table of Contents

1. What is ToxTwin?

ToxTwin is a toxicological prediction model that analyses the chemical structure of a molecule to estimate its probability of activity across 14 regulatory biological endpoints. It uses a graph neural network (GNN) trained on high-throughput screening data from the ChEMBL and Tox21 databases.

What ToxTwin does: - Estimate the preliminary toxicological risk of a molecule from its structure alone - Cover 12 Tox21 endpoints (nuclear receptors and cellular stress response) as well as hERG channel inhibition and Ames mutagenicity - Signal whether the molecule falls within the model's applicability domain - Generate a synthetic pharmacological interpretation

What ToxTwin does not do: - Replace mandatory biological tests (ICH S2, S7B) - Predict in vivo toxicity, pharmacokinetics or drug interactions - Provide a legally binding regulatory opinion - Guarantee an exhaustive toxicological profile for a drug candidate

V2.3 performance (5-fold scaffold CV strict):

Endpoint group	Mean AUC	Protocol
Tox21 (12 endpoints)	0.867 ± 0.043	5-fold scaffold CV
Ames mutagenicity	0.843 ± 0.029	5-fold scaffold CV
hERG inhibition	0.785 ± 0.053	5-fold scaffold CV

---

2. How to submit a molecule

2.1 Preparing the SMILES notation

ToxTwin accepts molecules in **SMILES notation** (Simplified Molecular Input Line Entry System) — the standard textual representation of chemical structure.

What is a SMILES?

A SMILES is a string of characters that describes the connectivity of a molecule. Examples:

Molecule	SMILES
Aspirin	`CC(=O)Oc1ccccc1C(=O)O`
Ibuprofen	`CC(C)Cc1ccc(cc1)C(C)C(=O)O`
Caffeine	`Cn1cnc2c1c(=O)n(c(=O)n2C)C`
Paracetamol	`CC(=O)Nc1ccc(O)cc1`

How to obtain a SMILES for your molecule:

- **PubChem** (free): search at pubchem.ncbi.nlm.nih.gov — the canonical SMILES is displayed in the compound record
- **ChEMBL** (free): ebi.ac.uk/chembl
- **ChemDraw / MarvinSketch**: draw the structure → Copy → SMILES
- **RDKit** (Python): `Chem.MolToSmiles(mol)`

> **Tip.** Always use the canonical SMILES rather than isomeric or arbitrary SMILES. ToxTwin applies RDKit canonicalisation upstream, but starting from a canonical SMILES reduces the risk of representation artefacts.

2.2 Demo interface

1. Open **twingital-ventures.com/en/achievements/toxtwin-demo/**
2. Paste your SMILES into the input field, or click a quick example (Aspirin, Ibuprofen, Caffeine, Pyrene)
3. Enter the molecule name (optional — for report readability)
4. Click **Analyse molecule** or press `Ctrl+Enter`
5. Wait for results (10–20 seconds depending on server load)

Freemium quota: 3 free analyses per visitor. Free registration unlocks 5 additional analyses. Professional API access available on request.

2.3 Valid and invalid SMILES

Situation	Behaviour
Valid SMILES, known molecule	Full prediction
Valid SMILES, out-of-domain molecule	Prediction with applicability domain warning
Invalid SMILES (syntax error)	Error message — verify with RDKit or PubChem
Very large molecule (> 500 heavy atoms)	Not recommended — likely out of domain
Mixture or salt (`.` in SMILES)	Largest fragment processed only

---

3. The 14 endpoints — what they mean

3.1 Nuclear Receptors (7 endpoints)

These endpoints assess a molecule's ability to activate or inhibit nuclear receptors — regulatory proteins involved in gene expression and endocrine disruption.

Endpoint	Receptor	Toxicological implication
NR-AR	Androgen receptor (full)	Androgenic endocrine disruption · developmental risk
NR-AR-LBD	Androgen receptor (LBD domain)	Structural specificity of the binding site
NR-AhR	Aryl hydrocarbon receptor	CYP1A1/1A2 induction · indirect genotoxicity
NR-Aromatase	Aromatase (CYP19A1)	Inhibition of oestrogen synthesis · endocrine disruption
NR-ER	Oestrogen receptor α (full)	Oestrogenic activity · hormonal risk

Endpoint	Receptor	Toxicological implication
NR-ER-LBD	Oestrogen receptor α (LBD domain)	Structural specificity of the binding site
NR-PPAR-γ	PPAR-gamma	Lipid metabolism · adipocyte differentiation

3.2 Cellular Stress Response (5 endpoints)

These endpoints assess a molecule's ability to trigger cellular stress pathways — indicators of genomic, mitochondrial or protein toxicity.

Endpoint	Pathway	Toxicological implication
SR-ARE	Nrf2-ARE (oxidative stress)	Electrophilicity · oxidative stress · antioxidant defence induction
SR-ATAD5	DNA replication (ATAD5)	Genomic instability · replication perturbation
SR-HSE	Heat shock response (HSP70)	Protein stress · denaturation · misfolding
SR-MMP	Mitochondrial membrane potential	Mitochondrial dysfunction · cytotoxicity
SR-p53	p53 pathway (genome guardian)	DNA damage · priority genotoxic signal

3.3 Pharmacotoxicology (2 endpoints)

Endpoint	Target	Regulatory reference
hERG	Potassium channel KCNH2 (hERG)	ICH S7B · responsible for > 30% of post-approval withdrawals
Ames	Salmonella mutagenicity	ICH S2(R1) · first mandatory genotoxicity test

> **Note on hERG and Ames.** These two endpoints have measured performance below the regulatory target in V2.3 (AUC 0.785 and 0.843 respectively). Their predictions are indicative and must be confirmed by the corresponding regulatory tests before any development decision.

---

4. Reading and interpreting results

4.1 Probability scores

Each endpoint returns a score between 0 and 1 representing the **calibrated probability** that the molecule is active on that endpoint.

Score	Risk level	Interpretation
< 0.25	**Low**	Non-significant signal in the model's context
0.25 – 0.50	**Moderate**	Signal to monitor — additional investigation recommended
0.50 – 0.75	**High**	Concerning signal — experimental investigation priority
> 0.75	**Critical**	Strong signal — mandatory experimental investigation

Important: these thresholds are indicative and must be contextualised relative to the molecule's structural profile, therapeutic class, and the overall toxicological profile.

4.2 Reading a complete profile

A complete toxicological profile is read at three levels:

1. Dominant signals — endpoints with score > 0.25. These are the priority points of attention. A high SR-p53 signal combined with a moderate SR-ARE signal suggests a coherent potential genotoxic activity (electrophilic reactivity, DNA damage, oxidative stress).

2. Structural coherence — signals should be consistent with the molecule's functional groups. A high hERG signal is expected for lipophilic molecules bearing a basic nitrogen. An NR-AhR signal is expected for planar aromatic compounds.

3. Applicability domain context — a high score on an out-of-domain molecule should be interpreted with increased caution (see section 5).

4.3 Interpretation example — Aspirin

SMILES: `CC(=O)Oc1ccccc1C(=O)O`

Endpoint	V2.3 Score	Level
NR-ER	0.031	Low
SR-ARE	0.008	Low
hERG	0.093	Low
Ames	0.029	Low

Endpoint	V2.3 Score	Level
All others	< 0.015	Low

Applicability domain: IN (Tanimoto 0.857 — molecule well represented in the training corpus)

Reading: Low risk profile across all 14 endpoints, consistent with Aspirin's well-documented pharmacological profile. The hERG signal (0.093) remains below the moderate concern threshold. The prediction is reliable (molecule within the applicability domain).

---

5. Applicability domain — when to trust the model

5.1 Definition

The **applicability domain** (AD) indicates whether the submitted molecule is structurally similar to molecules on which the model was trained. Outside this domain, predictions are **extrapolations** — potentially less reliable.

ToxTwin uses a composite AD score based on three signals:

Signal	Weight	What it measures
Tanimoto similarity	30%	Structural proximity to the nearest training corpus neighbour
k-NN distance (latent space)	40%	Distance in the learned representation space
KDE density	30%	Local density of the training corpus around the molecule

5.2 Interpreting the applicability domain badge

Badge	Meaning	Recommendation
✓ **In applicability domain**	Composite score > 0.40	Reliable prediction within the model's limits
⚠ **Outside applicability domain**	Composite score ≤ 0.40	Indicative prediction only — experimental validation priority

5.3 Common causes of out-of-domain results

- **Novel scaffolds** with no structural precedent in ChEMBL or Tox21

- **Unusual functional groups**: transition metals, charged halogens, highly strained rings
- **Peptides and oligonucleotides**: outside the model's scope (trained on small drug-like molecules)
- **Polymers and macromolecules**
- **Organometallic compounds**

> An out-of-domain score does not mean the molecule is toxic — it means the model has insufficient data to produce a reliable prediction.

---

6. The Phi-4 interpretation — what it is and what it is not

Each analysis is accompanied by a **textual interpretation** generated by Phi-4 14B, a scientific language model running locally on the Twingital Institute infrastructure.

What the interpretation provides

- Narrative synthesis of the most significant signals
- Correlation with the molecule's functional groups
- Overall risk level assessment (Low / Moderate / High / Critical)
- Regulatory context of the endpoints concerned

What the interpretation does not provide

- A regulatory opinion
- An in vivo prediction
- A guarantee of toxicological completeness
- A substitute for the judgement of a qualified toxicologist

> The interpretation is a starting point for reflection, not a conclusion. It must be reviewed and validated by a toxicology expert before any development decision.

---

7. Limitations and appropriate use cases

7.1 Appropriate use cases ✓

- **Early triage** of a compound library (computational screening before biological testing)

- **Mechanistic hypotheses** on the potential toxicity profile of a structure
- **Prioritisation** of endpoints to test first when designing experiments
- **Training and education** on structure-activity relationships in toxicology
- **Comparative monitoring** between structural analogues

7.2 Inappropriate use cases ✗

- **Regulatory decision** without experimental confirmation (ICH S2, S7B)
- **Substitute for the preclinical dossier** in a marketing authorisation application
- **Prediction for out-of-scope classes**: biologics, oligonucleotides, peptides > 50 amino acids
- **Assessment of mixtures or degradation products** without a valid individual SMILES

7.3 Endpoints not currently covered

ToxTwin V2.3 does not cover (planned for V3.0 and beyond):

- DILI (Drug-Induced Liver Injury)
- ClinTox (general clinical toxicity)
- Extended structural cardiotoxicity (hERG only currently)
- Phototoxicity
- Reproductive toxicity (DART)
- Metal complexes and organometallic compounds

---

8. Frequently asked questions

Q: My SMILES generates an "invalid" error. What should I do?

Verify the SMILES on PubChem or using the online tool cheminfo.org. The most common errors: unclosed parentheses, incorrect charges, unsupported atoms.

Q: The hERG score is high for my molecule. Should I abandon it?

No. A high score is a warning signal that justifies a patch-clamp test (ICH S7B), not a decision to stop development. Many approved drugs show moderate hERG scores without clinical QT interval prolongation.

Q: My molecule is outside the applicability domain. Are the scores useless?

No — they remain indicative. A strong signal (> 0.5) on a major endpoint warrants investigation even outside the domain. However, a weak signal outside the domain cannot be interpreted as an absence of toxicity.

Q: Do scores change between two identical submissions?

No. Inference is deterministic (dropout disabled in evaluation mode). Identical SMILES always produce identical scores.

Q: Can I submit a SMILES with defined stereochemistry?

Yes. ToxTwin takes stereochemistry into account in the molecular featurisation. Enantiomers may produce different scores if their biological profiles differ.

Q: How do I access the API to integrate ToxTwin into my pipeline?

Via the contact form. The REST API exposes the `/v1/score-toxicity` and `/v1/interpret` endpoints. Full technical documentation is available on request.

Q: Are submitted SMILES stored?

No. SMILES are used solely for real-time inference and are not recorded in our databases.

ToxTwin V2.3 · Twingital Institute / Qualees · April 2026

This document does not constitute a regulatory opinion. ToxTwin is a computational decision-support tool intended for pharmaceutical research and toxicology professionals.