Belfield, SJ, Basiri, H, Chavan, S, Chrysochoou, G, Enoch, SJ, Firman, JW, Gomatam, A, Hardy, B, Helmke, PS, Madden, JC, Maran, U, March-Vila, E, Nikolov, NG, Pastor, M, Piir, G, Sild, S, Smajić, A, Spînu, N, Wedebye, EB and Cronin, MTD

Moving towards making (quantitative) structure-activity relationships ((Q)SARs) for toxicity-related endpoints findable, accessible, interoperable and reusable (FAIR)

Research Article

Moving Towards Making (Quantitative) Structure-Activity Relationships ((Q)SARs) for Toxicity-Related Endpoints Findable, Accessible, Interoperable and Reusable (FAIR)

Samuel J. Belfield[1,2], Homa Basiri[1], Swapnil Chavan[3], Georgios Chrysochoou[1], Steven J. Enoch[1], James W. Firman[1], Anish Gomatam[1], Barry Hardy[4], Palle S. Helmke[5], Judith C. Madden[1], Uko Maran[6], Eric March-Vila[7], Nikolai G. Nikolov[8], Manuel Pastor[7], Geven Piir[6], Sulev Sild[6], Aljoša Smajić[5], Nicoleta Spînu[1], Eva B. Wedebye[8] and Mark T. D. Cronin[1]

[1]School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Liverpool, United Kingdom; [2]Department of Chemistry, University of Manchester, Manchester, United Kingdom; [3]Unit of Chemical and Pharmaceutical Toxicology, Research Institutes of Sweden (RISE), Södertalje, Sweden; [4]Edelweiss Connect, Basel, Switzerland; [5]Department of Pharmaceutical Sciences, University of Vienna, Vienna, Austria; [6]Institute of Chemistry, University of Tartu, Tartu, Estonia; [7]Research Programme on Biomedical Informatics (GRIB), Hospital del Mar Medical Research Institute (IMIM), Department of Medicine and Life Sciences (MELIS), Universitat Pompeu Fabra, Barcelona, Spain; [8]Technical University of Denmark, National Food Institute, Research Group for Chemical Risk Assessment and GMO, Lyngby, Denmark

Received November 16, 2024;
Accepted May 15, 2025;
Epub May 19, 2025;

Correspondence:
Mark Cronin, PhD,
School of Pharmacy and Biomolecular Sciences,
Liverpool John Moores University,
Byrom Street, Liverpool L3 3AF, UK
(m.t.cronin@ljmu.ac.uk)

OPEN ACCESS

ALTEX 42(4), 657-666.
doi:10.14573/altex.2411161

Abstract

(Quantitative) structure-activity relationships ((Q)SARs) are widely used in chemical safety assessment to predict toxicological effects. Many thousands of (Q)SAR models have been developed and published; however, few are easily available to use. This investigation applied previously developed findability, accessibility, interoperability, and reusability (FAIR) principles for *in silico* models to six published machine learning (ML) (Q)SARs for the same toxicity dataset (inhibition of growth of *Tetrahymena pyriformis*). The majority of principles were met; however, there are still gaps in making (Q)SARs FAIR. This study has enabled insights into, and recommendations for, the FAIRification of (Q)SARs, including areas where more work and effort may be required. For instance, there is still a need for (Q)SARs to be associated with a unique identifier and full data and/or metadata for toxicological activity or endpoints, molecular properties and descriptors, as well as model description to be provided in a standardized manner. A number of solutions to the challenges were identified, such as building on the QSAR Model Reporting Format (QMRF) and the application of the QSAR Assessment Framework (QAF). This study also demonstrated that resources such as the QSAR DataBank (QsarDB, https://www.qsardb.org) are valuable in storing ML QSARs in a searchable database and also provide a digital object identifier (DOI). Many activities related to FAIR are currently underway, and (Q)SAR modelers should be encouraged to utilize these to move towards easier access and use of models. Enabling FAIR computational toxicology models will support overall progress towards animal-free chemical safety assessment.

Plain language summary

This study relates to the availability of computational (termed *in silico*) models to predict harmful effects of substances based only on their chemical structure. The computational models referred to are (quantitative) structure-activity relationships (Q)SARs. Six machine learning models for toxicity were assessed against existing principles intended to make (Q)SARs findable, accessible, interoperable and reusable (FAIR principles). This highlighted several areas where improvements are needed to ensure the (Q)SARs are available for use. Currently there is no standard means to store (Q)SAR or provide a unique identifier; the QSAR DataBank (QsarDB) is illustrated as one possible solution.

1 Introduction

(Quantitative) structure-activity relationships ((Q)SARs) offer a means to predict biological activity and physico-chemical or pharmacokinetic properties from chemical structure alone. There is a particular need for, and focus on, the use of (Q)SARs with regard to animal-free chemical safety assessment under the umbrella of "computational toxicology" (Madden et al., 2020). These models can include approaches to predict adverse outcomes (AOs), molecular initiating events (MIEs), and key events (KEs) from ad-

verse outcome pathways (AOPs), as well as toxicokinetics and other properties (Cronin et al., 2022). As such, computational toxicology is an essential part of frameworks and strategies such as integrated approaches for testing and assessment (IATA) (Laroche et al., 2019) and next generation risk assessment (NGRA) (Yang et al., 2023), which aim to make risk and hazard assessment decisions based on various levels of information relating to exposure and hazard.

(Q)SARs, in particular, have found use in applications from rapid, high-throughput screening of molecules within inventories through to providing lines of evidence to support an overall weight-of-evidence (Barber et al., 2024a). As such, (Q)SARs and other computational methodologies are a key new approach methodology (often termed computational or *in silico* NAMs, meaning both computer simulations and computational models mimicking experiments or processes of physical laboratory work) and essential in non-animal chemical safety assessment (Westmoreland et al., 2022). Thus, the appropriate use of (Q)SAR and related technologies is seen as a fundamental component in the 3Rs (replacement, reduction, and refinement) of animal use (Laroche et al., 2019). (Q)SAR models have been demonstrated to have applicability to make predictions relevant for regulatory purposes (Bishop et al., 2024) and are applied in a variety of chemical legislations as alternatives to animal testing, one example being the European Union's Registration, Evaluation, Authorisation and restriction of CHemicals (REACH) regulation (ECHA, 2023).

The first use of the term "QSAR" is credited to Professors Hansch and Fujita and colleagues (Hansch et al., 1964). It is acknowledged, however, that the appreciation of a relationship between chemical structure, properties, descriptors and biological activity was known for over a century before modern definition of QSAR (Dearden, 2016). During their evolution, (Q)SARs have developed from the analysis of small data sets, either graphically or with linear regression analysis, through to the most recent approaches in machine learning (including recently deep learning) and other areas of artificial intelligence (Madden et al., 2020).

For use in chemical risk assessment, it is acknowledged that a (Q)SAR must fulfil established criteria for validity, such as the Organisation for Economic Cooperation and Development (OECD) principles for the validation of (Q)SARs (OECD, 2007). The OECD principles are the basis of the QSAR Model Reporting Format (QMRF), which provides a means of the textual documentation of a (Q)SAR[1]. From 2023, the OECD-adopted QMRF is part of the QSAR Assessment Framework (QAF), which has provided a means of documenting and assessing a prediction from a (Q)SAR, with a view to understanding uncertainties within the prediction (OECD, 2023; Gissi et al., 2024; Barber et al., 2024b). Whilst these methods, in addition to copious guidance from international agencies, provide a means of assessing models, it has been proposed that models should also adhere to the findability, accessibility, interoperability, and reusability (FAIR) principles (Wilkinson et al., 2016), originally intended to guide scientific data management and stewardship.

For a (Q)SAR to have practical application, it must be retrievable and usable. There are an unknown number of (Q)SARs in the published literature, possibly thousands of models across a myriad of endpoints. Most of these are, in reality, not usable (Piir et al., 2018). There are many reasons for the non-functionality of the majority of published models. For instance, whilst the QMRF provides a means of describing the model, not all modelers have adopted its use. A full, interpretable and transparent description of the model is not provided for many models, particularly those that precede the development of the QMRF. Neither is there a standardized method to store, search for and retrieve (Q)SARs. This suggests both a missed opportunity to use computational toxicology models but also insufficient attention and support by the modelling community and their practices to ensure the sustainability of their models.

In response to the issues related to making (Q)SAR models sustainable and usable, the existing principles to enable scientific data to be FAIR (Wilkinson et al., 2016) were adapted by Cronin et al. (2023). Eighteen FAIR principles for *in silico* toxicology models were proposed with the intention that they would allow models to be more usable, with a particular focus on improving regulatory applicability and acceptance. In addition, a number of resources were identified that may assist in making (Q)SAR models, in particular, FAIR. There are only limited resources that may be applied to make (Q)SAR models available, notable amongst these are the JRC QSAR Model Database[2]; BioModels (Glont et al., 2018; Malik-Sheriff et al., 2020); Harvard Dataverse[3]; the Danish (Q)SAR Database[4]; the ConcertREACH Gateway for predictive computational (QSAR) toxicology models[5], and the QSAR DataBank (QsarDB)[6] repository from the University of Tartu, Estonia (Ruusmann et al., 2015). Of these, the JRC QSAR Model Database is a static and historic resource, which only allows downloads of QMRF documentations of models, and the Harvard Dataverse and BioModels contain a paucity of (Q)SAR models. The Danish (Q)SAR Database is mainly a repository of predictions for 650,000 chemicals from freely available and commercial (Q)SAR models but also has the Danish (Q)SAR Models web application that contains in-house models. The ConcertREACH Gateway provides details of available software for toxicity prediction and links to the models. The QsarDB[6] provides a platform with the potential to meet many of the criteria of the FAIR principles, i.e., it is searchable in terms of chemistry, endpoint, etc. It is robust, interoperable, and provides access to (Q)SARs in a FAIR manner.

Further to limited resources for ensuring availability of (Q)SAR models, little is known about the FAIRness of existing (Q)SAR models for toxicity, especially those employing more complex modelling approaches, such as neural networks. The aims of this

[1] https://one.oecd.org/document/ENV/CBC/MONO(2023)32/ANN1/en/pdf

[2] http://data.europa.eu/89h/e4ef8d13-d743-4524-a6eb-80e18b58cba4

[3] https://dataverse.harvard.edu/

[4] https://qsar.food.dtu.dk/

[5] https://www.life-concertreach.eu/results/results-gateway/

[6] https://qsardb.org/

investigation were to evaluate existing (Q)SAR models for toxicity according to the previously published FAIR principles (Cronin et al., 2023) and identify areas where improvements are needed. The models chosen for assessment were six machine learning (ML) models developed on a single dataset and published previously by the first and corresponding authors of this manuscript (Belfield et al., 2023). From the analysis of these six existing models, a series of recommendations to promote the FAIRness of (Q)SAR models for toxicity were established.

2 Materials and methods

2.1 Retrieval of (Q)SAR models

Six machine learning QSAR models from Belfield et al. (2023) were selected for evaluation. The models were selected on the basis of being freely available, representing a variety of machine learning approaches, and having reasonable performance, with comparatively well-understood data quality and mechanistic interpretability. These models utilized six different ML algorithms (random forest, XGBoost, support vector machine, k-nearest neighbors, neural network, and deep neural network) to predict *Tetrahymena pyriformis* growth inhibition.

2.2 Evaluation of (Q)SARs according to FAIR principles

Models were assessed following the previously reported FAIR principles for *in silico* models (Cronin et al., 2023), with each principle being assessed qualitatively, accompanied by the rationale for the assessment as well as potential strategies for improving adherence to the FAIR principles, where appropriate. Due to the development and approach for each model being identical (except for the machine learning algorithm employed), models were evaluated against the FAIR principles as a singular set of models.

2.3 Uploading models to the QsarDB

(Q)SAR models were incorporated into the QsarDB. To achieve this, the original QSAR models were reproduced with the code from Jupyter notebooks in GitHub[7] and converted into Open Neural Network Exchange (ONNX) standardized format using the open-source Python library onnxmltools (version 1.11.2), based in part on the information collected by the QMRF. This conversion step is crucial for maintaining compatibility across various platforms and facilitating the integration of models into diverse software environments. It is acknowledged that conversion to ONNX from an original format may alter the original model, so there is a need to check the performance and statistics of the model. In addition, the ONNX capture does not necessarily go beyond or replace the QMRF but allows for a workable version of the model to be stored. Models in ONNX then became an integral part of the QsarDB data format. This involved preparing comprehensive metadata that describes the models' purpose, methodology, the specific endpoints it predicts, and complete structural data to derive and validate the model. The platform assigns a permanent digital object identifier (DOI) to the uploaded model. Appropriate licensing options were selected, i.e., Creative Commons licenses, which clarify the terms of use for the model.

3 Results and discussion

3.1 Evaluation of the six (Q)SARs for toxicity according to the FAIR principles

Eighteen principles covering all aspects of the FAIR concept have been developed and adapted for *in silico* models (Cronin et al., 2023). Each of the individual principles provides guidance and considerations for developers that will foster a model that has been produced, labelled and stored in a manner that fully promotes shareability and can be categorized as FAIR. Evaluation of models through application of the FAIR principles can highlight issues within a given workflow, which in turn may be hindering shareability. This study investigated the ability of the principles to be utilized in this manner to provide practical information on the models with regard to whether they are FAIR, and allowing for recommendations to be made.

The ability of the six previously developed machine learning models (Belfield et al., 2023) to satisfy the FAIR criteria is summarized in Tables 1-4. In total, 20 FAIR criteria were considered due to two of the original 18 FAIR principles (F2 and A1) having two parts. As the development of each model was identical, and they only differed regarding the algorithm utilized, all models were evaluated together as a single set. To this end, each principle was considered individually and a verdict recorded as to whether or not the principle was met (yes, no or partially) with the accompanying reasoning being recorded for transparency. As seen in Tables 1-4, the majority (11/20) of principles were sufficiently satisfied; however, six principles were not met, and the remaining three principles were met only partially. As it is essential that, for a model to be considered FAIR, all principles are sufficiently satisfied, strategies for improvement were suggested where appropriate.

The assessment of the models with regard to their "findability" is summarized in Table 1. The "Findable" principles relate, in part at least, to the ability to search for and retrieve a particular (Q)SAR, and it being adequately described. These principles are fundamental to being able to find relevant (Q)SARs for a particular purpose. When assessed in terms of being findable, the previously published models were found lacking due to the absence of unique identifiers and adequate metadata. Currently, there is no agreed and accepted unique identifier for a (Q)SAR, and most identifiers relate simply to the description of the model, which does not adhere to any standard format. A proposed solution to this issue would be the allocation of a DOI, a well-established standard for publications and other items. Whilst the DOI is non-descriptive, it is a globally unique and definitive identifier that can be directly accessed or searched for using an internet search engine. This identifier should allow the QSAR model to be findable even if its digital location changes.

[7] Datasets and Python source code employed for the processes of model construction, optimisation and performance assessment for the models evaluated were obtained from Belfield et al. (2023). https://github.com/LJMU-Chemoinformatics/Best-Practice-Supplementary (accessed 14.11.2024)

Tab. 1: Assessment of the "findability" of the set of six machine learning QSAR models with regard to the FAIR principles (Cronin et al., 2023) and associated improvement strategies

FAIR principle (from Cronin et al., 2023)	Verdict and reasoning	Comments, examples and improvement strategies, where applicable
F1. Each model is assigned a globally unique and persistent identifier, and different versions are assigned distinct identifiers.	**No.** Models were only assigned local identifiers that are associated throughout development. These identifiers are descriptive but would not be suitable for searching or cataloguing.	There is a clear need for the models to be assigned a unique global identifier such as a DOI. This would enable cataloguing and searching for the model.
F2. Models are described with rich metadata covering all aspects of the model.		
F2.1. Models are associated with searchable metadata for the property or endpoint to be predicted.	**Yes.** Models are developed with searchable metadata which is publicly available for the endpoint of interest.	All the data used in model development are available on GitHub as comma separated (.csv) files. This includes appropriate annotation and description of variables.
F2.2. Models are associated with searchable metadata or descriptions of the chemicals (e.g., InChI or SMILES), or chemical class(es), within the model, or a description of its applicability domain.	**Partially.** Unique chemical identifiers (CAS numbers, IUPAC name) are provided for each chemical in the metadata and used within the model. However, applicability domains were not defined, with the exception that descriptor ranges could be identified.	Models need to be associated with a clearly defined applicability domain. It is noted that the applicability domain may vary with different model types, so it would need to be defined for each machine learning QSAR.
F3. Models' (meta)data clearly and explicitly include the identifier of the model they describe and are registered or indexed in a searchable resource.	**No.** The metadata associated with the model (its identifier(s)) is minimal, providing only sparse information regarding the algorithm and data utilized. The type of ML model is described, but the model itself is not articulated.	Models should be associated with appropriate identifier(s) that describe the key characteristics including endpoint, modelling approach and type of data.
F4. Models are registered or indexed in a searchable resource.		
F4.1 Models' identifiers should be optimized to allow for use in multiple search engines.	**No.** The identifiers of the models are not available in a searchable resource.	Once the identifiers are established, they need to be stored within a searchable resource. A minimum is the provision of a DOI, but also description of the endpoint, e.g., species, test, duration, in addition to the types of chemicals tested, e.g., small organic molecules (non-pharmaceutical/biocide). Information on putative mechanisms of action may also be helpful.

There is also a need to standardize the description of the type of model and data contained within the model, for instance, building on the QMRF as implemented in the QAF. This could include a formalized approach to reporting, describing and providing the dependent and independent data as well as the modelling approach. Standardization of model reporting is crucial for ML approaches where terminology and descriptions can vary. In terms of making (Q)SARs findable, this would assist in being able to search for them using standard terms, e.g., an endpoint, species or type of (Q)SAR model. For a (non-commercial) model to be reproducible, many details need to be gathered and stored, especially relating to the creation of the model (see Cronin et al. (2023) and Piir et al. (2018) for details). In addition, there is an essential requirement to provide and store the applicability domain of a model. However, the applicability domain can be defined in a number of ways (Dimitrov et al., 2005; Netzeva et al., 2005) and may provide different metrics even for the same data set when modelled with different approaches. Applicability domains can be calculated automatically according to the OECD (Q)SAR Validation Principles if full data about the model are available (see the realization for regression models in the QsarDB repository).

The assessment of the QSAR models in terms of being "accessible" is given in Table 2. The models assessed were published (Belfield et al., 2023), and the (meta)data (for endpoint and descriptors) and Jupyter notebooks for reproducing the models are available via GitHub[7]; thus, the models can be considered to meet the "Accessibility" principles. It should be noted, however, that even when a model and its description are available, this does not necessarily mean that the model is functional and can be used to make a prediction. In addition, there may be limitations to the size of the model and associated data that can be stored, especially if free-to-use resources are utilized.

The assessment of the QSAR models in terms of their "interoperability" is summarized in Table 3. Whilst the machine learning models and metadata in this study were well described, there is no comprehensive and standardized ontology that is generally accepted for describing QSAR models, descriptors, performance statistics, etc. However, some efforts could be built upon and developed further to help resolve this issue, such as the OpenTox framework (Hardy et al., 2010, 2012), OntoQSAR (Angelo et al., 2020), and work towards interoperable QSAR datasets (Spjuth et

Tab. 2: Assessment of "accessibility" of the set of six machine learning QSAR models with regard to the FAIR principles (Cronin et al., 2023)

FAIR principle (from Cronin et al., 2023)	Verdict and reasoning	Comments, examples and improvement strategies
A1. Models are retrievable by their identifier using a standardized communications protocol.		
A1.1. The model (and any associated protocol represented by the model metadata) is openly accessible or reimplementable.	**Yes.** Models are openly accessible and stored within a public repository on GitHub from where they could be reimplemented.	The models are available on GitHub as .ipynb files that can be executed in a Jupyter notebook. All data are also available as .csv files, thus the models can be reimplemented and reproduced.
A1.2. The model (and any associated protocol) allows for an authentication and authorization procedure, where necessary.	**Yes.** Models' full developments are publicly available; GitHub platform implements access control mechanisms to protect this data from unauthorized changes.	In this context, GitHub implements access controls.
A2. Model (meta)data are accessible even when the model is no longer available, unless restricted for commercial, ethical or data protection reasons (e.g., blinding of confidential chemical structures).	**Yes.** Metadata are openly accessible and stored within a public repository on GitHub. However, the completeness and quality of the metadata should be assured.	There are no restrictions, e.g., confidentiality or ethical, on the data; therefore, all data are available. There are instances where the metadata could be more comprehensive, e.g., in describing the endpoint and descriptors in detail.

Tab. 3: Assessment of the "interoperability" of the set of six machine learning QSAR models with regard to the FAIR principles (Cronin et al., 2023) and associated improvement strategies

FAIR principle (from Cronin et al., 2023)	Verdict and reasoning	Comments, examples and improvement strategies
I1. The models and their (meta)data are described in a standardized manner, i.e., standards to define chemical structures, endpoints, molecular descriptors and modelling algorithms.	**Yes.** Models and their metadata are described and annotated. The data are described with terminologies for, e.g., endpoint, descriptors, that are well-defined and commonly used, whilst not officially standardized, e.g., in an accepted ontology. The model and data description are located within the associated documentation to URLs for the models in QsarDB as noted in Table 5.	Whilst this principle is met, it may be optimal to align the description with other accepted formats such as applied QMRF/QAF templates and QDB archive (Ruusmann et al. 2014). This principle also emphasizes the requirement for standardized approaches to all elements of a (Q)SAR.
I2. The model reads, writes and exchanges data in a way that meets domain-relevant community standards.	**No.** Models exchanging information did not follow domain-relevant community standards such as QMRF or QAF.	Community standards could be utilized, for instance as applied in the QMRF, QAF, OpenTox (Hardy et al., 2010; 2012). Currently, there is a paucity of community standards, so the model developer should use what is currently best practice, e.g., QMRF, QAF, etc.
I3. The model must be interoperable with other software, e.g., with a clearly defined input/output, i.e., with an appropriate application programming interface (API) for shared web services.	**Partially.** Clearly defined inputs and outputs for the models are outlined; however, no standardized, well-documented output for interoperability currently exists for them.	Models need to be machine actionable, i.e., further developed with clear consideration for how interoperability could be achieved, e.g., for shared web services. This will be essential to implement the models, make them available for, e.g., tiered testing strategies and/or automization.
I4. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.	**Partially.** Metadata are described using well utilized, although not always accepted, identifiers.	Identifiers need to be standardized, for instance aligning with QMRF, QAF. This will enable others to understand and utilize the model.
I5. (Meta)data use vocabularies that follow FAIR principles.	**Yes.** Metadata adhere to the FAIR principles and are provided in .csv files on GitHub.	As well as being good practice, this will form the basis for rational methods to search for the model, i.e., standardized search terms for endpoint, model type, number and type of descriptors, etc.
I6. The model includes qualified references to other objects, such as molecular descriptors.	**Yes.** Objects outside of the original metadata that have been produced are appropriately referenced to original sources with the information additionally being publicly available.	The other objects in the models are the molecular descriptor set. Not only are the descriptor values given, but the software is named. Since the chemical identifiers are provided, if required, the descriptors could be recalculated, updated or extended.

Tab. 4: Assessment of the "reusability" of the set of six machine learning QSAR models with regard to the FAIR principles (Cronin et al., 2023) and associated improvement strategies

FAIR principle (from Cronin et al., 2023)	Verdict and reasoning	Comments, examples and improvement strategies
R1. The model is available for use in some format (e.g., source code, executable, library or service).	**Yes.** The models are available for their intended use within an executable source code.	The models are available on GitHub as .ipynb files. Any similar approach or implementable file type would be suitable for this purpose.
R2. The usage license of the model should be clearly defined and appropriate to encourage its use.	**No.** No usage license for the models has been provided.	Models need to be accompanied by a usage license, such as Creative Commons, etc., that actively encourages their usage. Clear licensing establishes how and when the model can be used and whether it can be incorporated into further software / predictive toxicology tools.
R3. The storage of the model and (meta)data should be done on a sustainable and future-proofed platform, anticipating the impact on the availability of software changes over time.	**Yes.** Models and metadata are stored within a public repository on GitHub, which is globally accepted as a sustainable platform. It is appreciated that this can only be for a reasonable time period, though it should be made as robust as possible.	The use of .csv files means the data can be converted into formats suitable for other modelling tools. As the standards for future file types and software are unknown, storing generic file types is recommended.
R4. Software includes qualified references to other software, e.g., so that the correct molecular descriptors can be obtained, either as part of the model or storage of the molecular descriptors software or experimental protocol.	**Yes.** All software used throughout the development of the models and production of descriptors for the metadata are accurately referenced in the associated publication. Software version numbers are included.	Whilst software assessed in this investigation is available and calculations could be repeated, there should be consideration of being able to store the software itself for future use, e.g., for making predictions for chemicals at a later stage and outside of the model space.
R5. (Meta)data are richly described with a plurality of accurate and relevant attributes		
R5.1. The model and its (meta)data are associated with detailed provenance	**Yes.** The origins of the metadata are clearly provided, including version number, software provider, within the associated publication.	The metadata are clearly described with appropriate references.
R6. The model and its (meta)data meet domain-relevant community standards for documentation	**No.** Domain-relevant community standards for data documentation are unavailable.	Once community standards for data documentation are proposed, such procedures must be adhered to. These could utilize and/or extend the headings utilized in the QMRF template. Community standards are required and will increase the FAIRness of (Q)SARs as well as making them more usable in the future.

al., 2010), Chemical Information Ontology (Hastings et al., 2011), and QSAR descriptor ontologies[8,9]. The other key area to make QSAR models fully interoperable is to allow the models to communicate with other software, for instance, with an application programming interface (API) or standard output format.

FAIR principles for QSARs should also incorporate the concept of "machine actionability" (Principle I3). Wilkinson et al. (2016) used the term to describe a continuum whereby a digital object provides increasingly more detailed information to an autonomously acting computational resource that intends to utilize the data. Further, being "machine actionable" means that if a computational resource approaches data that it has not seen before, it is able to determine what the data (or object) are, the potential use of the data (or object) for the intended purpose, assess usability with regard to licenses, etc., and take appropriate action. FAIR QSARs being machine actionable indicates that other computational resources can use the data (both the modeled activity and descriptor set) and models automatically. This principle was demonstrated within the OpenTox framework (Hardy et al., 2010, 2012) and should allow data to be made available for purposes such as the automatic evaluation, development and extension of QSAR models. For this purpose, APIs can provide easy computational access to resources and for FAIR data and knowledge use. However, there are several caveats that should be borne in mind with regard to machine actionable APIs in QSAR development. This is an area where standardized vocabularies and ontologies are crucial to enable a third-party computational resource to access and utilize data. In addition, at least in the short term, there may be a requirement for human intervention in areas such as interpreting and assessing data reliability and quality.

The assessment of the QSAR models for being "reusable" is summarized in Table 4. The models published by Belfield et al.

8 http://edamontology.org/data_0847

9 http://qsar.sourceforge.net/dicts/qsar-descriptors/index.xhtml

Tab. 5: DOI and citations for the QSAR models for the prediction of the inhibition of growth of *Tetrahymena pyriformis* from Belfield et al. (2023) entered into QsarDB

Machine learning methods	QSAR Databank citation and URL
Random forest	Chrysochoou, G.; Sild, S. Model RF from: Guidance for good practice in the application of machine learning in development of toxicological quantitative structure-activity relationships (QSARs). QsarDB repository, QDB.264. 2024. doi:10.15152/QDB.264.RF Support vector machine
Support vector machine	Chrysochoou, G.; Sild, S. Model SVM from: Guidance for good practice in the application of machine learning in development of toxicological quantitative structure-activity relationships (QSARs). QsarDB repository, QDB.264. 2024. doi:10.15152/QDB.264.SVM Extreme gradient boosting
k-nearest neighbors	Chrysochoou, G.; Sild, S. Model KNN from: Guidance for good practice in the application of machine learning in development of toxicological quantitative structure-activity relationships (QSARs). QsarDB repository, QDB.264. 2024. doi:10.15152/QDB.264.KNN
Extreme gradient boosting	Chrysochoou, G.; Sild, S. Model XGB from: Guidance for good practice in the application of machine learning in development of toxicological quantitative structure-activity relationships (QSARs). QsarDB repository, QDB.264. 2024. doi:10.15152/QDB.264.XGB
Neural network	Chrysochoou, G.; Sild, S. Model SNN from: Guidance for good practice in the application of machine learning in development of toxicological quantitative structure-activity relationships (QSARs). QsarDB repository, QDB.264. 2024. doi.10.15152/QDB.264.SNN
Deep learning neural network	Chrysochoou, G.; Sild, S. Model DNN from: Guidance for good practice in the application of machine learning in development of toxicological quantitative structure-activity relationships (QSARs). QsarDB repository, QDB.264. 2024. doi:10.15152/QDB.264.DNN

(2023) were generally reusable, although specific information, such as the lack of a license for use, was absent. Again, this FAIR principle emphasizes the need for the use of a standard ontology. There may also need to be greater consideration of how third-party software may be versioned and can be made available for later use. For instance, if software is used to calculate descriptors, it will need to remain available for the future application of the (Q)SAR model. This may require that the model storage utilizes robust and sustainable software, or that descriptors can be obtained by other means while still providing the same values.

The reusability of a (Q)SAR should, ideally, go beyond simply downloading and using the model to make further predictions. Should the metadata and characteristics of the model be available but undefined, then it should be possible to recalculate, or recalibrate, a model. This would be highly valuable to ensure the model is reproducible and inevitably requires that the metadata are available. Having the (meta)data available also means that the model can be redeveloped, for instance if there is new information regarding mechanism of action or possible outliers, poor quality data, etc. The model can also be extended through enrichment of the training data set or descriptor set. For instance, this may become possible should further data become available when more testing is undertaken. Thus, reusability is a vital characteristic within the FAIR criteria, and this will allow for improved and extended models to be developed in the future, providing full metadata and description of the model is given.

3.2 Use of QsarDB to reduce non-compliance with the FAIR principles

In order to resolve the non-compliance with the FAIR principles, the models were entered in the QsarDB (Ruusman et al., 2015). QsarDB provides a means of organizing and archiving (Q)SAR models into QDB archive format (Ruusman et al., 2014). The structured approach to integrating (Q)SAR models into the QsarDB platform not only enhances the usability of these models but also promotes their sustainability and accessibility for future research. By ensuring compatibility, providing rich metadata, and assigning permanent identifiers, the QsarDB platform contributes to the advancement of (Q)SAR research and its applications in various fields. At the time of writing, QsarDB contained over 550 QSAR models. The models entered into QsarDB are assigned a DOI and can be used for data citations; Table 5 provides the DOI and citation information for the models considered in this investigation.

The entry of the models in QsarDB enabled a number of issues with the non-compliance with the FAIR principles to be resolved. QsarDB is dedicated to making (Q)SARs machine readable, interactive, predictive, and ultimately FAIR. It has received the CoreTrustSeal, meeting the requirements for trustworthy data repositories[10]. QsarDB enables not only the efficient storage of (Q)SAR models but also allows the user to search for models by various criteria such as endpoint, species, and chemistry, and importantly allows visualization of model content and to perform predictions. The QsarDB approach also provides technologies that assist in the resolution of the concerns raised in this investigation. For instance, the use of a DOI (e.g., as provided in QsarDB) provides a solution to the challenges raised with regard to F1, F3 and F4. DOIs are designed to be searchable, e.g., within an internet search engine, which increases the chance of a (Q)SAR model being found. QsarDB also provides a means to communicate with selected software, with an API being available (Principle I3). It can also provide a license for use of a model (Principle R2) with models being made available under the Creative Commons license.

[10] https://qsardb.org/blog/coretrustseal-certificate-awarded

QsarDB also provides a means of standardizing terminology with the platform (Principles I2 and R6), but this remains a significant issue for the broader (Q)SAR community, i.e., there is no formal agreement on community standards or ontology. QsarDB may also provide a sustainable platform for obtaining reliable descriptors and applying algorithms when other software may not be available.

QsarDB also supports uploading QMRF documents without fully disclosing the model, which helps the community by making documentation about commercial (Q)SAR models available. It should be noted that it is not the intention of the FAIR principles that all models should be transparently disclosed. This would be restrictive for complex models, such as generative AI, and specifically for commercial models. Instead, for these models, the FAIR principles intend to allow the model to be findable, described (as is possible with the QMRF), and identifiable.

While other resources, such as the large, cloud-based, storage sites Zenodo, GitHub, Figshare, etc. provide a means to store data and models, they do not ensure compliance with all FAIR principles, i.e., they do not provide an intuitive and searchable resource or encourage application of standard ontologies or community standards. Some of these issues are resolved elsewhere, e.g., BioModels and Harvard Dataverse; however, these are not bespoke to QSARs and have few representative models.

3.3 The importance of making QSARs FAIR

The FAIR concept has been applied broadly to data (Wilkinson et al., 2016) but also specifically in toxicologically relevant areas. For example, Wittwehr et al. (2024) demonstrated the importance of the FAIR principles to AOPs. Elsewhere, Ammar et al. (2024) developed a NanoSafety Data Reusability Assessment (NSDRA) framework to summarize the reusability of nanosafety datasets based on FAIR maturity indicators, including for human toxicity. Briggs et al. (2021) illustrated the benefits of data sharing for preclinical safety assessment in drug development following the FAIR principles, including reducing time spent curating, transforming, and aggregating datasets, allowing more time for data mining and analysis. Other practical recipes on how to implement FAIR for real-world settings include the FAIR Cookbook[11], created by biopharmaceutical and academic professionals, and guidance on data management practices, such as the RDMKit[12], which highlight how FAIR principles are the backbone for the creation of open systems leading to improved reproducibility and quality of data. In addition, organizations such as the GO FAIR Foundation[13] ctively promote the interpretation and implementation of FAIR principles to data as well as community standards, and training and tools to assess FAIR levels[14].

It seems logical, therefore, that (Q)SARs developed to support chemical safety assessment should be FAIR. This will enable models to be used across the community where ensuring (Q)SARs are reproducible, regardless of purpose, is essential. Reproducibility is particularly crucial with regard to regulatory assessment, where a third party (for instance a governmental agency) may wish to replicate, or better understand, a prediction. There is also a duty to make models available, particularly when they have been funded by public resources, and to make them sustainable. It should be remembered, however, that making a model FAIR does not imply all models should be open source – there is still a place for commercially sensitive models to be FAIR up to a level where commercial property is still protected but the necessary information is shared. For example, a model may be created using confidential information, which cannot itself be made publicly available. Alternatively, the model itself, or knowledge contained therein, may be commercially important. The FAIR principles do not imply that confidential knowledge and/or data or intellectual property should be freely accessible, but rather that there is sufficient access to the model information that is searchable.

Consideration of the FAIR principles in this investigation has revealed key areas for improvements. The first area is in the development of searchable databases to find relevant information and models. This would go beyond a simple internet or literature search, as it could incorporate not only the endpoint but also aspects of the chemistry, i.e., an intelligent search query could retrieve (Q)SARs for a particular endpoint and a particular chemistry. Further, there is an opportunity to make, where there are no commercial restrictions, models transparent, with data and algorithms presented according to relevant community standards. Means of capturing (Q)SARs are commonly applied, such as the QMRF, which go a long way to providing a framework that can be built upon and extended. Other approaches include the proposal of the OpenTox Framework (Hardy et al., 2010; 2012). Whilst the OpenTox Framework predates the FAIR principles, it is closely aligned by providing an interoperable standards-based framework supporting toxicology data management and reporting.

The implementation of the FAIR principles can utilize a commonly applied means of capturing and evaluating (Q)SARs (and their predictions). For instance, the OECD QAF (OECD, 2023) requires information on not only the model, e.g., goodness of fit, and mechanistic interpretation, but also on the assessment of predictions themselves, such as reliability and fitness for regulatory purposes. QAF was recently adapted for multiple predictions informed by multiple (Q)SARs (Gissi et al., 2024). Both QAF and FAIR are applicable irrespective of the modelling method used to build a model or the endpoint for which it was developed, which ensures real-world adoption. It is intended that the FAIR principles go beyond the documentation required by QAF to allow for full data and the model to be made available within a searchable form. As acknowledged by Barber et al. (2024b), trusting an *in silico* model needs to be questioned before the output is used to support a regulatory decision. FAIR principles allow us to evaluate this aspect as well as encourage the development of open benchmarks to facilitate the testing and increase the validity of such models.

Enabling FAIR (Q)SARs will require resources to be made available. (Q)SAR developers who wish to support FAIR models

11 https://faircookbook.elixir-europe.org

12 https://rdmkit.elixir-europe.org

13 https://www.gofair.foundation/

14 https://www.fairsfair.eu/

can do so as part of the modelling process. However, key elements of FAIR will require physical, financial and intellectual input. A resource to store and distribute (Q)SARs, such as any of those mentioned in this article, requires space on a hardware platform. The system requires funding for upkeep, maintenance, and as appropriate, training and dissemination. For robust sustainability, a reliable funding source is required with backup and contingency planning. Finally, intellectual input will be required in the creation and acceptance of community standards for topics such as ontologies. Utilizing working groups, such as those at OECD, will enable uptake of the outputs, but will require appropriate resources, both financially and in person months, and for member nations and other stakeholders to provide input.

3.4 Recommendations for making (Q)SARs FAIR

This investigation provides insights into making (Q)SAR models for toxicity prediction FAIR. There are considerable long-term benefits to this ambition, although for success there needs to be global vision and uptake – this may only be possible with appropriate international agreement and collaboration.

- There should be a move towards FAIR (Q)SAR models with greater emphasis on understanding what this means and a particular focus on how to make commercially or business-sensitive models FAIR.
- (Q)SAR models should be created using machine actionable descriptors and algorithms that can still be obtained in the future, i.e., when the software versions may have been superseded.
- Sustainable, searchable databases are required to store and retrieve models; one solution is QsarDB.
- There is a clear need for agreed community standards for reporting models, the underlying data and descriptors, as well as performance statistics. This could potentially include and further develop the QMRF.
- Journals and other means of reporting models should be encouraged to ensure that models are FAIR.
- The cost of providing sustainable and FAIR (Q)SAR models must be appreciated and budgeted.
- Greater efforts should be made to ensure functional models are interoperable with other software to enable their use.

4 Conclusions

The study has evaluated some existing machine learning (Q)SAR models for toxicity according to previously published FAIR principles. There were a number of key areas where the QSAR models did not comply with the FAIR principles. The models lacked a unique identifier, the full data and metadata were not presented, and licenses for use of the models were not available. Such limitations can be overcome using existing resources such as QsarDB (amongst others). A fundamental area that will require more concerted effort will be the definition of community standards. To retain credibility, (Q)SAR modelers should strive to make their models FAIR. However, addressing the concerns raised in this investigation will require a community effort to utilize current solutions and develop, where necessary, further solutions and allocation of resources.

References

Ammar, A., Evelo, C. and Willighagen, E. (2024). FAIR assessment of nanosafety data reusability with community standards. *Sci Data 11*, 503. doi:10.1038/s41597-024-03324-x

Angelo, R. M., Andreia, K. I., Almeida, M. P. et al. (2020). OntoQSAR: An ontology for interpreting chemical and biological data in quantitative structure-activity relationship studies. IEEE 14th International Conference on Semantic Computing (ICSC), San Diego, CA, USA. pp. 203-206. doi:10.1109/ICSC.2020.00042

Barber, C., Fowkes, A., Hanser, T. et al. (2024a). From model performance to decision support – The rise of computational toxicology in chemical safety assessments. *Comput Toxicol 31*, 100303. doi:10.1016/j.comtox.2024.100303

Barber, C., Heghes, C. and Johnston, L. (2024b). A framework to support the application of the OECD guidance documents on (Q)SAR model validation and prediction assessment for regulatory decisions. *Comput Toxicol 30*, 100305. doi:10.1016/j.comtox.2024.100305

Belfield, S. J., Cronin, M. T. D., Enoch, S. J. et al. (2023). Guidance for good practice in the application of machine learning in development of toxicological quantitative structure-activity relationships (QSARs). *PLoS One 18*, e0282924. doi:10.1371/journal.pone.0282924

Bishop, P. L., Mansouri, K., Eckel, W. P. et al. (2024). Evaluation of in silico model predictions for mammalian acute oral toxicity and regulatory application in pesticide hazard and risk assessment. *Regul Toxicol Pharmacol 149*, 105614. doi:10.1016/j.yrtph.2024.105614

Briggs, K., Bosc, N., Camara, T. et al. (2021). Guidelines for FAIR sharing of preclinical safety and off-target pharmacology data. *ALTEX 38*, 187-197. doi:10.14573/altex.2011181

Cronin, M. T. D., Enoch, S. J., Madden, J. C. et al. (2022). A review of in silico toxicology approaches to support the safety assessment of cosmetics-related materials. *Comput Toxicol 21*, 100213. doi:10.1016/j.comtox.2022.100213

Cronin, M. T. D., Belfield, S. J., Briggs, K. A. et al. (2023). Making in silico predictive models for toxicology FAIR. *Regul Toxicol Pharmacol 140*, 105385. doi:10.1016/j.yrtph.2023.105385

Dearden, J. C. (2016). The history and development of quantitative structure-activity relationships (QSARs). *Int J Quant Struct-Prop Relat 1*, 1-44. doi:10.4018/ijqspr.2016010101

Dimitrov, S., Dimitrova, G., Pavlov, T. et al. (2005). A stepwise approach for defining the applicability domain of SAR and QSAR models. *J Chem Inf Model 45*, 839-849. doi:10.1021/ci0500381

ECHA – European Chemicals Agency (2023). The use of alternatives to testing on animals for the REACH Regulation. Fifth report under Article 117(3) of the REACH Regulation. European Chemicals Agency, Helsinki, Finland. doi:10.2823/805454

Gissi, A., Tcheremenskaia, O., Bossa, C. et al. (2024). The OECD (Q)SAR assessment framework: A tool for increasing regulatory uptake of computational approaches. *Comput Toxicol 31*, 100326. doi:10.1016/j.comtox.2024.100326

Glont, M., Nguyen, T. V. N., Graesslin, M. et al. (2018). BioModels: Expanding horizons to include more modelling approaches and formats. *Nucl Acids Res 46, Issue D1*, D1248-D1253. doi:10.1093/nar/gkx1023

Hansch, C. and Fujita, T. (1964). p-σ-π analysis. A method for the correlation of biological activity and chemical structure. *J Am Chem Soc 86*, 1616-1626. doi:10.1021/ja01062a035

Hardy, B., Douglas, N., Helma, C. et al. (2010). Collaborative development of predictive toxicology applications. *J Cheminform 31*, 7. doi:10.1186/1758-2946-2-7

Hardy, B., Apic, G., Carthew, P. et al. (2012). Toxicology ontology perspectives. *ALTEX 29*, 139-156. doi:10.14573/altex.2012.2.139

Hastings, J., Chepelev, L. and Willighagen, E. (2011). The chemical information ontology: Provenance and disambiguation for chemical data on the biological semantic web. *PLoS One 6*, e25513. doi:10.1371/journal.pone.0025513

Laroche, C., Annys, E., Bender, H. et al. (2019). Finding synergies for the 3Rs – Repeated dose toxicity testing: Report from an EPAA partners' forum. *Regul Toxicol Pharmacol 108*, 104470. doi:10.1016/j.yrtph.2019.104470

Madden, J. C., Enoch, S. J., Paini, A. et al. (2020). A review of in silico tools as alternatives to animal testing: Principles, resources and applications. *Altern Lab Anim 48*, 146-172. doi:10.1177/0261192920965977

Malik-Sheriff, R. S., Glont, M., Nguyen, T. V. N. et al. (2020). BioModels – 15 years of sharing computational models in life science. *Nucl Acids Res 48*, D407-D415. doi:10.1093/nar/gkz1055

Netzeva, T. I., Worth, A., Aldenberg, T. et al. (2005). Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. The report and recommendations of ECVAM workshop 52. *Altern Lab Anim 33*, 155-173. doi:10.1177/026119290503300209

OECD – Organisation for Economic Cooperation and Development (2007). Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. *OECD Series on Testing and Assessment, No. 69*. OECD Publishing, Paris. doi:10.1787/9789264085442-en

OECD (2023). (Q)SAR Assessment Framework: Guidance for the regulatory assessment of (Quantitative) Structure Activity Relationship models and predictions. *OECD Series on Testing and Assessment, No. 386*. OECD Publishing, Paris. doi:10.1787/d96118f6-en

Piir, G., Kahn, I., García-Sosa, A. T. et al. (2018). Best practices for QSAR model reporting: Physical and chemical properties, ecotoxicity, environmental fate, human health and toxicokinetics endpoints. *Environ Health Perspect 126*, 126001. doi:10.1289/ehp3264

Ruusmann, V., Sild, S. and Maran, U. (2014). QSAR DataBank – An approach for the digital organization and archiving of QSAR model information. *J Chemoinform 6*, 25. doi:10.1186/1758-2946-6-25

Ruusmann, V., Sild, S. and Maran, U. (2015). QSAR DataBank repository: Open and linked qualitative and quantitative structure-activity relationship models. *J Chemoinform 7*, 32. doi:10.1186/s13321-015-0082-6

Spjuth, O., Willighagen, E. L., Guha, R. et al. (2010). Towards interoperable and reproducible QSAR analyses: Exchange of datasets. *J Cheminform 2*, 5. doi:10.1186/1758-2946-2-5

Westmoreland, C., Bender, H. J., Doe, J. E. et al. (2022). Use of new approach methodologies (NAMs) in regulatory decisions for chemical safety: Report from an EPAA Deep Dive Workshop. *Regul Toxicol Pharmacol 135*, 105261. doi:10.1016/j.yrtph.2022.105261

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Sci Data 3*, 160018. doi:10.1038/sdata.2016.18

Wittwehr, C., Clerbaux, L. A., Edwards, S. et al. (2024). Why adverse outcome pathways need to be FAIR. *ALTEX 41*, 50-56. doi:10.14573/altex.2307131

Yang, C., Rathman, J. F., Bienfait, B. et al. (2023). The role of a molecular informatics platform to support next generation risk assessment. *Comput Toxicol 26*, 100272. doi:10.1016/j.comtox.2023.100272

Disclaimer

The opinions expressed in this document reflect only the authors' views. The European Commission is not responsible for any use that may be made of the information it contains.

Conflict of interest

The QSAR Databank (QsarDB) has been developed and is maintained by Prof Uko Maran, Institute of Chemistry, University of Tartu, Tartu, Estonia.

Data availability statement

The QSARs and full data described in this article are freely available from QsarDB as specified in Table 5.

Acknowledgements

This project receives funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No. 964537 (RISK-HUNT3R), and it is part of the ASPIS cluster; the Horizon Europe Framework Programme project Partnership for the Assessment of Risks from Chemicals (PARC, grant 101057014) under Task 5.2; and the QUANTUM-TOX – Revolutionizing Computational Toxicology with Electronic Structure Descriptors and Artificial Intelligence (QUANTUM-TOX) HORIZON-EIC-2023-PATHFINDEROPEN-01 Project number: 101130724. UM, SS, GP acknowledge support by the Ministry of Education and Research, Republic of Estonia, through Estonian Research Council (grant number PRG1509), Ministry of Climate, Republic of Estonia (grant 4-4/22/19), Ministry of Social Affairs, Republic of Estonia (grant 3-4/1593-1).