

독성발현경로(Adverse Outcome Pathway)를 활용한 *In Silico* 예측기술 연구동향 분석이수진 , 박종서 , 김선미* , 서명원* 

한국화학연구원 화학분석센터

Trend of *In Silico* Prediction Research Using Adverse Outcome Pathway

Sujin Lee, Jongseo Park, Sunmi Kim*, and Myungwon Seo*

Chemical Analysis Center, Korea Institute of Chemical Technology (KRICT)

ABSTRACT

Background: The increasing need to minimize animal testing has sparked interest in alternative methods with more humane, cost-effective, and time-saving attributes. In particular, *in silico*-based computational toxicology is gaining prominence. Adverse outcome pathway (AOP) is a biological map depicting toxicological mechanisms, composed of molecular initiating events (MIEs), key events (KEs), and adverse outcomes (AOs). To understand toxicological mechanisms, predictive models are essential for AOP components in computational toxicology, including molecular structures.

Objectives: This study reviewed the literature and investigated previous research cases related to AOP and *in silico* methodologies. We describe the results obtained from the analysis, including predictive techniques and approaches that can be used for future *in silico*-based alternative methods to animal testing using AOP.

Methods: We analyzed *in silico* methods and databases used in the literature to identify trends in research on *in silico* prediction models.

Results: We reviewed 26 studies related to AOP and *in silico* methodologies. The ToxCast/Tox21 database was commonly used for toxicity studies, and MIE was the most frequently used predictive factor among the AOP components. Machine learning was most widely used among prediction techniques, and various *in silico* methods, such as deep learning, molecular docking, and molecular dynamics, were also utilized.

Conclusions: We analyzed the current research trends regarding *in silico*-based alternative methods for animal testing using AOPs. Developing predictive techniques that reflect toxicological mechanisms will be essential to replace animal testing with *in silico* methods. In the future, since the applicability of various predictive techniques is increasing, it will be necessary to continue monitoring the trend of predictive techniques and *in silico*-based approaches.

Key words: *In silico*, adverse outcome pathway, machine learning, computational toxicology, chemical hazard

Received April 2, 2024

Revised April 22, 2024

Accepted April 24, 2024

Highlights:

- ToxCast/Tox21 DB is widely used in toxicity research, with MIE often serving as predictive factors in developing an AOP-based predictive model.
- Machine learning was the most utilized relative to deep learning and molecular simulation techniques.
- Applying AOP components to develop predictive technologies that reflect accurate toxicity mechanisms is necessary.
- It is necessary to continue monitoring the trend of *in silico* predictive research to replace animal testing.

*Corresponding author:

Sunmi Kim: Chemical Analysis Center,
Korea Institute of Chemical Technology
(KRICT), 141 Gajeon-ro, Yuseong-gu
Daejeon 34114, Republic of Korea

Tel: +82-42-860-7182

Fax: +82-42-860-7183

E-mail: skim@kRICT.re.kr

Myungwon Seo: Chemical Analysis Center,
Korea Institute of Chemical Technology
(KRICT), 141 Gajeon-ro, Yuseong-gu
Daejeon 34114, Republic of Korea

Tel: +82-42-860-7102

Fax: +82-42-860-7183

E-mail: mwseo@kRICT.re.kr

I. 서 론

동물대체시험법은 동물을 사용하지 않거나 동물 개체 수 감소 또는 고통을 경감시키기 위한 모든 시험법으로 정의할 수 있다. 동물실험은 오랜 기간동안 화장품, 의약품, 식품 등의 화학제품 성분에 대한 연구개발을 위해 수행되어 왔으며, 실험 목적으로 사용되는 동물의 수가 수년에 걸쳐 증가함에 따라 국내외 여러 기관과 대중으로부터 윤리적 문제와 실효성에 대한 비판을 받아왔다.¹⁾ 1959년 Russel과 Burch의 '인도적인 실험기법에 관한 법칙'이라는 저서에 제시된 3Rs²⁾에 의해 동물대체시험법의 개발과 도입 필요성이 강조되며 화학물질 평가 시스템의 변화가 시작되었다.³⁾

동물대체시험법은 화학물질의 효능 또는 독성 평가실험을 위해 주로 약물 동태, 피부자극성 및 부식성, 질환모델 등 안전성 및 유효성 평가 분야에 활용되고 있다. 동물대체시험법은 세포 및 조직을 기반으로 하는 *in vitro* 시험법과, 살아있는 인자가 아닌 생체물질 또는 생체모사물질을 이용해 평가하는 비세포계 *in chemico* 시험법 및 화학물질 분자구조 특성 및 단백질 구조정보를 활용하여 독성을 예측하는 *in silico* 기반 시험법으로 구분할 수 있다.⁴⁾ 동물대체시험법은 동물을 활용한 기존 독성시험에 비해 시험동물의 희생을 줄일 수 있다는 인도적인 측면에서의 장점 뿐만 아니라 비용과 시간이 적게 든다는 장점^{5,6)}이 있으며 동물대체시험법 중 *in silico* 기반 시험법은 근본적으로 실험을 수행하지 않아 시간 및 비용 측면에서 장점이 있어 동물대체시험법의 주요한 영역으로 주목받고 있다.

동물시험 금지에 대한 국제적인 관심이 증가하면서 다양한 동물대체시험법 활용을 목표로 하는 국내외 규제가 증가하고 있는 추세이다. 우리나라에서는 2023년 12월, '동물대체시험법의 개발·보급 및 이용 촉진에 관한 법률안'과 '동물대체시험법의 개발 보급 및 이용 활성화에 관한 법률안'이 국회 보건복지위원회를 통과하며, 대체시험법 개발과 활성화를 위한 논의가 본격화되었다. 현재는 *in silico* 기반 시험법과 관련하여 2016년 화장품법에 '동물실험을 실시한 화장품 등의 유통 판매 금지' 조항이 신설되어 2017년부터 시행되고 있다.⁷⁾ 또한 「화학물질 등록 및 평가 등에 관한 법률」 제14조 및 시행령 13조 제3호에 따라 독성자료 생산 시 구조활성관계 예측 프로그램 모델(Qualitative or Quantitative Structure Activity Relationship model, QSAR) 활용을 허용하고 있어, 비시험자료 생산에서 *in silico* 기반 시험법이 활용되고 있다.⁸⁾ 국외에서는 유럽화학물질청(European Chemical Agency)과 유럽식품안전청(European Food Safety Authority)에서 살생물질 및 농약 성분의 내분비계 장애영향 스크리닝을 위해 동물실험의 대안으로 고속 대량 스크리닝 자료를 적극 활용하는 평가 전략을 제시하였다.⁹⁾ 또한, 2007년 시행된 European Union Registration, Evaluation, Authorization and Restriction of Chemicals (EU

REACH) 법령'에서 독성자료 생산 시 QSAR모델 사용을 허용하고 있어 *in silico* 시험법의 활용이 촉진되고 있는 추세이다.^{5,6,10-12)}

독성발현경로(Adverse Outcome Pathway, AOP)는 분자수준 초기현상(Molecular Initiating Event, MIE)으로부터 주요현상(Key Event, KE) 및 개체 수준에서 관찰할 수 있는 최종 독성영향(Adverse Outcome, AO)으로 구성, 연결되어 독성기전을 묘사한 일종의 생물학적 지도로, 위해성평가와 같은 규제 정책에 활용하기 위해 도입된 프레임워크이다. MIE는 화학물질과 생체 내 특정 분자와의 초기 상호작용을 의미하며, KE는 AOP 내에서 생물학적 상태의 변화를 나타내는 단계를 의미하고, AO란 개체 수준에서 관찰할 수 있는 최종 영향을 의미한다. 화학물질의 독성예측 시 독성기전을 보다 명확히 이해하기 위해서는 기존의 QSAR 모델과 같이 화학물질의 분자구조만을 활용한 예측보다는, 독성기전 정보를 포함하는 AOP 구성요소를 추가로 활용하는 것이 매우 중요하다. 특히 AOP 구성요소 중 MIE 수준에서 화학물질의 정확한 독성 예측이 가능하다면 보다 선제적으로 화학물질 관리에 활용할 수 있기 때문에, 관련 연구가 활발히 진행되고 있다.^{13,14)} AOP를 활용한 *in silico* 예측 연구로, Tan 등¹⁵⁾은 생식독성의 AOP와 연관된 화학물질을 활용하여 화학물질의 생식독성 발생 여부와 그 독성기전을 예측하는 모델을 개발하였으며, Troger 등¹⁶⁾은 파킨슨병 유발과 관련한 MIE 예측을 위해 Pharmacophore screening 기법과 머신러닝(Machine learning) 기법을 활용한 예측모델을 개발하였다. Gadaleta 등¹⁷⁾은 신경독성 유발에 관여하는 MIE에 대한 예측 모델을 개발하여 신경독성을 유발할 수 있는 잠재적인 화학물질을 선별하는데 활용하였다.

이와 같이 국내외 산업, 화학물질 관리 및 연구 분야에서 동물 실험에 대한 의존을 줄이고자 동물대체시험법에 대한 관심이 점차 높아지고 있으며 동물대체시험법 중 전산독성학(computational toxicology)을 기반으로 독성을 예측하는 *in silico* 기반 시험법이 주목받고 있으나, 현실적인 *in silico* 기반 예측 연구의 활용도는 여전히 낮은 실정이므로 *in silico* 기반 연구의 활성화가 필요하다. 본 논문에서는 *in silico* 기반 예측연구의 중요성 및 활용성 검토를 위해 *in silico* 기반 예측기술 및 AOP 구성요소를 활용한 최신 연구 동향을 조사하고 분석하였다. 이를 위해 문헌 조사에서 수집된 AOP 및 *in silico* 기반 시험법을 활용한 사전 연구 사례를 분석하였으며, 분석 결과를 바탕으로 향후 활용 가능한 데이터 및 예측기술 등을 제안하고자 한다.

II. 재료 및 방법

1. 문헌 데이터 수집 및 선정

국내외 *in silico* 기반 동물대체시험법에 대한 연구 동향을 확인하기 위해 Google 학술검색(Google scholar)을 통해 주요

어 ‘Computational toxicology’를 기준으로 게재된 문헌을 검색하고 수집하였다. 수집 문헌 중 ‘AOP (Adverse Outcome Pathway)’, ‘*in silico*’, ‘prediction’ 3가지 주요어 조합이 포함되어 있는 문헌으로 분류하였다. 분류된 문헌 중 AOP 및 *in silico* 예측기술을 활용한 모델 개발과 관련된 연구를 수행한 문헌을 최종적으로 선별하여 예측기술 동향분석에 활용하였다. 연구 동향 분석은 연구에 활용된 데이터와 *in silico* 예측기술을 중심으로 조사하였으며, 예측 대상물질 종류, AOP 요소, *in silico* 예측기법, 독성종말점 등 총 7개의 세부항목 분류를 통해 분석을 수행하였다. 전반적인 문헌 수집 및 선별 과정에 대한 흐름도는 Fig. 1과 같다.

2. 데이터 중심 분석

In silico 기반 동물대체시험법 연구에 사용된 데이터 분석을 위해 1) 대상물질 종류(단일/혼합), 2) 모델 개발에 사용된 AOP 구성요소(MIE, KE, AO)의 활용 빈도, 3) 데이터베이스 종류, 4) 예측모델 개발 시 활용된 데이터 종류 및 5) 데이터의 독성종말점(endpoint)으로 총 5개 항목으로 데이터에 대한 세부분석을 수행하였다. 대상물질 종류(단일/혼합) 정보를 통해 예측에 사용되는 대상이 단일물질 또는 혼합물인지를 확인하였으며, AOP 구성요소(MIE, KE, AO)의 활용 빈도 분석을 통

해 AOP 구성요소 중 독성 예측에 가장 많이 활용된 구성요소를 확인하였다. 연구에 활용된 데이터베이스 및 예측모델 개발 시 활용된 데이터 종류(화학물질, 단백질 등)를 분석하였으며, 마지막으로 모델 개발에 활용된 데이터의 독성종말점(endpoint)을 확인하였다. 독성종말점은 생식독성, 피부 과민성 및 호흡기 과민성 등을 포함한 다양한 유해성에 대해 분류를 의미하며 해당 항목에 대한 동향 분석을 수행하였다.

3. 예측기술 중심 분석

동물대체시험법에 적용 가능한 예측기술에 대한 분석을 위해 2가지 세부 항목으로 1) 연구에 사용된 *in silico* 예측 기법, 2) 모델 예측 종류(정량/정성)를 선정하여 분석을 수행하였다. 독성 예측모델 개발에 사용된 *in silico* 예측 기법 분석 시에는 머신러닝, 딥러닝(Deep learning) 등의 활용도를 확인하였고 연구에 사용된 세부 알고리즘에 대한 종류를 분석하였다. 독성 예측모델의 예측결과에 대해서 독성유무를 제공하는 정성적 방법인지, 정량적 방법(독성 농도, 반응 등)으로 제공하는지 등을 확인하여 예측기술 동향 분석을 수행하였다.

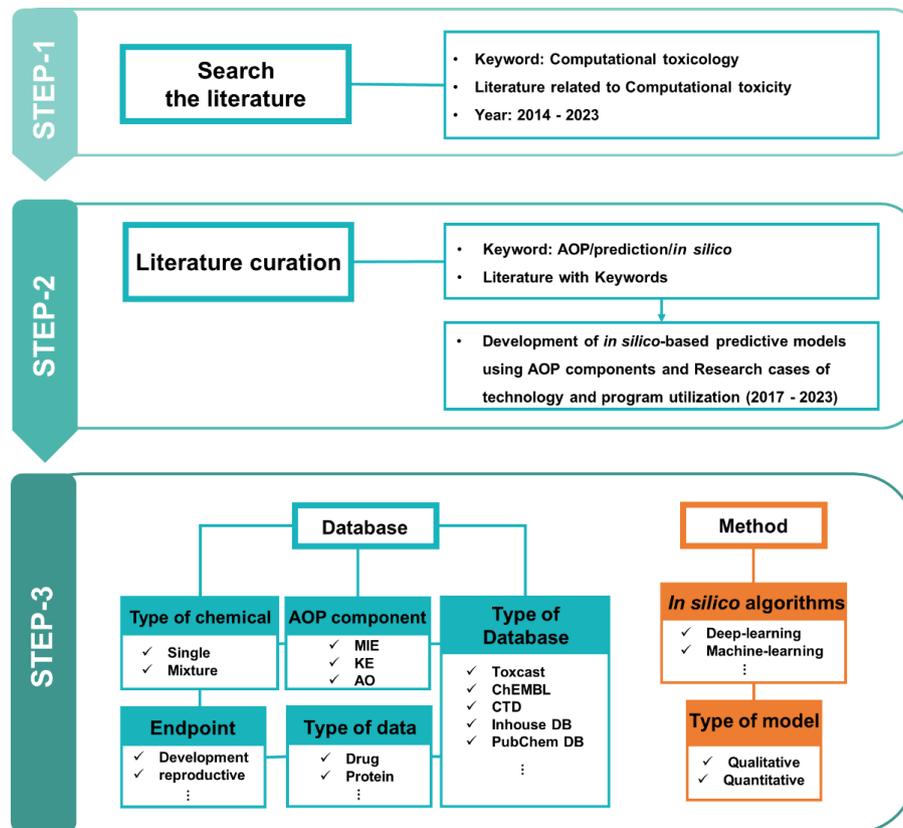


Fig. 1. Workflow of literature survey for *in silico* research trend analysis using AOPs

Table 1. Summary of *in silico* research using AOP components

No	Research content	Data type (database)	Algorithm type	Endpoint	AOP component	Reference no
1	Development of predictive models utilizing 2D and 3D structures for skin sensitization analysis	Chemical (PubChem)	ML (DT)	Skin sensitization	KE	18)
2	Development of MIE prediction models for correlation analysis with activation energy	Chemical (ChEMBL, Toxcast/Tox21)	DL (DNN)	Biological activity, structure features	MIE	19)
3	Development of toxicity prediction models based on MIE	Chemical (Toxcast/Tox21)	ML (XGBoost)	Endocrine disruption	MIE	22)
4	Development of toxicity prediction models based on MIE	Chemical (Toxcast/Tox21)	ML (MLR)	Proliferation inhibition	MIE	23)
5	Development of KE prediction models for AOP related to Parkinson's disease	Chemical (Toxcast/Tox21, ChEMBL, Inhouse DB)	ML (RF, XGBoost, DL (DNN)	Human mitochondrial respiratory complex I inhibition	KE	16)
6	Development of EDC prediction models using chemical structure images	Chemical (Toxcast/Tox21)	DL (CNN)	Nuclear receptor activity	MIE	24)
7	Development of MIE prediction models	Chemical (ChEMBL)	ML (PLS)	Glucocorticoid receptor, HERG, mu opioid receptor, ctclooxygenase-2, dopamine transporter	MIE	25)
8	Development of predictive models for skin hypersensitivity	Chemical (Inhouse DB)	ML (Kinetic method)	Skin sensitization	MIE	26)
9	Development of a pipeline for toxicity prediction	Chemical (Toxcast/Tox21)	ML (k-NN)	Androgen receptor antagonism	MIE	27)
10	Development of MIE prediction models related to thyroid hormone homeostasis	Chemical (Toxcast/Tox21)	ML (RF, SVM, XGBoost, ANN, LR)	Thyroid hormone homeostasis	MIE	28)
11	Development of predictive models using DNN	Chemical (Toxcast/Tox21)	ML (RF, SVM, BNB, k-NN), DL (DNN)	Nuclear estrogen receptor α and β (ER_{α} and ER_{β}) binding potentials	KE	29)
12	Development of AOP-based pulmonary fibrosis prediction models	Chemical (Toxcast/Tox21)	ML (MLR)	Lung tissue inflammation	KE (AOP 173)	30)
13	Development of toxicity prediction models using the open-source software framework 'Flame'	Chemical (NTP)	ML (RF, XGBoost)	Mitochondrial toxicity	MIE	31)
14	Development of skin sensitization prediction models using deep learning	Chemical (Inhouse DB)	DL (GCN)	Skin sensitization hazard	KE	32)

Table 1. Continued

No	Research content	Data type (database)	Algorithm type	Endpoint	AOP component	Reference no
15	Development of KE-specific activity prediction models using QSAR for each KE and stacked model using naïve Bayes	Chemical (Toxcast/Tox21)	ML (RF, SVM, NB, k-NN, DT)	Human reproductive toxicity	KE	15)
16	Development of various toxicity endpoint prediction models related to developmental and neurotoxicity	Chemical (Toxcast/Tox21, ChEMBL, Inhouse DB)	ML (k-NN, RF, ANN)	Developmental and neurotoxicity	MIE	17)
17	Prediction of the probability of AO related to developmental neurotoxicity using a Bayesian hierarchical model	Chemical (Inhouse DB)	ML (Bayesian hierarchical model)	Developmental and neurotoxicity	AO	33)
18	Improvement of AOP framework related to estrogen interference induced by TTP (Triphenyl phosphate)	Protein (PDB, UniProt)	Molecular simulation (molecular dynamics)	G protein-coupled estrogen receptor activation	MIE	34)
19	Development of EDC identification prediction models using toxicogenomic data	Chemical (CTD)	ML (GLM)	Endocrine disruption	MIE	35)
20	Construction of AOP for lung injury induced by AHR activation and verification of AOP network using prediction models	Chemical (Toxcast/Tox21, PubChem DB) Protein (AlphaFold protein structure database)	ML(RF), Molecular simulation (molecular docking)	AHR(aryl hydrogen receptor) activation to lung damages	KE	36)
21	Review of case studies on the development of qAOP using predictive technologies	Chemical (EFSA DB, OECD eChemPortal, DNT-DIVER)	ML(Bayesian networks)	Skin sensitization, neurotoxicity, carcinogenicity	AOPs	37)
22	Research review on the development of QSAR-based prediction models for mixture toxicity	Chemical (Toxcast/Tox21) Protein (PDB)	ML(ANN), Molecular simulation (molecular docking)	Acute/chronic toxicity	MIE	38)
23	Prediction of genetic regulation patterns of AOP using AI methods	Gene (AOP-DB, UK BioBank)	DL (GNN)	Hepatic steatosis, liver cancer	MIE	39)
24	Prediction of cellular-level MoA using protein data	Protein (PRIDE)	ML (kNN, RF)	Mode of action related to toxicity	MIE	40)
25	Development of MIE prediction models for AOP related to fatty liver	Chemical (AOPwiki, Toxcast/Tox21)	ML (RF)	Hepatic steatosis	MIE	20)
26	Development of MIE prediction models using DNN	Chemical (ChEMBL, Toxcast/Tox21)	DL (DNN)	AChE, ADORA2A, AR, hERG, SERT	MIE	21)

AO: Adverse Outcome, ANN: Artificial Neural Network, BNB: Bernoulli Naïve Bayes, CTD: Comparative Toxicogenomics Database, DNN: Deep Neural Network, DT: Decision Tree, GLM: General Linear Model, GNN: Graph Neural Network, KE: Key Event, k-NN: k-Nearest Neighbors, LR: Logistic Regression, MIE: Molecular Initiating Event, MLR: Multiple Linear Regression, NB: Naïve Bayes, NN: Neural Network, PLS: Partial Least-Squares, PRIDE: Proteomics Identifications Database, RF: Random Forest, SVM: Support Vector Machines.

III. 결과 및 고찰

1. 문헌 수집 및 분류 결과

국내외 최근 연구 동향 조사를 위해 주요어 'computational toxicology'가 포함된 문헌으로 1차 분류한 결과 2014년부터 2023년까지 최근 10년간 게재된 35,600개의 문헌이 검색되었다. 1차 분류를 통해 수집된 문헌을 대상으로 세 가지 주요어 ('AOP', 'in silico', 'prediction')를 포함하는 문헌을 추가로 선별한 결과 약 8,500개의 문헌이 분류되었다. 분류된 8,500개의 문헌 중 AOP 구성 요소를 대상으로 *in silico* 기반 예측모델 개발 또는 *in silico* 예측 기술 및 프로그램을 활용하여 분석한 연구사례에 대한 문헌을 선별한 결과 2017년부터 2023년까지 게재된 총 26편의 문헌이 최종 선정되었다(Table 1).¹⁵⁻⁴⁰ AOP 구성요소를 고려하여 *in silico* 기반 독성 예측연구를 수행한 시기는 2017년 이후로 연구된 것으로 확인되었다. 2017년부터 2023년까지 게재된 문헌의 수는 매년 증가 추세를 보이고 있으며 특히 2019년 이후부터 현재까지 머신러닝 및 딥러닝을 활용한 연구사례는 2019년부터 1년 단위로 최소 1편(4%)씩 증가하고 있으며 최근 3년간 누적 13편(54%)으로 지속적으로 증가하는 추세이다(Fig. 2A). 2010년 AOP가 처음 제시되고 2016년 The Organisation for Economic Co-operation and Development (OECD)에 의해 AOP를 활용한 독성 예측 연구의 필요성이 제시된 이후,¹³ 패러다임 변화에 따라 AOP를 활용한 독성 예측모델 연구에 대한 관심이 증가한 것으로 보인다. AOP 구성요소를 예측하는 연구사례는 2017년부터 현재까지 26건 정도 수준이나 독성 예측연구 관점에서 단순히 화학물질 분자 구조 중심의 특성(물리화학적 특성, 분자지문 등) 및 실험데이터를 활용하여 computational toxicology 기반 독성 예측모델 개발 연구사례는 2017년부터 현재까지 약 17,000건으로 많은 연구가 수행되었다. 독성기전을 반영할 수 있는 AOP 구성요소를 고려한 예측 연구사례는 전체 computational toxicology 기반 연구사례 중 약 0.15% 정도로 상대적으로 매우 미흡한 수준이다. 화학물질 독성에 대해 정확한 예측연구를 수행하기 위해서는 예측모델 개발 시 독성기전을 고려하는 것은 매우 중요한 요소이므로 AOP 정보를 활용한 예측 연구는 점차적으로 증가되어야 하며 추후 관련 연구동향을 지속적으로 주시해 나가는 것이 필요한 상황이다.

2. 데이터 기반 연구동향

In silico 기반 독성 예측모델 개발 사례를 보고한 최종 선정 문헌 26편 중 25편(96%)이 단일물질 예측을 중심으로 개발되었으며, 혼합물의 독성예측 연구를 수행한 문헌은 1편(4%)에 불과했다. 화학제품은 여러 구성성분이 혼합된 혼합물 형태이지만 화학제품 등록 시 단일물질 수준에서의 독성 평가 자료를 제출하고 있고, 혼합물 평가 시에는 일반적으로 각 성분의 독

성을 누적 계산하고 있어 다양한 혼합물의 구성성분 간 상호작용과 조합비에 따른 독성 변화를 고려한 평가는 수행되고 있지 않다. 우리나라는 2015년 제정된 「화학물질 등록 및 평가 등에 관한 법률」과 2017년 제정된 「생활화학제품 및 살생물제의 안전관리에 관한 법률」에 따라 화학제품의 안전관리 체계 선진화를 위해 노력하고 있으며, 특히 살생물물질의 경우 제품 내 타 성분과 상호 작용하여 일으키는 효과를 고려하도록 명시하고 있다. 그러나 세계적으로도 상승작용, 길항작용과 같은 상호작용에 대한 과학적 정보가 매우 부족하기 때문에 여전히 화학물질 및 화학제품 등록 시에는 단일물질 수준의 물리화학적 특성과 인체·환경 유해성을 고려하고 있다.⁴¹ 혼합물의 독성 정보와 상호 작용에 대한 정보는 모든 혼합물의 사례별로 도출하는 것이 쉽지 않기 때문에, 더욱 *in silico* 모델을 활용해 정보를 생산해야 할 필요성이 높다. 다만 현재까지는 *in silico* 기반 시험법 또한 미흡한 수준이므로 향후 제도의 효과를 극대화하기 위해 혼합독성 예측모델의 개발이 고려되어야 한다.^{42,43}

전체 문헌 중 AOP 구성요소 중 MIE를 예측지표로 사용한

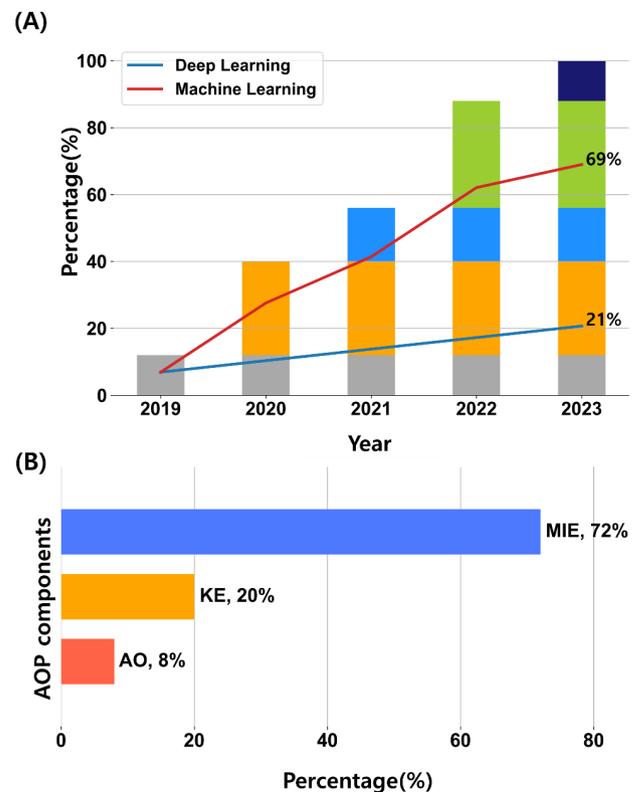


Fig. 2. Data-driven research trend results. This graph represents literature surveyed from before 2019 to 2023. (A) The annual cumulative number of articles and frequency of utilization of machine learning and deep learning. (B) Frequency of utilization of AOP components. MIE: Molecular Initiating Event, KE: Key Event, AO: Adverse Outcome.

문헌은 72% (18편)였으며, KE를 예측지표로 사용한 경우는 20% (7편), AO를 사용한 경우가 8% (1편)로 확인되었다. 이는 KE 및 AO에 대한 실험데이터가 부족하여 예측모델을 개발하기에는 어려움이 있어 관련 예측 연구가 많이 수행되지 못했음을 보여준다. KE를 활용한 예로, Ciallella 등²⁹⁾은 에스트로겐 수용체 활성화(KE) 예측을 위해 기식 기반 심층 신경망(k-DNN)을 활용한 예측 모델을 개발하였다. AO를 활용한 예로 Spinu 등³³⁾은 신경독성의 AO와 화학물질 노출과의 인과 요인을 이해하기 위한 베이지안 계측 모델 개발을 통해 신경독성 예측 연구를 수행하였다. 반면, MIE 수준에서 단백질과 화학물질 간의 상호작용을 중심으로 예측 모델 개발이 수행되므로 단백질 구조와 화학물질의 분자구조만 있다면 molecular docking,

molecular dynamics, pharmacophore screening 등 다양한 *in silico* 기반 예측 기술을 적용할 수 있다. 또한 MIE 지표는 결합 친화도(binding affinity), 저해(inhibition) 등 단백질과 화학물질 간의 상호작용 실험데이터가 다른 AOP 구성요소보다 상대적으로 많이 존재하기 때문에, 예측지표로 연구에 더 많이 활용되고 있다(Fig. 2B). MIE를 활용한 예시로 Gadaleta 등¹⁷⁾은 신경독성 유발에 관여하는 MIE에 대한 예측모델을 개발하여 신경독성을 유발할 수 있는 잠재적인 화학물질을 선별하는데 활용하였다. MIE는 AOP 순서 중 가장 첫 번째에 위치하며 독성기전의 시작을 설명할 수 있는 가장 중요한 단계이다. 명확하게 알려진 MIE를 대상으로 단백질 및 화학물질에 대한 구조정보만 있다면 독성 실험데이터가 없더라도 다양한 예측기술을 적용하여 연구를 수행할 수 있는 장점이 있어 동물대체시험법 중 독성기전 정보를 반영한 *in silico* 기반 시험법으로서 독성을 유발할 수 있는 물질을 초기에 선별하는 과정에서 활용될 수 있을 것이다.

예측 연구에 활용된 데이터베이스 중에서는 Toxcast 및 Tox21이 전체의 35%로 가장 많이 사용되었다(Fig. 3A, Table 2). Toxcast 및 Tox21은 일반세포독성뿐 아니라 내분비계 장애영향 등 다양한 독성지표에 대한 *in vitro* 시험법을 활용하여 고효율 대량 스크리닝(High-Throughput Screening, HTS)으로 실험데이터를 생산하고, 분석자료를 통합한 데이터베이스로 독성 연구에 활발히 활용되고 있다.⁴⁴⁻⁴⁶⁾ 다음으로는 *in vivo* 및 *in vitro* 실험을 통해 직접 생산한 데이터(In-house DB)와 다양한 화학물질 구조 및 생물학적 실험데이터를 포함한 ChEMBL Database가 각각 전체 문헌 중 12.5% (5건)로 사용된 것을 확인하였다. 또한, 생물학적 실험데이터를 제공하는 PubChem DB도 5% (2건) 정도 활용되고 있었으며, 그 외 AOP 구성요소 정보를 활용하기 위한 AOP-DB, AOP-Wiki 및 단백질 3차원 구조와 화학물질과 단백질 간의 상호작용정보를 사용하기 위한 Protein Data Bank (PDB), CTD 등 다양한 데이터베이스가 각각 2.5% (1건)의 활용도를 보였다. 그 외에도 AOP 구성요소 중 MIE 수준에서 molecular docking⁴⁷⁾과 같은 기술을 활용하여 단백질-화학물질 간의 상호작용을 고려하기 위해 단백질 실험데이터 및 3차원 구조 데이터를 포함하고 있는 PDB, AlphaFold Protein Structure Database, UniProt 등을 활용하고 있음을 보여주었다.

AOP 구성요소 예측 시 독성 실험데이터를 중심으로 화학물질 구조기반 독성 예측기술이 주로 사용되고 있기 때문에 다양한 화학물질 분자구조 및 독성 실험데이터를 포함하고 있는 Toxcast/Tox21의 활용도가 가장 높고, ChEMBL, PubChem DB 등의 데이터베이스에도 화학물질 및 단백질에 대한 실험데이터가 포함되어 있어 활용되고 있는 것으로 보인다. 대부분의 독성 예측 연구는 화학물질과 독성 간의 연관성을 포함하는 실험데이터 기반으로 연구가 수행되고 있으나 *in vitro*, *in vivo* 수

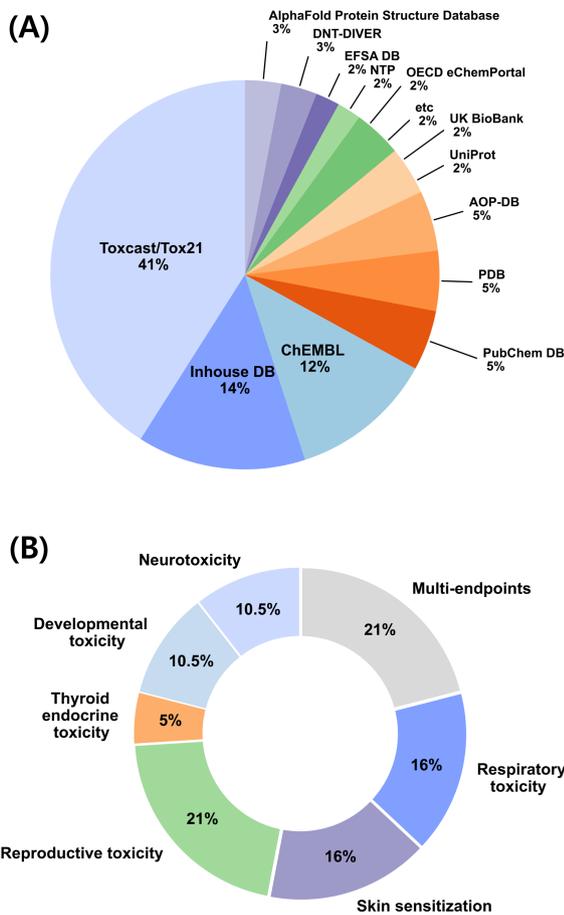


Fig. 3. Trend analysis of database and *in silico* method-driven research. (A) Databases utilized in prior studies and (B) frequency of toxicity endpoints used as predictors. DNT-DIVER: developmental neurotoxicity data integration and visualization enabling resource, EFSA-DB: European food safety authority database, NTP: national toxicology program, PDB: protein data bank, CTD: comparative toxicogenomics database, PRIDE: proteomics identifications database.

Table 2. Summary of database list

No	Name	Database content	Data	URL
1	Alphafold Protein Structure Database	A database containing predicted protein structures generated by the AlphaFold algorithm	• Protein: 1,508,588	https://alphafold.ebi.ac.uk/
2	AOP-DB	A database storing information on Adverse Outcome Pathways (AOPs), detailing the sequence of events leading to adverse outcomes due to chemical exposure	• Protein: 665,746 • Disease: 15,093 • Chemical: 164,640	https://aopdb.epa.gov/
3	AOP wiki	An online resource providing collaborative information and documentation on Adverse Outcome Pathways (AOPs)	• 465 AOPs	https://aopwiki.org/
4	ChEMBL	A database containing bioactivity data of small molecules, including compounds, drugs, and their targets	• Gene: 107 • Protein: 22,751 • Chemical: 1,920,643	https://www.ebi.ac.uk/chembl/
5	CTD	A database linking chemicals, genes, and diseases to elucidate molecular mechanisms underlying environmental health risks	• Chemical–gene interactions: 2,804,028 • Gene–disease associations: 32,234,482 • Chemical–disease associations: 3,381,685	https://ctdbase.org/
6	DNT-DIVER	A database integrating and visualizing data related to developmental neurotoxicity	• Chemical: 87	https://doi.org/10.22427/NTP-DATA-002-00062-0001-0000-1
7	EFSA DB	A database providing information on food safety assessments, including risk assessments and scientific opinions	• Food: 2,500	https://www.efsa.europa.eu/en/data-report/food-consumption-data
8	NTP	A database containing toxicological data, including results from toxicology studies and chemical testing	• Chemical: 1,060	https://ntp.niehs.nih.gov/data
9	OECD eChemPortal	A portal providing access to information on the properties and effects of chemicals, including toxicity data, collected by the OECD	• Chemical: 1,312,774	https://www.oecd.org/chemicalsafety/risk-assessment/ec-hemportalglobalportaltoinformat-iononchemicalsubstances.htm
10	PDB	A database containing experimentally determined 3D structures of proteins, nucleic acids, and complex assemblies	• Protein: 209,611	https://www.rcsb.org/
11	PRIDE	A database for sharing mass spectrometry-based proteomics data, including protein identification and quantification	• Protein: 26,338	https://www.ebi.ac.uk/pride/archive/
12	PubChem	A database providing information on chemical substances, including their structures, biological activities, and associated data	• Protein: 186,035 • Chemical: 115,728,101	https://pubchem.ncbi.nlm.nih.gov/
13	Toxcast/Tox21	Databases evaluating the potential toxicity of chemicals using high-throughput screening assays and computational models	• Chemical: 2,399,743	https://comptox.epa.gov/dashboard/
14	UK BioBank	A large-scale biomedical database containing health and genetic data from half a million participants in the UK	• Health-related outcomes: 2,650 • Biological samples: 993	https://www.ukbiobank.ac.uk/enable-your-research/
15	UniProt	A comprehensive resource providing information on proteins, including their sequences, functions, and annotations	• UniProtKB/Swiss-Prot: 569,793 • UniProtKB/TrEMBL: 48,272,897	https://www.uniprot.org/

CTD: Comparative Toxicogenomics Database, DNT-DIVER: Developmental Neurotoxicity Data Integration and Visualization Enabling Resource, EFSA-DB: European Food Safety Authority Database, NTP: National Toxicology Program, PDB: Protein Data Bank, PRIDE: Proteomics Identifications Database.

준에서의 다양한 독성 실험데이터를 지속적으로 생산하는 것은 한계가 있다. AOP를 활용하여 독성기전을 반영한 *in silico* 기반 시험법에서는 MIE 수준에서 연구가 많이 진행될 수 있어, 앞으로는 단백질 구조 정보 및 상호작용 정보를 포함하고 있는 CTD, PDB, AlphaFold Protein Structure Database 등의 데이터베이스 활용도가 상대적으로 높아질 것으로 보이며, 화학물질-단백질 상호작용에 대한 실험정보를 포함하는 Binding DB, PubChem DB, PRIDE 등 MIE에 관련된 실험데이터도 추가로 활용 가능할 것으로 예상된다.

독성 예측모델 개발에 고려된 독성종말점은 생식독성과 두 개 이상의 독성종말점을 활용하는 경우(Multi-endpoints)가 각각 21% (4건)로 가장 많았으며, 호흡기 독성과 피부 과민성이 16% (3건)로 두 번째로 많은 것으로 확인되었다. 이외에도 신경독성과 발달독성 각각 10.5% (2건)로 다양한 독성종말점이 예측된 것으로 나타났다(Fig. 3B). 각각의 독성종말점에 대한 독성을 예측하는 모델 개발 시 일반 합성화합물을 대상으로 *in vitro* 또는 *in vivo* 수준에서 독성실험을 통해 생산된 데이터를 활용하였으며 생활화학제품, 화장품 등에 포함된 구성성분을 활용하여 예측모델을 개발한 경우도 확인되었다(Table 1). 대부분의 예측모델들은 독성종말점에 대한 화학물질 및 독성정보(농도 또는 영향)를 포함하는 실험데이터를 기반으로 개발되었다. 공개된 독성 실험데이터가 많은 독성종말점의 경우 예측 모델 개발이 용이하여 많은 연구를 수행할 수 있으며 상대적으로 공개된 독성 실험데이터가 부족한 경우 관련 연구가 미흡할 수 밖에 없는 상황이다. 실험데이터를 기반으로 예측모델을 개발하는 것은 일반적인 방법이지만 실험데이터의 지속적인 생산은 현실적으로 제한적이며 정확한 예측모델을 만들기 위한 데이터 확보에 한계가 있다. 독성기전을 예측모델에 반영하기 위해서는 AOP 구성요소 중 MIE 수준에서 구조 중심의 단백질과 화학물질의 상호작용 부분에 대해 추가적으로 고려할 필요가 있다. 현 수준에서는 실험데이터가 상대적으로 많은 독성종말점에 대한 연구 빈도수가 높은 경향을 보이고 있으나 향후 다양한 독성종말점에 대한 명확한 AOP가 구성되어 있다면 해당 AOP의 MIE 수준에서 *in silico* 기반 예측연구가 많이 수행될 수 있을 것이다.

3. 예측기술 기반 연구동향

In silico 예측기법 분석결과 대부분은 독성 예측모델은 머신러닝 및 딥러닝 기반 세부 알고리즘을 사용하여 개발되었으며 molecular docking, molecular dynamics, pharmacophore screening 등을 포함하는 분자 시뮬레이션(molecular simulation) 기법도 활용된 것을 확인하였다. 2023년도를 기준으로 데이터를 활용한 머신러닝 및 딥러닝 기반 예측연구는 약 90%이며, 그중 머신러닝 알고리즘을 기반으로 한 연구는 전체의 약 69%로 예측 기술 중 가장 많은 비중을 차지하였으며 딥러

닝 알고리즘 기반 연구는 약 21%로 확인되었다. 머신러닝 기반 알고리즘 활용도가 딥러닝에 비해 약 48% 이상 많은 것을 확인하였다. 연구에 활용된 머신러닝 기반 알고리즘으로는 random forest (RF), support vector machine (SVM), logistic regression (LR), extreme gradient boost (XGBoost) 등이 있으며 대표적으로 RF, SVM 알고리즘이 많이 활용되었다(Table 1). RF는 Breiman⁴⁸⁾이 개발한 분류, 회귀분석 등에 사용되는 앙상블 학습 알고리즘이며 여러 의사 결정나무를 구성하여 다양한 부분집합에서 예측을 수행하여 과적합을 줄이고 안정적인 성능을 제공하는 알고리즘이다. SVM은 데이터를 고차원 공간으로 매핑 후 최적의 결정 경계를 찾아 여백(margin)을 최대화하는 기준 선을 정의하는 방식으로 Cortes와 Vapnik⁴⁹⁾이 도입한 알고리즘이다. 보통 신약개발 분야에서 활용되고 있으나 알고리즘 성능이 우수하여 독성 예측 분야에서도 활용되고 있으며 최근 독성 예측 모델 개발에 RF와 SVM의 활용이 증가하는 추세로¹²⁾ AOP를 활용한 독성 예측모델 연구에도 유사한 동향을 보여주고 있다. Gadaleta 등²⁰⁾은 지방간 관련 AOP의 MIE를 예측하기 위해 RF를 활용한 모델을 개발하였으며, 그 성능은 0.96 (Accuracy)으로 높은 수준이었다. 또한 Ciallella 등²⁹⁾은 딥러닝 기법의 DNN과 머신러닝의 RF, SVM 등의 기법을 활용한 KE 예측모델을 개발하였으며, 0.93 (Accuracy)의 높은 성능을 확인하였다.

딥러닝 기반 알고리즘으로는 deep neural network (DNN), convolutional neural network (CNN), graph convolution network (GCN) 등이 활용되었다. DNN⁵⁰⁾은 다층 신경망으로 구성된 모델로 입력층과 출력층 사이에 여러 은닉층을 포함하여 복잡한 학습에 활용되는 알고리즘이다. CNN⁵¹⁾은 이미지의 공간적인 구조를 고려하여 특징을 추출하고 분류 작업에 활용될 수 있는 알고리즘이며, GCN⁵²⁾은 그래프 데이터에서 노드 간 상호작용을 고려하여 학습할 수 있는 알고리즘이다. AOP 기반 독성 예측모델 개발 시 딥러닝 알고리즘을 활용한 연구 사례는 존재하지만, 독성 예측 연구에 사용하는 실험데이터 숫자가 딥러닝에 적용할 수준의 데이터 규모에 비해 매우 부족하기 때문에 머신러닝 알고리즘 기반 연구사례보다 상대적으로 미흡한 수준이다. 딥러닝 알고리즘을 사용한 연구사례로 Ciallella 등²⁹⁾은 DNN을 활용하여 개발한 KE 예측모델을 통해 잠재적인 유해화학물질을 선별하였으며, Matsuzaka와 Uesawa²⁴⁾는 기존 QSAR 기법을 보완하기 위해 화학물질의 3차원(3D) 입체 구조에서 생성된 분자 이미지를 사용하여 개발한 Deep-Snap-DL⁵³⁾을 사용하여 내분비계 장애물질(Endocrine disrupting chemicals, EDCs) 예측 모델을 개발하였다. 독성 데이터에서 Convolutional Neural Network (CNN)를 활용하여 분자 이미지의 특징을 추출하여 화학 구조에 대한 보다 자세한 정보를 사용할 수 있다는 점에서 높은 예측 성능을 보여주었다.

2019년부터 머신러닝 기반 알고리즘을 활용한 독성 예측모

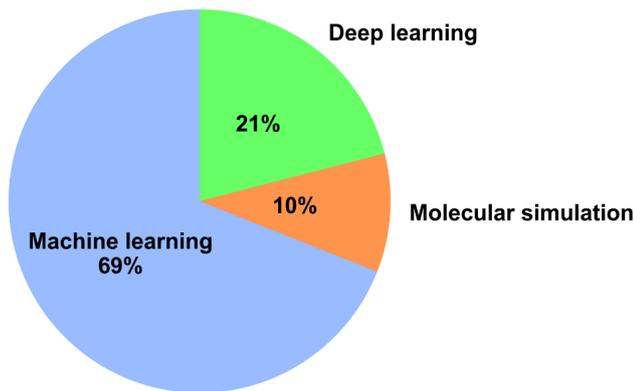


Fig. 4. The *in silico* techniques used to develop the predictive model

델 개발 연구는 점차적으로 증가하고 있으며 2021년부터는 급격한 증가 추세를 보여주고 있다. 딥러닝 기반 알고리즘을 활용한 독성 예측연구는 2019년부터 지속적으로 증가하는 경향을 보여주고 있다(Fig. 2A). 현재 제공되는 독성 실험데이터는 딥러닝 기반 예측모델 개발 연구에 적용하기에는 충분하지 않지만 복잡한 독성기전 정보를 반영하여 높은 정확도를 가지는 예측모델 개발을 위해서는 향후 딥러닝 기반 예측 알고리즘 적용은 필수적이다. 일부 데이터에 한하여 딥러닝 기반 예측 알고리즘이 분자 특성을 예측하는데 기존의 머신러닝 알고리즘보다 더 나은 성능을 나타낸다는 보고도 있다.^{54,55)} 이는 향후 딥러닝 기반의 알고리즘을 활용하여 다양한 연구가 수행되어 활용 사례가 점차 증가할 것으로 예상할 수 있다. 머신러닝 및 딥러닝 기반 예측 알고리즘 외에도 molecular docking, molecular dynamics, pharmacophore screening 기법 등 다양한 분자 시뮬레이션 예측기술이 사용되었다(Fig. 4, Table 1). 연구 동향에서 분자 시뮬레이션은 다른 예측기술에 비해 비중이 상대적으로 낮지만 AOP를 활용한 독성 예측연구에서 화합물과 단백질 간의 상호작용을 식별하는 것이 중요하기 때문에⁵⁶⁾ 향후 독성 기전을 반영한 AOP의 MIE를 활용한 연구를 수행하는데 있어서는 매우 중요한 기법 중 하나가 될 것이다. 새로운 실험데이터의 생산이 제한되는 부분을 고려한다면 향후 예측연구에 독성 기전을 반영하기 위해 분자 시뮬레이션 예측기술을 활용하여 MIE 예측 연구가 많이 수행될 것이다.

최종 선정된 26편의 문헌에서 개발된 모델들의 예측 종류는 정성예측이 54%로 가장 많았으며, 정량 및 정성 모두 확인된 문헌이 전체의 19%로 사용되었고 정량 형태의 예측이 가장 적게 확인되었다. AOP의 MIE를 활용하여 예측연구를 수행하면 단백질-화학물질 간의 결합력(affinity), 단백질 수용체의 agonism, antagonism에 대한 활성(positive), 비활성(negative) 여부를 판단하는 형태로 예측하기 때문에 정성 예측에 대한 연구사례가 상대적으로 더 많은 것으로 확인되었다. 정량 예측은 실험 데이터에 의존적이므로 실험데이터가 부족한 경우가 많아 예

측모델 개발이 어려워 연구사례가 상대적으로 부족함을 보여주었다.

IV. 결 론

본 연구에서는 AOP를 활용한 *in silico* 예측모델 개발 연구에 활용되고 있는 데이터 및 예측기술에 대한 분석을 통해 연구 동향을 살펴보았다. 데이터로는 AOP 요소 중 MIE가 예측지표로 사용된 경우가 가장 많았으며 데이터베이스는 독성 실험데이터를 많이 포함하고 있는 ToxCast/Tox21이 가장 많이 활용되고 있었다. 독성종말점으로는 생식독성 지표에 대한 연구가 가장 많이 수행되고 있음을 보여주었다. 활용된 예측기술 중에서는 머신러닝의 활용도가 딥러닝, 분자 시뮬레이션 기술보다 상대적으로 더 활용도가 높음을 확인하였다.

향후 동물대체시험법으로 활용할 수 있는 *in silico* 기반 시험법 개발을 위해서는 정확한 독성기전을 반영한 예측기술 개발이 필요하다. 독성기전 반영을 위해서는 AOP 구성요소를 적용하는 방안이 필요하며, 다양한 수준에서의 독성 실험데이터 생산에는 제한이 있기 때문에 화학물질 및 단백질 구조의 상호작용을 연구할 수 있는 MIE 수준에서의 연구가 지속적으로 증가할 것으로 보인다. 이에 따라 화학물질 및 단백질의 구조정보를 포함하는 데이터베이스와 상호작용을 분석할 수 있는 예측기술의 활용성이 높아질 것이며, 해당 기술들은 앞으로 동물대체시험법 중 *in silico* 기반 시험법에 활용도가 높아질 것으로 전망된다. 그러나 현재 예측기술은 단일물질 수준에서 예측을 수행하는 사례가 대부분이기 때문에, 혼합물과 그 혼합독성의 예측을 위해 AOP 구성요소를 활용하는 예측모델 개발이 요구되고 있어, 후속 연구가 필요할 것이다.

감사의 글

This study was supported by Korea Environmental Industry & Technology Institute (KEITI; Project no. RS-2023-00215857). We are grateful to Jiwon Choi at Korea Research Institute of Chemical Technology for assistance with data collection.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

- Reddy N, Lynch B, Gujral J, Karnik K. Alternatives to animal testing in toxicity testing: current status and future perspectives in food

- safety assessments. *Food Chem Toxicol.* 2023; 179: 113944.
- Russell WM, Burch RL. The principles of humane experimental technique. London: Methuen; 1959.
 - Korea Institute of Science & Technology Evaluation and Planning. Animal replacement test. Available: https://www.kistep.re.kr/board.es?mid=a10306010000&bid=0031&act=view&list_no=34987 [accessed 26 March 2024].
 - Doke SK, Dhawale SC. Alternatives to animal testing: a review. *Saudi Pharm J.* 2015; 23(3): 223-229.
 - National Institute of Environmental Research (NIER). A study on structural alerts and application of (Q)SARs for mutagenicity screening. Available: <https://scienceon.kisti.re.kr/srch/select-PORSrchReport.do?cn=TRKO201300007514> [accessed 26 March 2024].
 - National Institute of Environmental Research (NIER). Study on the improvement of genotoxicity prediction using QSARs models. Available: <https://doi.org/10.23000/TRKO201900002677> [accessed 26 March 2024].
 - Lee S. Trends in research on alternative to animal experiments for safety evaluation of cosmetic ingredients. *Korean Soc Cosmet Cosmetol.* 2023; 13(2): 359-368.
 - National Institute of Environmental Research (NIER). Guidelines for submitting QSAR data for industry. Available: <https://ecolibrary.me.go.kr/nier/#/search/detail/5704414> [accessed 26 March 2024].
 - European Chemical Agency (ECHA) and European Food Safety Authority (EFSA) with the technical support of the Joint Research Centre (JRC); Andersson N, Arena M, Auteri D, Barmaz S, Grignard E, Kienzler A, et al. Guidance for the identification of endocrine disruptors in the context of Regulations (EU) No 528/2012 and (EC) No 1107/2009. *EFSA J.* 2018; 16(6): e05311.
 - National Institute of Environmental Research (NIER). A study for the improvement of a prediction using QSAR models to find out hazardous substances(I). Available: <https://doi.org/10.23000/TRKO201900003829> [accessed 26 March 2024].
 - Kang D, Jang SW, Lee SW, Lee JH, Lee SH, Kim P, et al. Applicability of QSAR models for acute aquatic toxicity under the act on registration, evaluation, etc. of chemicals in the Republic of Korea. *J Environ Health Sci.* 2022; 48(3): 159-166.
 - Jeong J, Choi J. Artificial intelligence-based toxicity prediction of environmental chemicals: future directions for chemical management applications. *Environ Sci Technol.* 2022; 56(12): 7532-7543.
 - Organisation for Economic Co-operation and Development (OECD). Users' handbook supplement to the guidance document for developing and assessing adverse outcome pathways. Paris: OECD Publishing; 2016.
 - Lee I, Lee S, Ji K. A screening method to identify potential endocrine disruptors using chemical toxicity big data and a deep learning model with a focus on cleaning and laundry products. *J Environ Health Sci.* 2021; 47(5): 462-471.
 - Tan H, Wu J, Zhang R, Zhang C, Li W, Chen Q, et al. Development, validation, and application of a human reproductive toxicity prediction model based on adverse outcome pathway. *Environ Sci Technol.* 2022; 56(17): 12391-12403.
 - Troger F, Delp J, Funke M, Van Der Stel W, Colas C, Leist M, et al. Identification of mitochondrial toxicants by combined in silico and in vitro studies - a structure-based view on the adverse outcome pathway. *Comput Toxicol.* 2020; 14: 100123.
 - Gadaleta D, Spînu N, Roncaglioni A, Cronin MTD, Benfenati E. Prediction of the neurotoxic potential of chemicals based on modelling of molecular initiating events upstream of the adverse outcome pathways of (developmental) neurotoxicity. *Int J Mol Sci.* 2022; 23(6): 3053.
 - Wang CC, Lin YC, Wang SS, Shih C, Lin YH, Tung CW. SkinSens-DB: a curated database for skin sensitization assays. *J Cheminform.* 2017; 9: 5.
 - Allen TE, Goodman JM, Grayson MN, Gutsell S, Russell PJ. Computational approaches for predicting molecular initiating events. *Toxicol Lett.* 2018; 295(Suppl 1): S99-S100.
 - Gadaleta D, Manganelli S, Roncaglioni A, Toma C, Benfenati E, Mombelli E. QSAR modeling of ToxCast assays relevant to the molecular initiating events of aops leading to hepatic steatosis. *J Chem Inf Model.* 2018; 58(8): 1501-1517.
 - Allen TE, Gelzinyte E, Wedlake AJ, Goodman JM, Gutsell S, Russell PJ. Gaining insight into toxicity predicting machine learning algorithms. *Toxicol Lett.* 2019; 314: S280.
 - Kurosaki K, Wu R, Uesawa Y. A toxicity prediction tool for potential agonist/antagonist activities in molecular initiating events based on chemical structures. *Int J Mol Sci.* 2020; 21(21): 7853.
 - Uesawa Y. [AI-based QSAR modeling for prediction of active compounds in MIE/AOP]. *Yakugaku Zasshi.* 2020; 140(4): 499-505. Japanese.
 - Matsuzaka Y, Uesawa Y. Molecular image-based prediction models of nuclear receptor agonists and antagonists using the DeepSnap-Deep Learning approach with the Tox21 10K library. *Molecules.* 2020; 25(12): 2764.
 - Allen TEH, Goodman JM, Gutsell S, Russell PJ. Quantitative predictions for molecular initiating events using three-dimensional quantitative structure-activity relationships. *Chem Res Toxicol.* 2020; 33(2): 324-332.
 - Ivanova H, Dimitrova G, Kuseva C, Schultz TW, Mekenyan OG. Modeling hazard assessment of chemicals based on adducts formation. I. A basis for inclusion of kinetic factors in simulating skin sensitization. *Comput Toxicol.* 2020; 15: 100130.
 - Morger A, Mathea M, Achenbach JH, Wolf A, Buesen R, Schleifer KJ, et al. KnowTox: pipeline and case study for confident prediction of potential toxic effects of compounds in early phases of development. *J Cheminform.* 2020; 12(1): 24.
 - Garcia de Lomana M, Weber AG, Birk B, Landsiedel R, Achenbach J, Schleifer KJ, et al. In silico models to predict the perturbation of molecular initiating events related to thyroid hormone homeostasis. *Chem Res Toxicol.* 2021; 34(2): 396-411.
 - Ciallella HL, Russo DP, Aleksunes LM, Grimm FA, Zhu H. Revealing adverse outcome pathways from public high-throughput screening data to evaluate new toxicants by a knowledge-based deep neural network approach. *Environ Sci Technol.* 2021; 55(15): 10875-10887.
 - Jagiello K, Halappanavar S, Rybińska-Fryca A, Williams A, Vogel U, Puzyn T. Transcriptomics-based and AOP-informed structure-

- activity relationships to predict pulmonary pathology induced by multiwalled carbon nanotubes. *Small*. 2021; 17(15): e2003465.
31. Bringezu F, Gómez-Tamayo JC, Pastor M. Ensemble prediction of mitochondrial toxicity using machine learning technology. *Comput Toxicol*. 2021; 20: 100189.
 32. Jeon B, Lim MH, Choi TH, Kang BC, Kim S. A development of a graph-based ensemble machine learning model for skin sensitization hazard and potency assessment. *J Appl Toxicol*. 2022; 42(11): 1832-1842.
 33. Spînu N, Cronin MTD, Lao J, Bal-Price A, Campia I, Enoch SJ, et al. Probabilistic modelling of developmental neurotoxicity based on a simplified adverse outcome pathway network. *Comput Toxicol*. 2022; 21: 100206.
 34. Guan R, Li N, Wang W, Liu W, Li X, Zhao C. The adverse outcome pathway (AOP) of estrogen interference effect induced by triphenyl phosphate (TPP): integrated multi-omics and molecular dynamics approaches. *Ecotoxicol Environ Saf*. 2022; 234: 113387.
 35. Sakhteman A, Ghosh A, Fortino V. EDTox: an R Shiny application to predict the endocrine disruption potential of compounds. *Bioinformatics*. 2022; 38(7): 2066-2069.
 36. Jin Y, Qi G, Shou Y, Li D, Liu Y, Guan H, et al. High throughput data-based, toxicity pathway-oriented development of a quantitative adverse outcome pathway network linking AHR activation to lung damages. *J Hazard Mater*. 2022; 425: 128041.
 37. Paini A, Campia I, Cronin MTD, Asturiol D, Ceriani L, Exner TE, et al. Towards a qAOP framework for predictive toxicology - linking data to decisions. *Comput Toxicol*. 2022; 21: 100195.
 38. Belfield SJ, Firman JW, Enoch SJ, Madden JC, Erik Tollefsen K, Cronin MTD. A review of quantitative structure-activity relationship modelling approaches to predict the toxicity of mixtures. *Comput Toxicol*. 2023; 25: 100251.
 39. Romano JD, Mei L, Senn J, Moore JH, Mortensen HM. Exploring genetic influences on adverse outcome pathways using heuristic simulation and graph data science. *Comput Toxicol*. 2023; 25: 100261.
 40. Bahl A, Ibrahim C, Plate K, Haase A, Dengiel J, Nymark P, et al. PROTEOMAS: a workflow enabling harmonized proteomic meta-analysis and proteomic signature mapping. *J Cheminform*. 2023; 15(1): 34.
 41. Kim S. Research needs for mixture risk assessment and use of the mixture allocation factor (MAF). *J Environ Health Sci*. 2023; 49(6): i-ii.
 42. Kim S, Kang K, Kim J, Na M, Choi J. Toxicity of organophosphorus flame retardants (OPFRs) and their mixtures in *Aliivibrio fischeri* and human hepatocyte HepG2. *J Environ Health Sci*. 2023; 49(2): 89-98.
 43. Kim ML, Kim SM, Shin JH. Global emerging environmental issues. In: Kho YL, Kim KT, Kim MI, Kim JS, Kim CS, Kim HH, et al. editors. *Environmental health practice*. Seoul: Shinkwang; 2023. p.1523-1541.
 44. United States Environmental Protection Agency. ToxCast dashboard. Available: https://comptox.epa.gov/dashboard/chemical_lists/toxcast [accessed 20 August 2021].
 45. National Institute of Environmental Health Sciences. Tox21: toxicology in the 21st Century. Available: <https://ntp.niehs.nih.gov/whatwestudy/tox21/index.html> [accessed 20 August 2021].
 46. Rotroff DM, Dix DJ, Houck KA, Knudsen TB, Martin MT, McLaurin KW, et al. Using in vitro high throughput screening assays to identify potential endocrine-disrupting chemicals. *Environ Health Perspect*. 2013; 121(1): 7-14.
 47. Wang XH, Fan LY, Wang S, Wang Y, Yan LC, Zheng SS, et al. Relationship between acute and chronic toxicity for prevalent organic pollutants in *Vibrio fischeri* based upon chemical mode of action. *J Hazard Mater*. 2017; 338: 458-465.
 48. Breiman L. Random forests. *Mach Learn*. 2001; 45: 5-32.
 49. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995; 20: 273-297.
 50. Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Comput*. 2006; 18(7): 1527-1554.
 51. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE*. 1998; 86(11): 2278-2324.
 52. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv. 1609.02907 [Preprint]. 2016 [cited 26 March 2024]. Available: <https://doi.org/10.48550/arXiv.1609.02907>.
 53. Uesawa Y. Quantitative structure-activity relationship analysis using deep learning based on a novel molecular image input technique. *Bioorg Med Chem Lett*. 2018; 28(20): 3400-3403.
 54. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: toxicity prediction using deep learning. *Front Environ Sci*. 2016; 3: 80.
 55. Jain S, Siramshetty VB, Alves VM, Muratov EN, Kleinstreuer N, Tropsha A, et al. Large-scale modeling of multispecies acute toxicity end points using consensus of multitask deep learning methods. *J Chem Inf Model*. 2021; 61(2): 653-663.
 56. Tu W, Xu C, Jin Y, Lu B, Lin C, Wu Y, et al. Permethrin is a potential thyroid-disrupting chemical: in vivo and in silico evidence. *Aquat Toxicol*. 2016; 175: 39-46.

〈저자정보〉

이수진(연구원), 박종서(연구원), 김선미(책임연구원), 서명원(선임연구원)