

# Predicting the Fetotoxicity of Drugs Using Machine Learning

Myeonghyeon Jeong and Sunyong Yoo\*

Department of ICT Convergence System Engineering, Chonnam National University, Gwangju 61186, Korea

Received March 7, 2023 / Revised March 15, 2023 / Accepted March 28, 2023

Pregnant women may need to take medications to treat preexisting diseases or diseases that develop during pregnancy. However, some drugs may be fetotoxic and lead to, for example, teratogenicity and growth retardation. Predicting the fetotoxicity of drugs is thus important for the health of the mother and fetus. The fetotoxicity of many drugs has not been established because various challenges hinder the ability of researchers to determine their fetotoxicity. The need exists for *in silico*-based fetotoxicity assessment models, as they can modernize the testing paradigm, improve predictability, and reduce the use of animals and the costs of fetotoxicity testing. In this study, we collected data on the fetotoxicity of drugs and constructed fetotoxicity prediction models based on various machine learning algorithms. We optimized the models for more precise predictions by tuning the hyperparameters. We then performed quantitative performance evaluations. The results indicated that the constructed machine learning-based models had high performance (AUROC >0.85, AUPR >0.9) in fetotoxicity prediction. We also analyzed the feature importance of our model's predictions, which could be leveraged to identify the specific features of drugs that are strongly associated with fetotoxicity. The proposed model can be used to prescreen drugs and drug candidates at a lower cost and in less time. It provides a predictive score for fetotoxicity risk, which may be beneficial in the design of studies on fetotoxicity in human pregnancy.

**Key words :** Fetotoxicity, *in silico* toxicology, machine learning, molecular descriptors, toxicity prediction

## 서 론

임신 중 약물 복용은 태아에게 기형, 자궁 내 사망, 성장 지연, 가역적 또는 비가역적인 약물 효과 등을 나타낼 수 있는 태아 독성을 야기할 수 있다[3]. 하지만 임신부의 기존 질병 또는 임신 중 발생한 질병에 대한 치료를 위해 약물의 복용이 반드시 필요하다. 실제 임신부의 약물 복용 경험 조사 결과 미국에서 1976-2008년 동안 임신부의 60-90%가 1개 이상의 약물을 복용한 것으로 조사되었으며[27], 노르웨이, 캐나다 등의 국가에서 비슷한 연도에 수행된 조사 역시 비슷한 결과를 가지고 있다[2, 14, 18, 32, 38]. 이에 따라 약물의 태아 독성 발생 여부를 예측하는 것은 임신부와 태아의 안전한 치료 환경을 조성하는데 필수적이다. 하지만 임신 중 약물 복용에 대한 안전성이 확보된 약물은 소수에 불과한 데, 이는 약물의 임상시

험에서 유아 및 태아의 건강이 위협받을 수 있는 윤리적 문제로 임신부는 임상시험에서 제외되기 때문이다. 현재 약물의 태아 독성은 시판 후 조사(post-marketing surveillance) 또는 *in vitro*, *in vivo* 기반 독성 시험과 같은 비임상 시험을 기반으로 평가한다. 하지만, 신약의 시판 후 조사는 임신부 및 태아가 독성 위험에 노출된 이후 문제를 해결하는 후행적 접근 방안이기 때문에 임신부의 안전성을 보장할 수 없다. 따라서 약물의 승인 단계 이전에 임신 중 복용에 대한 안전성을 평가하는 선행적 접근 방안이 필요하다. 하지만 현재 약물의 임신부 또는 태아에 대한 잠재적 위험성을 평가하기 위해 수행하는 비임상시험의 경우 1960년대부터 도입된 포유류 종에 대한 생식 및 발달 독성 연구를 포함하는 시험 설계에서 크게 변하지 않았으며, 종에 따라 발생하는 약력학/약동학적 차이에 의해 인간 태아에 대한 독성을 명확하게 설명하기 어렵다[8, 11]. 또한, 비임상시험에서 일관성 있는 태아 독성 평가를 위해 필요한 설치류와 토끼의 새끼 수는 16~20마리이다[21]. 이는 윤리적인 문제에 대한 우려를 유발할 수 있으며[22], 상당한 시간과 비용을 투자해야 한다. 따라서 태아 독성 평가를 위한 시험 설계의 현대화, 예측성 개선, 동물 사용 및 투자 비용 감소를 위해 *in silico* 태아 독성 평가 모델에 대한 필요성이 대두되고 있다[11, 23].

### \*Corresponding author

Tel : +82-62-530-1761, Fax : +82-62-530-3439

E-mail : [syyoo@jnu.ac.kr](mailto:syyoo@jnu.ac.kr)

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

컴퓨터를 기반으로 수행하는 *in silico* 독성 평가 모델은 약물 또는 약물 후보 물질에서 화합물의 구조 또는 특성을 기반으로 다양한 독성 및 부작용에 대한 예측 또는 분석을 수행한다[16, 39]. 독성 평가를 위한 *in silico* 모델은 적은 비용과 시간으로 광범위한 화합물에 대한 조사가 가능하고 독성 및 안전성 위험이 높은 샘플의 우선 순위를 선별함으로써 *in vitro* 혹은 *in vivo* 시험에 사용되는 비용과 동물실험을 최소화함과 동시에 예측성을 개선할 수 있다[29, 30, 36]. 특히, 화합물을 합성 또는 추출하기 이전에 독성을 사전에 선별할 수 있다는 이점은 약물 개발 단계에서 발견되는 독성에 의한 시장 진출에 실패하는 비용을 최소화할 수 있다[15, 33]. 2000년대 초반의 *in silico* 모델은 주로 단일 또는 다중 물리화학적 매커니즘에 대한 지식 기반 시스템을 기반으로 독성 예측을 수행하였으나, 증가하는 대량의 데이터에 대한 자동화가 어렵다는 문제를 가지고 있다[4, 12]. 최근에는 이러한 문제를 극복하기 위해 대량의 데이터를 기반으로 학습하여 결과를 추론하는 기계학습 알고리즘을 적용하고 있다. 또한, 예측 모델의 입력으로 활용되는 화합물 특성 설명자에 대한 연구가 이루어지고 있으며, 이에 따라 기계학습 기반 독성 예측 모델에 대한 연구가 크게 발전하고 있다[43]. 기계학습 기반 독성 평가에 주로 사용되는 모델은 주로 Support Vector Machine, Random Forest, Artificial Neural Network 등의 알고리즘이 활용되고 있으며[5, 10, 41], 예측 모델의 입력으로 활용될 화합물 특성 설명자는 Extended Connectivity Fingerprints, Molecular Access System keys 와 같은 구조적 특성 설명자와 물리화학적, 생리학적 특성 설명자 등을 활용한다[17, 34, 43].

본 연구는 약물 또는 약물 후보 물질에서 태아 독성을 나타낼 위험성이 높은 물질을 제시하는 *in silico* 모델을 구축하기 위해 다양한 기계학습 알고리즘을 적용하였다. 구축한 모델의 태아 독성 예측 성능을 개선하기 위해 약물을 효과적으로 표현할 수 있는 구조적 특성 및 생화학적 특성을 선별했으며, 모델의 초매개변수(hyperparameter)를 최적화하였다. 또한, 각 기계학습 알고리즘 기반 모델의 정량적 성능 평가를 수행하고, 모델 예측 결과를 정성적으로 분석함으로써 구축한 모델의 태아 독성 예측의 유효성을 확인하였다.

## 재료 및 방법

### 태아 독성 정보 데이터셋 구축

기계학습 기반 태아 독성 예측 모델을 학습시키기 위해 약물의 태아 독성 정보와 구조적 정보를 포함한 데이터셋을 구축하였다. 각 약물의 태아 독성 정보는 한국의약품안전관리원, 호주의 Therapeutic Goods Administration (TGA)에서 제공하는 데이터베이스와 독성 정보를 포함한 약물

의 다양한 정보를 포함한 온라인 데이터베이스인 Drug Bank에서 수집하였다[1, 25, 42]. 각 데이터베이스에서 구조적 정보가 불분명한 약물, 바이러스, 백신, 독소, 다중복용 약물은 제외하였다. 약물의 태아 독성 예측을 위한 모델의 학습을 위해 입력 벡터를 생성해야 하므로, 이를 위해 각 약물의 구조 정보를 COCONUT (Compound Combination-Oriented Natural Product Database with Unified Terminology), PubChem, ChEBML에서 SMILES (simplified molecular-input line-entry system) 정보를 수집하였다[19, 24, 44]. 수집한 데이터는 각기 다른 데이터베이스에서 수집 되었으므로 데이터셋의 중복 발생을 방지하기 위해 SMILES 구조가 중복되는 경우 제거하였다. 중복되어 제거되는 약물의 태아 독성 등급은 TGA 데이터베이스를 기준으로 분류하였다. 본 연구는 약물의 인간 태아에 대한 독성 평가를 목적으로 하므로 각 약물의 태아 독성 여부는 인간 태아에서 명확한 위해가 관찰되는 경우를 기준으로 ‘태아 독성 있음’과 ‘태아 독성 없음’으로 분류하였다. 한국의약품안전관리원에서 제공하는 임부 급기 등급 약물 리스트의 경우 1등급으로 분류된 약물은 사람에서 태아에 대한 위해성이 명확하므로 ‘태아 독성 있음’, 2등급으로 분류된 약물은 잠재적인 위해성이므로 ‘태아 독성 없음’으로 분류하였다. 호주의 TGA에서 제공하는 데이터베이스의 경우 ‘category D, X’로 분류된 약물은 인간의 태아에서 명확한 위해가 관찰된 경우이므로 ‘태아 독성 있음’, ‘category A, B1, B2, B3’는 인간 태아에 대해서 위해가 관찰되지 않은 경우이므로 ‘태아 독성 없음’으로 분류하였으며, ‘category C’의 경우 실험 및 통계가 충분하지 않아 독성 여부가 명확하지 않은 경우이므로 제거하였다. DrugBank에서 수집한 약물은 ‘기형 발생’으로 분류되어 있는 약물이거나, 동물 실험에서 관찰된 결과도 함께 포함되어 있으므로 제공하는 독성 설명을 직접 확인하여 인간 태아에서 독성을 발현하는 경우는 ‘태아 독성 있음’, 그 외에는 ‘태아 독성 없음’으로 분류하였다. 결과적으로 ‘태아 독성 있음’ 약물은 285개, ‘태아 독성 없음’ 약물은 947개로 총 1,232개의 약물의 태아 독성 정보가 수집되었다(Fig. 1).

### 약물 특성 벡터 생성

모델의 입력 값으로 사용되는 특성 벡터는 각 약물의 SMILES 정보를 기반으로 구조적, 물리화학적, 생리활성 특성 벡터를 생성하였다. 구조적 특성 벡터의 경우, 분자가 가지는 각 원자의 특정 반경에 따른 구조적 환경 특성을 비트 벡터 형태로 표현하는 Morgan fingerprint를 생성하였으며, 각 비트는 해당 분자의 하위 분자구조를 설명할 수 있다. 하지만 두 개 이상의 다른 하위 분자구조가 동일한 비트로 생성되는 비트 충돌 문제가 발생할 수 있으므로 벡터의 길이를 늘려 비트 충돌 가능성을 줄일 필

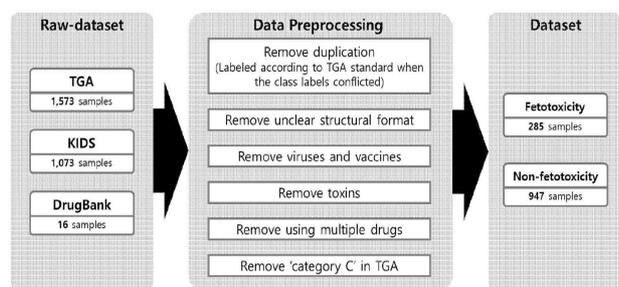


Fig. 1. Summary of dataset construction. We performed pre-processing on the collected datasets, including TGA, KIDS, and DrugBank. We removed duplicate data, unclear structures, viruses, vaccines, toxins, multiple drugs. In addition, "category C" in the TGA dataset was removed. The drugs were then labeled as 'fetotoxicity' or 'non-fetotoxicity' based on their toxicity to the human fetus.

요가 있다. 따라서, 본 연구는 Morgan fingerprint 벡터의 길이를 2,048로 설정하고 원자의 반경을 3으로 설정하여 비트 충돌 가능성을 줄이고 각 원자의 구조적 환경 특성을 최대한 반영하였다. 물리화학적 특성인 분자량 특성은 500Da 미만의 분자량을 가지는 약물은 수동 확산을 통해 태반을 쉽게 통과한다는 연구에 따라 500Da를 기준으로 생성하였다. 이에 따라, 분자량이 500Da 미만인 약물은 '0', 이상인 경우 '1'로 분자량 특성을 생성하였다[28]. 또한, 약물의 물리화학적 특징을 활용한 생리활성 특성 QED (quantitative estimate of druglikeness), OB (oral bio-availability), BBBP (blood brain barrier permeability), GIA (gastrointestinal absorption)을 생성하였다[6, 13, 40]. 분자의 구조적 특성 및 물리화학적 특성은 오픈소스 라이브러리인 'RDKit'을 통해 계산하였다[31].

#### 기계학습 기반 태아 독성 예측 모델 구축 및 최적화

약물의 태아 독성 예측을 위한 기계학습 모델은 선형 모델인 Logistic Regression, 트리 기반 앙상블 모델인 Random Forest, Extra Trees, Gradient Boosting Machine, XGBoost (eXtreme Gradient Boosting)으로 총 5가지의 기계학습 기반 모델을 구축하였다. 구축한 각 모델의 태아 독성 예측 정확도 향상을 위해서 입력되는 초매개변수를 조정할 필요가 있다. 우리는 학습 알고리즘의 최고의 성능을 위한 초매개변수를 탐색 및 적용하기 위해 Bayesian optimization 알고리즘을 적용하였다[37].

#### 예측 모델 정량적 성능 평가

본 연구에서 제안하는 기계학습 기반 태아 독성 예측 모델은 '태아 독성 있음'과 '태아 독성 없음'으로 분류하는 이진 분류임에 따라 이에 대한 오차 행렬(confusion matrix)을 활용해 성능 평가 지표를 계산한다(Table 1).

Table 1. Confusion matrix for binary classification

		Predicted	
		Positive	Negative
Actual	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Table 2. Performance metrics and its calculation formula

Performance metrics	Formula
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall (=TPR)	$\frac{TP}{TP + FN}$
FPR	$\frac{FP}{TN + FP}$
F1-Score	$\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

오차 행렬을 통해 계산한 정보를 바탕으로 정확도 (accuracy), 정밀도(precision), 재현도(recall), F1-score, AUROC (area under the receiver operating characteristic curve), AUPR (area under the precision-recall curve)을 성능 평가 지표로 활용하였다(Table 2).

AUROC를 계산하기 위한 ROC curve는 모든 임계값 (threshold)에서 분류 모델의 성능을 보여주는 그래프로 각 임계값에 따른 TPR (true positive rate)를 y축, FPR (false positive rate)를 x축으로 그린다. AUPR을 계산하기 위한 PR curve 역시 모든 임계값에 대한 정밀도를 y축, 재현도를 x축으로 그린다. 각 모델의 성능 평가를 위한 임계값 기준은 최고의 F1-score가 계산되도록 설정하였다.

#### 기계학습 모델 특성 중요도 계산

특성 중요도(feature importance)는 모델이 예측을 수행하는 데 각 특성이 미치는 영향을 정량적인 점수로 계산한 것이다. Logistic Regression 모델의 경우 각 특성의 변화에 따른 예측 값의 변화를 나타내는 계수(coefficient) 값을 특성 중요도로 활용하였다. 이 때 계수가 0에 가까울수록 해당 특성이 예측 값의 변화에 미치는 영향이 적다는 것을 의미한다. Random Forest, Extra Trees, Gradient Boosting Machine, XGBoost 기반 모델의 특성 중요도는 트리 기반 알고리즘의 각 특성이 트리를 분할하는 과정에서 불순도(impurity)를 얼마나 감소하는지에 대한 측정 값을 특성 중요도로 활용했다. 트리 기반 모델에서 불순도를 많이 감소하는 특성은 모델의 예측 성능을 개선하는 데

밀접한 관련성을 가지고 있다.

### 결과 및 고찰

#### 모델 학습 및 평가 데이터

약물의 태아 독성 정보와 특성 벡터를 결합한 데이터셋을 학습 셋, 테스트 셋으로 나눈 후, 학습 셋을 기계학습 모델이 학습하게 하였다. 구축한 데이터셋은 클래스가 불균형한 비대칭성을 가지고 있으며, 이는 모델의 성능을 저해할 수 있으므로 클래스 불균형을 해소한 후 학습 및 평가하도록 하였다. 모델 성능 평가를 위한 테스트 셋은 ‘태아 독성 있음’으로 분류된 약물의 20%를 선택하였다. 이후 ‘태아 독성 없음’으로 분류된 약물은 테스트 셋에 ‘태아 독성 있음’으로 분류된 약물의 수와 1:1로 맞추었다. 테스트 셋에 포함되지 않은 약물은 모두 학습 셋으로 지정하였다. 결과적으로 학습 셋의 ‘태아 독성 있음’ 약물의 수는 228개, ‘태아 독성 없음’ 약물의 수는 890개이며, 테스트 셋의 ‘태아 독성 있음’과 ‘태아 독성 없음’ 약물의 수는 각각 57개이다. 학습 셋의 클래스 균형은 대표적인 오버 샘플링 기법인 SMOTE (Synthetic Minority Over-sampling Technique) 기법을 적용하여 각 클래스의 비율을 1:1 이 되도록 조정하였다[9]. 모델의 초매개변수는 학습 셋을 활용해 Bayesian optimization 최적화 기법에 대해 5-fold

cross validation으로 평가한 후 최고의 성능을 보이는 초매개변수를 선택하였다(Table 3).

#### 모델 예측 성능 평가 결과

태아 독성 예측을 위한 5개의 기계학습 기반 모델에 대한 ROC curve와 PR curve를 그린 후 각 모델의 분류 성능을 비교하였다(Fig. 2). Random Forest, Gradient Boosting Machine 모델이 각각 AUROC, AUPR 지표에서 가장 높은 성능을 보였다. 구축한 모든 모델의 AUROC는 0.85 이상, 모든 모델의 AUPR은 0.9 이상이며, 이는 완전히 무작위적인 분류 모델의 AUROC 및 AUPR이 0.5를 가지는 것에 반해 1에 가까운 값을 가지므로 유효한 분류 모델임을 알 수 있다. 이후 가장 높은 F1-score를 가지는 각 모델의 임계값을 기준으로 설정하여 정량적 성능 평가를 수행한 결과, Random Forest 기반 모델이 정밀도, Extra Trees 기반 모델이 정확도, 재현도, F1-score 지표에서 높은 성능을 보였다(Table 4).

#### 기계학습 기반 모델 약물 예측 결과 분석

약물의 인간에 대한 태아 독성 여부는 동물 실험을 제외하고 임산부의 약물 복용에 따른 태아 독성 발생에 대한 관측 보고에 의해 정해지므로, 실제로는 인간에 대한 태아 독성을 포함하고 있으나 보고되지 않은 경우에 ‘태

Table 3. Bayesian optimization configuration and optimal values for each machine learning model

Model	Hyperparameter name	Range	Optimal value
Logistic regression	C	[0.001, 100]	0.34
Random forest	n_estimators	[10, 250]	105
	min_samples_split	[2, 25]	2
	max_features	[0.1, 0.999]	0.15
	max_depth	[30, 100]	68
	min_samples_leaf	[1, 10]	1
Extra tree	n_estimators	[10, 250]	130
	min_samples_split	[10, 50]	11
	max_features	[0.1, 0.999]	0.1
	max_depth	[20, 100]	82.47
	min_samples_leaf	[1, 20]	1
Gradient boosting machine	learning rate	[0.001, 0.1]	0.001
	n_estimators	[200, 500]	248
	max_depth	[1, 100]	92
	min_samples_split	[2, 25]	3
	min_samples_leaf	[1, 25]	1
	subsample	[0, 1]	0.58
XGBoost	eta	[0.001, 0.1]	0.02
	n_estimators	[300, 700]	347
	min_child_weight	[1, 2]	1
	gamma	[0, 2]	0.31
	max_depth	[5, 15]	13
	subsample	[0.5, 1]	0.65
	colsample_bytree	[0.5, 1]	0.52

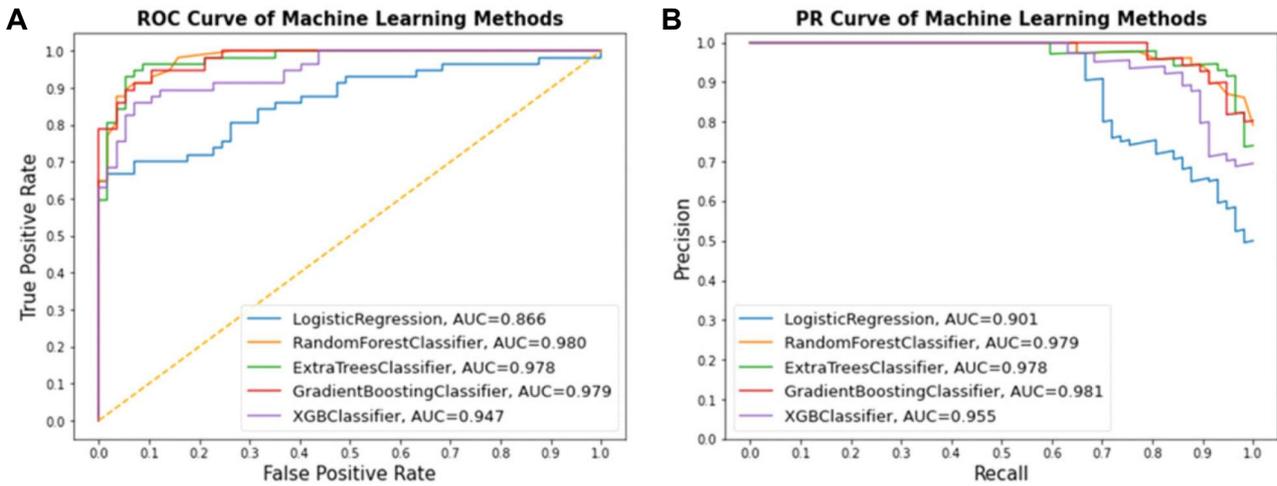


Fig. 2. ROC curve and PR curve for all models. (A) ROC curve graphed for the false positive rate and true positive rate according to all thresholds of each model. The closer the curve of the graph is to the top left corner, the better the classification performance. (B) PR curve graphed for precision and recall according to all thresholds of each model. The closer the curve of the graph is to the top right corner, the better the classification performance.

Table 4. Quantitative performance evaluation results for each fetal toxicity prediction model

Model	Accuracy	AUROC	AUPR	Precision	Recall	F1 score	Threshold
Logistic regression	0.816	0.866	0.901	0.909	0.702	0.792	0.1288
Random forest classifier	0.921	0.980	0.979	0.929	0.912	0.920	0.1619
Extra trees classifier	0.939	0.978	0.978	0.917	0.965	0.940	0.1000
Gradient boosting classifier	0.921	0.979	0.981	0.900	0.947	0.923	0.4168
XGBoost	0.895	0.947	0.955	0.925	0.860	0.891	0.1271

아 독성 없음'으로 분류될 수 있다. 이와 같이 실제 레이블이 잘못 분류된 경우에도 학습된 모델은 데이터를 올바르게 분류할 수 있으며, 이 경우 위양성(false positive) 결과 분석을 수행하여 데이터의 실제 분류가 잘못되었는지 확인할 수 있다. 따라서 우리는 각 모델이 약물의 표지가 '태아 독성 없음' 이지만 '태아 독성 있음'으로 분류한 약물(false positive)에 대한 결과를 분석하였다(Table 5). 모든 모델에서 위양성으로 분류된 약물인 'gadobenate'은 가돌리늄(gadolinium) 기반 MRI 조영제로 사용된다. 가돌리늄 기반 조영제의 경우 태반을 통과하고 양수 내에 장기간

남아 있을 수 있으며, 태아 독성을 유발할 수 있는 잠재적인 위험이 존재하는 것으로 밝혀져 있다[7]. 인간을 대상으로 한 통제 시험은 수행되지 않았으나, 소수 연구에서 태아에 대한 특별한 부작용이 보고되지 않았다[35]. 하지만 가돌리늄 기반 조영제는 여전히 태아의 생장 지연, 기형 발생 등의 잠재적인 위험을 포함하고 있으므로 임신 중 사용을 권장하지 않고 있다[7]. 해당 약물은 '태아 독성 없음'으로 분류되어 있으나 약물에 대해 태반을 통과하여 태아에게 독성이 나타날 수 있음을 본 연구에서 구축한 모델이 정상적으로 분류한 것을 확인하였다.

Table 5. List of drugs classified as 'fetotoxic' by machine learning models

Rank	Logistic regression	Random forest	Extra trees	Gradient boosting machine	XGBoost
1	Remdesivir	Pyrimethamine	Pyrimethamine	Pyrimethamine	Gadobenate
2	Gadobenate	Levocabastine	Levocabastine	Gadobenate	Fingolimod
3	Varenicline	Gadobenate	Gadobenate	Levocabastine	Cyclophosphamide
4	Fingolimod	Cobimetinib fumarate	Dutasteride	Fingolimod	Cobimetinib fumarate
5			Cyclophosphamide	Cobimetinib fumarate	
6				Lithium salts	

These drugs have been classified as 'non-fetotoxic' in the dataset, but can be potential candidates for fetotoxicity.

Table 6. Feature importance score for each machine learning model

Logistic regression			Random forest classifier		Extra trees classifier		Gradient boosting machine		XGBoost	
Rank	Feature	Score	Feature	Score	Feature	Score	Feature	Score	Feature	Score
1	1855	0.6750	80	0.0663	80	0.0719	1380	0.1288	1380	0.0103
2	1484	0.6604	1380	0.0507	1380	0.0479	80	0.1010	1873	0.0080
3	739	0.6452	1873	0.0449	1873	0.0392	1873	0.0492	80	0.0072
4	555	0.5769	1750	0.0274	1750	0.0343	QED	0.0345	1329	0.0050
5	75	0.5735	QED	0.0241	1	0.0229	715	0.0198	578	0.0046
6	533	0.5664	1	0.0230	807	0.0196	1	0.0171	1771	0.0042
7	1583	0.5649	650	0.0177	650	0.0190	1484	0.0135	1679	0.0040
8	MW	0.5369	807	0.0142	695	0.0158	114	0.0118	1395	0.0038
9	796	0.5323	715	0.0119	1199	0.0112	1154	0.0091	909	0.0035
10	1917	0.5158	695	0.0110	926	0.0110	1750	0.0076	1061	0.0033

### 각 모델의 특성 중요도

구축한 각 기계학습 기반 모델에서 계산된 상위 10개의 특성 중요도를 확인하였다(Table 6). 각 모델에서 높은 특성 중요도를 가지는 특성 중 가장 많은 빈도수를 보이는 것은 ‘80’, ‘1380’, ‘1873’번 특성으로, 태아 독성과 관련성이 높은 약물 하위 분자 구조로 보여진다. 이는 향후 연구에서 태아 독성과 연관된 분자 하위 구조를 탐색 및 시각화를 수행하여 태아 독성과 관련된 약물의 분자 하부구조를 탐색할 예정이다. 높은 특성 중요도를 가지는 물리화학적, 생리학적 특성은 ‘MW’와 ‘QED’이다. ‘MW’는 약물의 태반 통과에서 수동 확산과 높은 관련성을 가지는 물리화학적 특성이다[28]. ‘QED’는 분자의 물리화학적 특성의 분포를 정량화하여 나타내는 척도로[6], 계산하는 데 활용되는 물리화학적 특징 중 TPSA (topological polar surface area), logP, 수소 결합 부위 수는 태반 통과와 관련성이 있다[20, 26].

### 고 찰

약물의 잠재적인 태아 독성을 성공적으로 평가하기 위한 *in silico* 모델의 필요성이 대두되고 있다. 이에 따라 본 연구는 약물의 태아 독성 정보를 가진 약물 데이터셋을 활용하여 기계학습 기반 태아 독성 예측 모델을 구축하였으며, 높은 예측 정확도를 위해 모델을 최적화 하였다. 제안하는 모델에 대한 실제 예측 성능을 정량적으로 평가하였으며, 모든 모델의 성능 지표에서 유의한 결과를 보였다. 이를 통해 제안하는 기계학습 기반 태아 독성 예측 모델이 약물 및 약물 후보 물질에 대한 태아 독성 예측이 유효함을 증명하였다. 또한, 모든 모델이 ‘태아 독성 있음’으로 예측했으나 실제 분류는 ‘태아 독성 없음’인 약물에 대한 문헌 조사를 수행한 결과 해당 약물은 높은 태아 독성 위험으로 인해 임산부의 사용을 권장하지 않는 약물인 것을 확인하였다. 이는 제안하는 모델이 표지가

잘못된 약물에서도 약물의 특성을 추론하여 실제 태아 독성 가능성을 예측한다는 것을 확인할 수 있다. 구축한 모델은 약물의 합성, 추출 과정 이전 단계에서 태아 독성을 예측할 수 있으므로 수많은 약물 및 약물 후보 물질에 대한 태아 독성 위험도를 계산할 수 있을 것이다. 계산된 태아 독성 위험도는 약물 개발 단계에서 태아 독성 위험도가 높은 약물 후보 물질을 제외하여 신약의 시장 진출 실패 가능성을 줄일 수 있고, 약물의 태아 독성 시험에서 우선 순위를 부여하여 효율적인 약물의 태아 독성 식별이 가능할 것이다.

하지만, 기계학습 기반 독성 예측 모델은 실험 데이터의 오류가 존재하는 경우에도 이를 학습할 수 있으며 이는 모델에 대한 불확실성을 생성할 수 있다. 제안하는 모델 역시 기계학습 기반으로 수행하는 예측 모델이므로 이러한 불확실성에서 자유롭지 못하다. 이를 극복하기 위해 향후 연구에서는 본 연구에서 제시한 특성 중요도 등을 활용해 해석 가능한 인공지능(eXplainable Artificial Intelligence, XAI) 모델을 구축하여 태아 독성과 연관성이 높은 약물 분자 하위 구조를 선별하고 제시할 계획이다. 이를 통해 *in silico* 태아 독성 예측 모델이 단순한 약물의 태아 독성 예측에서 나아가 유의미한 분자 하위 구조를 제안함으로써 태아 독성 시험 및 연구에 도움이 될 것을 기대한다.

### 감사의 글

본 연구는 2021년도 식품의약품안전처의 연구개발비(21162MFDS045), 과학기술정보통신부 및 정보통신기획평가원의 지역지능화혁신인재 양성사업(IITP-2023-RS-2022-00156287), 과학기술정보통신부 및 정보통신기획평가원의 ICT혁신인재4.0사업(IITP-2023-00156385)의 연구 결과로 수행되었으며 이에 감사드립니다.

## Conflict of interest statement

The author declares no conflicts of interest.

## References

1. Australian Government Department of Health Therapeutic Goods Administration. Prescribing medicines in pregnancy database. [cited 2021 August 3]; Available from: <https://www.tga.gov.au/prescribing-medicines-pregnancy-database>.
2. Bakker, M. K., Jentink, J., Vroom, F., Van Den Berg, P. B., De Walle, H. E. and De Jong-Van Den Berg, L. 2006. Maternal medicine: Drug prescription patterns before, during and after pregnancy for chronic, occasional and pregnancy-related drugs in the Netherlands. *BJOG*. **113**, 559-568.
3. Bánhidly, F., Lowry, R. B. and Czeizel, A. E. 2005. Risk and benefit of drug use during pregnancy. *Int. J. Med. Sci.* **2**, 100-106.
4. Barratt, M. D. 2000. Prediction of toxicity from chemical structure. *Cell. Biol. Toxicol.* **16**, 1-13.
5. Baskin, I. I. 2018. Machine learning methods in computational toxicology, *Computational Toxicology*, pp. 119-139, Springer.
6. Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S. and Hopkins, A. L. 2012. Quantifying the chemical beauty of drugs. *Nat. Chem.* **4**, 90-98.
7. Bulas, D. and Egloff, A. 2013. Benefits and risks of MRI in pregnancy. *Semin. Perinatol.* **37**, 301-304.
8. Carter, A. M. 2007. Animal models of human placentation-a review. *Placenta* **28**, S41-S47.
9. Chawla, N. V., Bowyer, K. W., Hall, L. O. and Kegelmeyer, W. P. 2002. SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321-357.
10. Ciallella, H. L. and Zhu, H. 2019. Advancing computational toxicology in the big data era by artificial intelligence: data-driven and mechanism-driven modeling for chemical toxicity. *Chem. Res. Toxicol.* **32**, 536-547.
11. Clements, J., Hawkes, R., Jones, D., Adjei, A., Chambers, T., Simon, L., Stemplewski, H., Berry, N., Price, S. and Pirmohamed, M. 2020. Predicting the safety of medicines in pregnancy: A workshop report. *Reprod. Toxicol.* **93**, 199-210.
12. Compton, P., Preston, P., Edwards, G. and Kang, B. 1996. Knowledge based systems that have some idea of their limits. *Tenth Knowledge Acquisition for Knowledge-Based Systems Workshop*. November. Banff, Canada.
13. Daina, A. and Zoete, V. 2016. A boiled-egg to predict gastrointestinal absorption and brain penetration of small molecules. *ChemMedChem* **11**, 1117-1121.
14. Daw, J. R., Mintzes, B., Law, M. R., Hanley, G. E. and Morgan, S. G. 2012. Prescription drug use in pregnancy: a retrospective, population-based study in British Columbia, Canada (2001-2006). *Clin. Ther.* **34**, 239-249. e2.
15. Dearden, J. C. 2003. *In silico* prediction of drug toxicity. *J. Comput. Aided. Mol. Des.* **17**, 119-127.
16. Deeb, O. and Goodarzi, M. 2012. *In silico* quantitative structure toxicity relationship of chemical compounds: some case studies. *Curr. Drug. Saf.* **7**, 289-297.
17. Durant, J. L., Leland, B. A., Henry, D. R. and Nourse, J. G. 2002. Reoptimization of MDL keys for use in drug discovery. *J. Chem. Inf. Comput. Sci.* **42**, 1273-1280.
18. Engeland, A., Bjørge, T., Klungsoyr, K., Hjellvik, V., Skurtveit, S. and Furu, K. 2018. Trends in prescription drug use during pregnancy and postpartum in Norway, 2005 to 2015. *Pharmacoepidemiol. Drug. Saf.* **27**, 995-1004.
19. Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D. and Al-Lazikani, B. 2012. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic. Acids. Res.* **40**, D1100-D1107.
20. Giaginis, C., Theocharis, S. and Tsantili-Kakoulidou, A. 2012. Current toxicological aspects on drug and chemical transport and metabolism across the human placental barrier. *Expert Opin. Drug Metab. Toxicol.* **8**, 1263-1275.
21. Guideline, I. H. T. 2005. Detection of toxicity to reproduction for medicinal products & toxicity to male fertility S5 (R2).
22. Guideline, I. H. T. 2009. Guidance on nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals M3 (R2). *International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use*.
23. Ke, A. B., Greupink, R. and Abduljalil, K. 2018. Drug dosing in pregnant women: challenges and opportunities in using physiologically based pharmacokinetic modeling and simulations. *CPT. Pharmacometrics. Syst. Pharmacol.* **7**, 103-110.
24. Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A. and Yu, B. 2019. PubChem 2019 update: improved access to chemical data. *Nucleic. Acids. Res.* **47**, D1102-D1109.
25. Korea Institute of Drug Safety & Risk Management. Pregnancy taboo. [cited 2021 August 10]; Available from: <https://www.drugsafe.or.kr/iwt/ds/ko/useinfo/EgovDurInfoSerPn.do>.
26. Mathiesen, L., Buerki-Thurnherr, T., Pastuschek, J., Aengenheister, L. and Knudsen, L. E. 2021. Fetal exposure to environmental chemicals; insights from placental perfusion studies. *Placenta* **106**, 58-66.
27. Mitchell, A. A., Gilboa, S. M., Werler, M. M., Kelley, K. E., Louik, C., Hernández-Díaz, S. and Study, N. B. D. P. 2011. Medication use during pregnancy, with particular focus on prescription drugs: 1976-2008. *Am. J. Obstet. Gynecol.* **205**, 51. e1-e8.
28. Pacifici, G. M. and Nottoli, R. 1995. Placental transfer of drugs administered to the mother. *Clin. Pharmacokinet.* **28**, 235-269.
29. Pérez Santín, E., Rodríguez Solana, R., González García,

- M., García Suárez, M. D. M., Blanco Díaz, G. D., Cima Cabal, M. D., Moreno Rojas, J. M. and López Sánchez, J. I. 2021. Toxicity prediction based on artificial intelligence: A multidisciplinary overview. *Wiley. Interdiscip. Rev. Comput. Mol. Sci.* **11**, e1516.
30. Raies, A. B. and Bajic, V. B. 2016. *In silico* toxicology: computational methods for the prediction of chemical toxicity. *Wiley. Interdiscip. Rev. Comput. Mol. Sci.* **6**, 147-172.
31. RDKit: Open-Source Cheminformatics Software. [cited Septemeber 7 2021]; Available from: <https://www.rdkit.org>.
32. Refuerzo, J. S., Blackwell, S. C., Sokol, R. J., Lajeunesse, L., Firchau, K., Kruger, M. and Sorokin, Y. 2005. Use of over-the-counter medications and herbal remedies in pregnancy. *Am. J. Perinatol.* **22**, 321-324.
33. Reisfeld, B. and Mayeno, A. N. 2012. What is computational toxicology?, *Computational Toxicology*, 3-7, Springer.
34. Rogers, D. and Hahn, M. 2010. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **50**, 742-754.
35. Scarsbrook, A. F., Evans, A. L., Owen, A. R. and Gleeson, F. V. 2006. Diagnosis of suspected venous thromboembolic disease in pregnancy. *Clin. Radiol.* **61**, 1-12.
36. Segall, M. D. and Barber, C. 2014. Addressing toxicity risk when designing and selecting compounds in early drug discovery. *Drug Discov. Today* **19**, 688-693.
37. Snoek, J., Larochelle, H. and Adams, R. P. 2012. Practical bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems* 25. December 3-8. Nevada, USA.
38. Thorpe, P. G., Gilboa, S. M., Hernandez-Diaz, S., Lind, J., Cragan, J. D., Briggs, G., Kweder, S., Friedman, J. M., Mitchell, A. A. and Honein, M. A. 2013. Medications in the first trimester of pregnancy: most common exposures and critical gaps in understanding fetal risk. *Pharmacoepidemiol. Drug. Saf.* **22**, 1013-1018.
39. Valerio Jr, L. G. 2009. *In silico* toxicology for the pharmaceutical sciences. *Toxicol. Appl. Pharmacol.* **241**, 356-370.
40. Veber, D. F., Johnson, S. R., Cheng, H.-Y., Smith, B. R., Ward, K. W. and Kopple, K. D. 2002. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **45**, 2615-2623.
41. Wang, M. W., Goodman, J. M. and Allen, T. E. 2020. Machine learning in predictive toxicology: recent applications and future directions for classification models. *Chem. Res. Toxicol.* **34**, 217-239.
42. Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C. and Sayeeda, Z. 2018. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic. Acids. Res.* **46**, D1074-D1082.
43. Wu, Y. and Wang, G. 2018. Machine learning based toxicity prediction: from chemical structural description to transcriptome analysis. *Int. J. Mol. Sci.* **19**, 2358.
44. Yoo, S., Ha, S., Shin, M., Noh, K., Nam, H. and Lee, D. 2018. A data-driven approach for identifying medicinal combinations of natural products. *IEEE. Access* **6**, 58106-58118.

## 초록 : 기계학습 기반 약물의 태아 독성 예측 연구

정명현 · 유선용\*

(전남대학교 ICT융합시스템공학과)

임산부의 기존 질병 또는 임신 중 발생한 질병을 치료하기 위한 약물의 사용은 태아에게 잠재적인 위험이 될 수 있으므로 약물의 태아 독성 여부를 예측하는 것이 필수적이다. 하지만 약물의 태아 독성을 밝혀내는 것은 많은 시간과 비용을 필요로 하며 인간 태아에게서 독성 작용을 나타내는 근거가 불분명하다. 이에 따라 최근 태아 독성 평가를 위한 시험 설계의 현대화, 예측성 개선, 동물 사용 및 투자 비용 감소를 위한 *in silico* 태아 독성 평가 모델의 필요성이 대두되고 있다. 본 연구는 태아 독성 정보를 수집하고 다양한 기계학습 알고리즘을 적용하여 태아 독성 예측이 가능한 모델을 구축하였으며, 태아 독성 예측 모델의 입력 값으로 활용하기 위해 각 약물에 대한 구조적 및 생리학적 특성 벡터를 생성하였다. 이후 예측 정확도 개선을 위해 초매개변수를 조정하여 모델을 최적화 하였다. 개발한 태아 독성 예측 모델의 유효성을 검증하기 위해 학습 셋과 독립된 테스트 셋을 활용하여 정량적 성능 평가를 수행하였으며, 모든 모델의 약물 및 약물 후보 물질의 태아 독성 여부를 예측할 수 있는 것을 확인하였다(AUROC>0.85, AUPR>0.9). 나아가, 예측 모델의 특성 중요도를 분석하여 태아 독성과 관련성이 높은 약물의 특성을 제시하였다. 제한한 모델은 적은 비용과 시간으로 예측 점수를 제공함으로써 인간에 대한 태아 독성 연구를 설계하는 과정에 도움이 될 것을 기대한다.