

## TOPKAT<sup>®</sup>, Derek<sup>®</sup>, OECD toolbox를 활용한 화학물질 독성 예측 연구

이진욱 · 박선영 · 장석원 · 이상규 · 문상아 · 김현지 · 김필제 · 유승도 · 성창호<sup>†</sup>  
국립환경과학원

### Toxicity Prediction using Three Quantitative Structure-activity Relationship (QSAR) Programs (TOPKAT<sup>®</sup>, Derek<sup>®</sup>, OECD toolbox)

Jin Wuk Lee, Seonyeong Park, Seok-Won Jang, Sanggyu Lee, Sanga Moon,  
Hyunji Kim, Pilje Kim, Seung Do Yu, and Chang Ho Seong<sup>†</sup>

National Institute of Environmental Research

#### ABSTRACT

**Objectives:** Quantitative structure-activity relationship (QSAR) is one of the effective alternatives to animal testing, but its credibility in terms of toxicity prediction has been questionable. Thus, this work aims to evaluate its predictive capacity and find ways of improving its credibility.

**Methods:** Using TOPKAT<sup>®</sup>, OECD toolbox, and Derek<sup>®</sup>, all of which have been applied world-wide in the research, industrial, and regulatory fields, an analysis of prediction credibility markers including accuracy (A), sensitivity (S), specificity (SP), false negative (FN), and false positive (FP) was conducted.

**Results:** The multi-application of QSARs elevated the precision credibility relative to individual applications of QSARs. Moreover, we found that the type of chemical structure affects the credibility of markers significantly.

**Conclusions:** The credibility of individual QSAR is insufficient for both the prediction of chemical toxicity and regulation of hazardous chemicals. Thus, to increase the credibility, multi-QSAR application, and compensation of the prediction deviation by chemical structure are required.

**Key words:** QSAR, K-REACH, TOPKAT<sup>®</sup>, Derek<sup>®</sup>, OECD toolbox, alternative toxicity test

## I. 서 론

다양한 화학물질의 독성영향이 지속적으로 보고되고, 화학물질 노출 사고가 발생함에 따라, 국민의 건강과 환경보호를 위한 정부의 화학물질 규제관리 필요성이 대두되게 되었다.<sup>1,2,3,4</sup> 화학물질 관리의 일환으로 우리나라는 「화학물질 등록 및 평가 등에 관한 법률(이하 화학물질등록평가법)」을 시행하여 화학물질 제조·수입자의 유해성위해성 입증을 위한 법적 책임을 강화하였다.<sup>5</sup> 예로서, 화학물질등록평가법은 국내 유통되는 1 ton 이상 기준화학물질 및 100

kg 이상 신규화학물질을 제조·수입하는 자가 국가에 물리화학적 성질과 유해성에 관한 시험자료를 제출하도록 의무화 하였다.<sup>5</sup> 그러나, 현재 국내 유통되거나 유통된 적이 있는 화학물질의 수는 약 4만 여종에 이를 것으로 예상되는데 불구하고, 국내 시험자료 생산 연구기관의 수 및 국가기관 관리 인력은 법정업무를 수행하기에 부족한 상황이며, 시험자료 생산 비용도 산업체에 큰 부담이 될 것으로 예상된다.<sup>6,7,8</sup> 이에 따라, 국내 유통 화학물질에 대한 산업체와 관계 당국의 업무 효율을 증대시키기 위한 대체시험법 개발이 필요한 실정이다.<sup>7,9</sup>

<sup>†</sup>**Corresponding author:** Environmental Health Research Division, National Institute of Environmental Research, Hwankyung-ro 42, Seo-gu, Incheon, 22689, Republic of Korea, Tel: +82-32-560-7212, E-mail: Chsung@Korea.kr  
Received: 1 July 2019, Revised: 5 August 2019, Accepted: 7 August 2019

Structure-activity relationship (SAR)/Quantitative structure-activity relationship (구조-활성관계: QSAR) 모델은 화학물질의 물리화학적 성질 및 생물학적 활성을 구조에 기반하여 예측하는 수학적 모델의 일종이다. QSAR는 화학물질의 특성 예측 시 기존 동물 실험 및 물리화학 실험에 비해 비용과 소요시간이 적게 들고, 동물희생을 줄일 수 있는 장점이 있어 효과적인 대체시험법으로 여겨지고 있다.<sup>9,10,11,12,13</sup> 그러나, 여러 연구결과들이 QSAR 예측의 신뢰성이 높지 않다고 보고하고 있어 QSAR의 활용에 우려가 제기되고 있다. 예를 들어, 유전독성에 대한 QSAR 모델의 예측 일치율은 73.0% 이하로 확인되었으며, 발암성에 대한 36.8%의 높은 위음성율(실제는 독성이 있으나 독성이 없는 것으로 예측되는 비율)을 나타내었다.<sup>13,14</sup> 또한 국립환경과학원(과학원)에서 수행한 ECOSAR QSAR 프로그램의 수생태 독성예측 신뢰도 분석 연구결과, 모델의 예측 정확도가 32.7~80.0%로 확인되어 QSAR 프로그램의 수생태 독성 예측 신뢰성이 높지 않음이 확인되었다.<sup>15</sup> 따라서, 이미 상용/공용화 된 QSAR 모델을 활용한 화학물질 특성 예측자료의 올바른 활용을 위한 QSAR 예측결과의 신뢰성 향상이 매우 필요하다.

본 연구는 QSAR의 예측 신뢰성 향상을 위해, 먼저 기존 상용/공용화된 QSAR 모델의 급성 경구독성, 피부 과민성, 피부 자극성/부식성, 눈 자극성/부식성에 대한 예측 신뢰성을 검증하고 이어서, 복수 적용 및 구조 별 차이 분석을 통해 QSAR 신뢰성 향상 방안을 모색하고자 한다. 이를 위해 독성예측에 널리 활용되어온 TOPKAT®, Derek®, OECD toolbox를 연구대상 QSAR 프로그램으로 선정하고, 과학원에서 유해화학물질관리법(유해법)에 따라 등록 완료된 화학물질을 대상으로 시험대상 물질을 선정하였다. 시험대상 물질에 관하여 삼중 QSAR 프로그램의 신뢰성을 분석하고 QSAR 프로그램 복수 적용 등의 방법을 통해, QSAR예측 신뢰성 향상 방

안을 모색하였다. 본 연구 결과는 다양한 목적으로 화학물질의 유해성을 검토할 때 QSAR 활용을 위한 과학적 기초자료로서 활용이 예상된다.

## II. 재료 및 방법

### 1. 시험대상 화학물질 및 연구대상 특성 종합점 선정

1991년부터 2014년까지 과학원에 등록 및 심사가 완료된 화학물질을 대상으로 시험대상 물질을 선정하였다. 이 중 화학물질등록평가법에 따라 화학물질등록심사 시 폭넓게 검토되는 등록 완료 시험자료(급성 경구독성, 피부 과민성, 피부 자극성/부식성, 눈 자극성/부식성)를 보유한 화학물질을 1차로 선정하였다. 이 중 급성 경구독성에 관한 시험자료는 100 kg 이상 1 ton 미만 신규화학물질과 1 ton 이상 10 ton 미만인 기존/신규화학물질을 등록할 때 의무적으로 제출해야 하는 자료이다. 그리고 피부 과민성, 피부 자극성/부식성에 관한 시험자료는 제조·수입량이 1 ton 이상 10 ton 미만인 기존 및 신규화학물질을 등록할 때 의무적으로 제출해야 하는 자료이다.<sup>5</sup> 화학물질 등록평가법은 상위톤수를 등록하려는 제조·수입자가 하위 톤수의 시험자료를 함께 제출하게 하고 있으므로 해당 3종 독성에 대한 시험자료는 화학물질을 등록하기 위해 널리 활용되는 시험자료이다. 눈 자극성/부식성 시험자료 또한 10 ton 이상 100 ton 미만 화학물질 등록 시 제출해야 하는 인체유해성 시험자료로, 앞서 말한 시험자료들과 함께 화학물질 등록심사에 폭넓게 검토되고 있는 자료이다.<sup>5</sup>

등록 완료물질에 대해 제출된 급성 경구독성, 피부 과민성, 피부 자극성/부식성, 눈 자극성/부식성 시험자료가 1개 이상 있는 화학물질을 분류하고, 제출된 등록 시험자료가 우수시험실관리기준(GLP)을 준수하는지 여부와 OECD 시험지침과 같은 국제표준 시험법을 준수하였는지 여부를 평가하여 시험자료의 신뢰성을 평가하였다(Table 1). 구조활성관계 분석이

**Table 1.** The quality criteria for credibility of existing toxicological data

Toxicity	Test organisms	Test method	Description of effects
Acute oral toxicity	Rat	OECD test guideline 420/423/425	LD50 (mg/kg bw)
Skin irritation/corrosion	Rabbit	OECD test guideline 404	Positive/negative
Acute eye irritation/corrosion	Rabbit	OECD test guideline 405	Positive/negative
Skin sensitisation	Guineapig	OECD test guideline 406/429	Positive/negative

어려운 복잡다성분물질, 반용생성물, 고분자화합물, 이온성물질, 등은 시험대상 물질 군에서 배제하였다. 신뢰도가 동일한 시험결과에 대해서는 독성이 높은 결과 값을 연구대상 시험자료로 선정하였다.<sup>16)</sup>

선정물질 중 급성 경구독성 자료가 1개 이상 있는 화학물질 406종, 피부 과민성 시험자료가 1개 이상 있는 화학물질 412종, 피부 자극성/부식성 시험자료가 1개 이상 있는 물질 390종, 눈 자극성/부식성 자료가 1개 이상 있는 물질 370종을 각 독성에 대한 QSAR 연구에 활용하였다. 결과로서, 독성 종말점 시험자료가 1개 이상 있고, 시험자료의 신뢰성이 높은 화학물질 1,413종을 시험대상물질로 선정하였다.

## 2. 삼중 QSAR 프로그램 TOPKAT®, Derek®, OECD toolbox 선정

삼중 QSAR 중 Toxicity Prediction by Komputer Assisted Technology (TOPKAT®; Discovery Studio 2017R2, BIOVIA, US)은 machine learning 방식으로 기존 독성자료를 학습하고 화학물질 독성을 예측하는 QSAR 프로그램이다. 학습 데이터베이스에 부재한 자료의 경우 Optimal Predictive Space 법을 통해 신뢰성을 확보한다.<sup>9,17,18)</sup> Deductive Estimate of Risk From Existing Knowledge (Derek®; Derek Nexus v.2.1.0, Lhasa Limited, UK)은 화학물질에 대한 이화학적 성질 및 생물 활성관련 데이터베이스를 기반으로 독성유발 구조인 structural alert과 그 밖의 정보(예, 생물 중 특성 등)를 바탕으로 화학물질 독성을 예측하는 reasoning rule을 종합하여 독성영향을 예측한다.<sup>17)</sup> OECD toolbox (Toolbox, v.4.2, OECD)는 화학물질이 유도하는 독성 발현의 개념적 경로인 adverse outcome pathway에 근거한 Read-across를 통해 화학물질의 독성을 평가하는 사용자 중심 프로그램이다.<sup>17)</sup> 본 연구에서는 범용적으로 활용되어온 이 들 삼중 프로그램을 활용하여 QSAR

프로그램의 신뢰성 예측을 수행하였다. 각 프로그램을 활용한 화학물질 독성 예측은 각 프로그램의 매뉴얼을 따라서 수행하였다. 프로그램의 복수적용은 모든 프로그램에서 예측 결과가 일치된 화학물질을 대상으로 예측결과와 시험결과를 비교하여 복수적용 신뢰성을 분석하였다.

## 3. 삼중 QSAR 프로그램을 활용한 독성 예측 연구

예측 신뢰성은 기존 독성 시험자료와 QSAR 예측 결과의 일치 정도를 나타낸 일치 수준 지표와 기존 시험자료의 결과와 다르게 예측한 오류 수준 지표로 나누어 분석하였다. 일치 수준 지표는 시험자료와 예측 결과의 일치 정도의 백분율인 예측 일치율 (Accuracy: A (%)), 기존 시험자료의 양성반응(독성 있음)을 양성으로 예측한 정도인 민감도(Sensitivity: S (%)), 시험자료의 음성반응(독성 없음)을 음성으로 예측한 정도인 특이도(Specificity: SP (%))가 있다. 시험자료와 예측 결과의 불일치 도를 나타낸 오류 수준 지표로는 시험자료의 양성결과를 음성으로 예측한 정도를 나타내는 잘못된 음성(False negative: FN (%)), 시험자료의 음성결과를 양성으로 예측한 정도를 나타내는 잘못된 양성(False positive: FP (%))이 있다. A, S, SP, FP 및 FN의 화학물질 규제관리를 위한 수용 가능한 수준(Acceptable level) 판단기준은 국립환경과학원 QSAR 예측신뢰도 향상 방안 연구에서 활용하는 자체 평가 기준을 활용하였다.<sup>9)</sup>

피부 과민성은 3종 QSAR 모델을 활용하여 분석하였고 각 모델의 예측결과는 QSAR 모델간의 예측력 차이를 비교하는데 활용되었다. 급성 경구독성, 피부 자극성/부식성, 눈 자극성/부식성의 경우 TOPKAT®만을 활용하여 독성 예측 신뢰도를 분석하였다. 급성 경구독성에 관한 분석 시 예측 신뢰성 분석 후 Globally Harmonized System (GHS)에 따른 분류를 수행하고, 기존 시험자료 대비 화학물질 분류의 정확도를 검증하였다(Table 2). QSAR 예측

**Table 2.** Chemical classification based on GHS category for acute oral toxicity predicted by TOPKAT®

Categories	The range of LD <sub>50</sub> (mg/kg)	Hazard statement	Signal word
1-2	LD <sub>50</sub> ≤ 50	Fatal if swallowed	Danger
3	50 < LD <sub>50</sub> ≤ 300	Toxic if swallowed	Danger
4	300 < LD <sub>50</sub> ≤ 2,000	Harmful if swallowed	Warning
Non-Classified	LD <sub>50</sub> > 2,000	-	Warning

**Table 3.** The analysis results for skin sensitization using TOPKAT®, Derek®, OECD Toolbox

TOPKAT®			Derek®			OECD Toolbox		
A <sup>1)</sup>	FP <sup>2)</sup>	FN <sup>3)</sup>	A	FP	FN	A	FP	FN
(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
42.2	47.3	10.4	60.0	25.7	12.6	59.2	60.0	10.7

1) Accuracy, 2) false positive, 3) false negative

신뢰성 향상 방안 모색을 위하여 피부 과민성 시험 자료가 있는 화학물질에 대하여 QSAR 프로그램 단독 분석을 거친 후 복수의 QSAR의 분석을 수행하여 A, S, SP, FP, FN 값을 비교하였다. 또한, QSAR 예측 정확도의 화학물질 구조 별 차이를 연구하기 위하여, 2017년 국립환경과학원 연구사업에서 활용한 63종 구조카테고리를 기반으로 화학물질을 분류하고 각 구조 별 TOPKAT® 프로그램의 예측 신뢰성을 분석하였다.<sup>9,17)</sup>

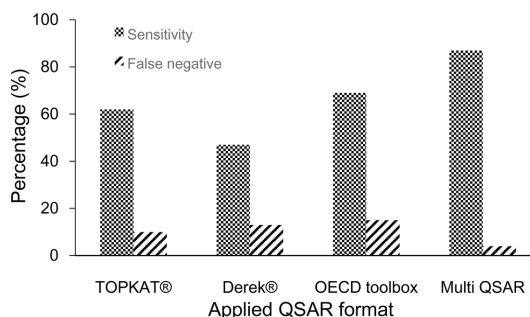
### III. 연구결과

#### 1. QSAR 모델(TOPKAT®, Derek®, OECD toolbox)의 피부 과민성(Skin sensitisation) 예측

삼중 QSAR 모델을 활용한 화학물질 독성 예측결과, TOPKAT®의 등록 시험자료 대비 A 수치는 42.2%로 나타났고, FP와 FN은 각각 47.3%과 10.4%로 확인되었다. Derek®의 경우 시험자료 대비 A 수치가 60.0%이었고, FP와 FN은 각각 25.7%과 12.6%로 확인되었다. OECD toolbox의 경우 A는 59.2%, FP는 60.0%, FN은 10.7%로 확인되었다(Table 3). TOPKAT®, Derek®, OECD toolbox의 S값은 각각 61.0, 47.0, 69.0%로 분석되었고, SP 값은 각각 35.0, 65.0, 51.0%로 분석되었다(Fig. 1).

#### 2. QSAR 모델(TOPKAT®)의 급성 경구독성(Acute oral toxicity) 예측

시험대상 화학물질 중 급성 경구독성이 있는 406종의 화학물질 시험자료에 대한 TOPKAT® 예측 신뢰성을 분석하였다. GHS 분류기준에 따라 구체적으로 급성 경구독성 예측의 GHS 분류 카테고리 별 예측 정확도를 확인하였다(Table 5). 연구결과 기존 시험자료 대비 예측 일치율 A는 47.1%로 나타났다. 화학물질 규제 목적으로 중요하게 활용되는 FN의 경우 8.6%의 값을 보여 피부 과민성과 비교하여 낮



**Fig. 1.** Comparison of QSAR prediction results based on application format. Multi QSAR means simultaneous application of TOPKAT®, Derek®, and OECD toolbox.

은 오차율을 나타내었다. 기타 잘못된 양성은 44.3%으로 확인되었다. 시험자료 상 GHS 분류군 1·2로 분류된 물질에 대한 TOPKAT® 비교 결과, A는 33.3%, FN은 66.7%로 확인되었다. 기존 시험자료에 따라 GHS 분류 군 3으로 분류된 물질에 대한 FP, A, FN 예측 결과는 각각 12.5, 18.7, 68.8%이었다. 시험자료 상 GHS 분류 군 4로 분류된 34종 물질에 대한 TOPKAT® 예측 결과, FP, A, FN은 각각 12.5, 56.2, 31.3%이었다. 시험 자료 상 유해성 분류 없음으로 분류된 물질 320종에 대한 분석결과 FP, A은 각각 53.1, 46.9%로 나타났다(Table 4).

#### 3. QSAR 모델(TOPKAT®)의 피부 자극성/부식성(Skin irritation/corrosion)·눈 자극성/부식성(Acute eye irritation/corrosion) 예측

피부 자극성/부식성 시험결과가 있는 390종 화학물질에 대한 TOPKAT® 예측 신뢰성 검토 결과, A, FP, FN, S, SP는 각각 57.2, 40.3, 2.6, 87.2, 49.7%로 확인되었다. 그리고, 370여 종의 눈 자극성/부식성 시험결과가 있는 화학물질에 대한 TOPKAT® 예측 신뢰도 분석결과, A, FP, FN, S, SP는 각각 50.5, 41.6, 7.8, 75.0, 39.4%로 확인되었다(Table 5).

**Table 4.** Globally Harmonized System (GHS) classification based on acute oral toxicity predicted by TOPKAT®

Experimental Categories	Number of chemicals	Prediction			N.C. <sup>4)</sup>	A <sup>1)</sup> (%)	FP <sup>2)</sup> (%)	FN <sup>3)</sup> (%)
		Category 1-2	Category 3	Category 4				
1-2	6	2	1	3	0	33.3		66.7
3	16	2	3	9	2	18.7	12.5	68.8
4	64	1	7	36	20	56.2	12.5	31.3
N.C. <sup>4)</sup>	320	1	31	138	150	46.9	53.1	
Sum	406	6	42	186	172	47.1	44.3	8.6

1) Accuracy, 2) false positive, 3) false negative, 4) non-classified

**Table 5.** TOPKAT® prediction results for the irritation/corrosion of skin/eyes

Toxicity	A <sup>1)</sup> (%)	S <sup>2)</sup> (%)	SP <sup>3)</sup> (%)	FP <sup>4)</sup> (%)	FN <sup>5)</sup> (%)
Skin irritation/corrosion	57.2	87.2	49.7	40.3	2.6
Eye irritation/corrosion	50.5	75	39.4	41.6	7.8

1) Accuracy, 2) Sensitivity, 3) specificity, 4) false positive, 5) false negative

**Table 6.** TOPKAT® prediction results for skin irritation/corrosion according to chemical structure

Chemical groups	Number of chemicals	S <sup>1)</sup> (%)	SP <sup>2)</sup> (%)	FP <sup>3)</sup> (%)	FN <sup>4)</sup> (%)
Natural organics	115	91.3	51.6	47.3	8.7
Esters	85	100	33.8	65	0
Aliphatic amines	31	90	31.3	68.8	10.0
Phenols	19	50	76.9	23.1	50
Amides	19	100	77.8	22.2	0

1) Sensitivity, 2) specificity, 3) false positive, 4) false negative

#### 4. QSAR 예측 신뢰도 향상 방안 연구

본 연구는 기존 화학물질 독성 및 물리화학적 성질 예측에 활용되어온 QSAR 프로그램들의 예측 신뢰성 향상 방안 모색의 일환으로 QSAR 프로그램 복수 적용 연구를 수행하였다. 피부 과민성 시험자료가 있는 412종 화학물질에 대한 3종 QSAR 프로그램의 복수적용 결과 S는 87.0%로 개별 적용에 비해 신뢰도가 증가하였다. 화학물질 규제목적의 지표로 주요하게 활용되는 FN의 경우 4%로 감소하여 개별 QSAR 예측 값에 비해 예측 신뢰도가 증가함을 확인하였다(Fig. 1).

**Table 7.** TOPKAT® prediction results for acute eye irritation/corrosion according to chemical structure

Categories	Number of chemicals	S <sup>1)</sup> (%)	SP <sup>2)</sup> (%)	FP <sup>3)</sup> (%)	FN <sup>4)</sup> (%)
Natural organics	115	78.4	31.6	68.4	21.6
Esters	85	50.0	54.2	45.8	50
Aliphatic amines	31	81.8	16.7	83.3	18.2
Phenols	19	100	64.3	35.7	0
Amides	19	80	15.4	84.6	20

1) Sensitivity, 2) specificity, 3) false positive, 4) false negative

**Table 8.** TOPKAT® prediction results for skin sensitisation according to chemical structure

Categories	Number of chemicals	S <sup>1)</sup> (%)	SP <sup>2)</sup> (%)	FP <sup>3)</sup> (%)	FN <sup>4)</sup> (%)
Natural organics	115	68.4	25	75	31.6
Esters	85	47.8	53.2	46.8	52.2
Aliphatic amines	31	28.6	29.4	70.6	71.4
Phenols	19	0	73.3	26.7	100
Amides	19	0	52.6	47.4	100

1) Sensitivity, 2) specificity, 3) false positive, 4) false negative

또한 예측 신뢰도에 화학물질의 구조가 미치는 영향을 파악하고 이에 대한 예측 신뢰도 향상 방안을 위해 63개 화학물질 구조 카테고리 별 예측 신뢰성 변화 양상을 분석하였다(Table 6, 7, 8). TOPKAT®의 구조 별 민감도는 피부 자극성/부식성, 눈 자극성/부식성, 피부 과민성에 관하여 각각 평균 86.26% (최소: 50.0%~최대: 100.0%), 78.04% (50.0~100.0%), 28.96% (0~68.4%)로 확인되었다 (Table 6, 7, 8). 피부 자극성/부식성, 눈 자극성/부식

성, 피부 과민성에 관하여 TOPKAT®의 구조 별 특이도는 각각 평균 54.28% (31.3~77.8%), 36.4% (15.4~64.3%), 46.7% (25.0~73.3%)로 나타났다(Table 6, 7, 8). 피부 자극성/부식성, 눈 자극성/부식성, 피부 과민성에 관하여 TOPKAT®의 구조 별 FP는 각각 평균 45.3% (22.2~68.8%), 63.6% (35.7~84.6%), 53.5% (26.7~75.0%)로 나타났다. 피부 자극성/부식성, 눈 자극성/부식성, 피부 과민성에 관하여 TOPKAT®의 구조 별 FN은 각각 평균 13.74% (0~50.0%), 21.96% (0~50.0%), 51.1% (0~100.0%)로 나타났다(Table 6, 7, 8).

#### IV. 고 찰

본 연구는 화학물질 독성 예측에 범용적으로 활용되어온 TOPKAT®, Derek®, OECD toolbox를 연구 대상 QSAR 프로그램으로 선정하고 QSAR의 예측 신뢰도 분석과 신뢰도 향상 방안을 연구하였다. 개별 모델을 활용하여 시험대상 물질의 급성독성을 예측한 후 해당 화학물질 등록 시 제출된 시험자료의 값을 기반으로 독성 예측 결과의 신뢰도를 분석하였다. 피부 과민성에 대한 삼 중 QSAR 프로그램 분석결과 개별 프로그램의 예측 일치 율(A)은 약 60% 이하의 값을 보였고 Derek®>OECD toolbox>TOPKAT® 순으로 일치 율이 높았다. 민감도의 경우 OECD toolbox>TOPKAT®>Derek®의 순서로 수치가 높았고, 특이도의 경우는 Derek®>OECD toolbox>TOPKAT®의 순서로 수치가 높았다. 오류 발생 수준 분석결과 FP는 Derek®<TOPKAT®<OECD toolbox 순서로 수치가 낮았고, FN은 TOPKAT®<OECD toolbox<Derek®순으로 수치가 낮았다. 종합해볼 때 삼중 프로그램 모두 쉰 지표에서 고르게 높은 예측 신뢰성을 보이는 프로그램은 없었다. 양성을 정확하게 예측하는 능력과 시험자료 대비 일치 율에서 Derek®의 신뢰성이 높은 반면 음성에 대한 예측 정확성은 OECD toolbox가 높았다. 오류 율을 분석한 결과 FN은 삼중 모두 유사하였으나, FP의 경우 Derek®의 음성자료를 양성자료로 잘 못 예측할 확률이 크게는 2배 정도 낮게 관찰되어, Derek®의 피부 과민성에 대한 예측 신뢰성이 다른 프로그램보다 다소 높은 것을 확인하였다. TOPKAT®의 경우 상대적으로 신뢰도 순위가 낮았으나, 수치의 차

이가 다른 프로그램과 크지 않아 신뢰성이 낮다고 단정할 수는 없었다. 상기한 결과는 QSAR 결과 값을 활용할 때 독성 종말점의 종류 및 분석지표를 고려해야 함을 의미한다. 또한, 본 연구에서의 프로그램 민감도 비교는 피부 과민성에 국한되어 있으므로 다른 독성 종말점에 대한 프로그램간 예측 신뢰성 차이를 추가 분석하여 각 프로그램의 차이를 보다 명확하게 규명할 필요성이 있다고 판단된다.

피부 자극성/부식성에 관하여 Natural organics, Aliphatic amines, Esters, Amides류 화학물질 및 눈 자극성/부식성 관련 phenol류 화학물질에 대하여 TOPKAT®의 예측결과 민감도는 90.0~100.0%로 높게 나타났다. 반면 피부 과민성에 관하여 TOPKAT®의 민감도는 68.0% 이하로 다른 독성에 비해 낮은 예측 정확도를 나타내었다. TOPKAT®은 피부 자극성/부식성과 눈 자극성/부식성에 관하여 평균 약 80%의 민감도를 나타낸 반면 피부 과민성에 대해서는 약 30% 미만의 민감도를 나타내어 독성에 따른 민감도가 동일 model상에서도 큰 차이가 있음이 확인되었다. TOPKAT®의 특이도는 피부 자극성/부식성 >피부 과민성 >눈 자극성/부식성 순서로 높은 수치가 확인되었는데, 민감도에 비해 독성 별 예측 값의 차이가 크지 않았다. FP는 눈 자극성/부식성 >피부 과민성 >피부 자극성/부식성의 순서로 차이가 나타났으나 특이도와 같이 독성 별 예측 차이가 크지 않았다. 반면 FN의 경우 피부 과민성의 경우 4에서 3배 정도 평균 수치의 차이가 관찰되었다. 특히 피부 과민성에서 phenol 및 amides 류 화학물질에 대하여 TOPKAT®은 양성 시험결과를 음성으로 100.0% 잘못 판단하는 예측을 하였다. 이 같은 결과는 QSAR가 일부 화학 구조에 대해 독성을 과소평가 할 수 있음을 의미한다. 따라서 phenol과 amides 류 화학물질에 대한 TOPKAT®의 규제 목적활용은 어렵다는 것을 확인하였다. 결과로서 동일한 프로그램을 활용하더라도 화학물질의 구조적 차이에 따라 독성 예측 일치율이 0~100.0%까지 차이가 날 수 있다는 것을 확인하였다. 따라서 TOPKAT®의 예측 값을 해석할 때 구조 별 신뢰도 차이를 반영한 예측 값 보정이 필요하며 이와 관련한 추가연구가 필요하다고 판단된다.

이번 연구는 다양한 분석조건을 설정하고 각 조건에 따른 기존 상용/공영화된 QSAR 프로그램의 예

측 신뢰성을 분석하였다. 분석 결과들은 QSAR의 예측 값은 개별 적용의 경우 신뢰성이 낮으며, 화학물질 구조, 프로그램 종류, 복수 조합 여부 등에 따라 크게 달라질 수 있음을 나타내었다. 앞서 말한 바와 같이 각 변수에 대한 심화연구를 통해 신뢰성에 영향을 주는 요인들을 적절히 통제하지 못 한다면 대체시험 법으로서 QSAR의 단독 활용은 한계가 있다고 판단된다. 실제로 OECD는 대체시험법의 신뢰성을 향상시키기 위해 *in vivo*, *in vitro* 연구결과를 물론 QSAR, read-across, 독성반응경로(AOP: adverse outcome pathway)의 종합적인 적용을 수행하는 Integrated Approaches to Testing and Assessment (IATA)을 연구하고 있다<sup>8)</sup>. 따라서, 대체시험법 활성화를 위해 QSAR의 단독 문제점 해결 방안연구는 물론, read-across, weight of evidence, AOP 등의 정보를 통합적으로 예측에 반영하는 연구의 수행이 병행 되어야 된다고 판단된다.

## V. 결 론

TOPKAT®, OECD toolbox, Derek®를 활용하여 화학물질 독성을 예측한 결과 프로그램의 개별적용 예측 신뢰성 보다 복수 적용 예측 신뢰성이 높았다. 따라서 QSAR 예측 신뢰성 향상을 위해 프로그램 복수적용이 필요함을 확인하였다. QSAR의 예측 정확도는 대상 화학물질의 구조에 따라 크게 상이함으로써, QSAR 예측 결과에 대한 화학물질 구조 별 해석이 필요함을 확인하였다. 끝으로, QSAR의 독성 예측결과의 신뢰도를 상승시키기 위해선, Read-across 등 다양한 비동물시험법의 종합적인 해석이 필요함을 확인하였다.

## 감사의 글

본 논문은 환경부의 재원으로 국립환경과학원의 지원을 받아 수행하였습니다(NIER-2018-01-01-011).

## References

1. Ganesan S, Vasudevan N. Impacts of perfluorinated compounds on human health. *Bull. Env. Pharmacol. Life Sci.* 2015; 4: 183-191.

2. Lee JW, Lee J-W, Kim K, Shin YJ, Kim J, Kim H, et al. n-Butyl acrylate-induced antioxidant system alteration through two generations in *Oryzias latipes*. 2018; doi: 10.1007/s10695-018-0584-z.
3. Lee JW, Shin YJ, Kim H, Kim H, Kim J, Min SA, et al. Metformin-induced endocrine disruption and oxidative stress of *Oryzias latipes* on two-generational condition. *J Hazard Mater.* 2019; 5:171-181.
4. Korean Society of Environmental Health (KSEH), Environmental Health, 1<sup>st</sup> ed, Seoul, Episteme Press, 2016, p. 19-33.
5. Ministry of Environment (MOE). Act on the Registration and Evaluation, etc. of Chemical Substance. Available: <http://www.law.go.kr/??/>[accessed 01 June 2019]
6. Kim K-Y, Shin SE, No KT. Assessment of quantitative structure-activity relationship of toxicity prediction models for Korean Chemical substance control legislation. *Environ Heal Toxicol.* 2015; 30: Article ID:s2015007, 10 page
7. Korea Development Institute (KDI). Analysis report: the effect of domestic industry according to EU-REACH legislation. Available: [http://www.prism.go.kr/homepage/researchCommon/retrieveResearchDetailPopup.do?sessionId=53B47AF8E1BB6B007845C668E9DAF279.node02?research\\_id=1480000-200700151\(Korean\)](http://www.prism.go.kr/homepage/researchCommon/retrieveResearchDetailPopup.do?sessionId=53B47AF8E1BB6B007845C668E9DAF279.node02?research_id=1480000-200700151(Korean)) [cited 2014 Dec 15].
8. Kim J, Seo, J-K, Kim T, Kim, H-K, Park, S, Kim P-J. Prediction of human health and ecotoxicity of chemical substances using the OECD QSAR application toolbox. *J Environ Health Sci.* 2013; 39: 130-137.
9. Sung CH, Park SY, Lee JW, Yu SD, Kim PJ, Jang S-W, et al. A study for the improvement of a prediction using QSAR models to find out hazardous substances (I) (2018). Available: <http://library.me.go.kr/search/DetailView.ax?sid=1&cid=5677755> [accessed 01 June 2019]
10. Jang T-S, Shin S-K, Kim H. Estimation of Physicochemical and Eco-toxicological properties of pharmaceuticals by QSAR approach. *J Korean Society for Environmental Analysis.* 2010; 13: 45-53
11. Mays C, Benfenati E, Pardoe S. Use and perceived benefits and barriers of QSAR models for REACH: findings from a questionnaire to stakeholders. *Chem Cent J.* 2012; 6(1): 159
12. Sung ND. Development of new agrochemicals by quantitative structure-activity relationship (QSAR) methodology, 3D QSAR methodologies and computer-assisted molecular design (CAMD). *Korean J Pestic Sci.* 2003; 03: 1-11.

13. Cariello NF, Wilson JD, Britt BH, Wedd DJ, Burlinson B, Gombar V. Comparison of the computer programs DEREK and TOPKAT to predict bacterial mutagenicity. *Mutagenesis*. 2002; 17(4):321-329.
14. Kim J, Choi K, Kim K, Kim D. QSAR approach for toxicity prediction of chemicals used in electronics industries. *J Environ Health Sci*. 2014; 40:105-113.
15. Kim HK, Kim JY, Cha HK, Park MY, Sung CH, Kim PL. A study on the best application of (Q)SARs to predict aquatic toxicity of organic chemicals (2010). Available: <http://library.me.go.kr/search/DetailView.ax?sid=1&cid=5256608> [accessed 01 June 2019]
16. Klimisch HJ, Andreae M, Tillmann U. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regul Toxicol Pharmacol*. 1997; 25: 1-5.
17. Sung CH, Park SY, Kim KT, Kim KH, Sung MH, Moon SA, et al. Study on the improvement of mutagenicity prediction using QSARs models (2016). Available: <http://library.me.go.kr/search/DetailView.ax?sid=1&cid=5638480> [accessed 01 June 2019]
18. Sung CH, Park SY, Choi YS, Kim HK, Sung MH, Moon SA, et al. Study on the improvement of genotoxicity prediction using QSARs models (2017). Available: <http://library.me.go.kr/search/DetailView.ax?sid=1&cid=5667023> [accessed 01 June 2019]

#### <저자정보>

이진욱(환경연구사), 박선영(공업연구사),  
장석원(전문위원), 이상규(전문위원),  
문상아(전문위원), 김현지(전문위원),  
김필제(과장), 유승도(부장), 성창호(공업연구관)