

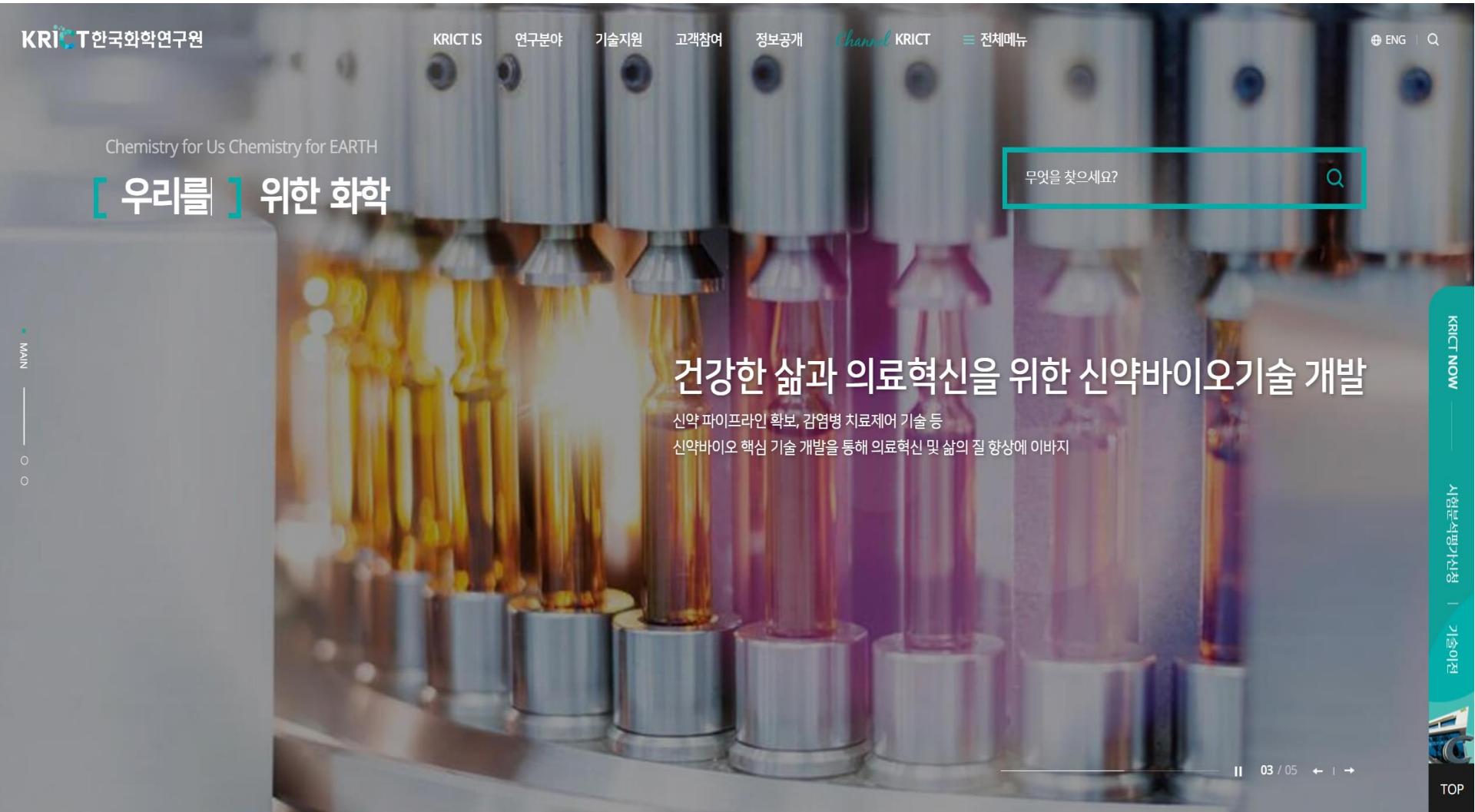
K-MELLODDY 사업
(분야2) 데이터 활용 신약개발 협력 체계구축
**K-MELLODDY 플랫폼 구축을 위한
데이터 지원**

2024. 08. 20.

한국화학연구원 오 광 석

K-MELLODDY 사업
연합학습기반
신약개발 가속화 프로젝트

01-1 한국화학연구원 (의약바이오연구본부)



01-2 한국화학연구원 (의약바이오연구본부 신약정보기술연구센터)

주요 연구분야



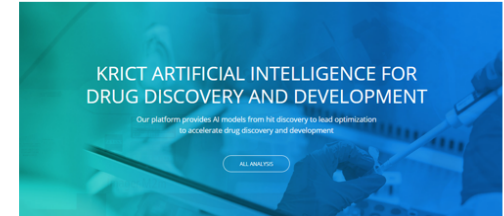
▶ 다중 오믹스 빅데이터 융합 기반 신약 파이프라인 연구

- 환자 유전체 데이터 기반 난치성 질환 신규 타겟 발굴
- 약물전사체 데이터 기반 치료용 화합물 탐색
- 차세대시퀀싱 데이터 기반 난치성 질환 후보물질의 작용기전 연구
- 정밀의료 혁신 신약 개발을 위한 동반진단 바이오마커 개발



▶ 인공지능 기반의 활성, 독성, 약물성 예측 알고리즘 개발







- 화합물 구조 기반의 독성, 약물성 예측 딥뉴럴네트워크 알고리즘 개발
- 실험적 검증을 위한 시험법 구축 및 이를 활용한 알고리즘 최적화 연구
- 개발중단 약물 대상 독성, 약물성 개선을 위한 drug rescue 플랫폼 구축



 Cardiotoxicity Prediction model of cardiotoxicity based on ensemble models VIEW MORE	 BBB Permeability Prediction model of blood-brain-barrier penetration based on Message Passing Neural Network VIEW MORE	 Cytochrome P450 Prediction model of Cytochrome P450 inhibition based on Message Passing Neural Network VIEW MORE
 Cardiotoxicity v2.0 Prediction model of cardiotoxicity based on Message Passing Neural Network VIEW MORE	 Metabolic stability Prediction model of metabolic stability based on Message Passing Neural Network VIEW MORE	 Derivative generation In silico chemical transformation module based on known chemical reaction patterns VIEW MORE
 Hepatotoxicity Prediction model of hepatotoxicity based on Message Passing Neural Network VIEW MORE	 Reproductive toxicity Prediction model of reproductive toxicity based on Message Passing Neural Network VIEW MORE	 LogBB Prediction model of blood-brain-barrier penetration based on Message Passing Neural Network VIEW MORE
 PredAOT Prediction model of acute oral toxicity based on random forest VIEW MORE	 PredPS Prediction model of plasma stability based on Message Passing Neural Network VIEW MORE	

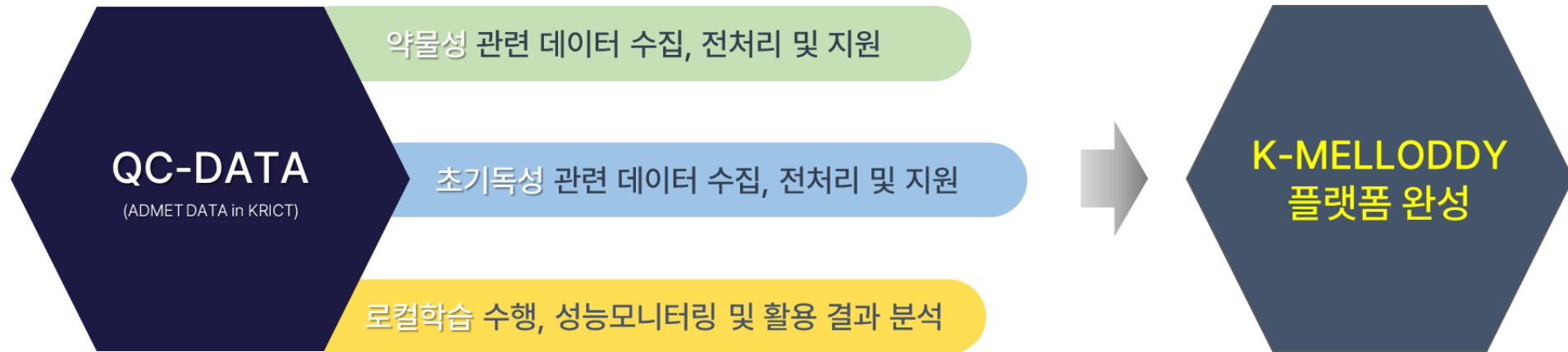
01-3 과제팀 참여연구원

· 데이터 전처리 전문가, 약리독성, 약동학 및 신약개발 전문가

 <p>오광석 Ph.D. 과제책임</p>	 <p>장우대 Ph.D. 참여</p>	 <p>장지돈 Ph.D. 참여</p>	 <p>안선주 Ph.D. 참여</p>	 <p>송진숙 Ph.D. 참여</p>	 <p>이병호 Ph.D. 참여</p>
<p>현 한국화학연구원 책임 연구원 전 정보융합신약연구센터 센터장</p>	<p>현 한국화학연구원 선임연구원 전 한국과학기술원 연수연구원</p>	<p>현 한국화학연구원 선임연구원 전</p>	<p>현 신약정보융합기술연구센터장 현 한국화학연구원 책임연구원</p>	<p>현 한국화학연구원 책임연구원 전</p>	<p>현 한국화학연구원 책임연구원 전 약리활성연구센터장</p>
<ul style="list-style-type: none"> 고려대학교 생명공학 박사 - 생리활성물질 효능 및 기전연구 2020 보건복지부 장관표창 2021 한국화학연구원장상 공로상 2021 UST 우수현장 연구상 2017년 양도 1억 유나이티드제약 기술이전 2019년 선급2억, 실적 21억 (㈜휴온스 기술이전) 2020년 선급3억, 총매출액 1% 카이노스메드 기술이전 생화학적 약효검색법 구축 및 개발, in vitro 약동력학평가, 인공지능기반 ADMET 예측모델관련 논문 72편, 특허 87건 	<ul style="list-style-type: none"> 한국과학기술원 시스템 생물학 박사 - 약물 가상 스크리닝 및 단백질 설계 연구 2023년 인공지능기반 혈장안정성 예측모델 개발 약물 가상 스크리닝, 인공지능 기반 신약개발, 시스템 생물학 분야 SCI급 논문 18 편(주저자 11편 포함), 특허 4건 	<ul style="list-style-type: none"> 한국과학기술원 화학정보학 박사 - 딥러닝 기반 소재 물성, 합성 가능성 예측모델 개발 - 분자동역학 시뮬레이션 연구 2022년 한국과학기술원 생명화학공학과 임천논문상 수상 화학합성, 약물 ADMET 예측 모델 개발 및 인공지능 플랫폼 고도화 JACS, Matter 등 SCI 저널에 인공지능 응용 논문 출판 	<ul style="list-style-type: none"> 오하이오주립대학교 약학박사 - 신규 항암제의 효능, 기전연구 - 신규 항암제 약동학 및 독성 연구 신약후보물질 발굴을 위한 약동학 연구 2023 과학기술정보통신부 장관표창 2023 한국화학연구원장상 공로상 FDA PhIII 유방암치료제 주개발자 10건의 저분자 후보물질을 국내 제약사에 기술이전 In vitro/In vivo 약동력학평가 및 약물의 기전연구 관련 논문 59편, 특허 출원 및 등록 41건 	<ul style="list-style-type: none"> 충남대학교 생물약학 박사 - 항당뇨제의 PK/PD 연구 신약후보물질 발굴을 위한 약동학 연구 2023년 파킨슨병치료제 기술이전 (선급3.5억원, 실적 107억원) 2018-2021년 폐암치료제 기술이전 3건 (선급 합 21억원) 2020년 파킨슨병 치료제 기술이전 (선급 3억원) 2012년 당뇨/비만 치료제 기술이전 (선급 3억원) 약물의 약동학 연구 관련 논문 57편, 특허 30건 	<ul style="list-style-type: none"> 중앙대학교 약리학 박사 - 신성 고혈압제의 약리/생리학적 연구 2021 산업통상자원부 장관표창 2022 한국화학연구원 공로상 2015년 선급2억, 실적16억, 총매출액 4% 카이노스메드 기술이전 2019년 선급2억, 실적 21억 (㈜휴온스 기술이전) 2020년 선급3억, 총매출액 1% 카이노스메드 기술이전 약리학적 약효검색법 구축 및 개발, in vivo 약효평가, 약물의 기전연구 관련 논문 142편, 특허 143건
<p>신약개발 전문가 약동학, 약리기전연구 전문가</p>	<p>시스템 생물학전문가 인공지능 모델구축 전문가</p>	<p>화학정보학전문가 인공지능 모델구축 전문가</p>	<p>약물동태학 전문가 플랫폼 구축 전문가</p>	<p>약물동태학 전문가 독성 전문가</p>	<p>신약개발 전문가 약동학, 약리학연구 전문가</p>

02 과제 목표

QC-DATA : 데이터 품질관리 시스템을 통한 **ADMET** 관련 데이터

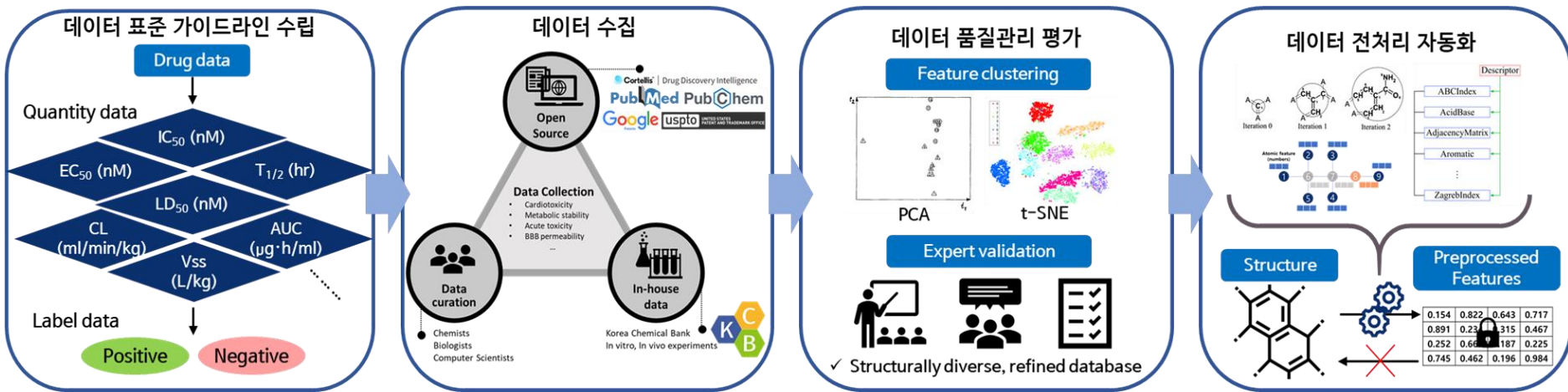


FAM에 효율적인 ADMET관련 데이터 지원을 통한 연합학습 구현

03 세부연구내용

1. 데이터 품질관리 시스템

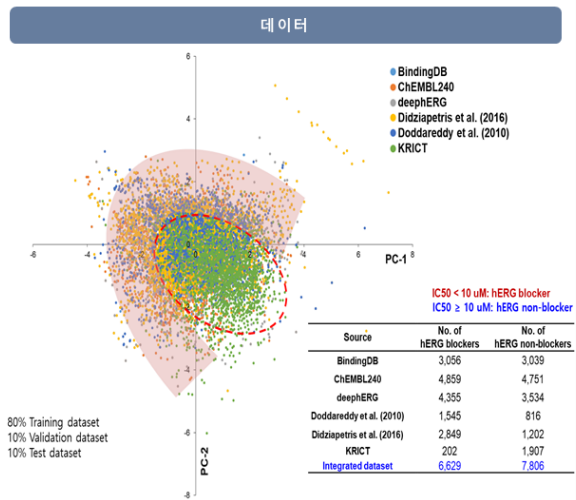
데이터 품질관리 시스템



04-1 화합물은행 대표화합물 대상 실측데이터 포함 데이터 분석 현황

심장독성 데이터

- 한국화합물은행의 대표화합물 대상 **2,130종**의 실측데이터.
- 텍스트 마이닝을 통한 문헌, 특허데이터 **12,331종** 수집.



> Bioinformatics. 2020 May 1;36(10):3049-3055. doi: 10.1093/bioinformatics/btaa075.

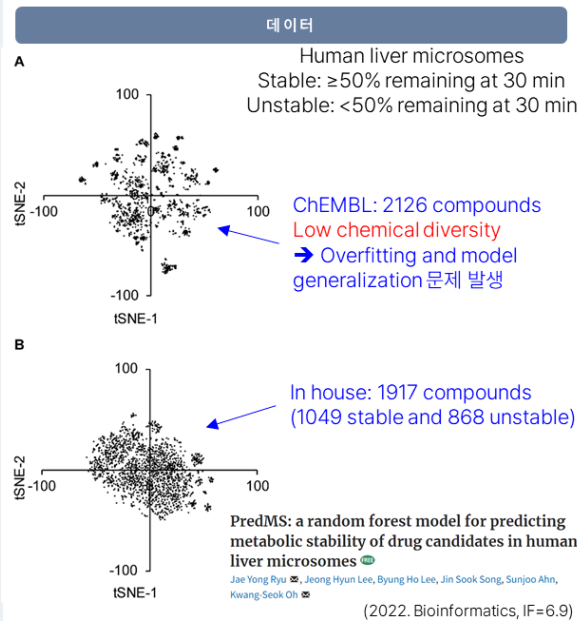
DeepHIT: a deep learning framework for prediction of hERG-induced cardiotoxicity

Jae Yong Ryu¹, Mi Young Lee¹, Jeong Hyun Lee¹, Byung Ho Lee¹, Kwang-Seok Oh^{1,2}

(2020. Bioinformatics, IF=6.9)

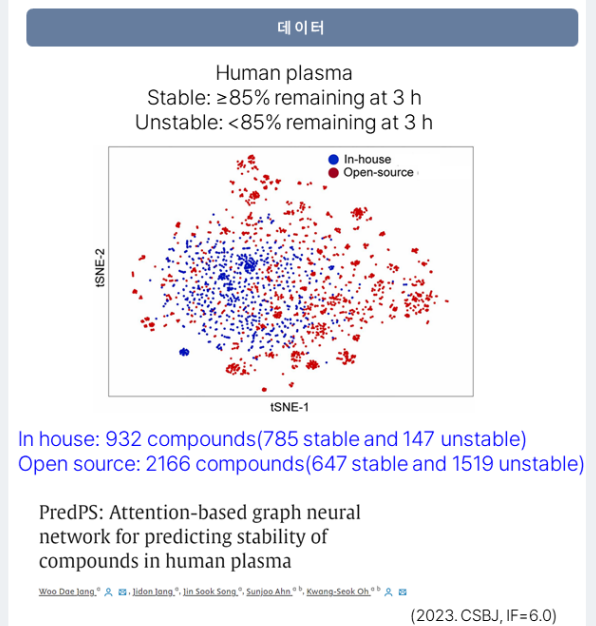
인간 대사 안정성 데이터

- 화합물은행 대표화합물 대상 **1,917종**의 실측데이터.
- ChEMBL 데이터 **2,126종** 외부데이터 수집.



인간 혈장 안정성 데이터

- 화합물은행 대표화합물 대상 **932종** 실측데이터.
- 텍스트 마이닝을 통한 문헌, 특허데이터 **2,166종** 수집.



04-2 데이터 보유현황

Status of Data used to build model

Module	Model type	Data size	Data type	In house	Literature Data
Cardiotoxicity DeepHIT (Bioinformatics, 2020)	Classification	14,440	IC ₅₀ ($< 10 \mu\text{M}$)	202 hERG blockers / 1907 hERG non-blockers	6430 hERG blockers / 5901 hERG non-blockers
Cardiotoxicity hERGpred (BMC Bioinformatics, 2019)	Classification	2,130	IC ₅₀ ($< 10 \mu\text{M}$)	221 hERG blockers / 1909 hERG non-blockers	-
BBB permeability (Bioinformatics, 2021)	Classification	7,162	Binary (0 or 1)	-	5453 BBB permeable / 1709 BBB non-permeable
LogBB_pred (Bioinformatics, 2023)	Regression	1,164	LogBB (LogBB > -1)	-	532 BBB permeable / 468 BBB non-permeable
Metabolic stability (Bioinformatics, 2022)	Classification	1,917	% Remaining ($\geq 50\%$ at 30 min)	1049 stable / 868 unstable	-
Hepatotoxicity (Current Bioinformatics, 2022)	Classification	2,170	Binary (0 or 1)	-	1146 toxic / 1024 non-toxic
Reproductive toxicity (BMC Bioinformatics, 2020)	Classification	7,468	Binary (0 or 1)	-	270 active / 7198 inactive
Acute oral toxicity (BMC Bioinformatics, 2023)	Classification & Regression	Mice: 6,226 Rats: 6,238	LD ₅₀ (5 category; $< 5 \text{ mpk}$, 5-50 mpk, 50-300 mpk, 300-2000 mpk, 2000-5000 mpk)	-	Mice (67, 282, 1183, 3413, 1281) Rats (173, 490, 1103, 2560, 1912)
Plasma stability (Comput. Struct. Biotechnol. J., 2023)	Classification	3,098	Binary (0 or 1) ($>85\%$ at 3 h)	785 stable / 147 unstable	647 stable / 1519 unstable
CYP1A2 CYP2C9 CYP2C19 CYPD6 CYP3A4	Classification	9,642 9,744 9,457 10,910 9,003	Binary (0 or 1)	-	5,992 active / 3,650 inactive 6,035 active / 3,709 inactive 7,167 active / 2,308 inactive 9,757 active / 1,153 inactive 6,282 active / 2,721 inactive

04-3 데이터 보유현황

Status of Data about in vivo PK

Parameter	Routes of drug administration	Species	Data size
Bioavailability (F%)	PO	Mouse	2074
		Rat	6882
AUC(ug.h/mL)	PO	Mouse	2562
		Rat	5360
AUC(ug.h/mL)	IV	Mouse	939
		Rat	2150
C _{max} (ug/mL)	PO	Mouse	2211
		Rat	4433
T _{max} (h)	PO	Mouse	1173
		Rat	2323
T _{1/2} (h)	PO	Mouse	1262
		Rat	2424
T _{1/2} (h)	IV	Mouse	1657
		Rat	5866
CL(mL/min/kg)	IV	Mouse	1906
		Rat	6930
V _{ss} (L/kg)	IV	Mouse	1607
		Rat	5043