

약물의 염전성 부정맥 유발 예측 지표로서 심장의 전기생리학적 특징 값들의 검증

유예담^{1†} · 정다운¹ · Aroli Marcellinus¹ · 임기무^{1,2*}

¹금오공과대학교 IT융복합공학과, ²금오공과대학교 메디컬IT융합공학과

Verification of Cardiac Electrophysiological Features as a Predictive Indicator of Drug-Induced Torsades de pointes

Yedam Yoo^{1†}, Da Un Jeong¹, Aroli Marcellinus¹ and Ki Moo Lim^{1,2*}

¹Dept of IT Convergence Engineering, Kumoh National Institute of Technology

²Dept of Medical IT Convergence Engineering, Kumoh National Institute of Technology

(Manuscript received 2 November 2021 ; revised 14 January 2022 ; accepted 3 February 2022)

Abstract: The Comprehensive in vitro Proarrhythmic Assay(CiPA) project was launched for solving the hERG assay problem of being classified as high-risk groups even though they are low-risk drugs due to their high sensitivity. CiPA presented a protocol to predict drug toxicity using physiological data calculated based on the in-silico model. In this study, features calculated through the in-silico model are analyzed for correlation of changing action potential in the near future, and features are verified through predictive performance according to drug datasets. Using the O'Hara Rudy model modified by Dutta et al., Pearson correlation analysis was performed between 13 features(dV_m/dt_{max} , AP_{peak}, AP_{resting}, APD₉₀, APD₅₀, APD_{tri}, Ca_{peak} , $Ca_{resting}$, CaD₉₀, CaD₅₀, CaD_{tri}, qNet, qInward) calculated at 100 pacing, and between dV_m/dt_{max_repol} calculated at 1,000 pacing, and linear regression analysis was performed on each of the 12 training drugs, 16 verification drugs, and 28 drugs. Indicators showing high coefficient of determination(R^2) in the training drug dataset were qNet 0.93, AP resting 0.83, APD_{tri} 0.78, Ca resting 0.76, dV_m/dt_{max} 0.63, and APD₉₀ 0.61. The indicators showing high determinants in the validated drug dataset were APD_{tri} 0.94, APD₉₀ 0.92, APD₅₀ 0.85, CaD₅₀ 0.84, qNet 0.76, and CaD₉₀ 0.64. Indicators with high coefficients of determination for all 28 drugs are qNet 0.78, APD₉₀ 0.74, and qInward 0.59. The indicators vary in predictive performance depending on the drug dataset, and qNet showed the same high performance of 0.7 or more on the training drug dataset, the verified drug dataset, and the entire drug dataset.

Key words: In silico, Electrophysiological features, CiPA, Torsades de pointes

1. 서 론

1990년부터 2001년 사이에 8개의 심혈관계와 관련이 없는 약물이 염전성 부정맥(Torsades de pointes; TdP)을 유도하여

미국 시장에서 퇴출되었다[1]. 2005년 International Council on Harmonization(ICH)는 신약 개발에서 TdP 위험성을 관리하기 위해 약물의 심장독성 평가 가이드라인을 제시하였다. 약물로 인한 TdP 부작용 사례는 줄어들었으나 엄격한 기준으로 인해 TdP를 유발하지 않는 약물임에도 시장에 유통되지 못하는 문제가 발생하였다[2,3]. ICH의 가이드라인 개정을 위해 Comprehensive of in vitro Proarrhythmia Assay(CiPA) 프로젝트가 13개 선진의료기관들과 함께 시작되었다. CiPA 프로젝트의 in silico working group은 약물의 TdP 유도 위험성을 미리 예측하는 목적으로 in silico 심근세포 모델 활용 지침을 제시하고 있다[4].

*Corresponding Author : Ki Moo Lim
Kumoh National Institute of Technology, Gumi 39177, Republic of Korea

Tel: +82-54-478-7780

E-mail: kmlim@kumoh.ac.kr

†Contributed equally to this work.

본 연구는 금오공과대학교 학술연구비에 의하여 연구된 논문임 (202001600001).

Li 그룹에서는 약물의 TdP 유도 위험군 분류기 개발을 위해 28개 약물의 TdP 유도 위험군을 고위험군, 중위험군, 저위험군으로 구분하여 선정하였다. 그들이 제시한 약물의 위험군 예측 지표는 qNet이라는 6개 이온채널(INaL, ICaL, IKr, IKs, IK1, Ito)을 통해 세포 내부와 외부로 이동하는 총 전하의 양 이다[5]. 각 12개 약물들의 qNet 분포를 기반으로 논리 회귀모델을 통해 계산되는 고위험군과 중위험군을 구별하는 qNet 임계 값과 중위험군과 저위험군을 구별하는 qNet 임계 값으로 세 개의 위험군을 분류 하였다[3]. qNet의 약물 TdP 위험군 분류 성능은 약물 위험군 평가에 주로 사용되는 활동전위 특성(APD90; Action potential duration 90%, APD50; Action potential duration 50%, Vm peak, Vm resting), 세포내 칼슘농도 변화 특성(Ca peak, diastolic Ca²⁺, CaD90; calcium transient duration 90%, CaD50; calcium transient duration 50%)과 비교했을 때 가장 높은 성능을 보인다[5].

현재까지 약물이 가지는 TdP 유도 위험성에 관한 연구들은 Early after depolarization(EAD) 발현에 대한 이온채널들의 민감성 분석 연구들과 전기생리학적 지표를 활용한 약물 위험군 분류에 대한 연구들이 진행되었다[5,6,8,14]. 전기생리학적 지표가 불연속적인 위험군(고위험군, 중위험군, 저위험군)을 예측하는 것 뿐만 아니라 활동전위 형상적인 변화를 예측하는 것 또한 약물이 가지는 약물의 독성을 유의미하게 확인 할 수 있어 검토할 필요가 있다. 또한 TdP는 세포 내 외부에 누적된 전하에 민감하게 반응하기 때문에 단기 시뮬레이션에서 나온 결과로 TdP 발생을 예단하기 힘들다. 그래서 단기적 시뮬레이션인 100회에서 계산된 전기생리학 지표(dVm/dtmax, APpeak, APresting, APD90, APD50, APDtri, Ca_{peak}, Ca_{resting}, CaD90, CaD50, CaDtri, qNet, qInward)와 장기적 시뮬레이션인 1,000회에서 계산된 활동전위 형상의 재분극 기간의 형상적 특성(dVm/dtmax_repol)과의 상관성 분석 그리고 회귀모델을 기반으로 정량적으로 지표의 예측성을 규명하고자 한다.

II. 연구 방법

O'Hara Rudy 그룹이 제안한 심실 심근세포모델을 기반으로 Dutta 그룹에 의해 5개 이온채널(INaL, ICaL, IKr, IKs, IK1)의 채널전도도를 수정한 in silico 모델을 사용하였다[5,9]. IKr 채널의 최대 전도도는 0.9에서 1.013으로, IKs 채널의 최대 전도도는 1.0에서 1.870으로, IK1의 최대 전도도는 1.0에서 1.698로, ICaL의 최대 전도도는 1.0에서 1.007로, 그리고 INaL의 최대 전도도는 1.0에서 2.661로 수정되었다. 전기생리학적 시뮬레이션은 심근세포에서 주요한 이온전류(INa, INaL, Ito, ICaL, IKr, IKs, IK1, INaCa, INaK)에 의해 변화되는 막전위를 계산한다(식 3). 각 이온 채널들의 전류식은 Hodgkin-Huxley 전류방정식과 게이트의 상미분 방정식의 결합으로 계산된다(식 2). 이온채널에 약물의 영향 구현은 약물에 의한 막힘(blockage)을 나타내는 막힘 상수(inhibition constant)를 통해 계산한다(식 1).

$$\text{Inhibition constant(IC)} = \left[1 + \left(\frac{D}{IC50} \right)^h \right]^{-1} \quad (1)$$

$$I_{ion} = IC \cdot G_{ion}(V_m - E_{ion}) \quad (2)$$

$$\frac{dV_m}{dt} = \frac{1}{C_m} (\sum_{ion\ channel} I_{ion} + I_{stim}) \quad (3)$$

위 식에서 I_{ion}은 하나의 이온채널의 전류 방정식이며 G_{ion}은 이온채널의 최대전도도이고 V_m은 막전위, E_{ion}은 이온채널의 평형 전위이다. C_m은 세포막의 커패시턴스이고 I_{stim}은 활동전위를 일으키는 외부자극이다. D는 약물의 농도(nM), IC50은 이온채널이 50% 막히는 지점에서의 농도, h는 Hill 계수이다.

Dutta 모델에서 약물 반응을 표현은 식 2와 같이 이온전류방정식에 막힘 상수를 곱하여 채널에 대해 단순 약물의 막힘 상태인 정적상태(static state)를 구현한다. 전체 실험 과정은 그림 1과 같이 진행된다. 약물 데이터는 CiPA에서 선정한 고위험군, 중위험군, 저위험군으로 구분되는 훈련약물 12개와 검증약물 16개를 사용하였다(표 1)[10]. 실험데이터는

20

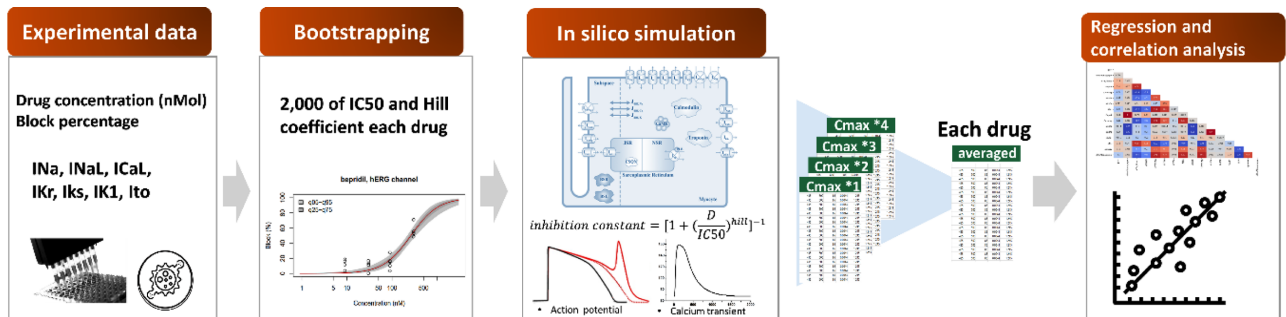


그림 1. 연구 방법 모식도

Fig. 1. The schematic diagram of the research method

표 1. Li 등에서 제시한 TdP 유발 위험군에 따른 28개 약물 목록[17]

Table 1. List of 28 drugs according to TdP-induced risk group suggested by Li et al.[17].

	High	Intermediate	Low
Training Drug Dataset	quinidine bepridil dofetilide sotalol	chlorpromazine cisapride terfenadine ondansetron	diltiazem mexiletine ranolazine verapamil
Validation Drug Dataset	ibutilide azimilide disopyramide vandetanib	risperidone domperidone clarithromycin astemizole droperidol pimozide clozapine	metoprolol tamoxifen loratadine nifedipine nitrendipine

Crumb 그룹이 제시한 전압고정 패치 클램프 실험(voltage-clamp)의 결과를 사용하였다. 실험데이터는 이온채널(INa, INaL, ICaL, IKr, IKs, IK1, Ito)들의 약물 농도 4개 조건에 따라 나타나는 이온전류 막힘 비율(ion channel block percentage) 정보를 사용하였다[11]. 실험데이터의 불확실성 정량화를 위해 Chang 그룹이 제시한 Markov chain Monte Carlo(MCMC) 모델을 사용하였다. MCMC 모델은 약물 농도에 따른 이온채널 막힘 비율을 나타내는 Hill curve를 유도한다. 최적의 Hill curve는 입력으로 사용한 실험데이터를 기반으로 최소제곱법(least square method)의 최소 오차를 가지는 시그모이드(sigmoid) 함수를 계산한다. 최적의 Hill curve를 기반으로 신뢰구간 95% 이내의 2,000개의 hill curve를 증폭하여 각 Hill curve에서 Hill 계수와 IC50를 추출한다[12]. IC50는 이온 전류가 50% 억제 되었을 때 약물 농도, Hill 계수는 IC50농도에서의 Hill curve의 기울기이

다. 2,000개의 데이터에서 각 샘플은 7개 이온채널(INa, INaL, ICaL, IKr, Ito, IK1, IKs)에 대한 IC50와 Hill 계수의 조합으로 구성된다. In vitro 약물 데이터 및 MCMC 모델은 CiPA에서 업로드한 자료를 사용하였다(github.com/CiPA).

In silico 모델 시뮬레이션의 파라미터 설정은 EAD가 발생하는 서맥을 구현하기 위해 자극주기는 2초로 설정하고 총 자극 횟수 1,000회로 시뮬레이션의 결과가 정상상태(steady state)에 도달하도록 설정하였다. 약물의 농도 설정은 약물의 고유 농도인 혈장내 최대 약물 농도 Cmax(nMol)를 기준으로 1배, 2배, 3배, 4배 농도에서 시뮬레이션을 진행하였다. 하나의 약물 당 농도에 따른 8,000개의 qNet, dVm/dt_{max_repol} 결과가 나오게 되며 하나의 샘플을 통해 계산된 4개 농도에서의 평균된 지표 값을 상관성 분석과 회귀 모델을 입력으로 사용한다. 최종적으로 한 약물의 하나의 지

표 2. In silico 시뮬레이션을 통해 계산되는 14개 전기생리학적 지표

Table 2. 14 electrophysiological indicators calculated through in silico simulations

Features	Description
dVm/dt _{max}	Maximum slope value in depolarization period
dVm/dt _{max_repol}	Maximum slope value between 30% and 90% below the maximum value in the repolarization period
APD90	The time interval between depolarization and repolarization corresponding to a membrane potential 90% below the maximum membrane potential value
APD50	The time interval between depolarization and repolarization corresponding to a membrane potential 50% below the maximum membrane potential value
APDtri	Difference between APD90 and APD50
APresting	Resting state of action potential Membrane potential
CaD90	Time interval at concentrations 90% below the maximum calcium concentration
CaD50	Time interval at concentrations 50% below the maximum calcium concentration
CaDtri	Difference Between CaD90 and CaD50
Capeak	Maximum Calcium Concentration Value
Caresting	Resting Calcium Concentration Values
qNet	INaL, ICaL, IKr, IKs, Ito, IK1 Sum of ionic charges moving through ion channels
qInward	Sum of ionic charges moving through INaL and ICaL channels

표 데이터 개수는 2,000개이다.

전기생리학적 지표는 약물에 의해 TdP 유도 위험성을 나타내는 주요한 특성들로 선택되었다[5,13]. In silico 심실 모델을 통해서 표 2에 기재된 14개의 전기생리학적 지표를 계산한다. 본 연구에서 사용한 14개 전기생리학적 지표는 7개의 활동전위 형상의 지표 (dV_m/dt_{max_repol} , dV_m/dt_{max} , AP peak, AP resting, APD90, APD50, APDtri), 5개의 세포 내 칼슘 농도 형상 지표(Ca peak, Ca resting, CaD90, CaD50, CaDtri) 그리고 2개의 이온전하 지표(qNet, qInward)로 구성된다. 활동전위 형상 지표의 dV_m/dt_{max_repol} 는 재분극 기간에 활동전위의 침도 값으로 부터 30%아래 지점과 90% 지점 사이에 가장 큰 값의 기울기이다. dV_m/dt_{max} 은 활동전위 형상에서 탈분극 기간에서 가장 큰 값의 기울기이며 약물로 인해 INa 채널의 막힘으로 인해 변화되는 활동전위 형상 변화를 나타내는 지표이다. Action potential peak potential(AP peak)는 활동전위의 최대 세포막 전위 값으로 INa 채널의 억제율에 따라 바뀌는 지표이다. Action potential resting potential(AP resting)은 활동전위의 휴식기 세포막 전위 값이다. Action potential duration 90%(APD90)은 AP peak에서 AP resting의 사이의 전위 차이에서 AP peak로부터 90% 아래의 전위 값 지점의 활동전위 탈분극과 재분극 기간의 시간 차이, Action potential duration 50%(APD50)은 AP peak로부터 50%아래의 전위 값 지점의 탈분극기간과 재분극 기간의 시간 차이, Action potential duration triangulation(APDtri)은 APD90과 APD50의 값 차이 이다(그림 2(a)). 활동전위 기간과 관련된 지표는 QT연장에 주요한 지표로 사용되는 지표이다[14]. 세포내 칼슘 농도형상 지표의 Ca peak는 최대 calcium transient(nM) 값이며, Ca resting은 휴식기 농도이다. Calcium transient Duration 90%(CaD90)은 Ca peak와 Ca resting을 차이에서 Ca peak에서 90% 아래 지점의 탈분극과 재분극 사이의 시간차이이다. Calcium Transient Duration 50%(CaD50)은 Ca peak와 Ca resting 차이의 50% 지점 사이의 탈분극과 재분극 사이의 간격이다. 세포내 칼슘농도 지표가 부정맥과 밀접한 관련이 있기 때문에 부정맥 예측 지표로서 사용하였다[15]. Calcium Transient Duration triangulation(CaDtri)는 CaD90과 CaD50의 차이이다(그림 2(b)). 알짜전하 특성인 qNet은 시간에 따라 변화되는 이온채널 전류의 합으로 만들어지는 $I_{net}(I_{net} = I_{NaL} + I_{CaL} + I_{Kr} + I_{Ks} + I_{K1} + I_{to})$ 의 아래 면적으로 6개 이온채널(I_{NaL} , I_{CaL} , I_{Kr} , I_{Ks} , I_{K1} , I_{to})을 통해 이동하는 전하들의 총합이다. qInward는 INaL 채널에서 약물이 없는 통제군 조건에서 채널을 통해 이동한 전하량과 약물 영향이 있는 환경에서 채널을 통해 이동한 전하의 비율과 ICaL 채널에서 약물이 없는 통제군에 대해 이온채널을 통해 이동하는 전하

량과 약물 영향 조건에서 이온채널을 통해 이동한 전하의 비율, 이 두가지 비율의 평균 값이다[16]. 이 13가지 지표는 자극횟수 100회째의 활동전위를 통해 계산되며, 추가적으로 자극횟수 1,000회에서 계산된 dV_m/dt_{max_repol} 를 추출한다.

100회 활동전위 에서 계산된 13개의 지표들과 1,000회 활동전위에서 계산된 dV_m/dt_{max_repol} 의 상관계수를 피어슨 상관성 분석을 통해 계산하였다. 상관성 분석에는 훈련용 약물 12개에 대해 14개 지표의 총 336,000개 데이터를 입력으로 사용하였고, 검증용 약물 16개에 대해 14개 지표의 448,000개 데이터가 사용되었으며, 28개 전체 약물에 대해 14개 지표의 784,000개 데이터가 사용되었다. 선형 회귀모델 12개의 훈련 약물, 16개의 검증약물, 28개 전체 약물군에 따라 수행하였으며 사용한 데이터 독립변수에 사용되는 하나의 지표는 훈련용 약물에 대해 24,000개, 검증용 약물에 대해 32,000개 그리고 28개 전체 약물에 대해 56,000개를 사용하였으며 종속변수는 각 조건에서 같은 수의 1,000 pacing에서 측정된 dV_m/dt_{max_repol} 이 사용되었다. 각 지표들이 EAD 발생 예측력을 확인하기 위해 선형회귀 모델을 통해 결정계수(R^2)와 평균제곱오차(Mean square error;MSE) 지표를 계산하였다. 독립변수는 100회 자극에서 계산된 각 지표이며 종속 변수는 1,000회 자극에서 계산된 dV_m/dt_{max_repol} 이다. 지표 계산에 사용된 선형회귀 모델은 10차 교차 검증으로 각 차수마다 추정되는 파라미터의 평균 값을 가지는 모델이다.

III. 연구 결과

훈련약물군, 검증약물군, 전체 약물군에 대해서 단기적 시뮬레이션(100번째 활동전위)에서 계산되는 13개의 TdP 예측 지표와 장기적 시뮬레이션(1,000번째 활동전위)에서 계산되는 dV_m/dt_{max_repol} 과의 상관성 분석을 하였다(그림 3). 0.5이상의 상관계수를 가지는 지표는 활동전위 특성에서 7개 지표 중에서 5개가 포함이 된다. dV_m/dt_{max} -0.86, AP peak가 -0.65, AP resting가 0.92, APD90이 0.78 그리고 APDtri가 0.89로 나타났다. 칼슘특성에서는 5개중에서 한 개의 지표 Ca resting가 0.91이며, 이온전하 특성인 qNet이 -0.97과 qInward 0.76으로 두 개다 0.5이상으로 나타났다. 13개 지표 중에서 0.5 이상의 높은 상관계수를 가지는 지표는 8개 지표이다. 검증약물군에 대해서 13개의 예측 지표와 1,000번째 활동전위에서 계산되는 dV_m/dt_{max_repol} 과의 상관성 분석을 하였다(그림 3(b)). 상관계수가 0.5 이상인 특성으로는 활동전위 특성에서 7개 지표 중에 4개의 지표가 포함되며 dV_m/dt_{max_repol} 이 0.99의 상관계수를 가지고 APD90이 0.97, APD50이 0.96, APDtri가 0.93의 상관계수를 가진다. 칼슘특성 지표 5개 중에서 3개가 0.5이상의 상관계수를

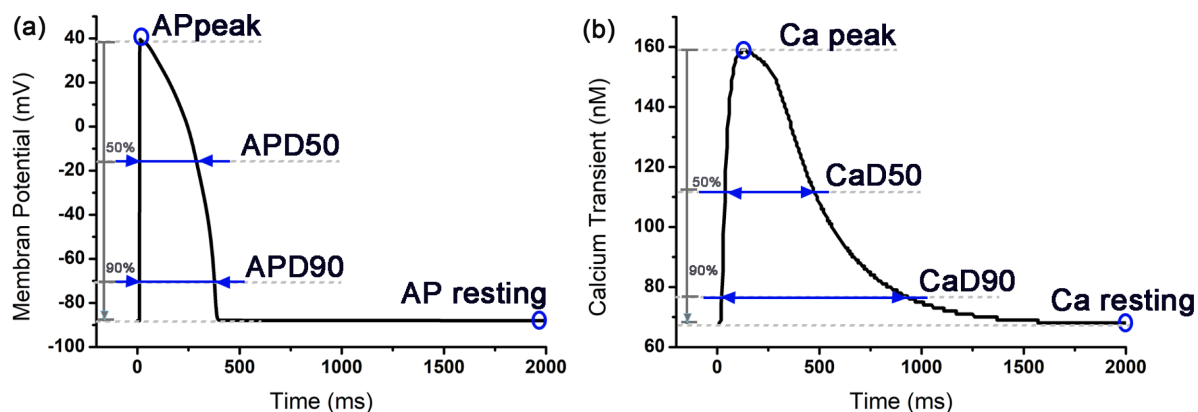


그림 2. *In silico* 심실모델을 통해 계산되는 지표. (a) 활동전위 형상의 지표, (b) 세포내 칼슘 농도 형상의 지표

Fig. 2. Features calculated through the *in silico* ventricular model. (a) Features of action potential shape, and (b) features of intracellular calcium concentration shape

가지며 CaD90이 0.94, CaD50이 0.94, CaDtri가 0.77의 값을 나타낸다. 알짜 전하 특성 2가지는 qNet -0.97, qInward 0.92의 상관계수를 가진다. 14개 지표 중 높은 상관성을 가지는 지표는 14개 중에서 9개가 해당한다. 전체 약물 28개에 대해서 예측 지표 14개와 1,000번째 활동전위에서 계산되는 dV_m/dt_{max_repol} 과의 상관성 분석을 하였다(그림 3(c)). 상관 계수가 0.5이상의 높은 상관성을 가지는 특성을 선별한다. 활동전위 특성에서 7개 지표 중에 한 개의 지표가 포함되며 dV_m/dt_{max_repol} 이 0.98의 상관계수를 가진다. 칼슘특성 지표 5개 중에서 3개가 높은 상관성을 가지며 Ca peak가 0.98, CaD90이 -0.98, CaD50이 -0.97 값을 나타낸다. 알짜 전하 특성은 높은 상관성을 가지는 지표가 없어서 포함되지

않는다. 전체 14개 예측 지표 중에서 높은 상관성을 가지는 지표는 4개 이다.

약물군에 따라 13개 지표들에 대해서 선형 회귀 모델을 통해서 결정계수를 계산하였다(표 2). 훈련약물군에서는 가장 높은 결정계수를 가지는 지표는 qNet으로 0.93이며 결정계수의 크기대로 AP resting이 0.83, APDtri가 0.78, Ca resting이 0.76, dV_m/dt_{max} 가 0.63, APD90이 0.57, qInward가 0.34, AP peak가 -0.088로 나타났다(표 2, 그림 3). 상관 계수가 0.5이상인 지표와 결정계수 0.5이상인 6개 지표 중에서 Ca resting과 APDtri 두개의 지표를 제외하고는 상관계수가 크면 결정계수도 크다(표 2, 그림 3a)). 검증약물군에서 크기대로 APDtri가 0.94, APD90이 0.92, APD50

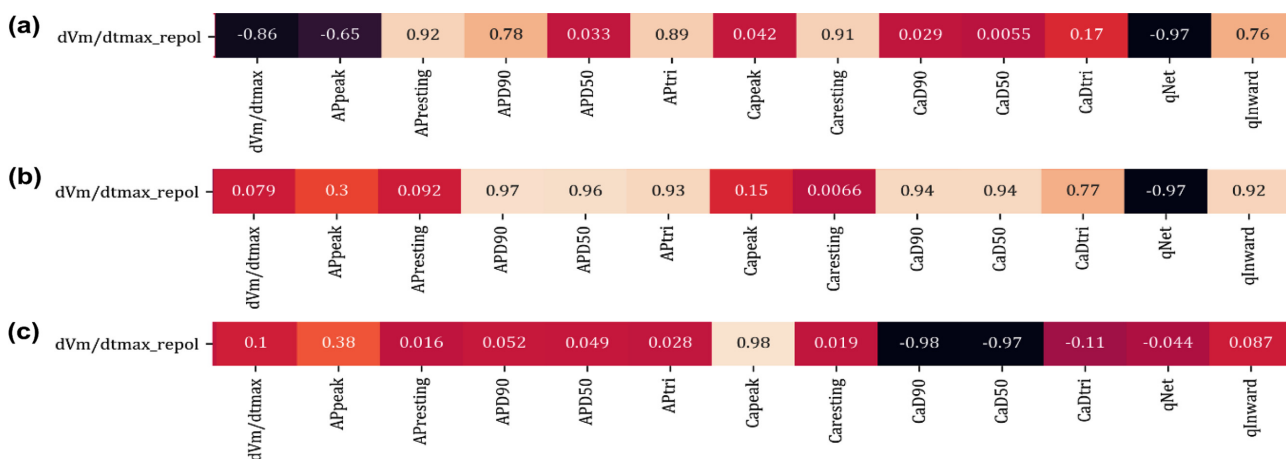


그림 3. 1,000pacing에서 계산된 dV_m/dt_{max_repol} 과 100 pacing에서 계산된 13개 지표(dV_m/dt_{max} , APpeak, APresting, APD90, APD50, APDtri, Ca_{peak}, Ca_{resting}, CaD90, CaD50, CaDtri, qNet, qInward)의 피어슨 상관분석 결과 (a) 12개 훈련약물군, (b) 16개 검증 약물군, (c) 28개 전체 약물군

Fig. 3. Pearson correlation analysis results between dV_m/dt_{max_repol} calculated at 1,000 pacings and 13 features (dV_m/dt_{max} , APpeak, APresting, APD90, APD50, APDtri, Ca_{peak}, Ca_{resting}, CaD90, CaD50, CaDtri, qNet, qInward) calculated at 100 pacing (a) 12 training drug datasets, (b) 16 verified drug datasets, and (c) 28 full drug datasets

표 3. 약물군에 따라 계산된 13개 지표들의 결정계수 비교

Table 3. Comparison of the coefficients of determination of 13 indicators calculated according to the drug dataset

Training drug dataset		Validation drug dataset		All drug dataset	
Feature	R ²	Feature	R ²	Feature	R ²
qNet	0.93	APDtri	0.94	qNet	0.78
Ap resting	0.83	APD90	0.92	APD90	0.74
APDtri	0.78	APD50	0.85	qInward	0.59
Ca resting	0.76	CaD50	0.84	AP resting	0.29
dVm/dtmax	0.63	qNet	0.76	APD50	0.19
APD90	0.61	CaD90	0.64	Ca resting	0.071
qInward	0.34	qInward	0.43	CaD50	-0.014
CaD90	-0.014	dVm/dtmax	-0.14	Ca peak	-0.2
AP peak	-0.088	CaDtri	-0.16	dVm/dtmax	-0.49
APD50	-0.17	Ca peak	-0.19	AP peak	-0.74
Ca peak	-0.31	Ca resting	-0.19	CaDtri	-7.9
CaD50	-0.31	AP peak	-0.2	APDtri	-8.8
CaDtri	-11	Ap resting	-8.7	CaD90	-83

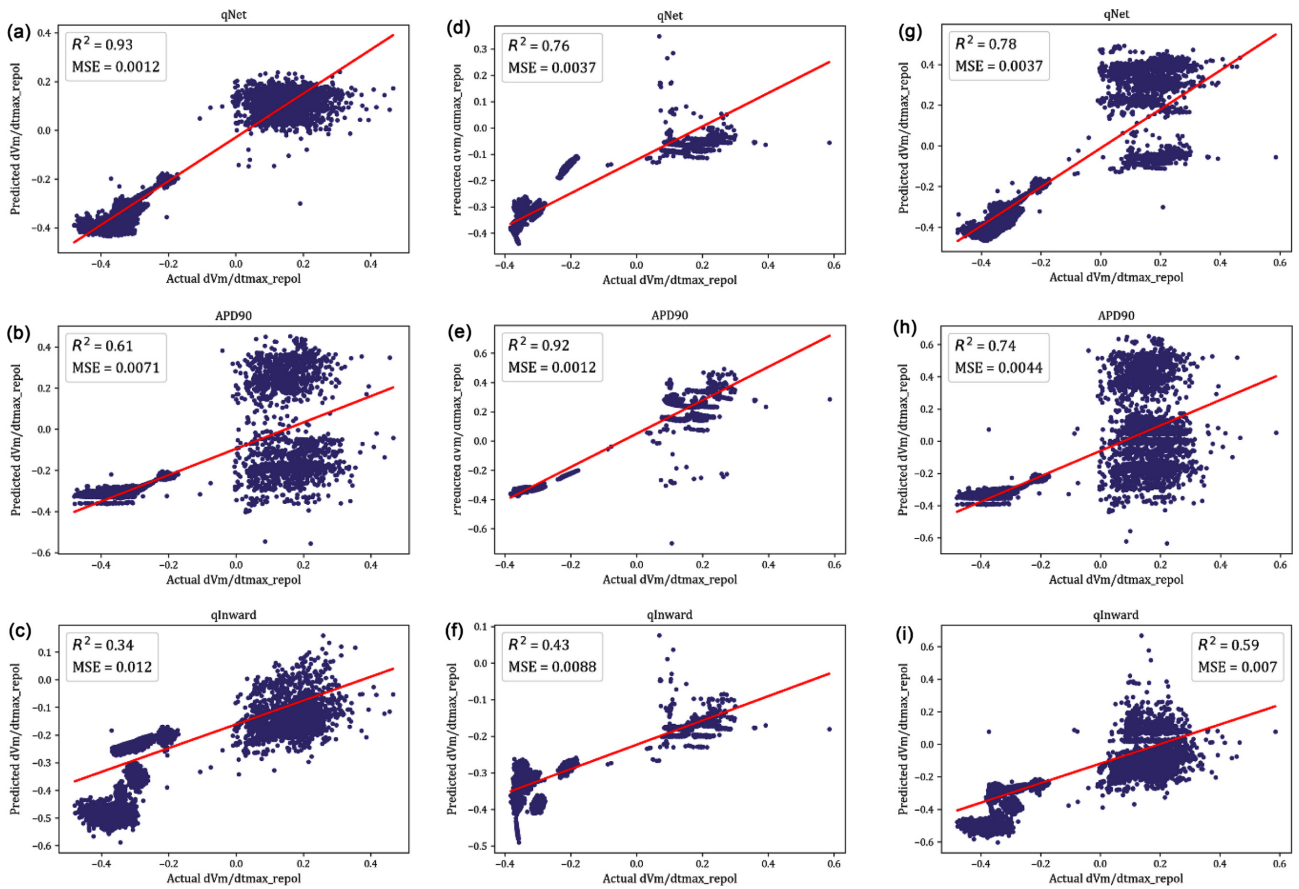


그림 4. qNet, APD90, qInward 3개 지표의군에 따른 선형회귀 결과 그래프. 훈련약물군의 (a) qNet, (b) APD90, (c) qInward, 검증약물군의 (d) qNet, (e) APD90, (f) qInward, 전체 28개 약물의 (g) qNet, (h) APD90, (i) qInward

Fig. 4. Graph of linear regression results according to datasets of qNet, APD90, and qInward. (a) qNet, (b) APD90, (c) qInward in the training drug dataset, (d) qNet, (e) APD90, (f) qInward in the validation drug dataset, (g) qNet, (h) APD90, (i) qInward in all 28 drugs

이 0.85, CaD50이 0.84, qNet이 0.76, CaD90이 0.64 으로 보여졌다(표 2). 전체약물 28개군에서 상관계수가 0.5이상의 결정계수는 qNet이 0.78, APD90이 0.74, qInward가 0.59이다(표 2). 세가지군에서 낮은 성능을 보이는 것은 AP peak, Ca peak와 APDtri, CaDtri이다. 반대로 세가지군에서 높은 성능을 보인 것은 qNet과 APD90이다(표 2). 훈련약물군과 검증약물군에 공통으로 높은 성능을 보이는 APDtri, APD90, qNet 각 결정계수를 비교 했을 때 qNet을 제외한 APDtri는 0.94, APD90 0.92로 훈련약물군에서의 성능보다 APDtri는 16%, APD90은 31% 높았다(표 2). 전체28개 약물군을 사용한 결과로 높은 결정계수를 가지는 지표는 qNet, APD90, qInward이다. 훈련약물군에서 qNet의 결정계수 0.93, 평균 제곱오차 값은 0.0012이고 검증약물군에서 결정계수 0.76, 평균제곱오차 0.0037이며 전체군에서는 결정계수 0.78, 평균 제곱오차 0.0037이며, 훈련약물군에서 가장 높은 성능을 보였다(그림 4(a), (d), (g)). APD90은 훈련 약물군에서 결정계수 0.61, 평균제곱오차 0.0071이고 검증약물군에서 결정계수 0.92, 평균제곱오차 0.0012이며 전체 약물군에서 결정계수 0.74, 평균제곱오차 0.0044이며 검증약물군에서 가장 높은 성능을 보였다(그림 4(b), (e), (h)). qInward는 훈련약물군에서 결정계수 0.34, 평균 제곱오차 0.012, 검증약물군에서 결정계수가 0.43, 평균 제곱오차가 0.0088, 전체 약물군에서 결정계수 0.59, 평균 제곱오차가 0.007로 전체 약물군에서 가장 높은 성능을 보였다(그림 4(c), (f), (i)).

IV. 토의 및 결론

본 연구에서는 약물로 인해 유도되는 TdP 예측 지표로서 사용되는 심실의 전기생리학적 지표(dV_m/dt_{max} , APD90, APD50, APDtri, AP peak, AP resting, CaD90, CaD50, CaD triangulation, Ca peak, Ca resting, qNet, qInward)들이 EAD의 형태적 특성인 dV_m/dt_{max_repol} 와의 상관성분석 그리고 TdP 예측 지표로서 성능을 검증하였다. 추가적으로 약물의 TdP 유도 위험군 분류 모델의 개발시에 사용되는 28개 약물을 훈련 약물과 검증 약물군에 따른 검토하였다. 이 연구를 통해서 qNet은 장기적 시뮬레이션을 예측하는 것에 대해서 훈련 데이터, 검증 데이터, 전체 데이터를 활용한 조건에서 높은 성능을 보였으며 CiPA에서 제시한 약물 데이터에 대해서 유의미한 지표임을 알 수 있었다.

훈련약물과 검증약물 그리고 전체 약물군에서 차이가 나타난 것은 12개약물을 포함하는 훈련약물군은 세개 위험군(고위험군, 중위험군, 저위험군)에 대해서 균형적으로 4개씩 분포하지만 16개 약물을 포함하는 검증약물군은 고위험군 4개, 중위험군 7개, 저위험군 5개로 전기생리학적 지표 데이터 개수 차이가 고위험군의 데이터에 비해 중위험군이

6,000개 그리고 저위험군이 2,000개 더 많다. 전체 데이터 조건에서의 상관계수가 높은 순서와 검증약물의 상관계수가 높은 순서가 비슷한 경향을 보인다. 이러한 결과는 전체 약물군의 상관성 분석 결과에 대해 검증 약물군의 영향력이 훈련약물 보다 더 큰 것을 알 수 있으며 결정계수를 도출하는 데에는 영향력이 미미하다. 전체 약물군은 훈련약물군과 검증약물군을 포함하지만 전체약물군의 결과는 훈련, 검증 약물군과 다른 결과를 보인다.

Dutta 그룹에서는 12개 훈련 약물들을 기반으로 최적화된 TdP 예측 지표로 qNet을 제시하였다[5]. Dutta 그룹은 우리가 사용한 동일한 Ord 모델에서 IKr 채널의 약물 동적결합 모델이 결합된 것을 사용하여 계산된 13개 지표(dV_m/dt_{max} , peak V_m , resting V_m , APD90, APD50, APDtri, peak Ca, diastolic Ca, CaD90, CaD50, CaDtri, qNet, qInward)를 개별적으로 회귀 모델의 입력으로 사용하여 훈련약물군 내에서 하나의 약물을 검증 약물로 선정하여 검증약물의 위험군과 회귀 모델을 통해 예측된 위험군과의 오차를 비교하였다. 오차가 가장 작은 지표순으로 qNet, qInward, APD90, APD50, dV_m/dt_{max} , peak V_m , APDtri, diastolic Ca, peak Ca, Ca tri, resting V_m , CaD90, CaD50이다. 우리의 결과와 비교하였을 때 qNet이 가장 높은 성능을 가지는 것은 동일하나 다른 지표의 순위는 차이가 났다[5].

0보다 큰 dV_m/dt_{max_repol} 범위에서는 회귀모델의 예측 정확도가 0보다 작은 범위에서보다 더 낮았다(그림 3). 0보다 큰 범위는 훈련약물군에서는 quinidine과 검증약물군에서는 ibutilide 약물조건하에서 발생한 EAD에 의한 영향이다. 회귀모델은 안정한 것 보다 상대적으로 불안정한 약물 예측이 어려움을 알 수 있다. 훈련약물군 24,000개 중에서 quinidine은 2,000개로 약 8.33%이고 검증약물군 32,000개 중에 ibutilide는 2,000개로 6.25%를 차지한다. 전체군 56,000개 데이터 중에서 ibutilide와 quinidine 약물의 데이터는 4,000개로 전체 데이터의 7%에 불과하다. 따라서 CiPA 에서 제시한 약물군 내에서 전기생리학적 지표들의 성능을 비교하고 예측 지표로서 실효성을 검증하는 본 연구의 목적에는 큰 영향을 미치지 않을 것으로 보인다.

본 연구의 한계점으로는 약물이 가지는 이상치에 대한 제거 절차를 하지 않아서 지표의 성능이 과소평가 될 수 있다는 점이다. 하지만 CiPA 프로토콜을 제한한 Li 그룹에서도 약물의 반응에 대한 이상치 제거를 하지않고 약물의 위험군 분류기에 활용을 했다[17]. 이러한 제한점이 본 연구의 결과에 영향을 미치지 않을 것으로 예상된다. 추후 연구로는 검증된 지표를 활용한 약물 독성 분류기의 성능을 확인할 필요할 것이다. 또한 EAD발생을 유도하는 약물을 추가하여 EAD이 발생했을 때의 예측 정확성을 더 면밀하게 확인해 볼 필요가 있다.

본 연구의 결과는 약물의 TdP 유발 예측으로 활용되는 단기적 시뮬레이션에서 계산된 전기생리학적 지표가 장기적 시뮬레이션에서 계산되는 활동전위의 형상을 예측하는 성능에 대해 정량적으로 비교하여 지표 검증을 하였다. 결론적으로 지표는 약물군에 따라서 성능이 달라지며 qNet은 혼련약물군, 검증약물군, 전체약물군에서 동일하게 0.7이상의 높은 성능을 보였다.

References

- [1] Fermini B, Fossa AA. "The impact of drug-induced QT interval prolongation on drug discovery and development," *Nat. Rev. Drug Discov.*, 2003;2(6):439-447. doi: 10.1038/nrd1108.
- [2] Strauss DG *et al.*, "Comprehensive In Vitro Proarrhythmia Assay (CiPA) Update from a Cardiac Safety Research Consortium / Health and Environmental Sciences Institute / FDA Meeting," *Ther. Innov. Regul. Sci.*, 2019;53(4):519-525, doi: 10.1177/2168479018795117.
- [3] Fermini B. *et al.*, "A new perspective in the field of cardiac safety testing through the comprehensive in vitro proarrhythmia assay paradigm," *J. Biomol. Screen.*, 2016;21(1):1-11, doi: 10.1177/1087057115594589.
- [4] Colatsky T. *et al.*, "The Comprehensive in Vitro Proarrhythmia Assay (CiPA) initiative — Update on progress," *J. Pharmacol. Toxicol. Methods*, 2016;81:15-20, doi: 10.1016/j.vascn.2016.06.002.
- [5] Dutta S. *et al.*, "Optimization of an in silico cardiac cell model for proarrhythmia risk assessment," *Front. Physiol.*, 2017;8:1-15, doi: 10.3389/fphys.2017.00616.
- [6] McMillan B, Gavaghan DJ, Mirams GR. "Early afterdepolarisation tendency as a simulated pro-arrhythmic risk indicator," *Toxicol. Res. (Camb)*, 2017;6(6):912-921, doi: 10.1039/c7tx00141j.
- [7] Mirams GR. *et al.*, "Simulation of multiple ion channel block provides improved prediction of compounds' clinical torsadogenic risk," *Cardiovasc. Res.*, 2011;91(1):53-61, doi: 10.1093/cvr/cvr044.
- [8] Llopis-Lorente J, Gomis-Tena J, Cano J, Romero L, Saiz J, Trenor B. "InSilico Classifiers for the Assessment of Drug Proarrhythmicity," *J. Chem. Inf. Model.*, 2020;60(10):5172-5187, doi: 10.1021/acs.jcim.0c00201.
- [9] O'Hara T, Virág L, Varró A, Rudy Y. "Simulation of the undiseased human cardiac ventricular action potential: Model formulation and experimental validation," *PLoS Comput. Biol.*, 2011;7(5), doi: 10.1371/journal.pcbi.1002061.
- [10] Han X. *et al.*, "A general procedure to select calibration drugs for lab-specific validation and calibration of proarrhythmia risk prediction models: An illustrative example using the CiPA model," *J. Pharmacol. Toxicol. Methods*, 2020;105:106890, doi: 10.1016/j.vascn.2020.106890.
- [11] Crumb WJ, Vicente J, Johannesen L, Strauss DG, "An evaluation of 30 clinical drugs against the comprehensive in vitro proarrhythmia assay (CiPA) proposed ion channel panel," *J. Pharmacol. Toxicol. Methods*, 2016;81:251-262, doi: 10.1016/j.vascn.2016.03.009.
- [12] Chang KC. *et al.*, "Uncertainty quantification reveals the importance of data variability and experimental design considerations for in silico proarrhythmia risk assessment," *Front. Physiol.*, 2017;8:1-17, doi: 10.3389/fphys.2017.00917.
- [13] Parikh J, Di Achille P, Kozloski J, Gurev V. "Global sensitivity analysis of ventricular myocyte model-derived metrics for proarrhythmic risk assessment," *Front. Pharmacol.*, 2019;10:1-18, doi: 10.3389/fphar.2019.01054.
- [14] Ahmad K, Dorian P. "Drug-induced QT prolongation and proarrhythmia: an inevitable link?," *Europace*, 2007;9:4, doi: 10.1093/europace/eum167.
- [15] Lancaster MC, Sobie EA. "Improved Prediction of Drug-Induced Torsades de Pointes Through Simulations of Dynamics and Machine Learning Algorithms," *Clin. Pharmacol. Ther.*, 2016;100(4):371-379, doi: 10.1002/cpt.367.
- [16] Li Z. *et al.*, "Improving the in silico assessment of proarrhythmia risk by combining hERG (Human Ether-à-go-go-Related Gene) channel-drug binding kinetics and multichannel pharmacology," *Circ. Arrhythmia Electrophysiol.*, 2017;10(2):1-12, doi: 10.1161/CIRCEP.116.004628.
- [17] Li Z. *et al.*, "Assessment of an In Silico Mechanistic Model for Proarrhythmia Risk Prediction Under the CiPA Initiative," *Clin. Pharmacol. Ther.*, 2019;105(2):466-475, doi: 10.1002/cpt.1184.