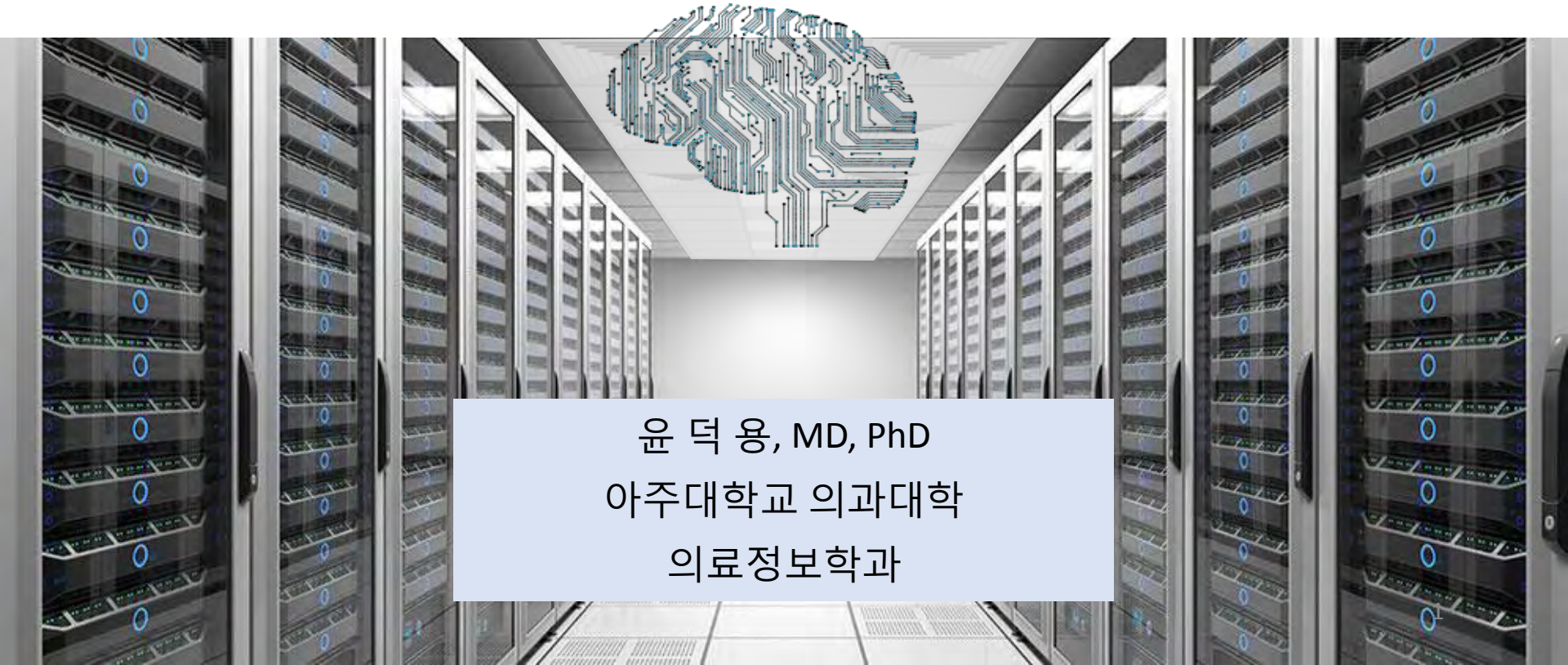
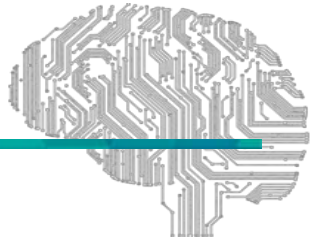
의료 데이터 이해와 활용 및 분석 사례



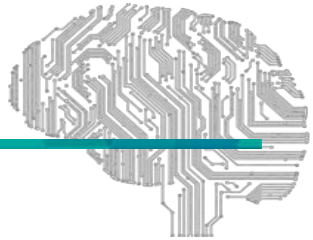
윤 덕 용, MD, PhD
아주대학교 의과대학
의료정보학과

About myself



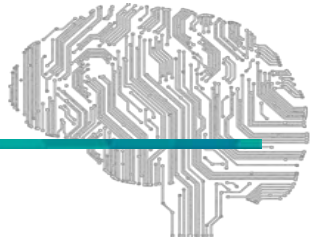
- EDUCATION
  - PhD
    - Department of Biomedical informatics, Ajou University School of Medicine, Suwon, Korea (Mar/2011 – Feb/2016)
  - MS
    - Department of Biomedical informatics, Ajou University School of Medicine, Suwon, Korea (Mar/2009 – Feb/2011)
  - MD
    - Department of Medicine, Ajou University School of Medicine, Suwon, Republic of Korea (Mar/2002 – Feb/2008)

About myself (cont.)



- PROFESSIONAL APPOINTMENTS
  - Assistant professor
    - Department of Biomedical informatics, Ajou University School of Medicine, Suwon, Korea (Mar/2016 – present)
  - Member of Board of Directors
    - The Korean Society of Medical Informatics (Feb/2016 – present)
  - Research assistant
    - Department of Biomedical informatics, Ajou University School of Medicine, Suwon, Korea (Mar/2009 – Feb/2016)
  - Internship
    - Ajou University Medical Center (Mar/2008 – Feb/2009)

Overview of research topics

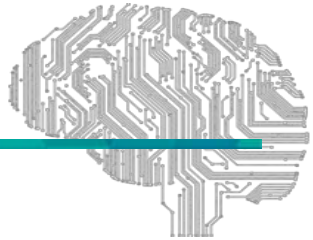


● Published / Accepted ● Finished / under review ● On-going

	Data (Infrastructure)			Knowledge (Discovery)			
	Status survey	**Database construction**	**Data integration & cleaning**	**Disease marker (Bioinfo)**	**Drug safety (EMR/Claim)**	**Disease prediction (Bio-signal)**	**Disease biomarker (Actigraphy)**
2011	● EHR adoption			● HIV	● CERT		
2012		● ECG-ViEW			● CLEAR		
2013				● CMT	● PACE ● ARB-Serum K		
2014-2015	● EHR adoption (2015 update)	● PFT ● BMD ● CAG	● CDM ● Normali-zation		● MAE ✓ ● Statin-NOD ● Statin-GIB ✓ ● DES-MACE		
2016		✓ ● Bio-signal database	● DRN		● Olmesartan ● DPP4i	✓ ● Sepsis prediction	
2017							✓ ● Alzheimer's dz biomarker

Analysis Algorithms

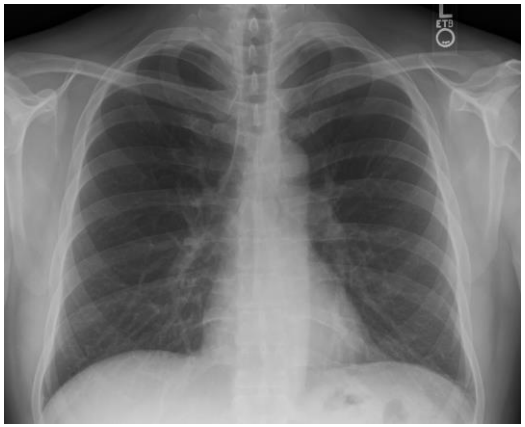
목차



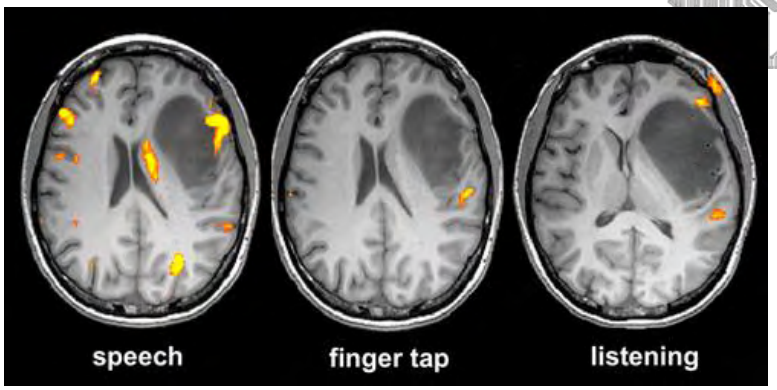
- 의료데이터의 종류와 특징
  - Photographs
  - Narrative textual data
  - Numerical measurement
  - Recorded signal

- 의료데이터 처리 및 분석 사례
  - scan된 이미지 처리
  - 자유기술문 처리
  - EMR로 부터 clinical event 정의
  - 다기관 정보 공유 네트워크 구축 사례
  - 생체신호 분석 사례

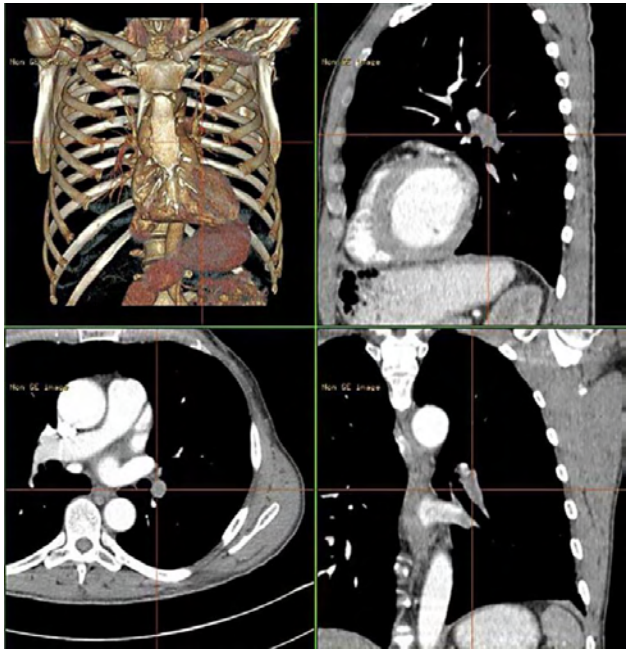
의료데이터의 이해 - Photographs



Chest X-ray

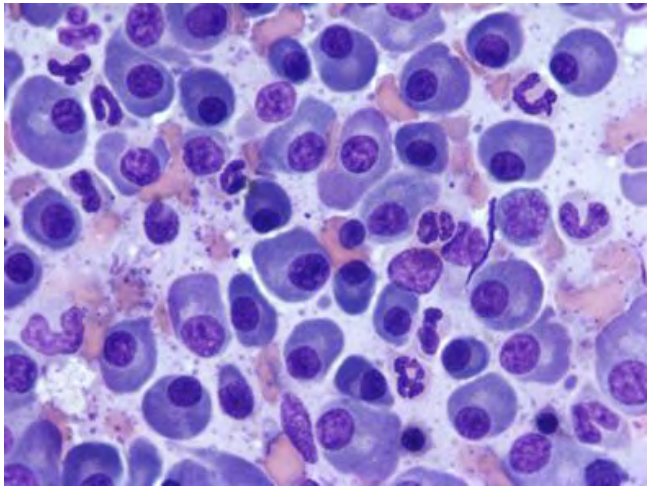


Functional MRI



CT &
3D reconstruction

Cytology image

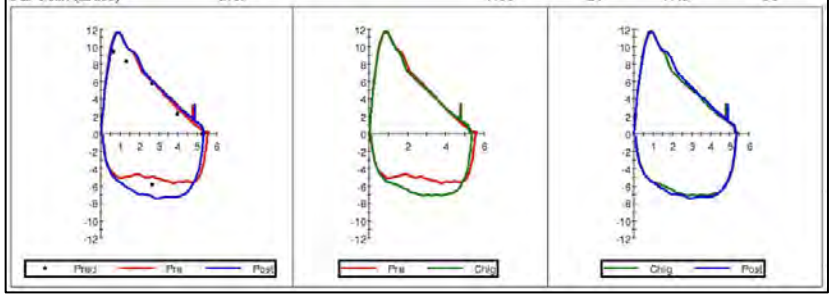


의료데이터의 이해 - Photographs

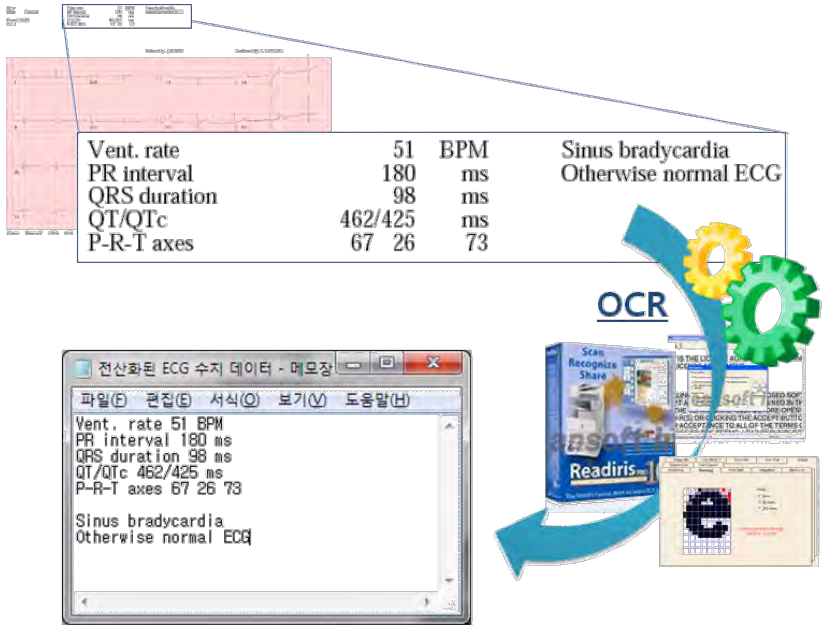
Scanned reports - PFT

Dyspnea: Cough: Wheeze:
Tbco Prod: Yrs Smk: Pks/Day: Yrs Quit:
PF Reference
Pre Test Comments:
Post Test Comments:

	Pre-Bronch			Challenge		Post-Bronch	
	Actual	Pred	%Pred	Actual	%Chng	Actual	%Chng
--- SPIROMETRY --							
FVC (L)	5.59	5.27	106	5.35	-4	5.37	-4
FEV1 (L)	4.77	4.44	107	4.82	1	4.92	3
FEV1/FVC (%)	85	84	102	90	5	92	7
FEF 25% (L/sec)	9.89	8.32	119	9.75	-1	9.87	-0
FEF 75% (L/sec)	2.48	2.33	106	2.55	3	2.77	12
FEF 25-75% (L/sec)	4.98	4.75	105	4.94	-1	5.34	7
FEF Max (L/sec)	11.58	9.49	122	11.74	1	11.64	0
FIVC (L)	5.52			5.24	-5	5.29	-4
FIF Max (L/sec)	5.69			7.08	24	7.45	31



Scanned reports - ECG



의료데이터의 이해 - Narrative textual data

입원기록지

Admission note [내과] [ ]

나이 [ ] 입원 회차 2 입원과 [ ]

성별 [ ] 재실 [ ] 주치의 [ ]

입원일 [ ] 입원경로 외래

주호소

R/O	주증상 및 내원사유	duration	onset	기타
	HD			

현병력

상기 [ ]세 [ ]환 6년전 DM 진단 받고 medication 중인 환자로 타병원에서 cough 에 대한 evaluation 하던중 발견된 serum Cr elevation 에 대해 evaluation 하기 위해 [ ] 입원하였고 당시 CKD 진단 받고 안과에서 avastin injection 받았으며 BP control 및 renal Bx. 시행하고퇴원하였고 이후 impending ESRD로 [ ] 진료 하던 중 [ ] perm catheter insertion 및 hemodialysis 시작하기 위해 입원함

방사선 판독 보고서

```
Slightly enlarged both kidneys with suspicious mild increased cortical
echogenecities.
Both kidney shape are within normal range.
No evidence of hydronephrosis.
Bladder is not remarkable.
  - with urine debris.
--
====== [Conclusion] ======
1. R/O Bilateral renal parenchymal disease.
2. Urine debris.
```

의료데이터의 이해 - Narrative textual data

- CC, Present illness, Past Hx, Social Hx, Family Hx, ROS, P/E
- surgical procedure, consult, 방사선판독보고서, 병리결과보고서, 퇴원요약지

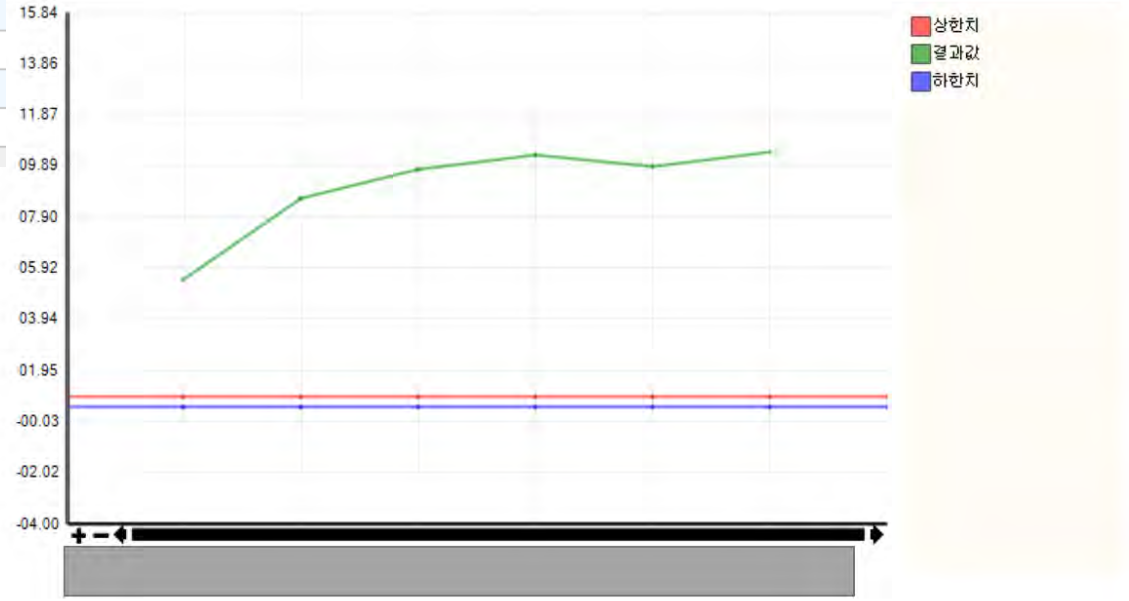
- 자유기술문 데이터의 특징
  - Loosely coded
  - 약자 혹은 임의의 약자
    - WNL: within normal limit, ROM: Range of motion
    - 표준화되지않거나중복되는약자
      - MI: myocardial infarction or myocardial insufficiency
      - (CS/DM)
  - Complete phrases: loose standards expression
    - Mild dyspnea on exertion, failure to thrive, soft and flat
    - 환자의 이질적 상황을 요약하여 단순한 개념으로 전달하기 위함

의료데이터의 이해 – Numerical measurement

검사결과 | 누적검사결과

검사코드	검사명	단위	참고치(하)	참고치(상)				
C3730001	BUN	mg/dL	6.0	20.0	77.1	89.7	89.4	92.8
C3750001	Creatinine	mg/dL	0.50	0.90	5.47	8.60	9.76	10.33

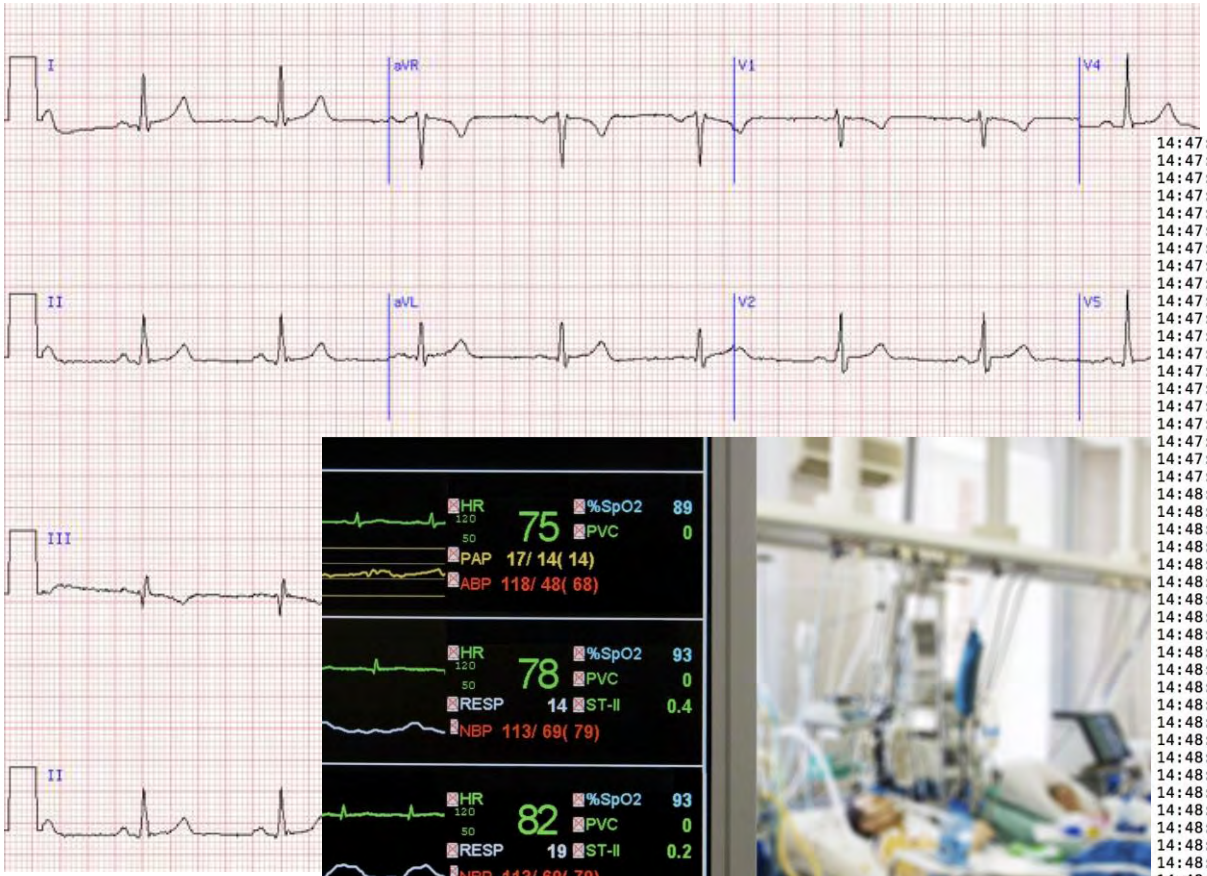
검사명	결과(수치)	단위	정상구
BUN	98.5	mg/dL	H
Creatinine	10.42	mg/dL	H
estimated GFR	4.56		
Na	139	mMol/	
K	5.0	mMol/	
Cl	105	mMol/	
CO2	14	mMol/	L



의료데이터의 이해 – Numerical measurement

- Laboratory tests (임상병리검사)
  - vital signs (temperature, pulse rate)
- Physical examination (진찰)
- Precision Issue
  - 복부진찰에서간크기9cm와10cm 차이를구별할수있는가?
  - serum sodium: 128.94mEq/L ?
  - 몸무게1kg fluctuation/week?

의료데이터의 이해 – Recorded signal



14:47:45.239	128	90	107	81	0.742	21	1.72	0.275	3.732
14:47:45.981	128	91	106	81	0.745	22	1.75	0.280	3.642
14:47:46.726	126	91	104	80	0.750	22	1.78	0.275	3.514
14:47:47.476	124	89	105	81	0.745	24	1.93	0.270	3.256
14:47:48.221	127	92	107	82	0.735	25	2.02	0.280	3.163
14:47:48.956	126	91	106	79	0.758	26	2.01	0.285	3.149
14:47:49.714	123	90	105	81	0.742	25	2.03	0.280	3.088
14:47:50.456	125	90	106	81	0.743	26	2.13	0.280	2.992
14:47:51.199	127	92	107	81	0.745	26	2.11	0.290	3.035
14:47:51.944	124	91	107	80	0.751	26	2.10	0.290	3.050
14:47:52.695	124	91	107	78	0.765	26	2.06	0.290	3.109
14:47:53.460	124	91	107	79	0.755	26	2.09	0.290	3.064
14:47:54.215	124	91	107	83	0.725	26	2.18	0.290	2.938
14:47:54.940	124	91	107	80	0.747	26	2.11	0.290	3.035
14:47:55.687	129	90	107	81	0.739	24	1.90	0.280	3.371
14:47:56.426	132	92	109	82	0.732	24	1.97	0.280	3.312
14:47:57.158	131	91	108	79	0.755	26	2.05	0.295	3.153
14:47:57.913	128	91	106	81	0.740	26	2.07	0.285	3.065
14:47:58.653	129	90	107	81	0.740	28	2.25	0.280	2.853
14:47:59.393	130	91	107	82	0.733	28	2.28	0.285	2.822
14:48:00.126	128	91	106	80	0.750	27	2.18	0.275	2.910
14:48:00.876	126	89	105	80	0.750	28	2.20	0.275	2.850
14:48:01.626	128	90	107	81	0.740	28	2.27	0.275	2.821
14:48:02.366	128	91	107	81	0.744	28	2.25	0.275	2.846
14:48:03.110	128	91	107	78	0.765	28	2.20	0.275	2.911
14:48:03.875	128	91	107	80	0.750	28	2.24	0.275	2.859
14:48:04.625	128	91	107	84	0.715	28	2.35	0.275	2.725
14:48:05.340	128	91	107	80	0.752	28	2.24	0.275	2.859
14:48:06.092	131	91	107	80	0.750	23	1.84	0.280	3.480
14:48:06.842	131	91	107	81	0.739	24	1.96	0.275	3.267
14:48:07.581	126	86	202	28	2.175	27	0.74	0.290	16.337
14:48:09.756	111	77	95	79	0.757	35	2.74	0.285	2.074
14:48:10.513	119	82	99	82	0.735	28	2.28	0.290	2.611
14:48:11.248	116	83	99	83	0.727	28	2.31	0.290	2.577
14:48:11.975	116	83	99	82	0.730	28	2.30	0.290	2.589
14:48:12.705	116	83	99	80	0.750	28	2.24	0.290	2.658
14:48:13.455	116	83	99	78	0.765	28	2.20	0.290	2.706
14:48:14.220	116	83	99	84	0.715	28	2.35	0.290	2.534
14:48:14.935	116	83	99	82	0.730	28	2.30	0.290	2.589
14:48:15.665	116	83	99	78	0.773	28	2.17	0.290	2.744
14:48:16.438	126	86	102	80	0.750	24	1.88	0.290	3.247
14:48:17.188	127	86	104	81	0.740	25	2.05	0.290	3.029
14:48:17.928	127	87	104	81	0.743	26	2.13	0.295	2.915
14:48:18.671	126	87	101	79	0.755	27	2.13	0.295	2.845
14:48:19.426	123	85	101	81	0.742	29	2.33	0.280	2.600
14:48:20.168	125	87	103	82	0.736	29	2.34	0.295	2.628
14:48:20.904	125	87	102	79	0.755	29	2.31	0.290	2.642
14:48:21.659	122	85	100	80	0.750	29	2.34	0.280	2.570
14:48:22.409	124	86	102	81	0.740	30	2.43	0.290	2.518

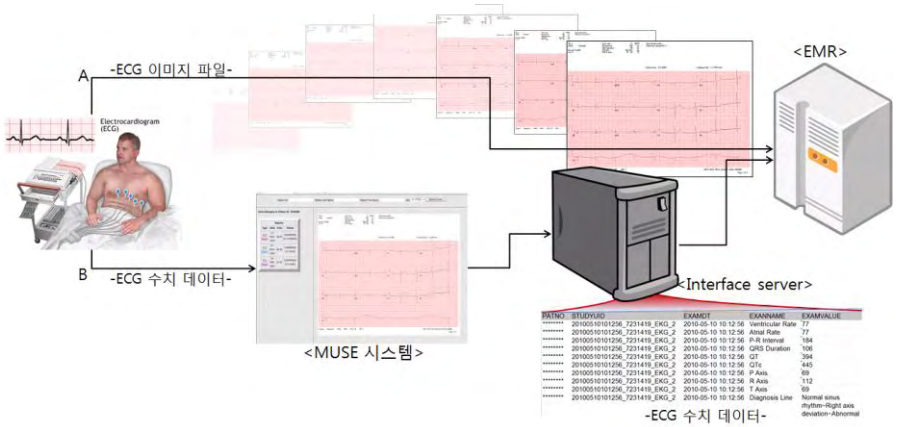
의료데이터의 이해 - Time

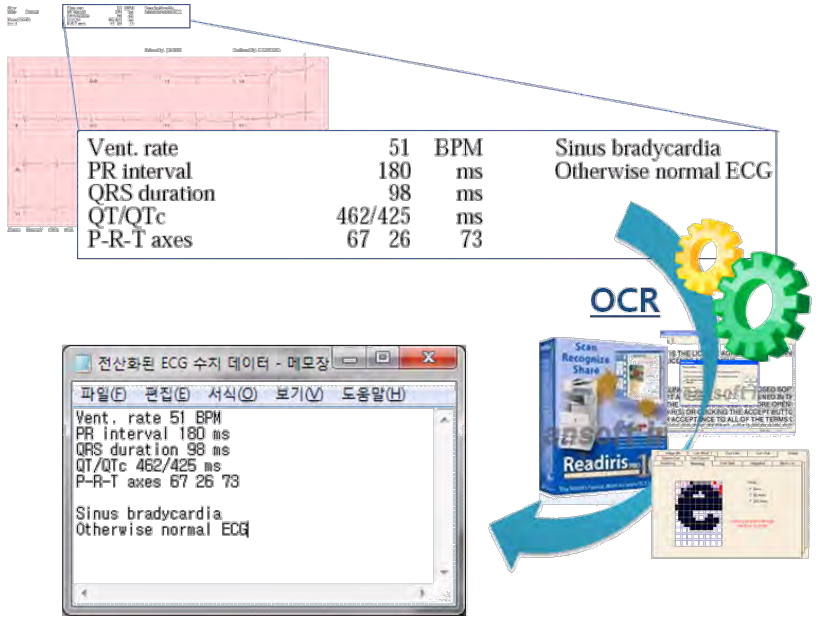
- 상황에 따라 정확성에 대한 요구도가 달라짐
  - 외래: 날짜 정도만 있어도 충분
  - 응급상황:
    - Diabetic ketoacidosis (당뇨병성케톤혈증): 당뇨병의급성합병증 → 분단위로 혈당 측정
    - 심장성쇼크 → MAP(평균동맥혈압) 연속측정

Database construction

- Outcome database 1: ECG
  - The QT interval data is essential for surveillance of the proarrhythmia potential of drugs (the second most common cause of withdrawal)
  - However, many ECG records are still stored as printed documents

*ECG: electrocardiogram

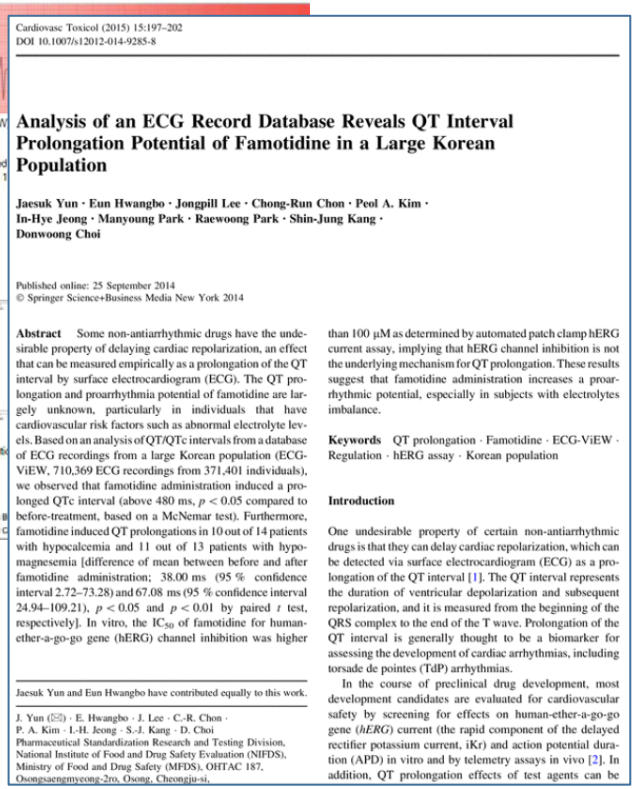




Database construction

- Outcome database 1: ECG (cont.)

Characteristics	Value
Patients, n	371401
Age, years*	42.4 ± 19.2
RR interval, ms*	853.6 ± 176.4
QT interval, ms*	385.2 ± 41.6
QTc (Bazett), ms*	419.4 ± 27.1
Male	414.9 ± 26.3
Female	423.9 ± 27.1
QTc (Fridericia), ms*	407.6 ± 26.2
QTc (Framingham), ms*	385.3 ± 41.6
QTc (Bazett) prolongation*	30168 (8.1%)
Department*	
Health examination	62576 (16.8%)
Outpatient	194219 (52.3%)
Emergency	59899 (16.1%)
Inpatient	54707 (14.7%)
Observation period (days)†	502.0 ± 1008.2
Number of ECG/patient	1.9 ± 2.0
Medications§	
No. of classes	911
No. of prescriptions	37874129
Laboratory test	
No. of serum potassium	1328621
No. of serum magnesium	520817
No. of serum calcium	1063795



Cardiovasc Toxicol (2015) 15:197–202
DOI 10.1007/s12012-014-9285-8

Analysis of an ECG Record Database Reveals QT Interval Prolongation Potential of Famotidine in a Large Korean Population

Jaesuk Yun · Eun Hwangbo · Jongpill Lee · Chong-Run Chon · Peol A. Kim · In-Hye Jeong · Manyoung Park · Raewoong Park · Shin-Jung Kang · Donwoong Choi

Published online: 25 September 2014
© Springer Science+Business Media New York 2014

Abstract Some non-antiarrhythmic drugs have the undesirable property of delaying cardiac repolarization, an effect that can be measured empirically as a prolongation of the QT interval by surface electrocardiogram (ECG). The QT prolongation and proarrhythmia potential of famotidine are largely unknown, particularly in individuals that have cardiovascular risk factors such as abnormal electrolyte levels. Based on an analysis of QT/QTc intervals from a database of ECG recordings from a large Korean population (ECG-ViEW, 710,369 ECG recordings from 371,401 individuals), we observed that famotidine administration induced a prolonged QTc interval (above 480 ms, $p < 0.05$ compared to before-treatment, based on a McNemar test). Furthermore, famotidine induced QT prolongations in 10 out of 14 patients with hypocalcemia and 11 out of 13 patients with hypomagnesemia [difference of mean between before and after famotidine administration; 38.00 ms (95 % confidence interval 2.72–73.28) and 67.08 ms (95 % confidence interval 24.94–109.21), $p < 0.05$ and $p < 0.01$ by paired t test, respectively]. In vitro, the IC_{50} of famotidine for human-ether-a-go-go gene (hERG) channel inhibition was higher than 100 μM as determined by automated patch clamp hERG current assay, implying that hERG channel inhibition is not the underlying mechanism for QT prolongation. These results suggest that famotidine administration increases a proarrhythmic potential, especially in subjects with electrolytes imbalance.

Keywords QT prolongation · Famotidine · ECG-ViEW · Regulation · hERG assay · Korean population

Introduction

One undesirable property of certain non-antiarrhythmic drugs is that they can delay cardiac repolarization, which can be detected via surface electrocardiogram (ECG) as a prolongation of the QT interval [1]. The QT interval represents the duration of ventricular depolarization and subsequent repolarization, and it is measured from the beginning of the QRS complex to the end of the T wave. Prolongation of the QT interval is generally thought to be a biomarker for assessing the development of cardiac arrhythmias, including torsade de pointes (TdP) arrhythmias.

In the course of preclinical drug development, most development candidates are evaluated for cardiovascular safety by screening for effects on human-ether-a-go-go gene (*hERG*) current (the rapid component of the delayed rectifier potassium current, iKr) and action potential duration (APD) in vitro and by telemetry assays in vivo [2]. In addition, QT prolongation effects of test agents can be

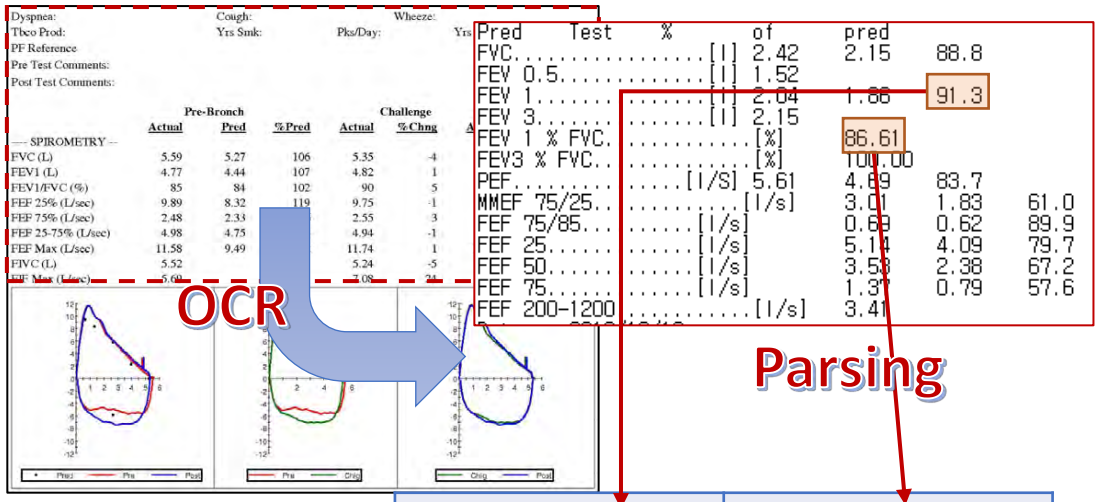
Jaesuk Yun and Eun Hwangbo have contributed equally to this work.

J. Yun (✉) · E. Hwangbo · J. Lee · C.-R. Chon · P. A. Kim · I.-H. Jeong · S.-J. Kang · D. Choi
Pharmaceutical Standardization Research and Testing Division, National Institute of Food and Drug Safety Evaluation (NIFDS), Ministry of Food and Drug Safety (MFDS), OHTAC 187, Osongsaengmyeong-2ro, Osong, Cheongju-si,

Database construction

- Outcome database 2: PFT

Scanned paper results



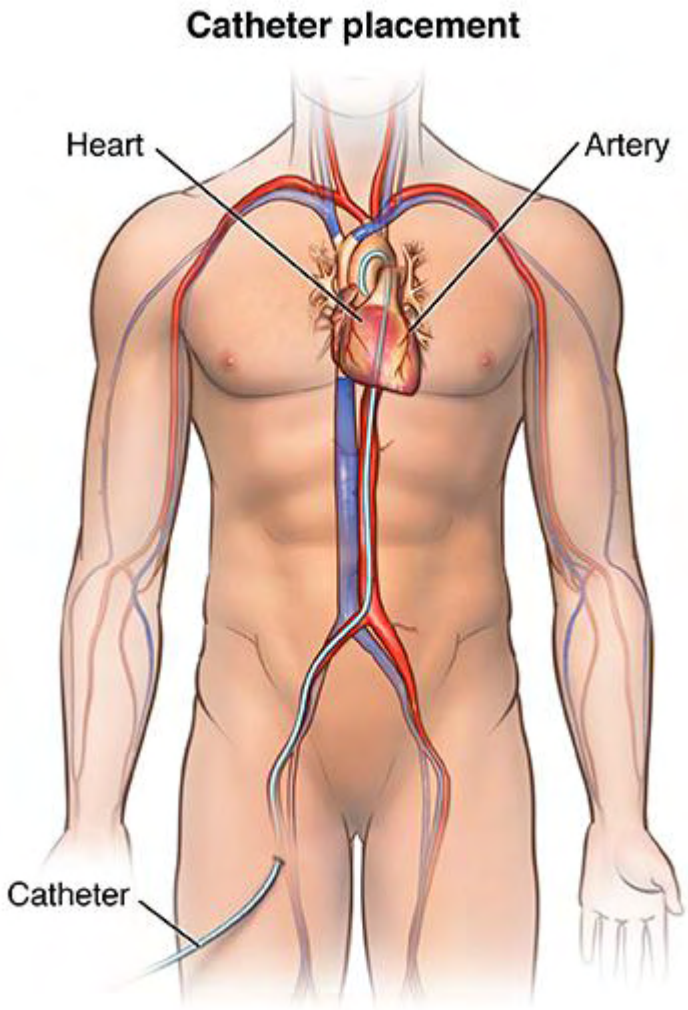
Paper results:
86,327 results

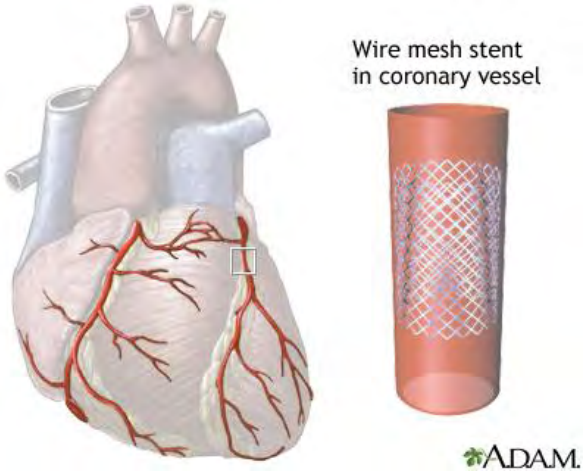
FEV1 (% pred)	FEV1/FVC
91.3	86.61

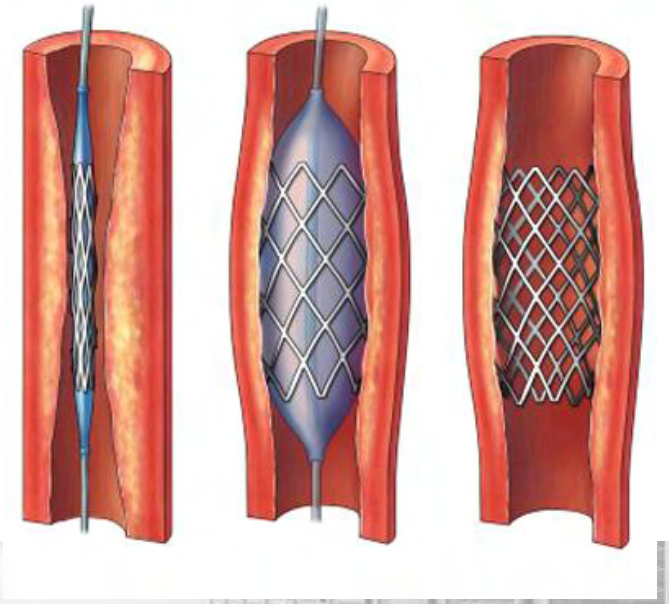
Total : 405,149 results
(including health examination data)

Characteristics	Value
Patients, n	162,255
Age, years*	44.15±10.29
Sex (male, %)	63.0 %
FEV1 (% pred)	101.58±16.26
FEV1/FVC	85.16±6.83
Smoking	3.43 %
Comorbidity	
Asthma	2.04%
COPD	1.29%
Hypertension	5.37%
Diabetes mellitus	3.13%
Medications	
No. of ATC codes	1,191
No. of prescriptions	1,023,529
Laboratory test	
No. of CRP	5,838
No. of magnesium	16,749
No. of total bilirubin	243,463
No. of uric acid	242,828
No. of HbA1c	18,568
No. of HDL	251,368

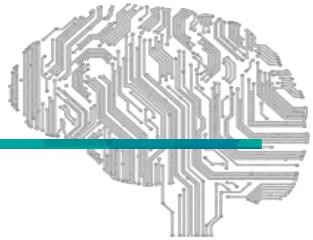
Extracting information from PCI reports (1)







Extracting information from PCI reports (2)



Diagnostic CAG
1. Rt. Femoral artery approach
2. JL 6/4F and JR 6/3.5F catheters were used
3. Findings
  1) LM: normal
  2) LAD: pLAD: tubular eccentric 48%
    mLAD: No ISR at previous stented site
    distal edge: tubular concentric 52%
  3) LCX: p-dLCX: diffuse irregular up to 68%
    (small vessel)
    OM: total occlusion
    (TIMI 0, small vessel)
  4) RCA: p-mRCA: diffuse irregular up to 55%
    dRCA: diffuse irregular up to 39%

*Diagnostic angiography 시행 후 acute onset chest pain 및 ECG상 ST elevation 소견 보여 angiography 다시 시행함
 -> proximal edge에 total occluion 소견 보임 (TIMI 0, TMPG 0)

PCI
1.EBU 7/3.5F guiding catheter was used.
2.Primary PCI was performed.
3.PCI Descriptions
 mLAD: predilation with Lacrosse 2.5/20mm(6 atm)
  -> total occlusion 소견 여전히 보임
  -> suction with Thrombuster catether
    (aspiration material: red clot)
  -> distal edge-dLAD에 diffuse stenotic
    lesion 보임 (TIMI 1, TMPG 0)
 distal edge-dLAD: Resolute integrity 2.75/26mm(9 atm)

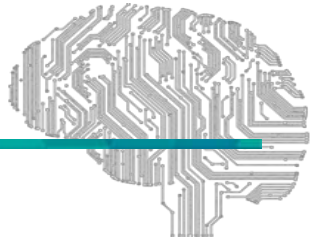
(*spastic component가 동반된 것으로 판단되어 stent diameter를 작은 것으로 삽입,
overlapped with previous stent)
-> slow flow was seen (TIMI 2, TMPG 1)
-> IC adenosine 후 호전
-> Final flow TIMI 3, TMPG 2
-> far distal LAD에 intraluminal filling defect (R/O thrombus) 소견 보임
far dLAD: PTCA with miniTrek 1.5/15mm(12 atm)
OM: PTCA with miniTrek 1.5/15mm(10 atm)
-> final flow TIMI 3, TMPG 1

4.Successful PCI without complications

Extracting information from PCI reports (3)

- Diagnostic CAG/r/n/r/n/r/n 1. Rt. Femoral artery approach/r/n/r/n/r/n 2. JL 6/4F and JR 6/3.5F catheters were used/r/n/r/n/r/n 3. Findings/r/n/r/n/r/n 1) LM: normal/r/n/r/n/r/n 2) LAD: pLAD: tubular eccentric 48%/r/n/r/n/r/n mLAD: No ISR at previous stented site/r/n/r/n/r/n distal edge: tubular concentric 52%/r/n/r/n/r/n 3) LCX: p-dLCX: diffuse irregular up to 68%/r/n/r/n/r/n (small vessel)/r/n/r/n/r/n OM: total occlusion /r/n/r/n/r/n (TIMI 0, small vessel)/r/n/r/n/r/n 4) RCA: p-mRCA: diffuse irregular up to 55%/r/n/r/n/r/n dRCA: diffuse irregular up to 39%/r/n/r/n/r/n/r/n/r/n/r/n *Diagnostic angiography 시행 후 acute onset/r/n/r/n/r/n chest pain 및 ECG상 ST elevation 소견 보여 /r/n/r/n/r/n angiography 다시 시행함 /r/n/r/n/r/n -> proximal edge에 total occluion 소견 보임/r/n/r/n/r/n (TIMI 0, TMPG 0)/r/n/r/n/r/n/r/n/r/n/r/n/r/nPCI/r/n/r/n/r/n 1.EBU 7/3.5F guiding catheter was used. /r/n/r/n/r/n 2.Primary PCI was performed./r/n/r/n/r/n 3.PCI Descriptions/r/n/r/n/r/n mLAD: predilation with Lacrosse 2.5/20mm(6 atm)/r/n/r/n/r/n -> total occlusion 소견 여전히 보임 /r/n/r/n/r/n -> suction with Thrombuster catether/r/n/r/n/r/n (aspiration material: red clot)/r/n/r/n/r/n -> distal edge-dLAD에 diffuse stenotic /r/n/r/n/r/n lesion 보임 (TIMI 1, TMPG 0)/r/n/r/n/r/n distal edge-dLAD:/r/n/r/n/r/n Resolute integrity 2.75/26mm(9 atm)/r/n/r/n/r/n (*spastic component가 동반된 것으로 판단되어/r/n/r/n/r/n stent diameter를 작은 것으로 삽입,/r/n/r/n/r/n overlapped with previous stent)/r/n/r/n/r/n -> slow flow was seen (TIMI 2, TMPG 1)/r/n/r/n/r/n -> IC adenosine 후 호전 /r/n/r/n/r/n -> Final flow TIMI 3, TMPG 2/r/n/r/n/r/n -> far distal LAD에 intraluminal filling/r/n/r/n/r/n defect (R/O thrombus) 소견 보임 /r/n/r/n/r/n far dLAD: PTCA with miniTrek 1.5/15mm(12 atm)/r/n/r/n/r/n/r/n/r/n/r/n/r/n OM: PTCA with miniTrek 1.5/15mm(10 atm)/r/n/r/n/r/n -> final flow TIMI 3, TMPG 1/r/n/r/n/r/n 4.Successful PCI without complications/r/n/r/n/r/n

Extracting information from PCI reports (4)



```
library(stringr)
PCI<-str_extract(Text,"PCI [Dd]e.*")
PCI<-gsub("(/r/n *)|(/r/n( */r/n)*)","/r/n ",PCI)
print(PCI)
```

```
## [1] "PCI Descriptions/r/n   mLAD: predilation with Lacrosse 2.5/20mm(6 atm)/r/n
-> total occlusion 소견 여전히 보임/r/n      -> suction with Thrombuster catether/r/n
(aspiration material: red clot)/r/n       -> distal edge-dLAD에 diffuse stenotic /r/n
lesion 보임 (TIMI 1, TMPG 0)/r/n   distal edge-dLAD:/r/n     Resolute integrity
2.75/26mm(9 atm)/r/n    (*spastic component가 동반된 것으로 판단되어/r/n     stent
diameter를 작은 것으로 삽입,/r/n      overlapped with previous stent)/r/n     -> slow
flow was seen (TIMI 2, TMPG 1)/r/n     -> IC adenosine 후 호전/r/n     -> Final flow
TIMI 3, TMPG 2/r/n     -> far distal LAD에 intraluminal filling/r/n      defect (R/O
thrombus) 소견 보임 /r/n   far dLAD: PTCA with miniTrek 1.5/15mm(12 atm)/r/n
OM: PTCA with miniTrek 1.5/15mm(10 atm)/r/n      -> final flow TIMI 3, TMPG 1/r/n
4.Successful PCI without complications/r/n "
```

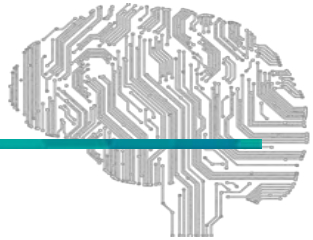
Extracting information from PCI reports (4)

(Step 1) Extracting all words between "new-line (\n)" and ":" in order to find the target vessels of the Percutaneous Coronary Intervention

```
tv.loc<-list()
tv2<-list()
for(x in 1:length(PCI)){
   if(is.na(PCI[[x]])==T){
   tv.loc<-append(tv.loc,list(NULL))
   tv2<-append(tv2,list(NA))
  }
   if(is.na(PCI[[x]])==F){
   tv<-rbind(c(0,0),str_locate_all(PCI[[x]],

"(LM.{0,5}LAD|LM.{0,5}LCx|Ramus|RI|Dx|LAD|L[Cc][Xx]|RCA|PDA|PLV|LM|OM|D1|D2|PLB|[Dd]iagonal)(.{0,8}[:;])")[[1]])
   if(nrow(tv)!=1){
     tv2<-append(tv2,list(sapply(2:nrow(tv),function(y)substr(PCI[[x]],tv[y,1],tv[y,2]))))
     tv.loc<-
append(tv.loc,str_locate_all(PCI[[x]],"(LM.{0,5}LAD|LM.{0,5}LCx|Ramus|RI|Dx|LAD|L[Cc][Xx]|RCA|PDA|PLV|LM|OM|D1|D
2|PLB|[Dd]iagonal)(.{0,8}[:;])"))
   }
   if(nrow(tv)==1){
     tv2<-append(tv2,NA)
     tv.loc<-append(tv.loc,list(NULL))
   }
  }
}
tv3<-unlist(tv2)[complete.cases(unlist(tv2))]
print(tv3)
```

```
## [1] "LAD:" "LAD:" "LAD:" "OM:"
```

Extracting information from PCI reports (5)



(Step 2) Matching the extracted words with pre-defined vessel categories

```
lesions<-tv3
RCA<-grepl("PDA|PLB|PLV|RCA",lesions)
LM<-grepl("LM",lesions)
LAD<-grepl("LAD|D[12xX]|[Dd]?iagonal",lesions)
LCx<-grepl("OM|RI|OM1|OM2|Ramus|Raus|LCX|LCx|Lcx",lesions)


lesions[RCA]<-"RCA"
lesions[LCx&!LM]<-"LCx"
lesions[LAD&!LM]<-"LAD"
lesions[LAD&LM]<-"LM-LAD"
lesions[LCx&LM]<-"LM-LCx"
lesions[LM&!LCx&!LAD]<-"LM"
print(lesions)
```

```
## [1] "LAD" "LAD" "LAD" "LCx"
```

Extracting information from PCI reports (6)

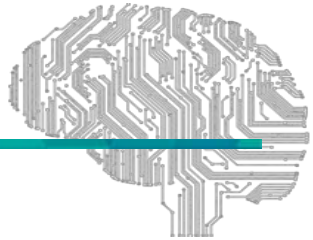
(Step 3-1) Extracting all words between the detected vessel names or between the detected vessel name and the end of report

```
str.loc<-tv.loc
good<-lapply(str.loc,function(x)length(x)>0)
str.loc[unlist(good)]<-lapply(str.loc[unlist(good)],invert_match)
str.ext<-list()
for(i in 1:length(str.loc)){
  if(length(str.loc[[i]])==0){
    str.ext<-append(str.ext,list(NULL))
  }
  if(length(str.loc[[i]])!=0){
    map<-str.loc[[i]]
    map[nrow(map),2]<-10000
    str.ext<-append(str.ext,list(sapply(
  }
}
lstring.N<-str.ext

lengths.N<-sapply(str.ext,length)

strings<-unlist(lstring.N)
print(strings)
```

[1] "PCI Descriptions/r/n mLAD: **predilation with Lacrosse 2.5/20mm(6 atm)/r/n -> total occlusion 소견 여전히 보임/r/n -> suction with Thrombuster catether/r/n (aspiration material: red clot)/r/n -> distal edge-d**LAD에 diffuse stenotic /r/n lesion 보임 (TIMI 1, TMPG 0)/r/n distal edge-dLAD:/r/n Resolute integrity 2.75/26mm(9 atm)/r/n (*spastic component가 동반된 것으로 판단되어/r/n stent diameter를 작은 것으로 삽입,/r/n overlapped with previous stent)/r/n -> slow flow was seen (TIMI 2, TMPG 1)/r/n -> IC adenosine 후 호전/r/n -> Final flow TIMI 3, TMPG 2/r/n -> far distal LAD에 intraluminal filling/r/n defect (R/O thrombus) 소견 보임 /r/n far dLAD: PTCA with miniTrek 1.5/15mm(12 atm)/r/n OM: PTCA with miniTrek 1.5/15mm(10 atm)/r/n -> final flow TIMI 3, TMPG 1/r/n 4.Successful PCI without complications/r/n "

[1] : " predilation with Lacrosse 2.5/20mm(6 atm)/r/n -> total occlusion 소견 여전히 보임/r/n -> suction with Thrombuster catether/r/n (aspiration material: red clot)/r/n -> distal edge-dLAD에 diffuse stenotic /r/n lesion 보임 (TIMI 1, TMPG 0)/r/n distal edge-d"

Extracting information from PCI reports (7)



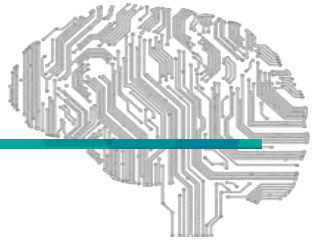
(Step 3-2) The terms that come before "mm"

```
unlist(str_extract_all(PCI,"([[:alnum:]]* ){3}[[:digit:]]\\.[[:digit:]]{0,2}/[[:digit:]]{2}mm"))
```

[1] "PCI Descriptions/r/n mLAD: **predilation with Lacrosse 2.5/20mm**(6 atm)/r/n -> total occlusion 소견 여전히 보임/r/n -> suction with Thrombuster catether/r/n (aspiration material: red clot)/r/n -> distal edge-dLAD에 diffuse stenotic /r/n lesion 보임 (TIMI 1, TMPG 0)/r/n distal edge-dLAD:/r/n Resolute integrity 2.75/26mm(9 atm)/r/n (*spastic component가 동반된 것으로 판단되어/r/n stent diameter를 작은 것으로 삽입,/r/n overlapped with previous stent)/r/n -> slow flow was seen (TIMI 2, TMPG 1)/r/n -> IC adenosine 후 호전/r/n -> Final flow TIMI 3, TMPG 2/r/n -> far distal LAD에 intraluminal filling/r/n defect (R/O thrombus) 소견 보임 /r/n far dLAD: PTCA with miniTrek 1.5/15mm(12 atm)/r/n OM: PTCA with miniTrek 1.5/15mm(10 atm)/r/n -> final flow TIMI 3, TMPG 1/r/n 4.Successful PCI without complications/r/n "

```
## [1] "predilation with Lacrosse 2.5/20mm"
## [2] " Resolute integrity 2.75/26mm"
## [3] "PTCA with miniTrek 1.5/15mm"
## [4] "PTCA with miniTrek 1.5/15mm"
```

Extracting information from PCI reports (7)

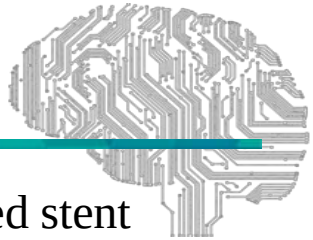


(Step 4) Matching the extracted words with pre-defined stent names

```
coroflex.isar<-"[Cc]oroflex [Ii][Ss][Aa][Rr]"
desyne<-"[Dd]esyne"
osiro<-"[Oo]siro|[Oo]rsiro"
vision<-"[Vv]ision"
zeta<-"[Zz]eta"
coroflex.blue<-"[Cc]oro[Ff]l?ex ?[Bb]lue|Cofoflex Blue|Corofelx blue"
driver<-"[Dd]river"
genoss<-"[Gg]enoss|GENOSS"
resolute.integrity<-"[Rr]esolute [Ii]ntegrity|[Rr]\\. ?[Ii]ntegrity|[Rr]esolutei
[Ii]ntegrity|[Rr]esolute ?\\.[Ii]ntegrity|[Rr]esolute [^iI]|[Ii]ntegrity|[Rr]esolute intergrity|[Rr]esolute
[Ii]|Resolute intetrity|Resolute integriyt“
. . . . . .
```

```
##   resolute.integrity biomatrix coroflex.please xience cypher nobori
## 1                  0         0                0      0      0      0
## 2                  1         0                0      0      0      0
## 3                  0         0                0      0      0      0
```

Extracting information from PCI reports (8)

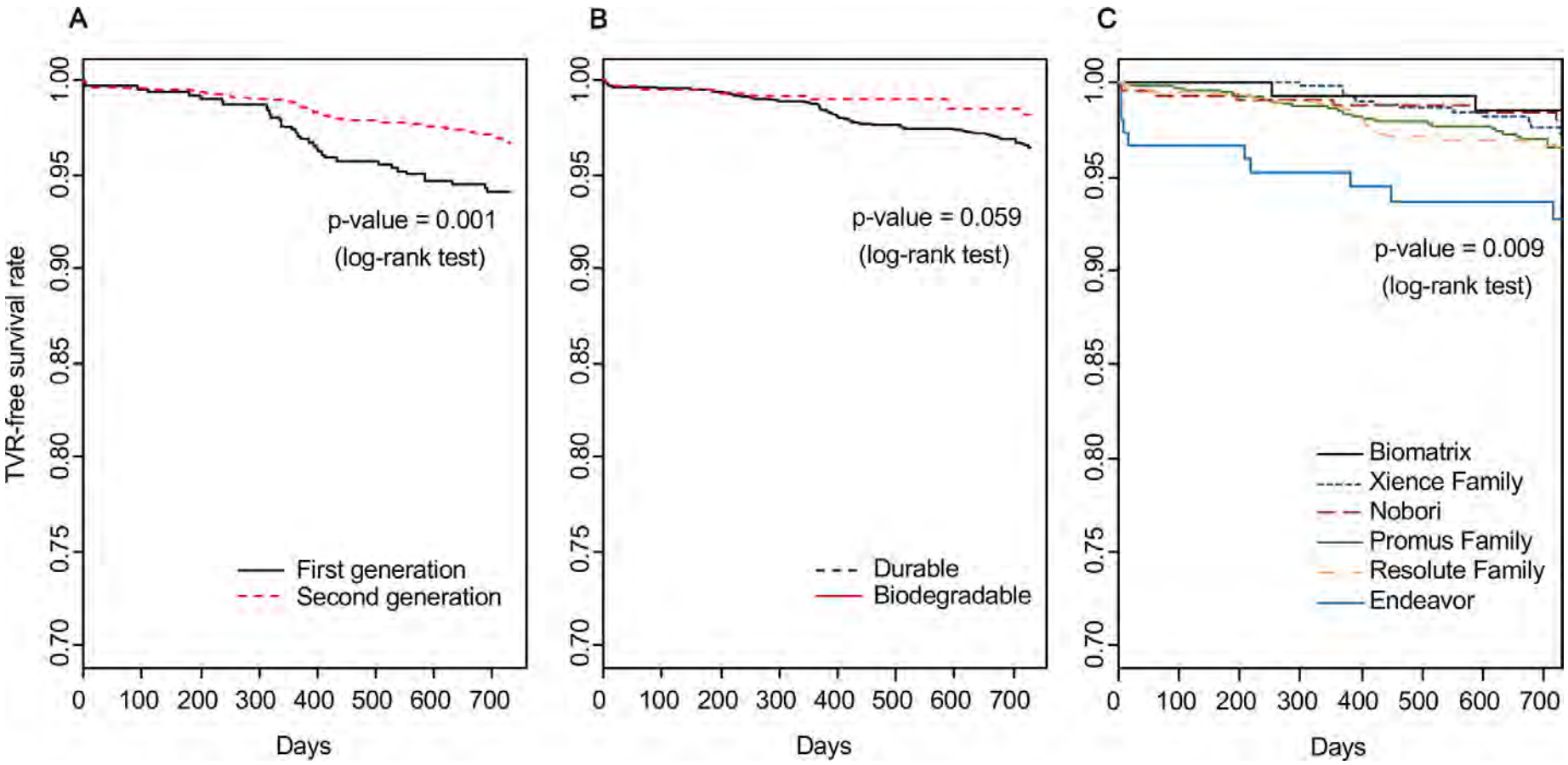


(Step 5) Extracting the diameter and length information following the detected stent names (two numbers followed by "mm", which are separated by "/").

```
stent.info.l<-NULL
for(i in 1:length(strings)){
  new.l<-str_extract_all(strings[i],paste0("(",names,")","
*?(/r/n){0,3}","(.{0,8}[[:digit:]]{0,1}\\.[[:digit:]]{1,2}.{1,2}[[:digit:]]{1,2})",
                                    "(@(.{0,8}[[:digit:]]{0,1}\\.[[:digit:]]{1,2}.{1,2}[[:digit:]]{1,2}))?"))
  stent.info.l<-c(stent.info.l,list(new.l))
}

stent.info<-lapply(stent.info.l,unlist)
stent.info2<-lapply(stent.info,function(x)str_extract(x,"[[:digit:]].*"))
stent.info2[sapply(stent.info2,length)==0]<-0
s.diameter<-unlist(lapply(stent.info2,function(x)str_extract(x,"^[[:digit:]]\\.[[:digit:]]{1,2}")))
s.length<-unlist(lapply(stent.info2,function(x)str_extract(x,"[[:digit:]]{2}$")))
s.diameter[is.na(s.diameter)==1]<-0
s.length[is.na(s.length)==1]<-0
info<-data.frame(as.numeric(s.diameter),as.numeric(s.length))
colnames(info)<-c("Diameter","Length")
print(info)
```

```
##   Diameter Length
## 1     0.00      0
## 2     2.75     26
## 3     0.00      0
## 4     0.00      0
```

Extracting information from PCI reports (9)



특정 질병 유무 및 발생 시기 정의

Observational Study

Medicine

OPEN

Statins and risk for new-onset diabetes mellitus

A real-world cohort study using a clinical research database

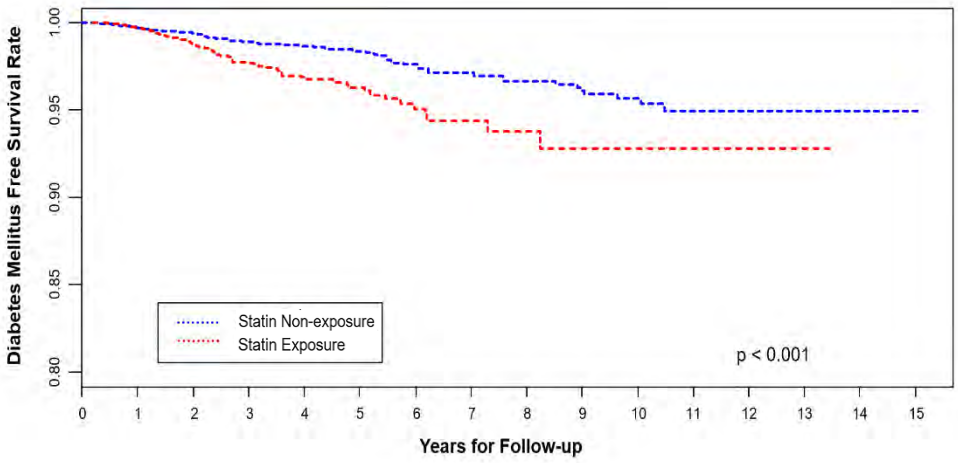
Dukyong Yoon, MD, PhD[a], Seung Soo Sheen, MD[b], Sukhyang Lee, PharmD, PhD[c], Yong Jun Choi, MD[d], Rae Woong Park, MD, PhD[a,*], Hong-Seok Lim, MD, PhD[e,*]

Abstract

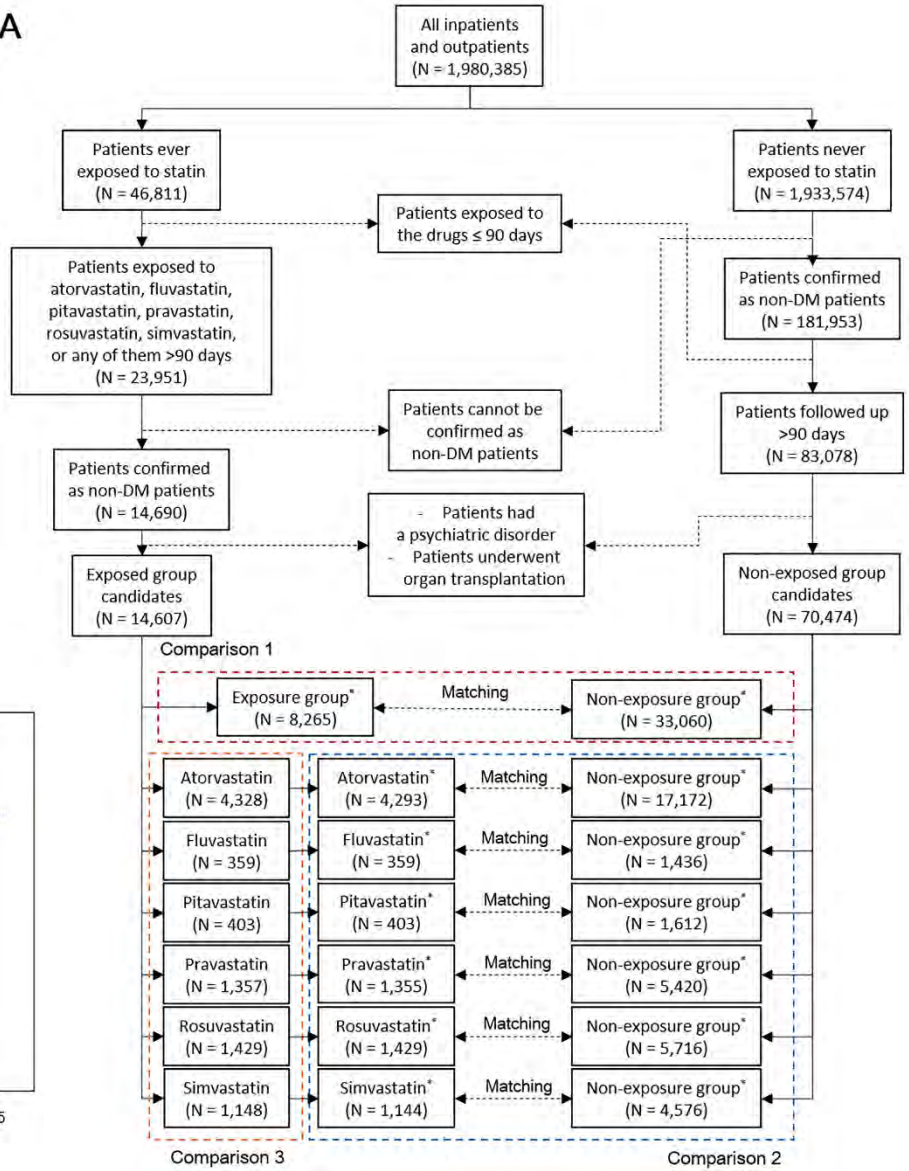
Although concern regarding the increased risk for new-onset diabetes mellitus (NODM) after statin treatment has been raised, there has been a lack of evidence in real-world clinical practice, particularly in East Asians. We investigated whether statin use is associated with risk for NODM in Koreans. We conducted a retrospective cohort study using the clinical research database from electronic health records. The study cohort consisted of 8265 statin-exposed and 33,060 matched nonexposed patients between January 1996 and August 2013. Matching at a 1:4 ratio was performed using a propensity score based on age, gender, baseline glucose levels (mg/dL), and hypertension. The comparative risks for NODM with various statins (atorvastatin, fluvastatin, pitavastatin, pravastatin, rosuvastatin, and simvastatin) were estimated by both statin exposure versus matched nonexposed and within-class comparisons. The incidence of NODM among the statin-exposed group (6.000 per 1000 patient-years [PY]) was higher than that of the nonexposed group (3.244 per 1000 PY). The hazard ratio (HR) of NODM after statin exposure was 1.872 (95% confidence interval [CI], 1.432–2.445). Male gender (HR, 1.944; 95% CI, 1.497–2.523), baseline glucose per mg/dL (HR, 1.014; 95% CI, 1.013–1.016), hypertension (HR, 2.232; 95% CI, 1.515–3.288), and thiazide use (HR, 1.337; 95% CI, 1.081–1.655) showed an increased risk for NODM, while angiotensin-converting enzyme inhibitor or angiotensin II receptor blocker showed a decreased risk (HR, 0.774; 95% CI, 0.668–0.897). Atorvastatin-exposed patients showed a higher risk for NODM than their matched nonexposed counterparts (HR, 1.939; 95% CI, 1.278–2.943). However, the risk for NODM was not significantly different among statins in within-class comparisons. In conclusion, an increased risk for NODM was observed among statin users in a practical healthcare setting in Korea.

Abbreviations: ACEi = angiotensin-converting enzyme inhibitor, ARB = angiotensin II receptor blocker, BMI = body mass index, CCI = Charlson comorbidity index, CI = confidence interval, DM = diabetes mellitus, EHR = electronic health record, HR = hazard ratio, ICD-10 = the International Classification of Diseases 10th Revision, NODM = new-onset diabetes mellitus.

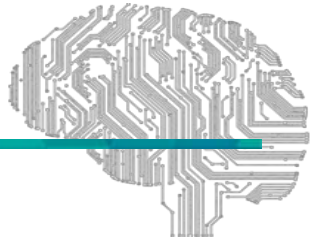
Keywords: East Asians, electronic health record, new-onset diabetes mellitus, statin



A

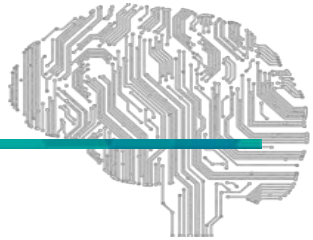


Detecting new-onset DM patients (1)



- Inclusion
  - Patients who visited the subject hospital more than once, regardless of outpatient visits or hospitalization
  - Patients who had more than 1 fasting glucose measurement before the start of observation.
- Exclusion
  - Patients with abnormal random glucose levels (≥200mg/dL)
  - Patients with abnormal fasting glucose levels (≥126mg/dL)
  - Patients with abnormal hemoglobin A1c (HbA1c) results (≥6.5%)
  - Patients with ICD-10 diagnosis codes related to diabetes (E10–E14)
  - Patients who had received a prescription for diabetes medication(s) (acarbose, gemigliptin, glibenclamide, gliclazide, glimepiride, linagliptin, metformin, mitiglinide, nateglinide, pioglitazone, repaglinide, saxagliptin, sitagliptin, vildagliptin, and voglibose) including insulin before the start of observations.

Detecting new-onset DM patients (2)



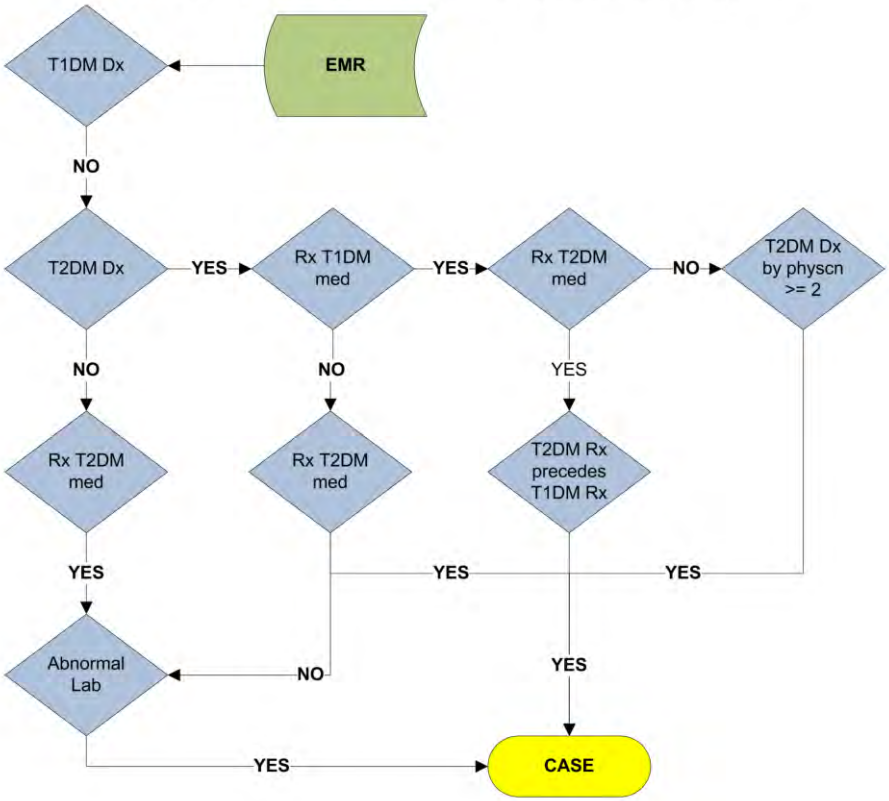
- Outcome (NODM)
  - Excludes patients who have T1DM diagnosis codes (E10).
  - If patients have T2DM diagnosis codes (E11)
    - the algorithm checks whether their medication history met the T2DM treatment standard.
    -
  - In cases without T2DM diagnosis codes
    - patients who received medication(s) for T2DM
    - and had abnormal glucose or HbA1c results were identified as T2DM patients
  - The earliest time at which patients met the algorithm was considered the time the event occurred.

Detecting new-onset DM patients (3)

PheKB a knowledgebase for discovering phenotypes from electronic medical records

https://phekb.org/

Figure 1: Algorithm for identifying T2DM cases in the EMR.



1 **if** T1DM-DX-DT-CNT(pt) == 0 ⇐ Algorithm 2
AND T2DM-DX-DT-CNT(pt) > 0 ⇐ Algorithm 3
AND T2DM-RX-DT(pt) ≠ NULL ⇐ Algorithm 4
AND T1DM-RX-DT(pt) ≠ NULL ⇐ Algorithm 5
AND T2DM-RX-DT(pt) < T1DM-RX-DT(pt)
status = CASE

2 **elseif** T1DM-DX-DT-CNT(pt) == 0
AND T2DM-DX-DT-CNT(pt) > 0
AND T1DM-RX-DT(pt) == NULL
AND T2DM-RX-DT(pt) ≠ NULL
status = CASE

3 **elseif** T1DM-DX-DT-CNT(pt) == 0
AND T2DM-DX-DT-CNT(pt) > 0
AND T1DM-RX-DT(pt) == NULL
AND T2DM-RX-DT(pt) == NULL
AND ABNORMAL-LAB(pt) == TRUE ⇐ Algorithm 6
status = CASE

4 **elseif** T1DM-DX-DT-CNT(pt) == 0
AND T2DM-DX-DT-CNT(pt) == 0
AND T2DM-RX-DT(pt) ≠ NULL
AND ABNORMAL-LAB(pt) == TRUE
status = CASE

5 **elseif** T1DM-DX-DT-CNT(pt) == 0
AND T2DM-DX-DT-CNT(pt) > 0
AND T1DM-RX-DT(pt) ≠ NULL
AND T2DM-RX-DT(pt) == NULL
AND T2DM-PHYSCN-DX-DT-CNT(pt) ≥ 2 ⇐ Algorithm 7
status = CASE

return *status*

Detecting new-onset DM patients (4)

. . . . . .

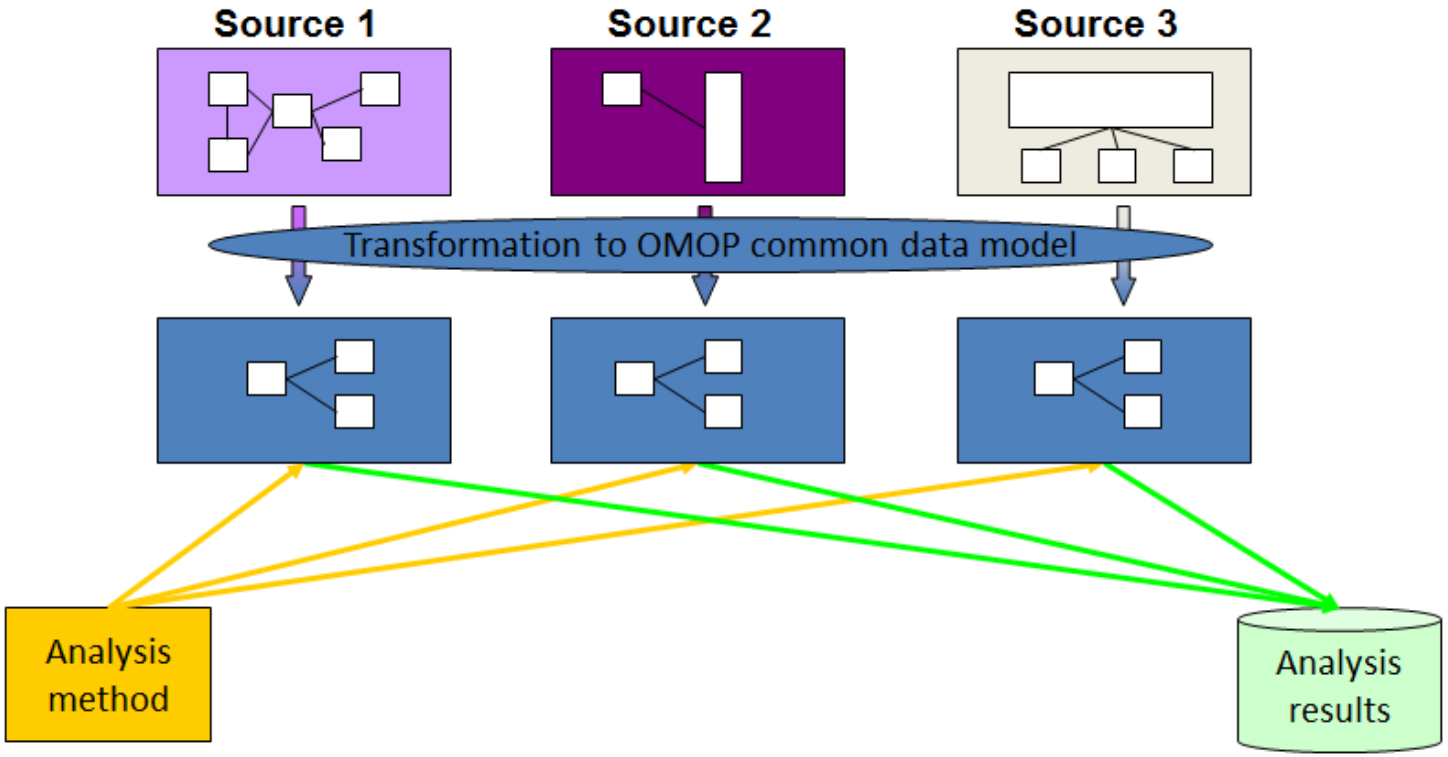
1,940 lines

Data integration

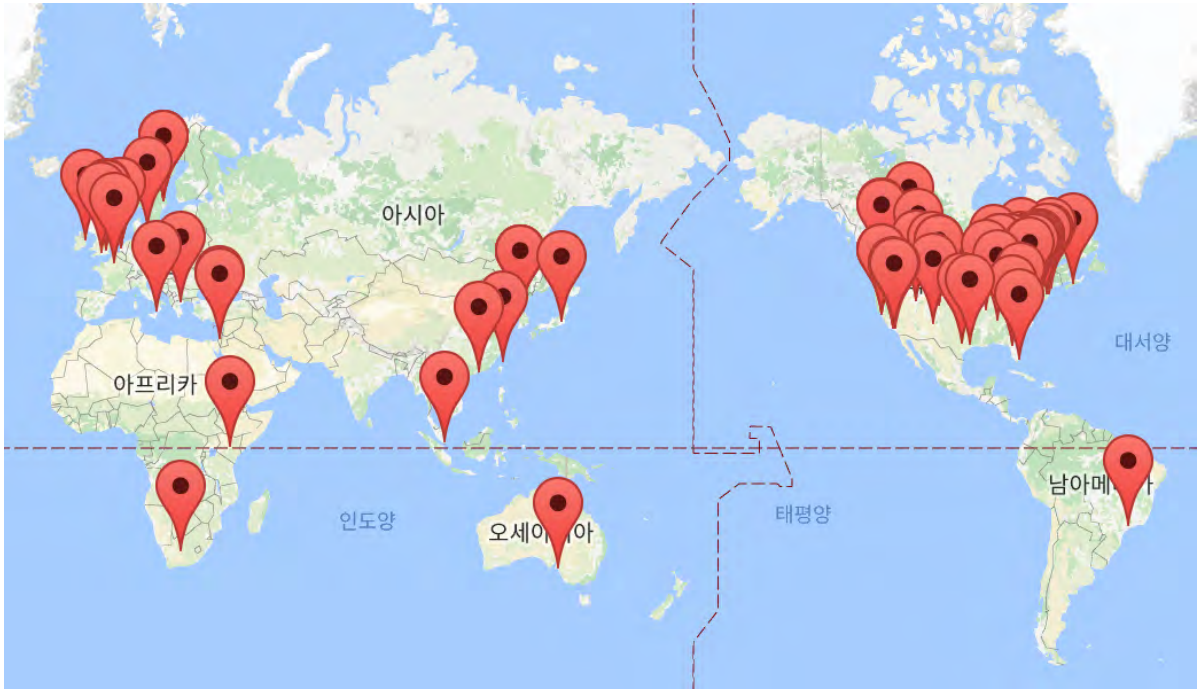
- Distributed Research Network

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

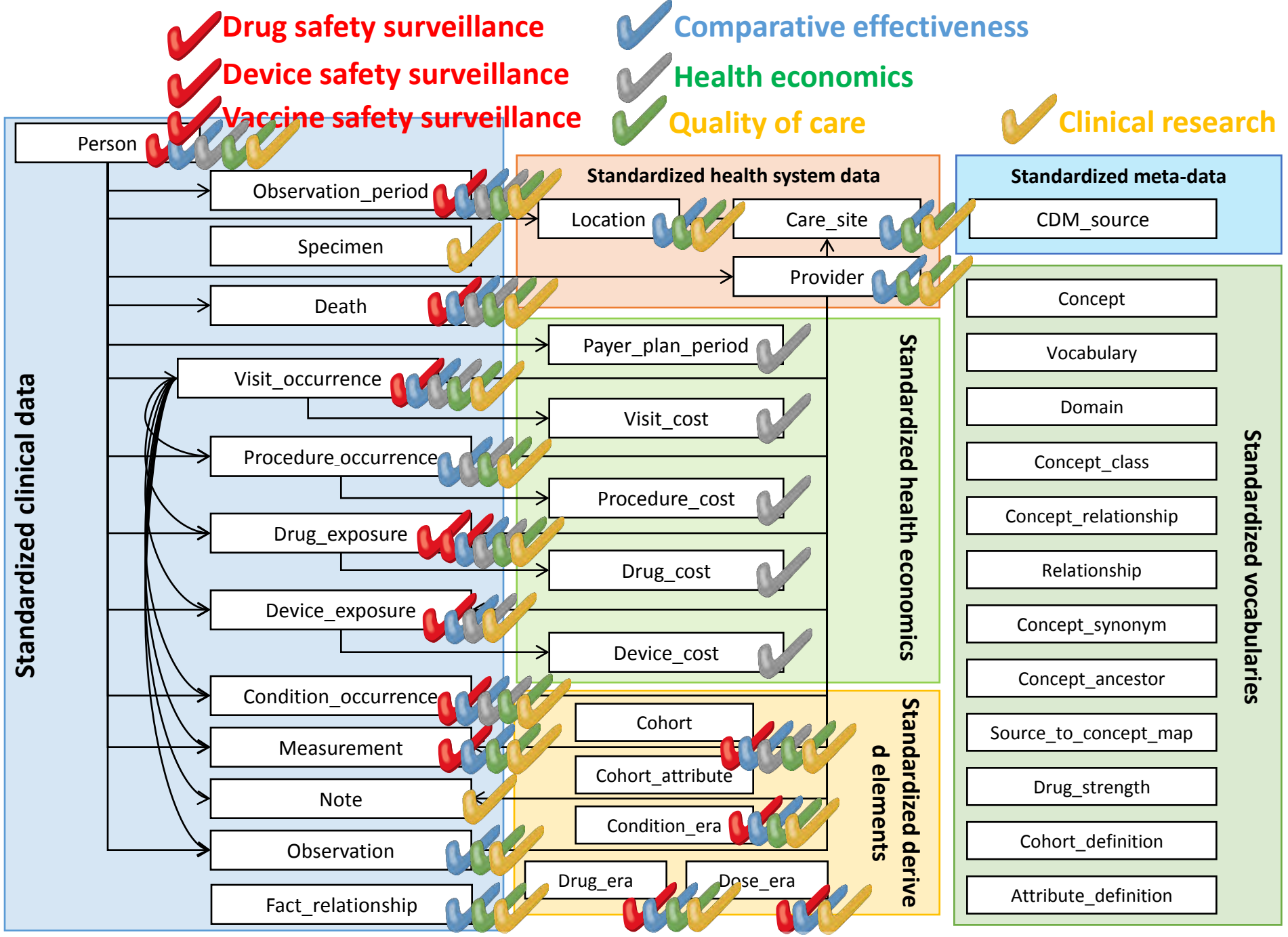
http://www.ohdsi.org/



OHDSI collaborators

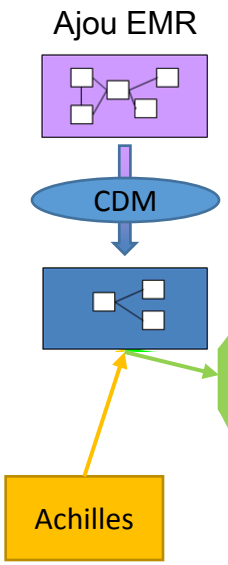


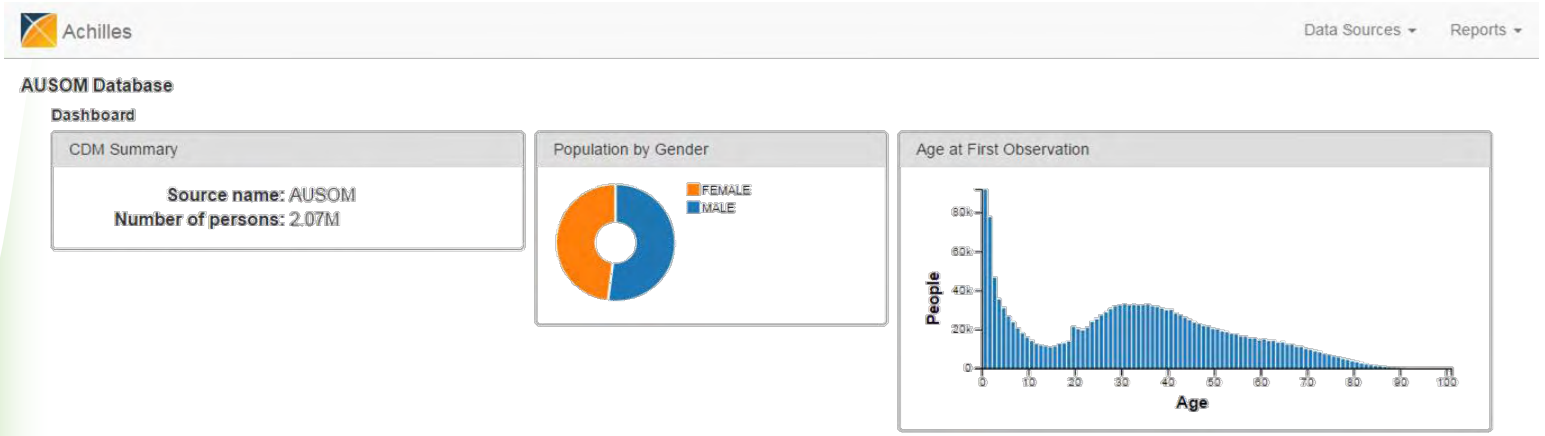
Ajou University School of Medicine
Albert Einstein College of Medicine
Ben-Gurion University
Children's Hospital of Philadelphia
Columbia University
Department of Veterans Affairs
Drug Safety Research Unit
Ephir, Inc.
Erasmus University Medical Center
Evidera
IMS Health
Indiana University School of Medicine
INSERM
Intermountain Healthcare
Janssen Research and Development
Johnson & Johnson
LTS Computing LLC
Maine Medical Center Research Institute
Massachusetts Institute of Technology
Medical University of South Carolina
Merck Research Laboratories
Montana State University
National Institutes of Health
New York University
Oracle Health Sciences
Outcomes Insights
Paxata
Regenstrief Institute
Rutgers University
Siemens Health Services
Stanford University School of Medicine
Taipei Medical University
The Ohio State University College of Medicine
University of California, Los Angeles
University of Colorado Denver
University of Florida
University of Hong Kong
University of Messina
University of New Mexico
University of Pittsburgh
University of Pittsburgh School of Medicine
University of South Australia
University of Southern California
University of Texas at Austin
University of Utah
University of Utah School of Medicine
University of Washington
Vanderbilt University
WHO Uppsala Monitoring Centre



Slide source: courtesy of Dr. Patrick Ryan's Presentation: 'Welcome to the journey: OHDSI Symposium 2015'

http://ami.ajou.ac.kr:8080/

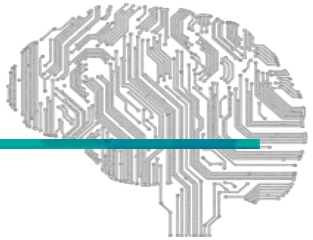




Diagnosis

Laboratory test results

Github.com/OHDSI/



Observational Health Data Sciences and Informatics

http://ohdsi.org

Repositories People 5

Filters ▾ Find a repository

Achilles

R ★28 ⑂32

Automated Characterization of Health Information at Large-scale Longitudinal Evidence Systems (ACHILLES) - descriptive statistics about a OMOP CDM v4 database

Updated 4 hours ago

Iris

R ★2 ⑂2

IRIS provides a high-level descriptive summary of a population within a OMOP CDM-compliant database

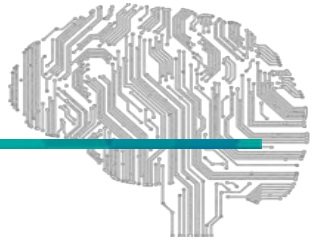
Updated 4 hours ago

Atlas

JavaScript ★2 ⑂8

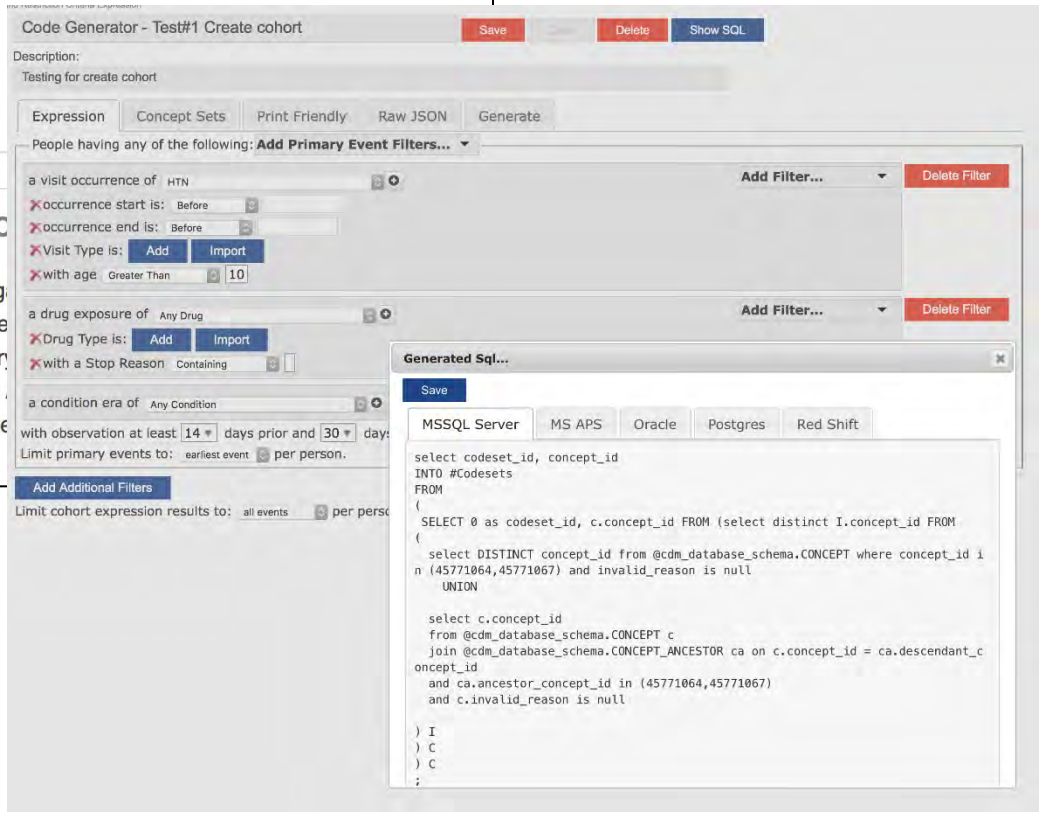
People 5 >

Github.com/OHDSI/



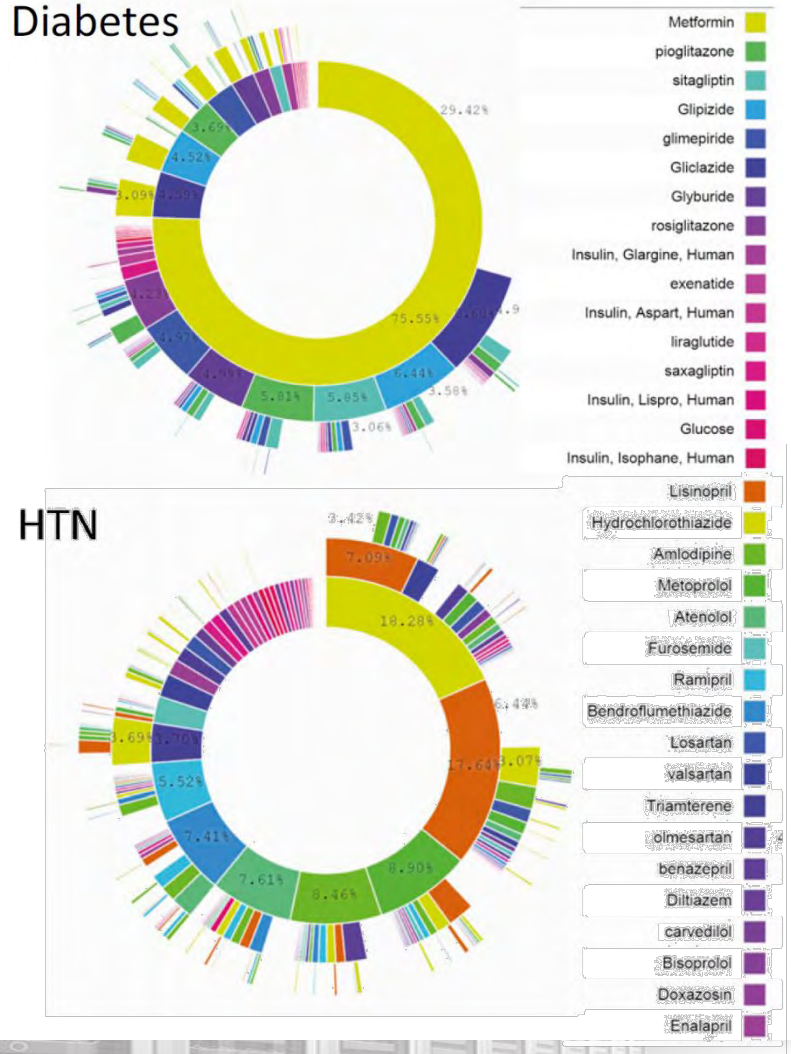
- 69 Repositories (2016-05-13 기준)
  - 17 suspended/stopped or invalid projects
  - 52 valid repositories
    - 9 web apps, 1 server tier project
    - 2 unified web app projects
    - 23 methods and libraries
    - 7 tools for ETL process
    - 10 major repositories

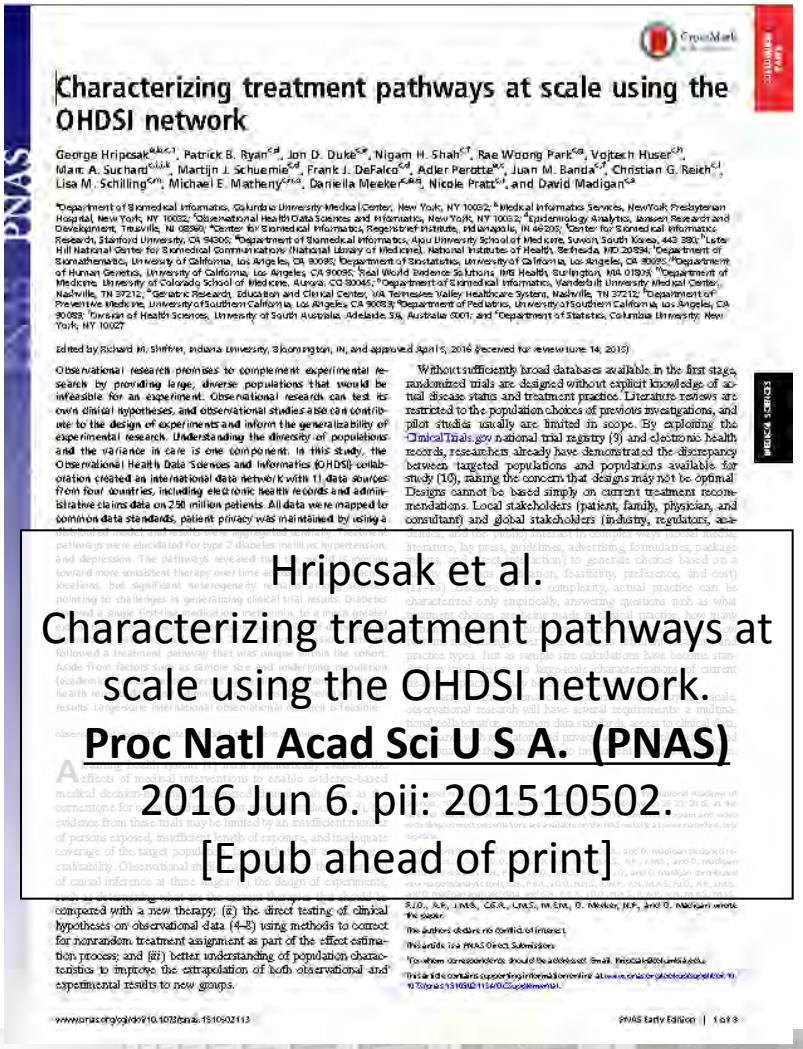
http://ami.ajou.ac.kr:8080/AURORA/



국제 공동 연구 사례

(NIDDM) Treatment Pathway / 7 institutes (2 hundred million)





Characterizing treatment pathways at scale using the OHDSI network

George Hripcsak[a,b,c,1], Patrick B. Ryan[c,d], Jon D. Duke[c,e], Nigam H. Shah[c,f], Rae Woong Park[c,g], Vojtech Huser[c,h], Marc A. Suchard[c,i,j,k], Martijn J. Schuemie[c,d], Frank J. DeFalco[c,d], Adler Perotte[c,l], Juan M. Banda[c,f], Christian G. Reich[c], Lisa M. Schilling[c,m], Michael E. Matheny[c,n,o], Daniella Meeker[c,p,q], Nicole Pratt[c,r], and David Madigan[c,s]

[a]Department of Biomedical Informatics, Columbia University Medical Center, New York, NY 10032; [b]Medical Informatics Services, NewYork Presbyterian Hospital, New York, NY 10032; [c]Observational Health Data Sciences and Informatics, New York, NY 10032; [d]Epidemiology Analytics, Janssen Research and Development, Titusville, NJ 08560; [e]Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, IN 46202; [f]Center for Biomedical Informatics Research, Stanford University, CA 94305; [g]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, South Korea, 443 380; [h]Lister Hill National Center for Biomedical Communications (National Library of Medicine), National Institutes of Health), Bethesda, MD 20894; [i]Department of Biomathematics, University of California, Los Angeles, CA 90095; [j]Department of Biostatistics, University of California, Los Angeles, CA 90095; [k]Department of Human Genetics, University of California, Los Angeles, CA 90095; [l]Real World Evidence Solutions, IMS Health, Burlington, MA 01803; [m]Department of Medicine, University of Colorado School of Medicine, Aurora, CO 80045; [n]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN 37212; [o]Geriatric Research, Education and Clinical Center, VA Tennessee Valley Healthcare System, Nashville, TN 37212; [p]Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90033; [q]Department of Pediatrics, University of Southern California, Los Angeles, CA 90033; [r]Division of Health Sciences, University of South Australia, Adelaide SA, Australia 5001; and [s]Department of Statistics, Columbia University, New York, NY 10027

Edited by Richard M. Shiffrin, Indiana University, Bloomington, IN, and approved April 5, 2016 (received for review June 14, 2015)

Observational research promises to complement experimental research by providing large, diverse populations that would be infeasible for an experiment. Observational research can test its own clinical hypotheses, and observational studies also can contribute to the design of experiments and inform the generalizability of experimental research. Understanding the diversity of populations and the variance in care is one component. In this study, the Observational Health Data Sciences and Informatics (OHDSI) collaboration created an international data network with 11 data sources from four countries, including electronic health records and administrative claims data on 250 million patients. All data were mapped to common data standards, patient privacy was maintained by using a

Without sufficiently broad databases available in the first stage, randomized trials are designed without explicit knowledge of actual disease status and treatment practice. Literature reviews are restricted to the population choices of previous investigations, and pilot studies usually are limited in scope. By exploring the ClinicalTrials.gov national trial registry (9) and electronic health records, researchers already have demonstrated the discrepancy between targeted populations and populations available for study (10), raising the concern that designs may not be optimal. Designs cannot be based simply on current treatment recommendations. Local stakeholders (patient, family, physician, and consultant) and global stakeholders (industry, regulators, aca-

compared with a new therapy; (ii) the direct testing of clinical hypotheses on observational data (4–8) using methods to correct for nonrandom treatment assignment as part of the effect estimation process; and (iii) better understanding of population characteristics to improve the extrapolation of both observational and experimental results to new groups.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Hripcsak et al.
Characterizing treatment pathways at scale using the OHDSI network.
Proc Natl Acad Sci U S A. (PNAS)
2016 Jun 6. pii: 201510502.
[Epub ahead of print]

생체 신호 데이터 분석 사례

Background - I

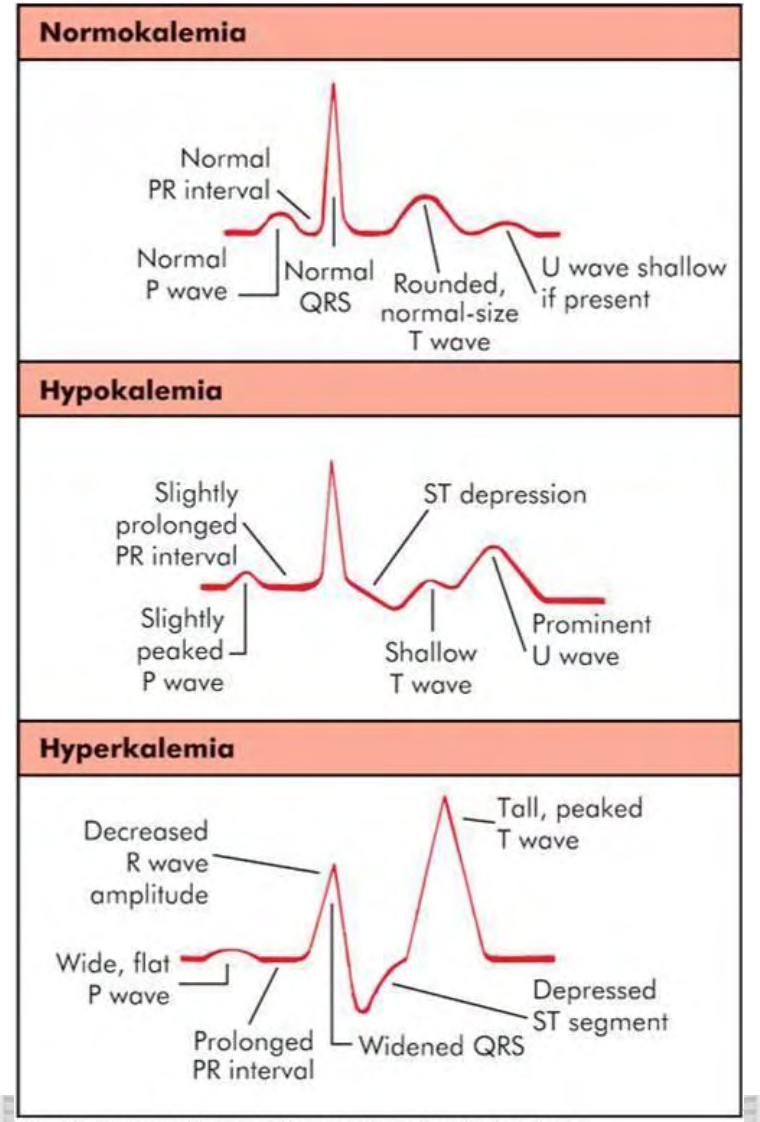
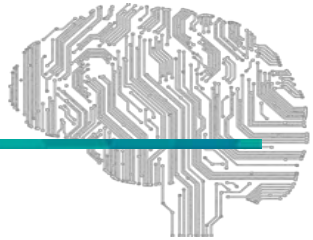


Fig. 4-7. Electrocardiogram Changes with Potassium Imbalance

Copyright © 2008 by Mosby, Inc., an affiliate of Elsevier Inc.

생체 신호 데이터 분석 사례



Background - II

Ann Emerg Med. 1991 Nov;20(11):1229-32.

The ability of physicians to predict hyperkalemia from the ECG.

Wrenn KD[1], Slovis CM, Slovis BS.

Author information

Abstract

STUDY OBJECTIVE: To determine whether physicians blinded to the serum potassium level can predict hyperkalemia (potassium concentration of more than 5.0 mmol/L) from the ECG.

DESIGN: ECGs of patients at high risk for hyperkalemia were interpreted retrospectively by two physicians blinded not only to the specific clinical diagnosis of the patient and to their serum potassium measurement but also to each other's interpretation. The physicians predicted the presence or absence of hyperkalemia as well as the severity of hyperkalemia on a nominal scale (mild, moderate, or severe).

SETTING: The emergency department of a university-affiliated urban county hospital.

PATIENTS: Two hundred twenty consecutive patients admitted to the hospital from the ED with a diagnosis of renal failure or hyperkalemia. Eighty-seven patients had hyperkalemia, and 133 did not.

RESULTS: The sensitivities of the readers for predicting hyperkalemia were .43 and .34, respectively (best positive predictive value, .65). The respective specificities for detecting hyperkalemia were .85 and .86 (best negative predictive value, .69). When only patients with moderate-to-severe hyperkalemia (potassium of more than 6.5 mmol/L) were analyzed, sensitivities were .62 and .55. The readers' ability to predict the severity of hyperkalemia was equally poor.

CONCLUSION: The ECG is not a sensitive method of detecting hyperkalemia, even in high-risk patients. The specificity of the ECG is better for hyperkalemia, but empiric treatment of hyperkalemia based on the ECG alone will lead to mistreatment of at least 15% of patients.

생체 신호 데이터 분석 사례

Background – III-(1)

Novel Bloodless Potassium Determination Using a Signal-Processed Single-Lead ECG

Zachi I. Attia, BSc; Christopher V. DeSimone, MD, PhD; John J. Dillon, MD; Yehu Sapir, BSc; Virend K. Somers, MD, PhD; Jennifer L. Dugan, CRC; Charles J. Bruce, MD; Michael J. Ackerman, MD; Samuel J. Asirvatham, MD; Bryan L. Striemer, BS; Jan Bukartyk, MS; Christopher G. Scott, MS; Kevin E. Bennet, BS, MBA; Dorothy J. Ladewig, BS; Emily J. Gilles, MS; Dan Sadot, PhD; Amir B. Geva, PhD; Paul A. Friedman, MD

Background—Hyper- and hypokalemia are clinically silent, common in patients with renal or cardiac disease, and are life threatening. A noninvasive, unobtrusive, blood-free method for tracking potassium would be an important clinical advance.

Methods and Results—Two groups of hemodialysis patients (development group, n=26; validation group, n=19) underwent high-resolution digital ECG recordings and had 2 to 3 blood tests during dialysis. Using advanced signal processing, we developed a personalized regression model for each patient to noninvasively calculate potassium values during the second and third dialysis sessions using only the processed single-channel ECG. In addition, by analyzing the entire development group's first-visit data, we created a global model for all patients that was validated against subsequent sessions in the development group and in a separate validation group. This global model sought to predict potassium, based on the T wave characteristics, with no blood tests required. For the personalized model, we successfully calculated potassium values with an absolute error of 0.36±0.34 mmol/L (or 10% of the measured blood potassium). For the global model, potassium prediction was also accurate, with an absolute error of 0.44±0.47 mmol/L for the training group (or 11% of the measured blood potassium) and 0.5±0.42 for the validation set (or 12% of the measured blood potassium).

Conclusions—The signal-processed ECG derived from a single lead can be used to calculate potassium values with clinically meaningful resolution using a strategy that requires no blood tests. This enables a cost-effective, noninvasive, unobtrusive strategy for potassium assessment that can be used during remote monitoring. **(*J Am Heart Assoc*. 2016;5:e002746 doi: 10.1161/JAHA.115.002746)**

Key Words: electrophysiology • potassium • waves

생체 신호 데이터 분석 사례

Background – III-(2)

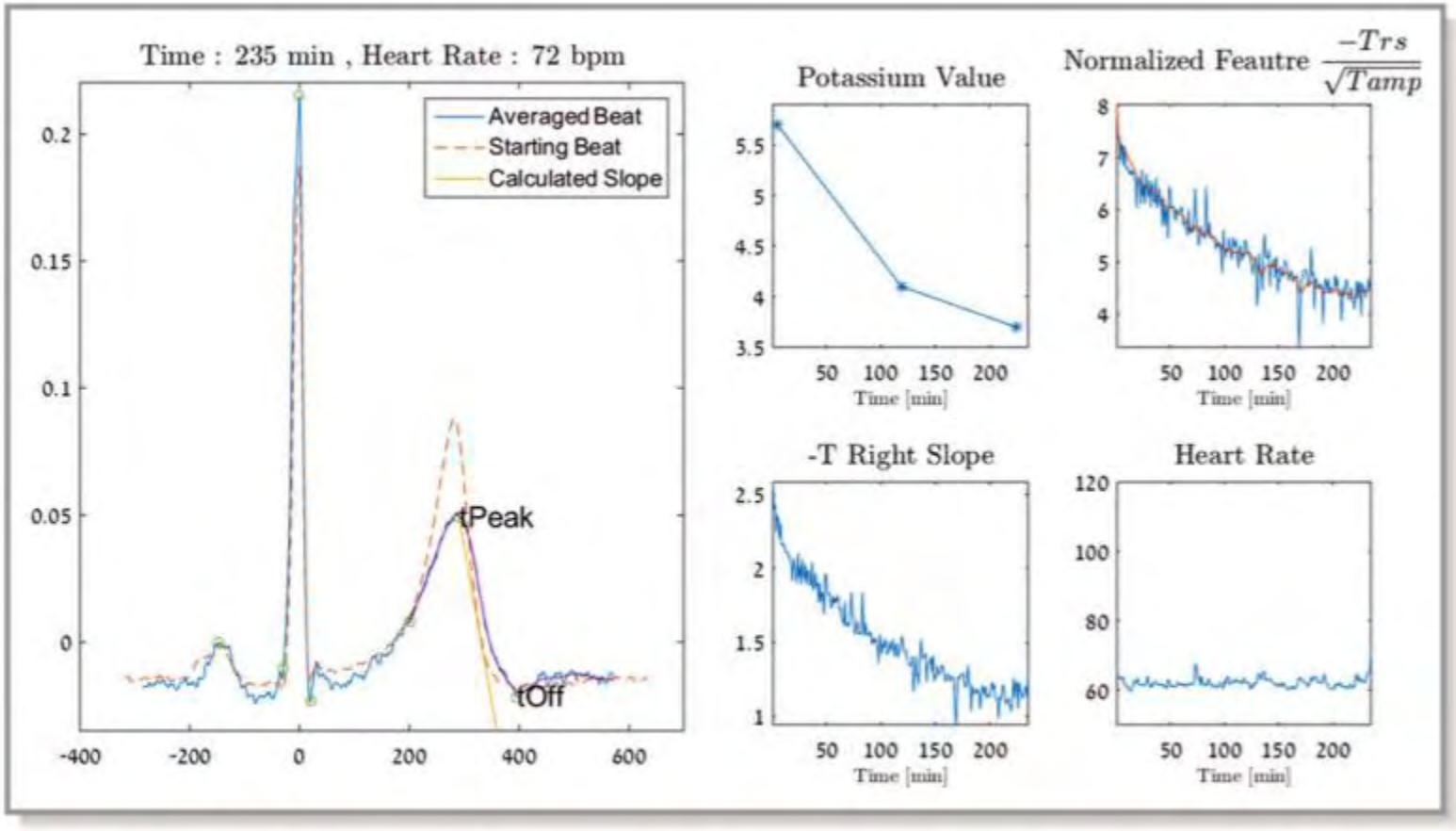
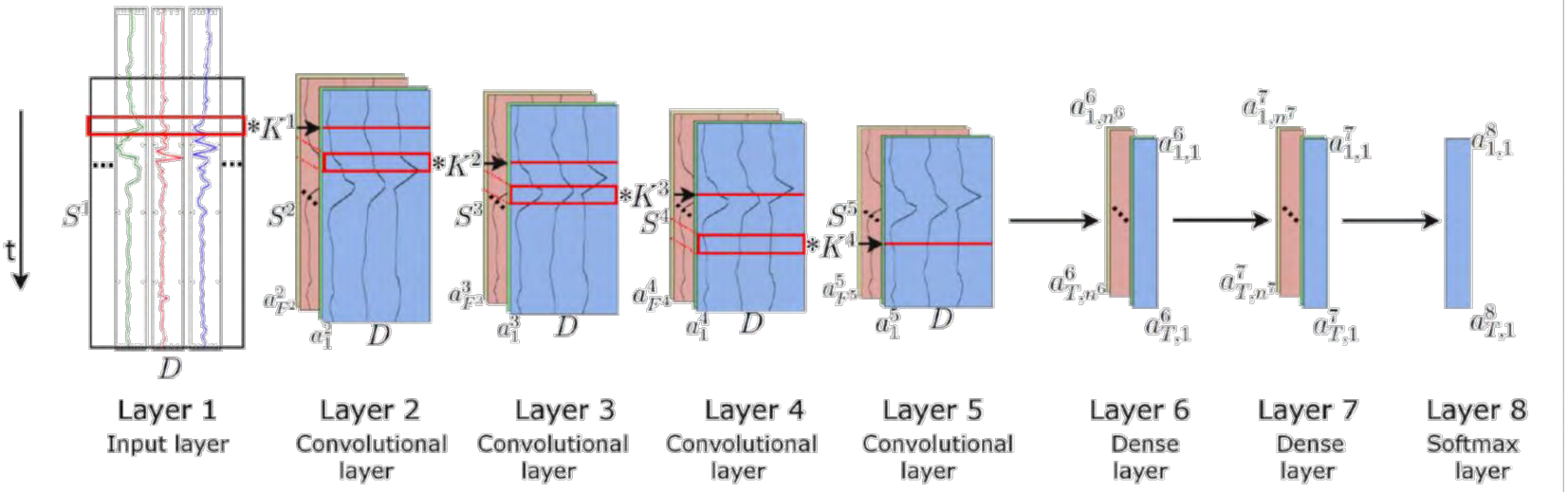
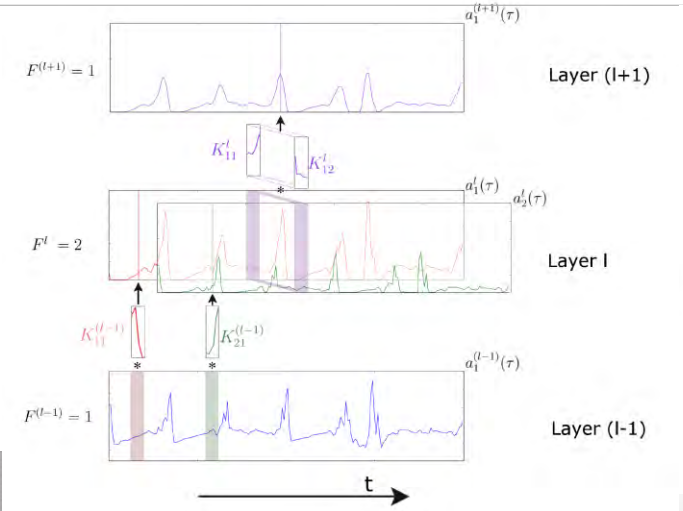


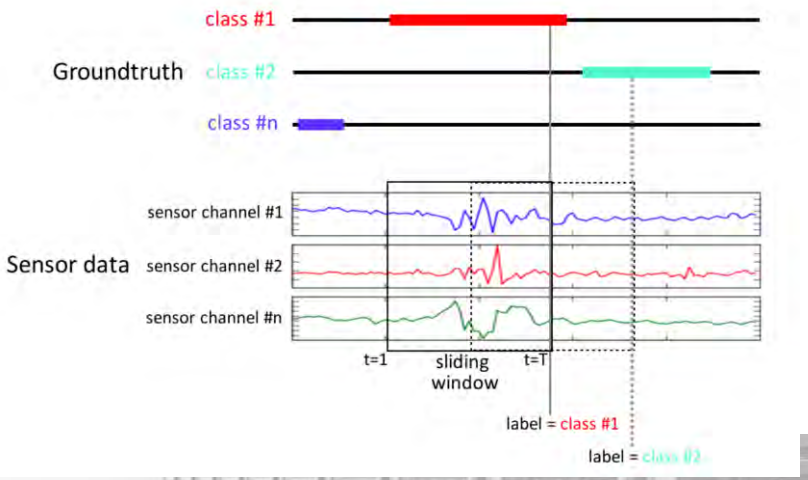
Figure 3. Temporal change of potassium using the temporal progression tool. This image is a still frame taken from Video S1. The left panel shows a representative ECG complex that has been processed, filtered, and displayed. The dashed ECG complex is an initial processed ECG acquired before dialysis commenced.

생체 신호 데이터 분석 사례

- A deep learning framework composed of convolutional and LSTM recurrent layers
- That is capable of automatically learning feature representations and modelling the temporal dependencies between their activation.

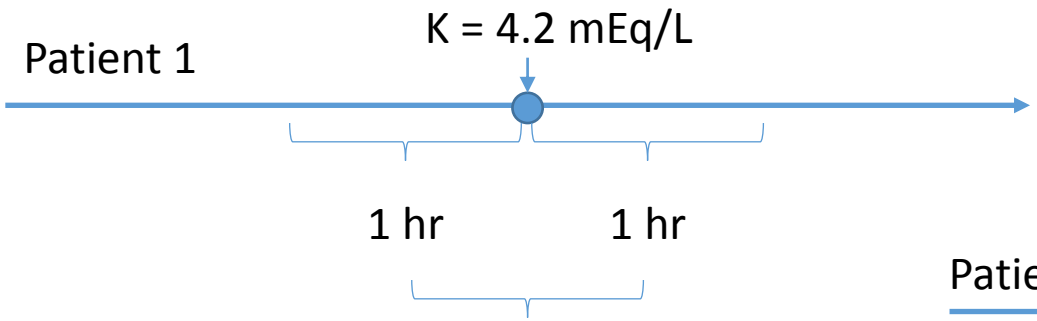




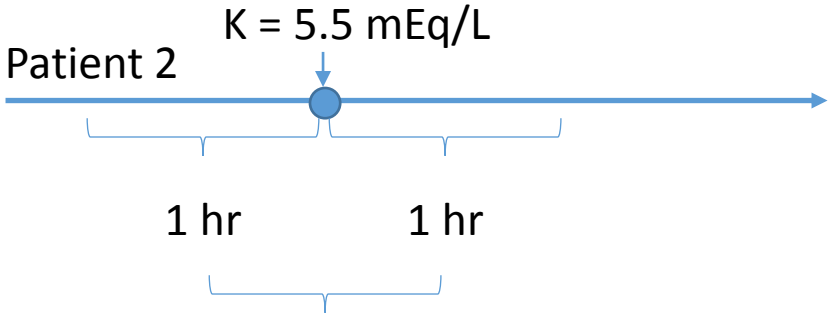


Ref: Deep Convolutional and LSTM Recurrent Neural Networks for MultimodalWearable Activity Recognition. Sensors 2016, 16, 115

생체 신호 데이터 분석 사례



720 * 10s waveform data



720 * 10s waveform data

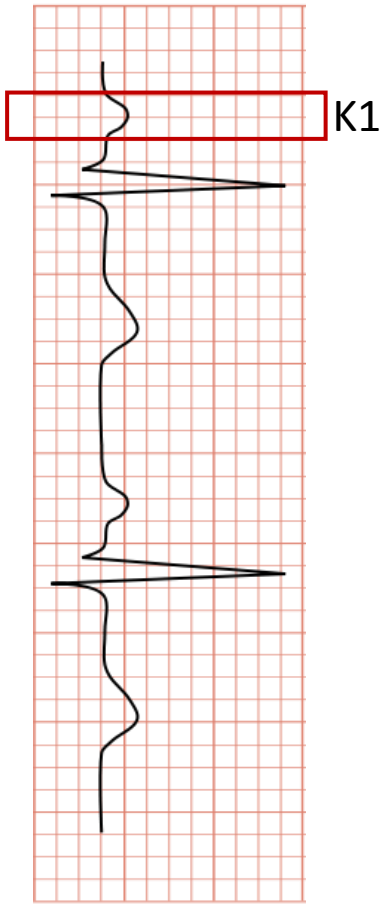
.
.
.
.

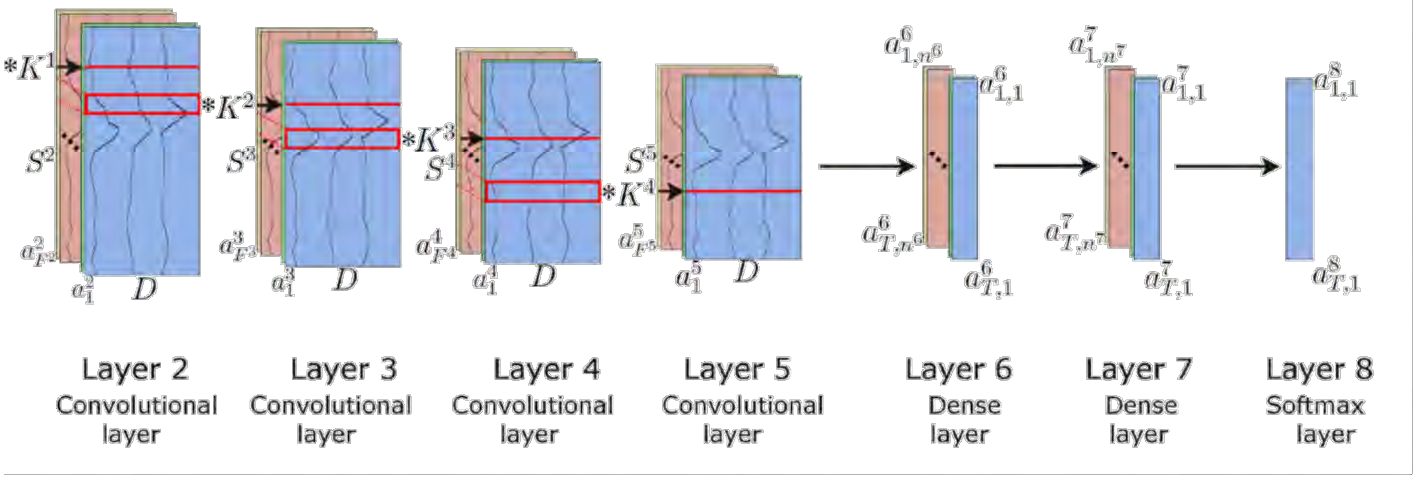
About 96,000 labeled waveform data

Random → Train

Random → Test

생체 신호 데이터 분석 사례



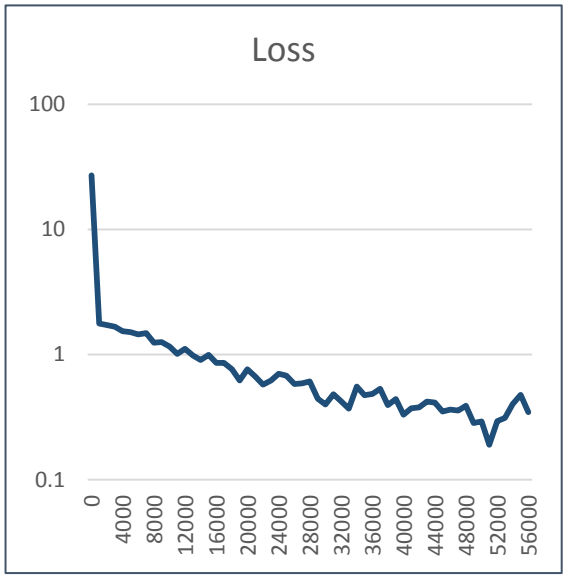


Labeling

- <3.5 : 0
- >=3.5 and <4.0 : 1
- >=4.0 and <4.5 : 2
- >=4.5 and <5.0 : 3
- >=5.0 and <5.5 : 4
- >=5.5 : 5

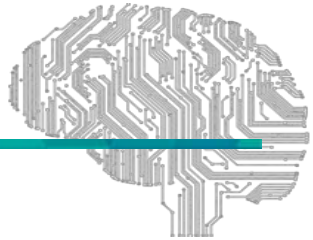
Accuracy?

생체 신호 데이터 분석 사례



Testing Accuracy: 최대 92%

QnA



- 윤덕용, MD, PhD
- www.d-yoon.com
- d.yoon.ajou@gmail.com