

미국 중환자실 오픈데이터베이스 MIMIC-III

출처 : 빅데이터 임상활용연구회 3차 세미나 자료

1. MIMIC-III란?



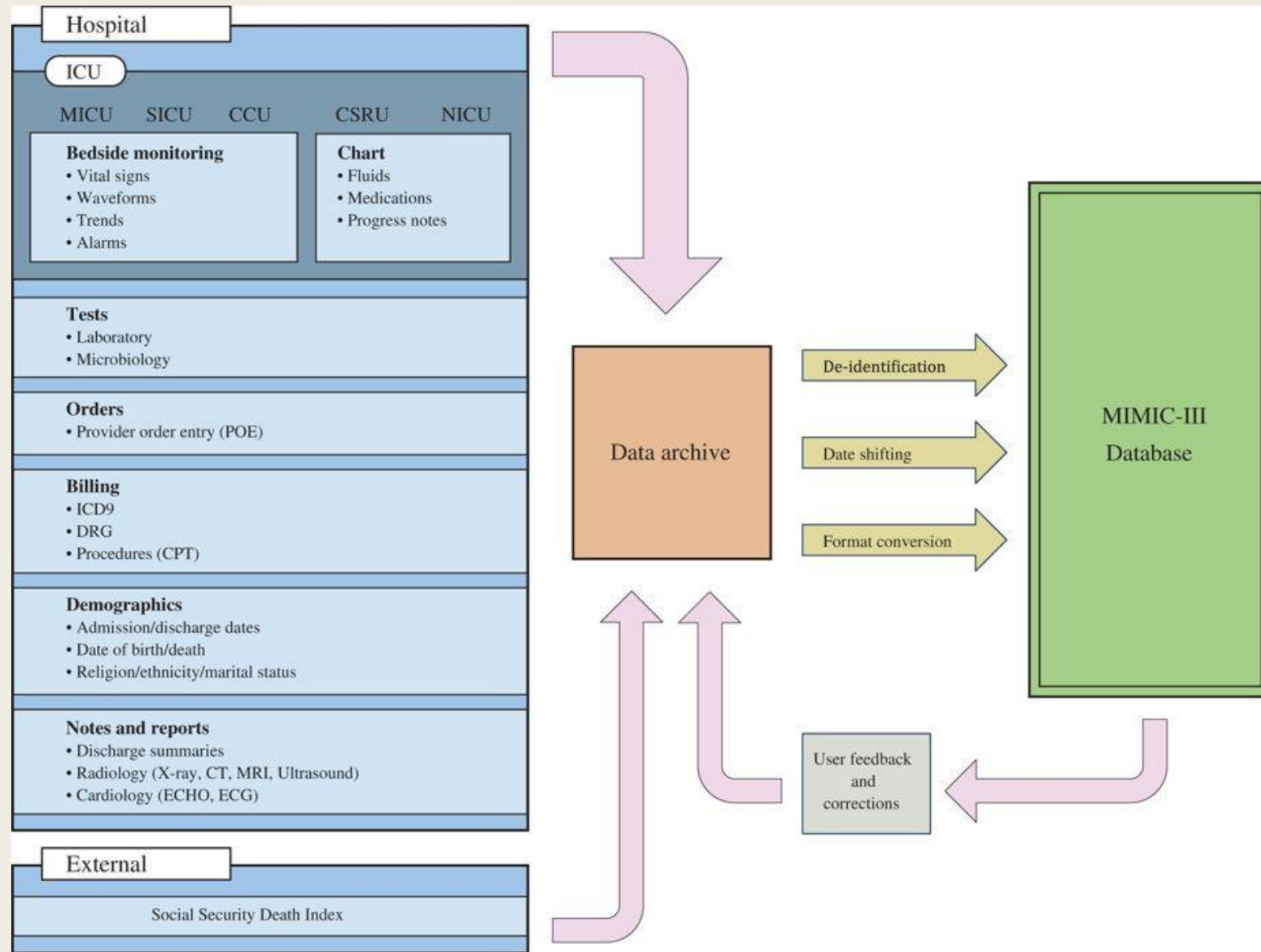
- <https://mimic.physionet.org/>
- MIMIC-III (**M**edical **I**nformation **M**art for **I**ntensive **C**are **III**)
- the Beth Israel Deaconess Medical Center between 2001 and 2012
- 전세계 모든 연구자에게 무료로 공개된 데이터
- 다양하고 방대한 데이터 세트
- 랩수치, 전자의무기록, 생체신호와 같은 중환자 데이터가 높은 해상도로 존재

2. MIMI-III 데이터 들여다보기

- 대상 : 성인 38,645명 / 신생아 7,875명
- 기간 : 2001.6 ~ 2012.10
- 환자 개인 정보는 비식별화
- 입원일 진료일 등 시간정보는 모두 time-shifted, 그러나 같은 환자 내에선 시간순서 보존

테이블 이름	설명	총 row수	예시 항목
PATIENTS	환자 정보	46,520건	성별, 출생일, 사망일 등
ADMISSIONS	환자의 입원 당시의 기록 정보	58,976건	입원일, 퇴원일, 민족, 종교, 사망 시간 등
TRANSFER	중환자실 입원 시 병동 이동에 대한 정보	261,897건	이전 병동, 현재 병동, 이동시간
PRESCRIPTION	중환자실 내의 모든 처방에 대한 정보	4,156,450	투여약(GSN, NDC 코드), 투여 기간, 투여 량, 단위 등
CHART EVENT	환자의 차트 데이터에 대한 모든 정보	330,712,483건	Heart Rate, Blood Pressure, GCS, 키, 몸무게, 체온 등
LAB EVENT	랩 검사 데이터에 대한 정보	27,864,055건	potassium, Glucose, WBC_count, Hb, Troponin 등
INPUT_EVENT	입원 기간 내 환자에게 주입되는 모든 수액에 관한 정보	21,146,926건	투여액(Crystalloid, Colloid, etc.), 투여량 등
OUTPUT_EVENT	입원 기간 내 환자 체외로 배출되는 모든 fluid에 대한 정보	4,349,218건	소변량 등
NOTE_EVENT	입원 기간 중 담당의의 TEXT NOTE 정보	2,083,180건	기록 일시, 카테고리(퇴원 기록, 진단 기록 등), 텍스트 내용

2. MIMI-III 데이터 들여다보기



2. MIMIC-III 데이터 들여다보기

데이터 구조

SUBJECT_ID

- PATIENTS table의 고유키
- 성별, 출생일, 사망일 등 존재

HADM_ID

- ADMISSION table의 고유키
- 입원시간, 퇴원시간, 사망시간, 입원 형태 등 존재

ICUSTAY_ID

- ICUSTAYS table의 고유키
- ICU 입원 시간, 퇴원 시간 등 존재
- 하나의 ADMISSION에 다수의 ICUSTAY가 존재 할 수 있음



SUBJECT_ID : 환자 별 고유 식별자



HADA_ID : 환자 입원 식별자

ICUSTAY_ID

입원 중
ICU 입원
식별자



ICUSTAY_ID

입원 중
ICU 입원
식별자



MIMIC-III 데이터 구조

2. MIMIC-III 테이블

- 각 테이블 별 상세 설명 존재
- MIMIC-III의 전체적인 데이터베이스 구조 확인

<https://mimic.physionet.org/mimictables/>



테이블 설명 상세

- 테이블 간 관계도

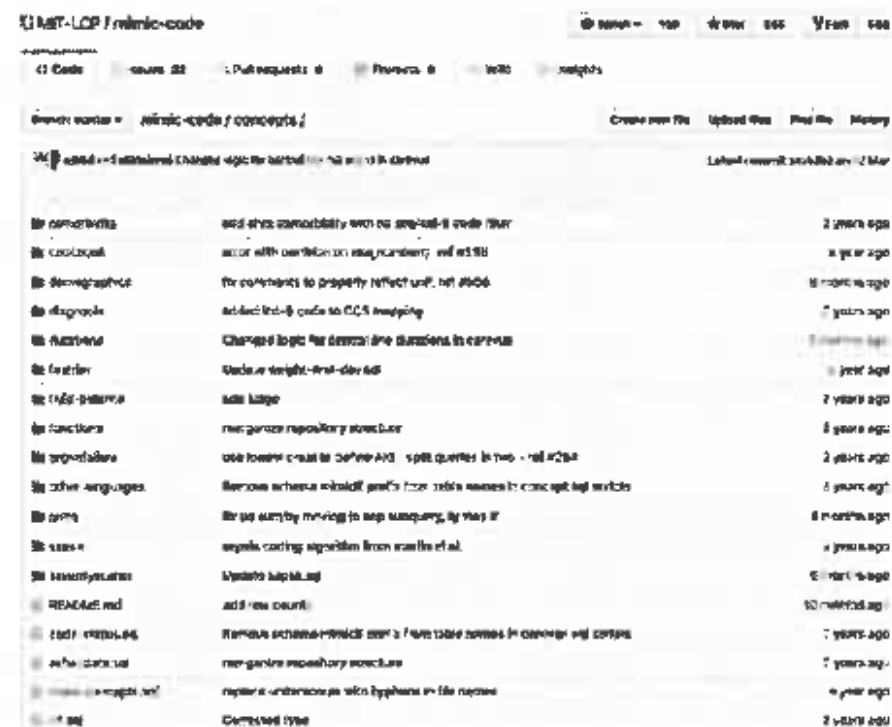
<https://mit-lcp.github.io/mimic-schema-spy/relationships.html>



테이블 관계도

2. MIMI-III script

- 연구에 활용된 다양한 쿼리 예시 존재
- PostgreSQL로 작성되어 있음
 - 인구학적 정보 추출 쿼리
 - Lab test 정보 추출 쿼리 (Potassium 등)
 - Sepsis 환자군 정의 기준 별 추출 쿼리



<https://github.com/MIT-LCP/mimic-code/tree/master/concepts>

2. MIMIC-III 연구예시

- 패혈증 발생 후 패혈 쇼크까지 일어나기 전에 미리 사전 위험을 감지하는 딥러닝 모델 개발
- 다양한 Numeric 변수들이 최소 약 1시간 간격으로 수집되어 저장되어 있음
- MIMIC-III는 중환자의 시간적 변화를 추적 할 수 있으며, 병동이 바뀌어도 추적이 가능함
- 다양한 Sepsis기준에 맞춰 sepsis 환자군을 추출하는 쿼리 존재

<https://github.com/MIT-LCP/mimic-code/tree/master/concepts/sepsis>

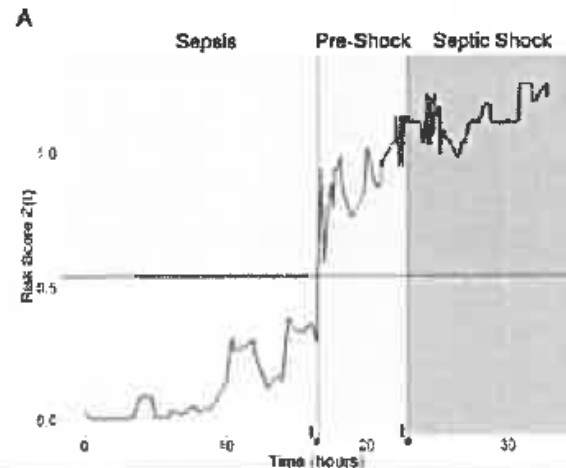
SCIENTIFIC REPORTS

Article OPF% Published: 16 April 2019

Data-driven discovery of a novel sepsis pre-shock state predicts impending septic shock in the ICU

Feature	Mean hours/data point	% datasets with at least 1 data point
HR	0.9	99.4%
Respiratory Rate	0.9	99.4%
Temperature	5.7	51.7%
Systolic BP	1.2	98.9%
Diastolic BP	1.2	98.9%
Mean BP	1.2	98.9%
CVP	4.1	32.0%
PaO ₂	16.2	53.3%
FiO ₂	3.9	54.3%
GCS	4.6	86.5%

Availability of EHR Data in MIMIC-III



3. 데이터 신청

CITI 수료 사이트

<https://about.citiprogram.org/en/homepage/>

The screenshot displays the CITI Program website. At the top, the CITI PROGRAM logo is on the left, and navigation links for Subscriptions, Courses, CEC/CEEs, Resources, and Support are in the center. A search icon and a Register button are on the right. A language dropdown menu is open, showing options for English, Español, 한국어 (Korean), 中文 (Chinese), Русский (Russian), Français (French), and Tiếng Việt (Vietnamese). The main heading is "Research Ethics and Compliance Training". Below this, four course cards are featured: "Protocol Registration and Results Summary Disclosure in Clinical Trials.gov", "Semiannual Evaluations In Depth", "Protocol Builder: Protocol-writing technology", and "Essentials of Grant Proposal Development". Each card includes a brief description and a "View Course" button. At the bottom, there is a login section with fields for "Email (or Username)" and "Password", a "Forgot your password?" link, and a "Log In" button. The CITI PROGRAM logo is also present at the bottom center.

CITI 수료 사이트
<https://about.citiprogram.org/en/homepage/>

Research Ethics and Compliance Training

Protocol Registration and Results Summary Disclosure in Clinical Trials.gov
A video-enhanced practical guide to compliance with protocol registration and summary results reporting.
View Course

Semiannual Evaluations In Depth
Maintain a high-quality animal care and use program at your organization.
View Course

Protocol Builder: Protocol-writing technology
Learn how leading institutions are using this tool to create more consistent, complete, compliant research study protocols.
Sign up for a Webinar

Essentials of Grant Proposal Development
Take the mystery out of grant writing with in-depth instruction that includes videos, exercises, resources, and a sample.
View Course

CITI PROGRAM

로그인

회원가입

이메일 (or Username) ID/비밀번호 찾기

비밀번호 ID/비밀번호 찾기

로그인

CITI 프로그램에 처음 오셨습니까? 시작 안내서 홈 열거나 시작하기 비디오를 시청하십시오.

3. 데이터 신청-CITI 수료

1. 개인정보 입력 - 기관 선택

2. ID/비밀번호 입력

3. 거주 국가 입력

4. 설문 조사 - 연구 목적

5. MIT의 설문조사 & 수강신청

6. 이메일 인증 가입 완료

소속기관(영문명) 검색

CITI Program 사이트 : 수료증들을 위한 기관 검색

학습자는 소속기관(영문명)을 검색하여 선택하시거나, 소속기관의 검색되지 않을 경우 "개인"에 해당합니다. 해당
우선해 소속기관 대신 "CITI Program Korea(CITI KOREA)"를 선택하시기 바랍니다.

Massachusetts Institute of Technology Affiliates

Massachusetts Institute of Technology Affiliates 로그인ID/PW 사용은 선택한 기관에 접속할 때만 사용됩니다. 회원가입 2단계
에서 로그인ID/PW를 설정합니다.

- ✓ 본인은 CITI 프로그램 자료에 액세스하기 위해 선택한 기관의 접근권 제한 정보도도 정책에 동의합니다
- ✗ 본인은 선택한 기관의 교육지침대상자 또는 소속학습자명이 아닙니다

2단계로 넘어가기

기관 검색 : **Massachusetts Institute of Technology Affiliates 선택**

Question 1

Human Subjects

For new trainees requiring IRB courses, select the group most appropriate to your research activity

This question is required. Choose one answer.

- ☐ Biomedical Research Investigators
- ☐ Social & Behavioral Research Investigators
- ☒ Data or Specimens Only Research

IRB Members: The purpose of this group is for reference purposes. It should not be selected to achieve CME/CEU Eligibility credits or completion reports. You may change your Learner Group status later to "IRB Reference Resource" for ongoing access and resource use of all CITI modules.

연구 목적 : **Data or Specimens Only Research 선택**

3. 데이터 신청-CITI 수료

7. Course & Quiz

팁 : Quiz를 모두 틀리면, 정답을 보여준다고 합니다.

8. 수료증 저장

- PhysioNet에 업로드할 수료증은 Complete Report

The screenshot shows the 'Massachusetts Institute of Technology Affiliates' course completion interface. It includes sections for 'Active Courses' (showing no active courses), 'Courses Ready to Begin' (listing 'Data or Specimens Only Research Stage 2 - Refresher Course' with a 'Start Now' button), and a 'Completion Report' table. The table lists the user 'JONG-KWAN JUNG' who completed the course on 30-Apr-2019. Below the table are links for 'View / Print' and 'Copy Link' for both the report and certificate.

Massachusetts Institute of Technology Affiliates

Active Courses [Learner Tools](#)

You have no active courses for this institution.

Courses Ready to Begin [Learner Tools](#)

Massachusetts Institute of Technology Affiliates
Data or Specimens Only Research
Stage 2 - Refresher Course
0 / 15 modules completed [Start Now](#)

Massachusetts Institute of Technology Affiliates
ID 1911

Data or Specimens Only Research

Name	Stage	Completion Date	Expiration Date	Record ID
JONG-KWAN JUNG	2 - Refresher Course	30-Apr-2019	29-Apr-2022	31458996

Completion Report
Completion Reports are transcripts of your course work, and include all quiz scores. Part 1 shows scores "frozen" at the time you completed and passed the course. Part 2 reflects scores for any subsequent quiz attempts.

Completion Certificate
Completion Certificates are "diplomas" that reflect course completion, but do not include quiz scores. Certificates are suitable for sharing with persons who do not need to see your quiz results, or posting online.

[View / Print](#) [Copy Link](#) [View / Print](#) [Copy Link](#)

3. 데이터 신청-CITI 수료

7. Course & Quiz

팁 : Quiz를 모두 틀리면, 정답을 보여준다고 합니다.

8. 수료증 저장

- PhysioNet에 업로드할 수료증은 Complete Report

The screenshot shows the 'Massachusetts Institute of Technology Affiliates' course completion interface. It includes sections for 'Active Courses' (showing no active courses), 'Courses Ready to Begin' (listing 'Data or Specimens Only Research Stage 2 - Refresher Course' with a 'Start Now' button), and a 'Completion Report' table. The table lists the user 'JONG-KWAN JUNG' who completed the course on 30-Apr-2019. Below the table, there are links for 'View / Print' and 'Copy Link' for both the 'Completion Report' and 'Completion Certificate'.

Massachusetts Institute of Technology Affiliates

Active Courses [Learner Tools](#)

You have no active courses for this institution.

Courses Ready to Begin [Learner Tools](#)

Massachusetts Institute of Technology Affiliates
Data or Specimens Only Research
Stage 2 - Refresher Course
0 / 15 modules completed

[Start Now](#)

Massachusetts Institute of Technology Affiliates
ID: 1911

Data or Specimens Only Research

Name	Stage	Completion Date	Expiration Date	Record ID
JONG-KWAN JUNG	2 - Refresher Course	30-Apr-2019	29-Apr-2022	31458996

Completion Report

Completion Reports are transcripts of your course work, and include all quiz scores. Part 1 shows scores "frozen" at the time you completed and passed the course. Part 2 reflects scores for any subsequent quiz attempts.

[View / Print](#) [Copy Link](#)

Completion Certificate

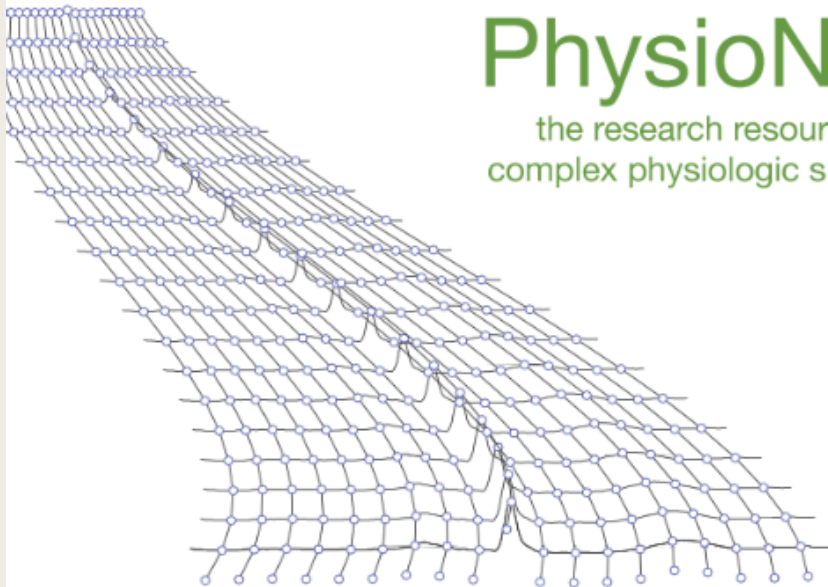
Completion Certificates are "diplomas" that reflect course completion, but do not include quiz scores. Certificates are suitable for sharing with persons who do not need to see your quiz results, or posting online.

[View / Print](#) [Copy Link](#)

3. 데이터 신청- PhysioNet

PHYSIONET ▼

Search



PhysioNet

the research resource for
complex physiologic signals

PhysioNet offers free web access to large collections of recorded physiologic signals ([PhysioBank](#)) and related open-source software ([PhysioToolkit](#)).

[PhysioNetWorks](#) workspaces are available to members of the PhysioNet community for works in progress that will be made publicly available in PhysioBank and PhysioToolkit when complete.

Use the **PhysioNet** button (top left) or start from the [site map](#) to explore this web site, or use the **Search** tool (top right) to locate pages of interest. The answers to many common questions are in the [FAQ](#)

TOP LINKS

[PhysioNet/CinC Challenge 2019](#)
[Recent News](#)
[Other Data Resources](#)

PhysioBank - Data

[PhysioBank Database Index](#)
[PhysioBank Record Search](#)
[MIMIC-III Database](#) ↗

PhysioToolkit - Software

[PhysioToolkit Software Index](#)
[WFDB Software Package](#)
[WFDB Python Package](#) ↗
[WFDB Matlab Toolbox](#)

PhysioNetWorks - Workspaces

[About PhysioNetWorks](#)
[PhysioNetWorks User Home](#)
[PhysioNetWorks Active Projects](#)

Help and Guides

3. 데이터 신청- PhysioNet

1. PhysioNet 회원가입

<https://physionet.org/pnw/login>

2. 이메일 인증

3. 가입완료

LOGIN TO PHYSIONETWORKS

Please note that you are not required to login in order to use PhysioBank, PhysioToolkit, or the PhysioNet Library, all of which can be accessed freely. Use of PhysioNetWorks is also free, but it requires logging in.

By using PhysioNetWorks, you agree that:

- You will not use it to post or copy anything that you do not have permission to copy and redistribute.
- You will maintain up-to-date copies of anything you deposit here, to be used in the event of data loss from your PhysioNetWorks workspace.
- You will not allow anyone else to use your personal PhysioNetWorks account.

If you do not agree to these terms and conditions, click here: [Return to Copyright/Inquiry page](#)

If you know your user name and password, click here: [Login](#)

If you do not have a PhysioNetWorks account, or if you forgot your password, enter your email address and choose the appropriate option.

Email address: [Create account](#) [Reset password](#)

You will not be able to set up your account or reset your password unless you can receive mail at the address you enter. Choose a permanent address you do not share with anyone else.

PhysioNetWorks login [발문페이지](#)

?

DoNotReply@physionet.org

To complete your PhysioNetWorks login, click on the link below, or copy and paste it into

?

DoNotReply@physionet.org

나에게 -

영어 → 한국어 → 매일 번역

To complete your PhysioNetWorks login, click on the link below, or copy and paste it into the location bar of your web browser:

4. 데이터 접근 권한 얻기

5. 입력 항목 & 수료증 업로드

- 만약 학생일 경우 교수님 혹은 연구 책임자의 보증을 받아야 함
- 책임자의 이메일 요구

[illegible]

3. 데이터 신청- PhysioNet

6. 승인 메일

7. physioNet works

<https://physionet.org/works/MIMICIII/ClinicalDatabase>

8. 원하는 테이블 다운로드

1. [checksum_md5_zipped.txt](#)
2. [checksum_md5_unzipped.txt](#)
3. [ADMISSIONS.csv.gz](#)
4. [CALLOUT.csv.gz](#)
5. [CAREGIVERS.csv.gz](#)
6. [CPTEVENTS.csv.gz](#)
7. [DATETIMEEVENTS.csv.gz](#)
8. [DIAGNOSES_ICD.csv.gz](#)
9. [DRGCODES.csv.gz](#)
10. [D_GPT.csv.gz](#)
11. [D_ICD_DIAGNOSES.csv.gz](#)
12. [D_ICD_PROCEDURES.csv.gz](#)
13. [D_ITEMS.csv.gz](#)
14. [D_LABITEMS.csv.gz](#)
15. [ICUSTAYS.csv.gz](#)
16. [INPUTEVENTS_CV.csv.gz](#)
17. [INPUTEVENTS_MV.csv.gz](#)[LABEVENTS.csv.gz](#)
18. [LABEVENTS.csv.gz](#)
19. [MICROBIOLOGYEVENTS.csv.gz](#)
20. [NOTEVENTS.csv.gz](#)

21. [OUTPUTEVENTS.csv.gz](#)
22. [PATIENTS.csv.gz](#)
23. [PRESCRIPTIONS.csv.gz](#)
24. [PROCEDUREEVENTS_MV.csv.gz](#)
25. [PROCEDURES_ICD.csv.gz](#)
26. [SERVICES.csv.gz](#)

PhysioNet DUA for

참고문헌

— approved

bioethics@physionet.org

이메일

3A. 영어 > 한국어 > 재일본어

Thank you for your interest in the PhysioNet Clinical Database (MIMIC-III, MIMIC-IV, eICU Collaborative Research Database, MIMIC-CXR).

Your request for access to PhysioNet restricted databases has been approved. Your access should be activated shortly.

For access, log in to PhysioNetWorks and click on the relevant link on your PhysioNetWorks home page.

Information about building databases from downloaded data files is available on the databases' respective PhysioNetWorks pages.

You may also log in to the MIMIC Query Builder (<http://mimic.org/mimic3/mimic3querybuilder/>) using your PhysioNetWorks username and password.

Role	Active Project
Reviewer	Deidentified Medical Text
Reviewer	eICU Collaborative Research Database
Reviewer	MIMIC-CXR
Reviewer	MIMIC-III Clinical Database
Reviewer	MIMIC-III arterial line dataset (HST-953)
Reviewer	MIMIC-III Clinical Database
Reviewer	MIMIC-III Derived Data Repository

Downloads

The MIMIC-III dataset is a collection of 34 tables organized in a CD format. The data is distributed across the files and is not restricted to any table. The data is provided in the MIMIC-III format. The data is provided in the MIMIC-III format. The data is provided in the MIMIC-III format.

For more information, please visit the MIMIC-III website.

For more information, please visit the MIMIC-III website.

For more information, please visit the MIMIC-III website.

For more information, please visit the MIMIC-III website.

For more information, please visit the MIMIC-III website.

1. [checksum_md5_zipped.txt](#) (MD5 checksum for zipped files)
2. [checksum_md5_unzipped.txt](#) (MD5 checksum for unzipped files)
3. [ADMISSIONS.csv.gz](#) (MIMIC-III admission data)
4. [CALLOUT.csv.gz](#) (MIMIC-III callout data)
5. [CAREGIVERS.csv.gz](#) (MIMIC-III caregiver data)
6. [CPTEVENTS.csv.gz](#) (MIMIC-III CPT event data)
7. [DATETIMEEVENTS.csv.gz](#) (MIMIC-III datetime event data)
8. [DIAGNOSES_ICD.csv.gz](#) (MIMIC-III ICD diagnosis data)
9. [DRGCODES.csv.gz](#) (MIMIC-III DRG code data)
10. [D_GPT.csv.gz](#) (MIMIC-III GPT data)
11. [D_ICD_DIAGNOSES.csv.gz](#) (MIMIC-III ICD diagnosis data)
12. [D_ICD_PROCEDURES.csv.gz](#) (MIMIC-III ICD procedure data)
13. [D_ITEMS.csv.gz](#) (MIMIC-III item data)
14. [D_LABITEMS.csv.gz](#) (MIMIC-III lab item data)
15. [ICUSTAYS.csv.gz](#) (MIMIC-III ICU stay data)
16. [INPUTEVENTS_CV.csv.gz](#) (MIMIC-III CV input event data)
17. [INPUTEVENTS_MV.csv.gz](#) (MIMIC-III MV input event data)
18. [LABEVENTS.csv.gz](#) (MIMIC-III lab event data)
19. [MICROBIOLOGYEVENTS.csv.gz](#) (MIMIC-III microbiology event data)
20. [NOTEVENTS.csv.gz](#) (MIMIC-III note event data)
21. [OUTPUTEVENTS.csv.gz](#) (MIMIC-III output event data)
22. [PATIENTS.csv.gz](#) (MIMIC-III patient data)
23. [PRESCRIPTIONS.csv.gz](#) (MIMIC-III prescription data)
24. [PROCEDUREEVENTS_MV.csv.gz](#) (MIMIC-III MV procedure event data)
25. [PROCEDURES_ICD.csv.gz](#) (MIMIC-III ICD procedure data)
26. [SERVICES.csv.gz](#) (MIMIC-III service data)

4. DB구축 - MIMIC III

0. 요구되는 프로그래밍 스킬

SQL (Structured Query Language)

- MIMIC-III 데이터베이스에서 원하는 정보를 추출하기 위해 필요

- PostgreSQL

공식사이트) <https://www.tutorialspoint.com/postgresql/index.htm>

한글) MSSQL 한글 강좌 (1-7강) - SQL은 문법은 대부분 유사함

아래 강좌 수강 후, 위 공식 사이트 참조.

<http://www.d-yoon.com/database-mssql-for-medical-data-analysis>

Python

- 추출된 MIMIC-III 데이터 전처리 및 분석 시 활용
- MIMIC-III의 Waveform 데이터 처리를 위해 필요

<https://www.edwith.org/pythonforeverybody>

