

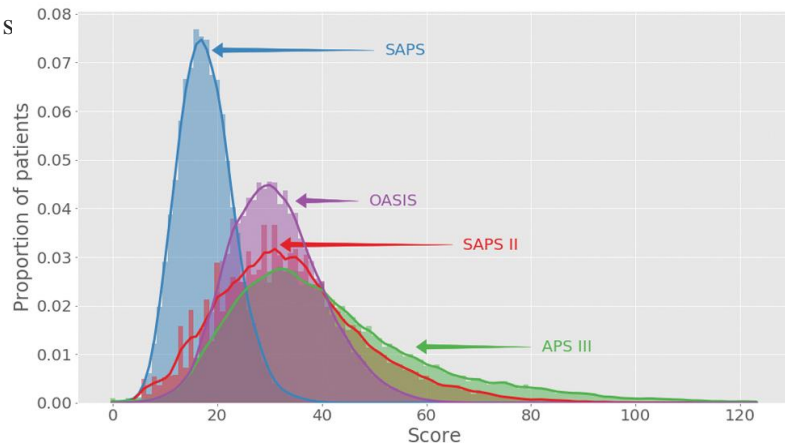
# The MIMIC Code Repository: enabling reproducibility in critical care research

# MIMIC code repository

- Concept: 중요한 컨셉을 추출하는 코드
- Executable documents: 튜토리얼 노트
- community: 공개토론장

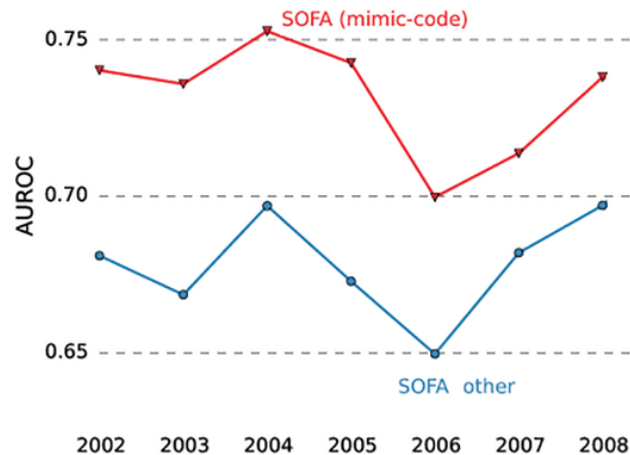
# Concepts: Severity of illness scores

- 연구 목적에 맞는 population, risk 조정
- 수집 데이터의 요소가 부족한 경우: 신경 장애의 일반적인 마커 인 Glasgow Coma Scale (GCS)은 진정 작용으로 인해 평가할 수 없는 환자의 경우 일반적으로 15라고 가정.
- 5가지의 severity를 가짐
  - acute physiology score (APS)-III, simplified acute physiology score (OASIS).
  - 환자가 처음 머문 24시간 동안의 데이터를 사용



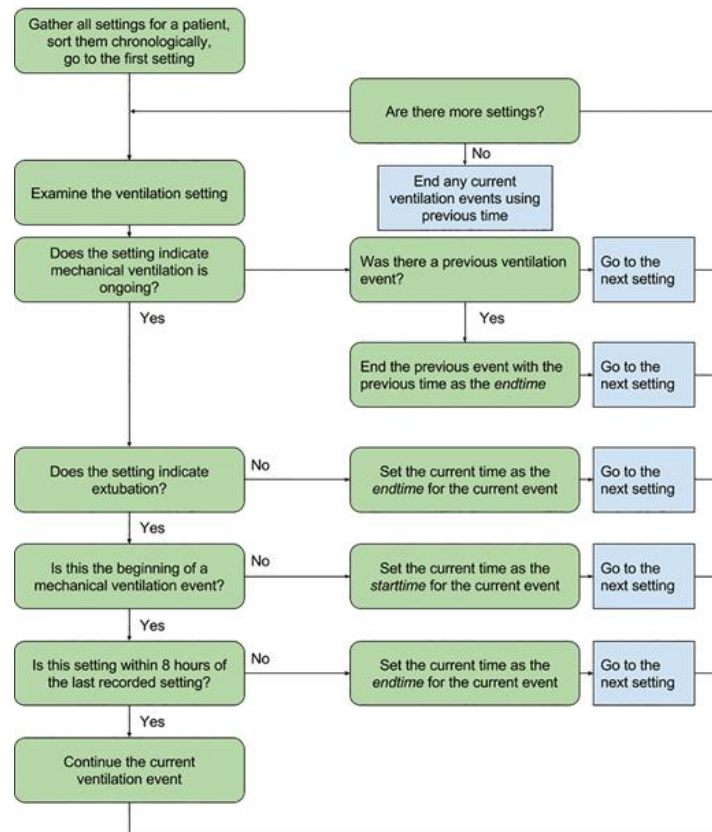
# Concepts: organ dysfunction scores

- SOFA, LODS : 6개의 organ system을 평가
- 불일치: GCS 때문?



# Concepts: Timing of treatment

- 병원 데이터 수집의 제한이 있음.



# Concepts: Sepsis

- 패혈증: 2011년 병원 비용 5.2%, 2013년 병원 비용 6.2%
- 패혈증의 정확한 발병은 EHR로 설명 어려움
- 항생제 교차 사용으로 환자 확인 후 미생물 평가
- Sepsis-3 guideline: SOFA>2
- billing code: (International Classification of Diseases, Ninth Revision [ICD-9] codes 785.52 and 995.92)

# Concepts: Comorbidities

# Concepts: Concept road map

Category	Concepts
Severity of illness scores	APS III, SAPS, SAPS II, OASIS
Organ dysfunction scores	SOFA, qSOFA, LODS, SIRS, MELD, KDIGO, AKIN
Treatments	Continuous renal replacement therapy, intermittent hemodialysis, vasopressors, mechanical ventilation
Sepsis	Suspicion of infection, Angus et al. criteria, Martin et al. criteria, explicit ICD-9 coding of sepsis, <i>CMS sepsis criteria</i> , <i>CDC sepsis criteria</i>
Comorbid burden	Elixhauser et al. (AHRQ), Quan et al., <i>Charlson et al.</i>
First 24 h aggregates	Vital signs, laboratory values, blood gas values, urine output
Diagnosis groups	Certified Coding Specialist groups
Demographics	Weight, height, age, gender, <i>service type</i>
<i>Hourly data</i>	<i>Vasopressor doses, vital signs, laboratory values, blood gas values</i>
<i>Fluid balance</i>	<i>Total fluid intake, total fluid output</i>



# Executable documents

Cohort selection will begin with three tables: *patients*, *admissions*, and *icustays*:

- *patients*: information about a patient that does not change - e.g. date of birth, genotypical sex
- *admissions*: information recorded on hospital admission - admission type (elective, emergency), time of admission
- *icustays*: information recorded on intensive care unit admission - primarily admission and discharge time

As MIMIC-III is primarily an intensive care unit (ICU) database, the focus will be on patients admitted to and discharged from the ICU. That is, rather than selecting our cohort based off the individual patient (identified by `subject_id` in the database), we will usually want to select our cohort based off the ICU stay (identified by `icustay_id`). Thus, it is sensible to begin with the *icustays* table:

```
In [2]: query = query_schema + """
SELECT subject_id, hadm_id, icustay_id
FROM icustays
LIMIT 10
"""
df = pd.read_sql_query(query, con)
df.head()
```

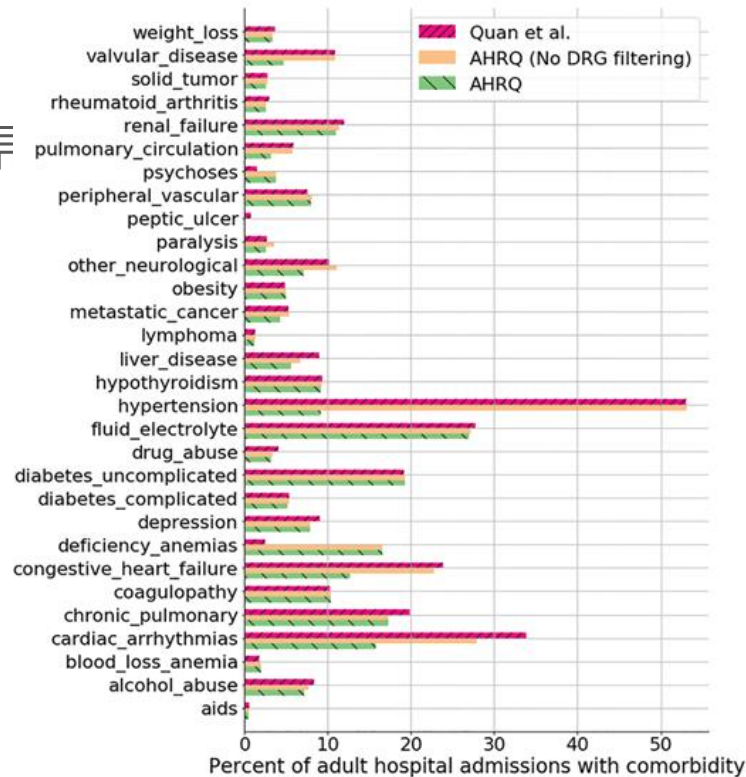
```
Out[2]:
```

	subject_id	hadm_id	icustay_id
0	268	110404	280836
1	269	106296	206613
2	270	188028	220345
3	271	173727	249196
4	272	164716	210407

Note: in the above we use `LIMIT 10`: this limits our results to only 10 rows. It's nice to include this statement when prototyping as it speeds up queries immensely. Later on when we are doing full data extraction, we would remove this statement.

# Community

- 많은 응급실 환자는 만성 질환이 있음.
- AHRQ(Elixhauser): ICD-9코드 29개 범주로 분류
- Diagnosis related group을 사용하여 향상
- Quan: 불일치를 검사하여 향상



# 결론

- 목표: Transparent research process
- 제한사항:
  - technical tools(git, SQL)
  - standard SQL
  - CDM 적용
  - 계속 개발됨