```sql
WITH demographic AS (
    SELECT
        p.subject_id,
        a.hadm_id,
        p.gender,
        FLOOR(TIMESTAMP_DIFF(a.admittime, p.dob, YEAR)) AS age,
        a.ethnicity
    FROM `physionet-data.mimiciii_clinical.patients` p
    JOIN `physionet-data.mimiciii_clinical.admissions` a ON p.subject_id =
a.subject_id
),

comorbidities AS (
    SELECT
        d.subject_id,
        d.hadm_id,
        MAX(CASE WHEN d.icd9_code LIKE '250%' THEN 1 ELSE 0 END) AS diabetes,
        MAX(CASE WHEN d.icd9_code LIKE '401%' THEN 1 ELSE 0 END) AS hypertension,
        MAX(CASE WHEN d.icd9_code LIKE '410%' THEN 1 ELSE 0 END) AS
myocardial_infarction,
        MAX(CASE WHEN d.icd9_code LIKE '430%' OR d.icd9_code LIKE '431%' OR
d.icd9_code LIKE '432%' THEN 1 ELSE 0 END) AS cerebral_vascular_disease,
        MAX(CASE WHEN d.icd9_code LIKE '272%' THEN 1 ELSE 0 END) AS hyperlipidemia,
        MAX(CASE WHEN d.icd9_code LIKE '414%' THEN 1 ELSE 0 END) AS
coronary_heart_disease,
        MAX(CASE WHEN d.icd9_code LIKE '493%' THEN 1 ELSE 0 END) AS asthma,
        MAX(CASE WHEN d.icd9_code LIKE '571%' THEN 1 ELSE 0 END) AS
chronic_liver_disease,
        MAX(CASE WHEN d.icd9_code LIKE '585%' THEN 1 ELSE 0 END) AS
chronic_renal_disease,
        MAX(CASE WHEN d.icd9_code LIKE '140%' OR d.icd9_code LIKE '141%' THEN 1
ELSE 0 END) AS cancer
    FROM `physionet-data.mimiciii_clinical.diagnoses_icd` d
    GROUP BY d.subject_id, d.hadm_id
),

vital_signs AS (
    SELECT
        ce.subject_id,
        ce.hadm_id,
        MAX(CASE WHEN ce.itemid = 220045 THEN ce.valuenum END) AS heart_rate,
        MAX(CASE WHEN ce.itemid = 220210 THEN ce.valuenum END) AS systolic_bp,
        MAX(CASE WHEN ce.itemid = 220277 THEN ce.valuenum END) AS gcs
    FROM `physionet-data.mimiciii_clinical.chartevents` ce
    GROUP BY ce.subject_id, ce.hadm_id
),

lab_tests AS (
    SELECT
        le.subject_id,
        le.hadm_id,
```

```sql
        MAX(CASE WHEN le.itemid = 51221 THEN le.valuenum END) AS white_blood_cell,
        MAX(CASE WHEN le.itemid = 51265 THEN le.valuenum END) AS hemoglobin,
        MAX(CASE WHEN le.itemid = 51277 THEN le.valuenum END) AS platelet
    FROM `physionet-data.mimiciii_clinical.labevents` le
    GROUP BY le.subject_id, le.hadm_id
),

procedures AS (
    SELECT
        pe.subject_id,
        pe.hadm_id,
        MAX(CASE WHEN pe.itemid = 225792 THEN 1 ELSE 0 END) AS tracheostomy,
        MAX(CASE WHEN pe.itemid = 220210 THEN 1 ELSE 0 END) AS vasopressor
    FROM `physionet-data.mimiciii_clinical.procedureevents_mv` pe
    GROUP BY pe.subject_id, pe.hadm_id
),

outcomes AS (
    SELECT
        a.subject_id,
        a.hadm_id,
        CASE WHEN a.deathtime IS NOT NULL AND TIMESTAMP_DIFF(a.deathtime,
a.admittime, DAY) <= 30 THEN 1 ELSE 0 END AS mortality_30d,
        TIMESTAMP_DIFF(a.dischtime, a.admittime, DAY) AS hospital_stay_length,
        icu.icustay_length
    FROM `physionet-data.mimiciii_clinical.admissions` a
    LEFT JOIN (
        SELECT
            icu.subject_id,
            icu.hadm_id,
            TIMESTAMP_DIFF(icu.outtime, icu.intime, DAY) AS icustay_length
        FROM `physionet-data.mimiciii_clinical.icustays` icu
    ) icu ON a.subject_id = icu.subject_id AND a.hadm_id = icu.hadm_id
),

comorbidity_percentages AS (
    SELECT
        COUNT(CASE WHEN com.diabetes = 1 THEN 1 END) / COUNT(*) * 100 AS
diabetes_percentage,
        COUNT(CASE WHEN com.hypertension = 1 THEN 1 END) / COUNT(*) * 100 AS
hypertension_percentage,
        COUNT(CASE WHEN com.myocardial_infarction = 1 THEN 1 END) / COUNT(*) * 100
AS myocardial_infarction_percentage,
        COUNT(CASE WHEN com.cerebral_vascular_disease = 1 THEN 1 END) / COUNT(*) *
100 AS cerebral_vascular_disease_percentage,
        COUNT(CASE WHEN com.hyperlipidemia = 1 THEN 1 END) / COUNT(*) * 100 AS
hyperlipidemia_percentage,
        COUNT(CASE WHEN com.coronary_heart_disease = 1 THEN 1 END) / COUNT(*) * 100
AS coronary_heart_disease_percentage,
        COUNT(CASE WHEN com.asthma = 1 THEN 1 END) / COUNT(*) * 100 AS
asthma_percentage,
```

```sql
        COUNT(CASE WHEN com.chronic_liver_disease = 1 THEN 1 END) / COUNT(*) * 100
AS chronic_liver_disease_percentage,
        COUNT(CASE WHEN com.chronic_renal_disease = 1 THEN 1 END) / COUNT(*) * 100
AS chronic_renal_disease_percentage,
        COUNT(CASE WHEN com.cancer = 1 THEN 1 END) / COUNT(*) * 100 AS
cancer_percentage
    FROM comorbidities com
)

SELECT
    demo.subject_id,
    demo.hadm_id,
    demo.gender,
    demo.age,
    demo.ethnicity,
    com.diabetes,
    com.hypertension,
    com.myocardial_infarction,
    com.cerebral_vascular_disease,
    com.hyperlipidemia,
    com.coronary_heart_disease,
    com.asthma,
    com.chronic_liver_disease,
    com.chronic_renal_disease,
    com.cancer,
    vs.heart_rate,
    vs.systolic_bp,
    vs.gcs,
    lab.white_blood_cell,
    lab.hemoglobin,
    lab.platelet,
    proc.tracheostomy,
    proc.vasopressor,
    out.mortality_30d,
    out.icustay_length,
    out.hospital_stay_length,
    perc.diabetes_percentage,
    perc.hypertension_percentage,
    perc.myocardial_infarction_percentage,
    perc.cerebral_vascular_disease_percentage,
    perc.hyperlipidemia_percentage,
    perc.coronary_heart_disease_percentage,
    perc.asthma_percentage,
    perc.chronic_liver_disease_percentage,
    perc.chronic_renal_disease_percentage,
    perc.cancer_percentage
FROM demographic demo
LEFT JOIN comorbidities com ON demo.subject_id = com.subject_id AND demo.hadm_id =
com.hadm_id
LEFT JOIN vital_signs vs ON demo.subject_id = vs.subject_id AND demo.hadm_id =
vs.hadm_id
```

```
LEFT JOIN lab_tests lab ON demo.subject_id = lab.subject_id AND demo.hadm_id =
lab.hadm_id
LEFT JOIN procedures proc ON demo.subject_id = proc.subject_id AND demo.hadm_id =
proc.hadm_id
LEFT JOIN outcomes out ON demo.subject_id = out.subject_id AND demo.hadm_id =
out.hadm_id
LEFT JOIN comorbidity_percentages perc ON 1 = 1;  -- This adds the percentage data
for comorbidities
```

THE QUERY for the Target variable "VAP"

```
WITH tbi_patients AS (
    SELECT DISTINCT p.subject_id, a.hadm_id
    FROM `physionet-data.mimiciii_clinical.patients` p
    JOIN `physionet-data.mimiciii_clinical.admissions` a ON p.subject_id =
a.subject_id
    JOIN `physionet-data.mimiciii_clinical.diagnoses_icd` d ON a.hadm_id =
d.hadm_id
    WHERE
        (d.icd9_code BETWEEN '80000' AND '80199') OR
        (d.icd9_code BETWEEN '80300' AND '80499') OR
        (d.icd9_code BETWEEN '8500' AND '85419')  -- TBI ICD-9 Codes
),

vap_patients AS (
    SELECT DISTINCT subject_id, hadm_id, 1 AS vap
    FROM `physionet-data.mimiciii_clinical.diagnoses_icd`
    WHERE icd9_code = '99731' -- VAP ICD-9 Code
)

-- Combine the two to add VAP status for TBI patients
SELECT
    t.subject_id,
    t.hadm_id,
    COALESCE(v.vap, 0) AS vap -- 1 if VAP is present, 0 if not
FROM tbi_patients t
LEFT JOIN vap_patients v ON t.subject_id = v.subject_id AND t.hadm_id = v.hadm_id
ORDER BY t.subject_id, t.hadm_id;
```