Study population:

```
-- Step 1: Identify patients with DKA using ICD-9 codes
WITH dka_patients AS (
  SELECT DISTINCT
       di.subject_id,
       di.hadm_id,
       di.icd9_code
  FROM `physionet-data.mimiciii_clinical.diagnoses_icd` di
  WHERE di.icd9_code IN ('25010', '25011', '25012', '25013') -- ICD-9 codes for DKA
),
-- Step 2: Identify patients with CKD stage 5
ckd_stage_5_patients AS (
  SELECT DISTINCT
       di.subject_id,
       di.hadm_id
  FROM `physionet-data.mimiciii_clinical.diagnoses_icd` di
  WHERE di.icd9_code IN ('5855', '5856') -- ICD-9 codes for CKD stage 5
),
-- Step 3: Find the first ICU stay for each hospital admission
first_icu_stay AS (
  SELECT
       icu.subject_id,
       icu.hadm_id,
       icu.icustay_id,
       icu.intime,
       ROW_NUMBER() OVER (PARTITION BY icu.hadm_id ORDER BY icu.intime) AS rn
  FROM `physionet-data.mimiciii_clinical.icustays` icu
),
-- Step 4: Combine DKA patients with their first ICU stay, excluding CKD stage 5
patients
final_population AS (
  SELECT
       dka.subject_id,
```

```
dka.hadm_id,
       icu.icustay_id,
      icu.intime
   FROM dka_patients dka
   JOIN first_icu_stay icu
      ON dka.subject_id = icu.subject_id
      AND dka.hadm_id = icu.hadm_id
  LEFT JOIN ckd_stage_5_patients ckd
      ON dka.subject_id = ckd.subject_id
      AND dka.hadm_id = ckd.hadm_id
  WHERE icu.rn = 1 -- Keep only the first ICU stay
    AND ckd.hadm_id IS NULL -- Exclude patients with CKD stage 5
)
SELECT *
FROM final_population
ORDER BY subject_id, intime;
Variables
Demographics: subject_id, gender, age, ethnicity, hadm_id, weight, icustay_id.
CREATE OR REPLACE TABLE `proud-archery-435803-j0.healthcare_650.demographics` AS
SELECT
  p.subject_id,
  p.gender,
  DATE_DIFF(CAST(a.admittime AS DATE), CAST(p.dob AS DATE), YEAR) AS age, --
Calculate age at admission
  a.ethnicity,
   a.hadm_id, -- Include hadm_id
  wf.weight AS weight, -- Weight from weightfirstday table
  hf.height AS height, -- Height from heightfirstday table
   i.icustay_id
FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
JOIN `physionet-data.mimiciii_clinical.icustays` i
   ON fp.icustay_id = i.icustay_id
JOIN `physionet-data.mimiciii_clinical.admissions` a
  ON i.hadm_id = a.hadm_id
JOIN `physionet-data.mimiciii_clinical.patients` p
```

```
ON a.subject_id = p.subject_id
LEFT JOIN `physionet-data.mimiciii_derived.weightfirstday` wf
  ON i.icustay_id = wf.icustay_id
LEFT JOIN `physionet-data.mimiciii_derived.heightfirstday` hf
   ON i.icustay_id = hf.icustay_id
ORDER BY p.subject_id, a.admittime;
Vitals: subject_id, hadm_id, icustay_id, Mean Heart Rate, Systolic BP,
Diastolic BP, Mean Respiratory Rate, Temperature.
CREATE OR REPLACE TABLE `proud-archery-435803-j0.healthcare_650.vitals` AS
SELECT
  fp.subject_id,
   fp.hadm_id,
  fp.icustay_id,
  vf.HeartRate_Mean AS heart_rate,
  vf.SysBP_Mean AS systolic_bp,
  vf.DiasBP_Mean AS diastolic_bp,
  vf.RespRate_Mean AS respiratory_rate,
  vf.TempC_Mean AS temperature,
   -- vf.SpO2_Mean AS spo2 -- Oxygen saturation, if needed
FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
JOIN `physionet-data.mimiciii_derived.vitalsfirstday` vf
   ON fp.icustay_id = vf.icustay_id
-- WHERE vf.HeartRate_Mean IS NOT NULL
     OR vf.SysBP_Mean IS NOT NULL
   OR vf.DiasBP Mean IS NOT NULL
     OR vf.RespRate_Mean IS NOT NULL
     OR vf.TempC_Mean IS NOT NULL
ORDER BY fp.subject_id, fp.hadm_id;
Complications: 450
CREATE OR REPLACE TABLE `proud-archery-435803-j0.healthcare_650.complications` AS
SELECT
   fp.subject_id,
  fp.hadm_id,
   fp.icustay_id,
```

```
CASE
      WHEN di.icd9_code IN ('25040', '25041', '25042') THEN 'Diabetic Nephropathy'
      WHEN di.icd9_code IN ('36201', '36202', '36203') THEN 'Diabetic Retinopathy'
      WHEN di.icd9_code IN ('3569', '3572') THEN 'Diabetic Peripheral Neuropathy'
      WHEN di.icd9_code IN ('41401', '41402', '41403') THEN 'Coronary Heart Disease'
      WHEN di.icd9_code IN ('4370', '4371', '4372') THEN 'Cerebral Atherosclerosis'
      WHEN di.icd9_code IN ('44381', '44389', '4439') THEN 'Peripheral Artery
Disease'
      ELSE 'Other'
  END AS complication_type
FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
JOIN `physionet-data.mimiciii_clinical.diagnoses_icd` di
   ON fp.hadm_id = di.hadm_id
WHERE di.icd9_code IN (
   -- Microangiopathy
   '25040', '25041', '25042', -- Diabetic Nephropathy
   '36201', '36202', '36203', -- Diabetic Retinopathy
   '3569', '3572',
                       -- Diabetic Peripheral Neuropathy
   -- Macroangiopathy
   '41401', '41402', '41403', -- Coronary Heart Disease
   '4370', '4371', '4372', -- Cerebral Atherosclerosis
   '44381', '44389', '4439' -- Peripheral Artery Disease
)
ORDER BY fp.subject_id, fp.hadm_id;
Comorbidities: 584
CREATE OR REPLACE TABLE `proud-archery-435803-j0.healthcare_650.comorbidities` AS
SELECT
  fp.subject_id,
  fp.hadm_id,
  fp.icustay_id,
  CASE
      WHEN di.icd9_code IN ('5851', '5852', '5853', '5854') THEN 'CKD' -- Chronic
Kidnev Disease
      WHEN di.icd9_code IN ('5990') THEN 'UTI' -- Urinary Tract Infection
```

```
WHEN di.icd9_code IN ('486', '481', '4820', '4821', '4822') THEN 'Pneumonia' --
Pneumonia
      WHEN di.icd9_code IN ('5712', '5715', '5716') THEN 'Liver Disease' -- Liver
Disease
       WHEN di.icd9_code IN ('4011', '4019', '40210') THEN 'Hypertension' -- History
of Hypertension
      WHEN di.icd9_code IN ('4280', '4281', '4289') THEN 'CHF' -- History of
Congestive Heart Failure
      ELSE 'Other'
  END AS comorbidity
FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
JOIN `physionet-data.mimiciii_clinical.diagnoses_icd` di
   ON fp.hadm_id = di.hadm_id
WHERE di.icd9_code IN (
   -- Comorbidities ICD-9 codes
   '5851', '5852', '5853', '5854', -- CKD
   '5990', -- UTI
   '486', '481', '4820', '4821', '4822', -- Pneumonia
   '5712', '5715', '5716', -- Liver Disease
   '4011', '4019', '40210', -- Hypertension
   '4280', '4281', '4289' -- CHF
)
ORDER BY fp.subject_id, fp.hadm_id;
Lab Vitals:
CREATE OR REPLACE TABLE `proud-archery-435803-j0.healthcare_650.lab_vitals` AS
SELECT
  fp.subject_id,
  fp.hadm_id,
  fp.icustay_id,
   1.BICARBONATE_min AS bicarbonate, -- Bicarbonate
   1.WBC_min AS wbc,
                                   -- White Blood Cell count
   1.HEMOGLOBIN_min AS hemoglobin, -- Hemoglobin
   1.PLATELET_min AS platelets,
                                  -- Platelet count
   1.SODIUM_min AS sodium,
                                  -- Sodium

    CHLORIDE_min AS chloride,

                                 -- Chloride
```

```
-- Blood Urea Nitrogen
  1.BUN_min AS bun,
  1.CREATININE_min AS creatinine, -- Serum Creatinine
  1.POTASSIUM_min AS potassium, -- Potassium
  1.GLUCOSE_min AS glucose, -- Blood Glucose
  1.ANIONGAP_min AS anion_gap, -- Anion Gap
FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
JOIN `physionet-data.mimiciii_derived.labsfirstday` 1
  ON fp.icustay_id = l.icustay_id
-- WHERE 1.BICARBONATE_min IS NOT NULL
     OR 1.WBC_min IS NOT NULL
     OR 1.HEMOGLOBIN_min IS NOT NULL
     OR 1.PLATELET_min IS NOT NULL
     OR 1.SODIUM_min IS NOT NULL
     OR 1.CHLORIDE_min IS NOT NULL
     OR 1.BUN_min IS NOT NULL
     OR 1. CREATININE min IS NOT NULL
     OR 1.POTASSIUM_min IS NOT NULL
     OR 1.GLUCOSE_min IS NOT NULL
     OR 1.ANIONGAP_min IS NOT NULL
ORDER BY fp.subject_id, fp.hadm_id;
Scoring Systems:
CREATE OR REPLACE TABLE `proud-archery-435803-j0.healthcare_650.scoring_systems` AS
SELECT
  fp.subject_id,
  fp.hadm_id,
  fp.icustay_id,
  s.sapsii AS saps_ii, -- Simplified Acute Physiology Score II
  o.oasis AS oasis,
                              -- Oxford Acute Severity of Illness Score
  so.SOFA AS sofa,
                              -- Sequential Organ Failure Assessment Score
  g.GCSEyes AS gcs_eyes, -- Eyes Glasgow Coma Scale
  g.GCSMotor AS gcs_motor,
                                  -- Motor Glasgow Coma Scale
  g.GCSVerbal AS gcs_verbal
                                   -- Verbal Glasgow Coma Scale
FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
LEFT JOIN `physionet-data.mimiciii_derived.sapsii` s
  ON fp.icustay_id = s.icustay_id
```

```
LEFT JOIN `physionet-data.mimiciii_derived.oasis` o
   ON fp.icustay_id = o.icustay_id
LEFT JOIN `physionet-data.mimiciii_derived.sofa` so
  ON fp.icustay_id = so.icustay_id
LEFT JOIN `physionet-data.mimiciii_derived.gcsfirstday` g
   ON fp.icustay_id = g.icustay_id
WHERE s.sapsii IS NOT NULL
 OR o.oasis IS NOT NULL
 OR so.SOFA IS NOT NULL
 OR g.GCSEyes IS NOT NULL
 OR g.GCSMotor IS NOT NULL
 OR g.GCSVerbal IS NOT NULL
ORDER BY fp.subject_id, fp.hadm_id;
Other Vital Information: Diabetes Melitus type, Infusion volume, Urine output
- 9,093 rows
CREATE OR REPLACE TABLE `proud-archery-435803-j0.healthcare_650.other_vital_info` AS
WITH dm_type AS (
  -- Step 1: Determine DM type from ICD-9 codes
  SELECT
      fp.subject_id,
      fp.hadm_id,
      fp.icustay_id,
      CASE
          WHEN di.icd9_code IN ('25001', '25003', '25011', '25013', '25021', '25023')
THEN 'T1DM' -- Type 1 Diabetes Mellitus
           WHEN di.icd9_code IN ('25000', '25002', '25010', '25012', '25020', '25022')
THEN 'T2DM' -- Type 2 Diabetes Mellitus
           ELSE 'Unknown'
       END AS dm_type
   FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
  LEFT JOIN `physionet-data.mimiciii_clinical.diagnoses_icd` di
      ON fp.hadm_id = di.hadm_id
),
infusion_volume_combined AS (
   -- Step 2: Combine infusion volumes from inputevents_mv and inputevents_cv
  SELECT
```

```
fp.subject_id,
       fp.hadm_id,
       fp.icustay_id,
       CASE
           WHEN SUM(iv_mv.amount) IS NOT NULL THEN SUM(iv_mv.amount) -- Take
inputevents_mv if available
           WHEN SUM(iv_cv.amount) IS NOT NULL THEN SUM(iv_cv.amount) -- Otherwise,
take inputevents_cv
           ELSE NULL
       END AS infusion_volume
  FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
  LEFT JOIN `physionet-data.mimiciii_clinical.inputevents_mv` iv_mv
       ON fp.icustay_id = iv_mv.icustay_id
  LEFT JOIN `physionet-data.mimiciii_clinical.inputevents_cv` iv_cv
       ON fp.icustay_id = iv_cv.icustay_id
   GROUP BY fp.subject_id, fp.hadm_id, fp.icustay_id
),
urine_output AS (
   -- Step 3: Calculate total urine output from outputevents
   SELECT
       fp.subject_id,
       fp.hadm_id,
       fp.icustay_id,
       SUM(oe.value) AS total_urine_output
  FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
  LEFT JOIN `physionet-data.mimiciii_clinical.outputevents` oe
       ON fp.icustay_id = oe.icustay_id
   WHERE oe.itemid IN (40055, 43175, 40069, 40094, 40428, 40405, 40413) -- Urine
output ITEMIDs
       AND oe.value IS NOT NULL
   GROUP BY fp.subject_id, fp.hadm_id, fp.icustay_id
-- Combine DM type, infusion volume, and urine output
SELECT
  fp.subject_id,
   fp.hadm_id,
```

```
fp.icustay_id,
   dm.dm_type,
   ivc.infusion_volume, -- Single column for infusion volume
  uo.total_urine_output
FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
LEFT JOIN dm_type dm
  ON fp.subject_id = dm.subject_id
   AND fp.hadm_id = dm.hadm_id
LEFT JOIN infusion_volume_combined ivc
  ON fp.subject_id = ivc.subject_id
  AND fp.hadm_id = ivc.hadm_id
LEFT JOIN urine_output uo
   ON fp.subject_id = uo.subject_id
  AND fp.hadm_id = uo.hadm_id
ORDER BY fp.subject_id, fp.hadm_id;
Final Population Combined (EXCLUDING creatinine 48hr and 7 day values)
CREATE OR REPLACE TABLE
`proud-archery-435803-j0.healthcare_650.final_population_combined` AS
WITH dm_type AS (
   -- Step 1: Determine DM type from ICD-9 codes
  SELECT
       fp.subject_id,
       fp.hadm_id,
       fp.icustay_id,
       CASE
           WHEN di.icd9_code IN ('25001', '25003', '25011', '25013', '25021', '25023')
THEN 'T1DM' -- Type 1 Diabetes Mellitus
           WHEN di.icd9_code IN ('25000', '25002', '25010', '25012', '25020', '25022')
THEN 'T2DM' -- Type 2 Diabetes Mellitus
           ELSE 'Unknown'
       END AS dm_type
  FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
  LEFT JOIN `physionet-data.mimiciii_clinical.diagnoses_icd` di
       ON fp.hadm_id = di.hadm_id
)
SELECT
```

```
fp.subject_id,
fp.hadm_id,
fp.icustay_id,
-- Demographics
ANY_VALUE(d.gender) AS gender,
ANY_VALUE(d.age) AS age,
ANY_VALUE(d.ethnicity) AS ethnicity,
ANY_VALUE(d.weight) AS weight,
ANY_VALUE(d.height) AS height,
-- DM Type
ANY_VALUE(dm.dm_type) AS dm_type,
-- Complications: Aggregate to a single row per key
STRING_AGG(c.complication_type, '; ') AS complications,
-- Comorbidities: Aggregate to a single row per key
STRING_AGG(co.comorbidity, '; ') AS comorbidities,
-- Vitals
ANY_VALUE(v.heart_rate) AS heart_rate,
ANY_VALUE(v.systolic_bp) AS systolic_bp,
ANY_VALUE(v.diastolic_bp) AS diastolic_bp,
ANY_VALUE(v.respiratory_rate) AS respiratory_rate,
ANY_VALUE(v.temperature) AS temperature,
-- Lab Vitals
ANY_VALUE(lv.bicarbonate) AS bicarbonate,
ANY_VALUE(lv.wbc) AS wbc,
ANY_VALUE(lv.hemoglobin) AS hemoglobin,
ANY_VALUE(lv.platelets) AS platelets,
ANY_VALUE(lv.sodium) AS sodium,
ANY_VALUE(lv.chloride) AS chloride,
ANY_VALUE(1v.bun) AS bun,
ANY_VALUE(lv.creatinine) AS creatinine,
```

```
ANY_VALUE(lv.potassium) AS potassium,
   ANY_VALUE(lv.glucose) AS glucose,
   ANY_VALUE(lv.anion_gap) AS anion_gap,
   -- Scoring Systems
   ANY_VALUE(ss.saps_ii) AS saps_ii,
   ANY_VALUE(ss.oasis) AS oasis,
   ANY_VALUE(ss.sofa) AS sofa,
   ANY_VALUE(ss.gcs_eyes) AS gcs_eyes,
   ANY_VALUE(ss.gcs_motor) AS gcs_motor,
   ANY_VALUE(ss.gcs_verbal) AS gcs_verbal,
   -- Other Vital Info
   ANY_VALUE(ovi.infusion_volume) AS infusion_volume,
   ANY_VALUE(ovi.total_urine_output) AS total_urine_output
FROM `proud-archery-435803-j0.healthcare_650.final_population` fp
-- Join Demographics
LEFT JOIN `proud-archery-435803-j0.healthcare_650.demographics` d
   ON fp.subject_id = d.subject_id
   AND fp.hadm_id = d.hadm_id
   AND fp.icustay_id = d.icustay_id
-- Join DM Type
LEFT JOIN dm_type dm
   ON fp.subject_id = dm.subject_id
   AND fp.hadm_id = dm.hadm_id
   AND fp.icustay_id = dm.icustay_id
-- Join Complications: Aggregate to ensure one row per hadm_id and icustay_id
LEFT JOIN (
   SELECT
       subject_id,
       hadm_id,
       icustay_id,
```

```
STRING_AGG(complication_type, '; ') AS complication_type
   FROM `proud-archery-435803-j0.healthcare_650.complications`
  GROUP BY subject_id, hadm_id, icustay_id
) c
   ON fp.subject_id = c.subject_id
  AND fp.hadm_id = c.hadm_id
  AND fp.icustay_id = c.icustay_id
-- Join Comorbidities: Aggregate to ensure one row per hadm_id and icustay_id
LEFT JOIN (
  SELECT
       subject_id,
       hadm_id,
       icustay_id,
       STRING_AGG(comorbidity, '; ') AS comorbidity
  FROM `proud-archery-435803-j0.healthcare_650.comorbidities`
  GROUP BY subject_id, hadm_id, icustay_id
) co
  ON fp.subject_id = co.subject_id
  AND fp.hadm_id = co.hadm_id
   AND fp.icustay_id = co.icustay_id
-- Join Vitals
LEFT JOIN `proud-archery-435803-j0.healthcare_650.vitals` v
  ON fp.subject_id = v.subject_id
  AND fp.hadm_id = v.hadm_id
  AND fp.icustay_id = v.icustay_id
-- Join Lab Vitals
LEFT JOIN `proud-archery-435803-j0.healthcare_650.lab_vitals` lv
  ON fp.subject_id = lv.subject_id
  AND fp.hadm_id = lv.hadm_id
  AND fp.icustay_id = lv.icustay_id
-- Join Scoring Systems
LEFT JOIN `proud-archery-435803-j0.healthcare_650.scoring_systems` ss
```

```
ON fp.subject_id = ss.subject_id
   AND fp.hadm_id = ss.hadm_id
   AND fp.icustay_id = ss.icustay_id
-- Join Other Vital Info
LEFT JOIN `proud-archery-435803-j0.healthcare_650.other_vital_info` ovi
  ON fp.subject_id = ovi.subject_id
  AND fp.hadm_id = ovi.hadm_id
   AND fp.icustay_id = ovi.icustay_id
GROUP BY fp.subject_id, fp.hadm_id, fp.icustay_id
ORDER BY fp.subject_id, fp.hadm_id;
Creatinine values at 48hr and 7 day intervals:
CREATE OR REPLACE TABLE `proud-archery-435803-j0.healthcare_650.creatinine_intervals`
WITH first_icu_admissions AS (
  SELECT
       icu.subject_id,
       icu.hadm_id,
       icu.icustay_id,
       icu.intime,
       icu.outtime,
       ROW_NUMBER() OVER (PARTITION BY icu.hadm_id ORDER BY icu.intime) AS rn
  FROM
       `physionet-data.mimiciii_clinical.icustays` icu
   JOIN
       `proud-archery-435803-j0.healthcare_650.final_population_combined` fp
   ON icu.icustay_id = fp.icustay_id -- Ensure we only include ICU stays from
final_population_combined
),
filtered_icu AS (
  SELECT *
  FROM first icu admissions
  WHERE rn = 1 -- Keep only the first ICU admission per hospital admission
),
```

```
creatinine_data AS (
  SELECT
       le.subject_id,
       le.hadm_id,
       icu.icustay_id,
      le.charttime,
       le.valuenum AS creatinine,
       TIMESTAMP_DIFF(le.charttime, icu.intime, HOUR) AS hours_from_admission
  FROM
       `physionet-data.mimiciii_clinical.labevents` le
   JOIN
       filtered_icu icu
   ON le.hadm_id = icu.hadm_id
   AND le.subject_id = icu.subject_id
   WHERE
       le.itemid = 50912 -- Item ID for creatinine in MIMIC-III
       AND le.valuenum IS NOT NULL
       AND le.charttime BETWEEN icu.intime AND TIMESTAMP_ADD(icu.intime, INTERVAL 7
DAY) -- Ignore data beyond 7 days
),
split_intervals AS (
  SELECT
       subject_id,
       hadm_id,
       icustay_id,
       -- First creatinine value within 48 hours
       MIN(CASE
           WHEN hours_from_admission <= 48 THEN creatinine
       END) AS creatinine_48hrs,
       -- First creatinine value beyond 48 hours and within 7 days
       MIN(CASE
           WHEN hours_from_admission > 48 AND hours_from_admission <= 168 THEN
creatinine
       END) AS creatinine_7days
  FROM
       creatinine_data
```

```
GROUP BY
       subject_id, hadm_id, icustay_id
)
SELECT
   subject_id,
   hadm_id,
   icustay_id,
   creatinine_48hrs,
   creatinine_7days
FROM
   split_intervals
ORDER BY
   Hadm_id;
Join Creatinine intervals table to final population combined table:
CREATE OR REPLACE TABLE
`proud-archery-435803-j0.healthcare_650.final_population_combined` AS
SELECT
   -- All columns from final_population_combined
   fp.*,
   -- Columns from creatinine_intervals
   ci.creatinine_48hrs,
  ci.creatinine_7days
FROM
   `proud-archery-435803-j0.healthcare_650.final_population_combined` fp
LEFT JOIN
   `proud-archery-435803-j0.healthcare_650.creatinine_intervals` ci
ON
   fp.subject_id = ci.subject_id
   AND fp.hadm_id = ci.hadm_id
   AND fp.icustay_id = ci.icustay_id;
```