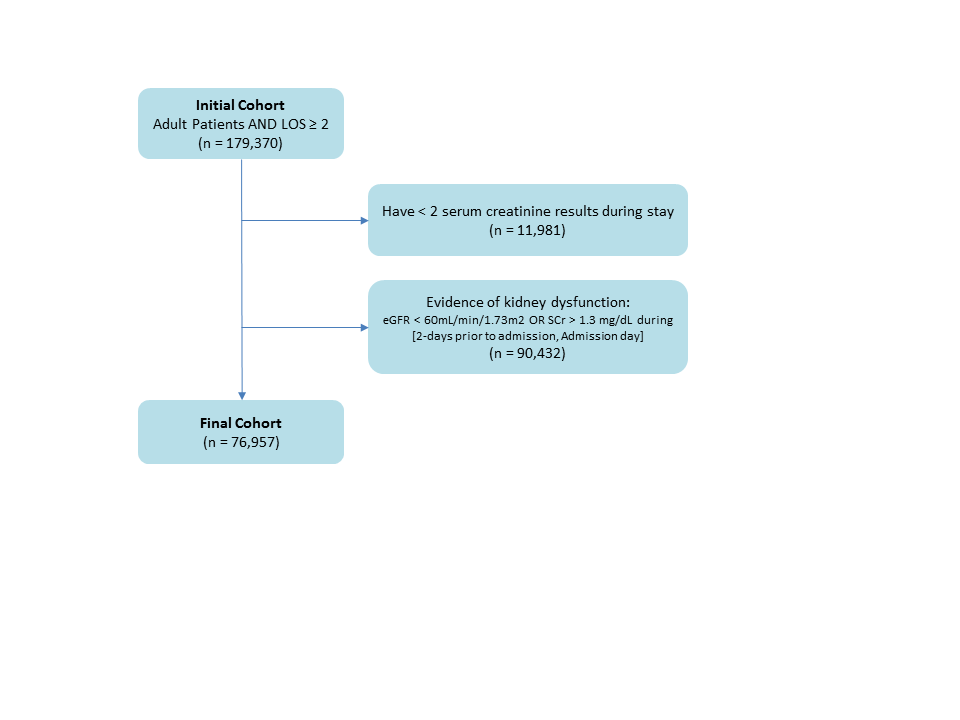
**Study Population**

All adult patients (age at visit >18) hospitalized for at least 2 days at a tertiary care, academic hospital (University of Kansas Health System – KUH) from November 2007 to December 2016 were included in the observational cohort study (*n* = 96,590 patients). Given that a patient may have multiple admissions (encounters) of at least 2 days and develop AKI during one but not another, this study is conducted at the encounter level with a total of 179,370 encounters. From the 179,370 encounters, we excluded those (a) missing necessary data for outcome determination, i.e. less than two serum creatinine measurements and (b) had evidence of moderate or severe kidney dysfunction, i.e. estimated Glomerular Filtration Rate (eGFR) less than 60 mL/min/1.73 m2 or abnormal serum creatinine (SCr) level of >1.3 mg/dL within 24 hours of hospital admission. A summary of patient encounter exclusions is shown in Figure 1. The final analysis cohort consisted of 76,957 encounters.



**Figure 1**. Summary of the study cohort after each step of exclusion

Then for each encounter, KUH’s de-identified clinical data repository HERON (Health Enterprise Repository for Ontological Narration) [cite1] was queried to obtain structured data corresponding to the encounter. HERON integrates data from KUH’s electronic health records (EHR), billing, clinical registries, and national data sources. The structured data extracted included demographic information, admission and discharge dates, medications, laboratory values, comorbidities, and admission diagnosis.

**AKI and Baesline Creatinine Definition**

AKI was defined according to the Kidney Disease Improving Global Outcomes (KDIGO) criteria [cite2] as described in Table 1.

**Table 1**. The KDIGO staging system for AKI

|  |  |
| --- | --- |
| **AKI Stage** | **Serum Creatinine (SCr) Criteria** |
| 1 | Increase >26.4 µmol/L (0.3 mg/dL) or 1.5-1.9 times baseline |
| 2 | Increase 2.0-2.9 times baseline |
| 3 | Increase creatinine >354 µmol/L (4.0 mg/dL) or 3 times baseline |

Baseline SCr level was defined as either the last measurement within 2-day time window prior to hospital admission or the first SCr measured after hospital admission. All SCr levels measured between admission and discharge were evaluated to determine the occurrence of in-hospital AKI. Out of total 76,957 encounters in the final analysis cohort, 7,259 encounters had any AKI of stage 1, 2, or 3 (9.43%) and 69,698 had no AKI events.

**AKI Risk Factors**

A summary of clinical variables used in building the AKI prediction models is described in Table 2. We referred to Matheny et al[31](#_ENREF_31) to select laboratory tests that may represent potential presence of a comorbidity that is correlated with in-hospital AKI. For example, an elevated white blood cell count (WBC) is associated with bacterial infection that may cause AKI. SCr was not included as a predictive variable as it was used to determine the AKI vs non-AKI encounters. For laboratory tests and vitals, only the last recorded value before a prediction point was used and their values were categorized. Laboratory values were categorized as either “present and normal”, “present and abnormal”, or “unknown” according to standard reference ranges. Vitals were categorized into groups as shown in Table 3. Missing values in vitals and lab tests were captured as “unknowns” because information may be contained in the choice to not perform the measurement.

Medication variable included inpatient (i.e. dispensed during stay) and outpatient medications (i.e. historical meds). All medication names were normalized by mapping to RxNorm ingredient. Comorbidity and admission diagnosis, i.e., all patient refined diagnosis related group (APR-DRG) variables were collected from the University Healthsystem Consortium (UHC) data source in HERON. Patient medical history was captured as major diagnoses (ICD-9 codes grouped according to the Clinical Classifications Software (CCS) diagnosis categories by the Agency for Healthcare Research and Quality). Medical history, medication, comorbidity and admission diagnosis variables took either “yes” or “no” values.

**Table 2.** Clinical variables considered in building predictive models for hospital-acquired AKI

|  |  |  |
| --- | --- | --- |
| Feature Category | # of Variables | Details |
| Demographics | 3 | Age, gender, race |
| Vitals | 5 | BMI, diastolic BP, systolic BP, pulse, temperature |
| Lab tests | 14 | Albumin, ALT, AST, Ammonia, Blood Bilirubin, BUN, Ca, CK-MB, CK, Glucose, Lipase, Platelets, Troponin, WBC |
| Comorbidities | 29 | UHC comorbidity |
| Admission diagnosis | 315 | UHC APR-DRG |
| Medications | 1271 | All medications are mapped to RxNorm ingredient |
| Medical History | 280 | ICD9 codes mapped to CCS major diagnoses |

Vitals, labs, medical history and medication variables were time-stamped relative to the admission date, referred here as time-dependent variables. Comorbidities, admission diagnosis, and demographics were presumed to be available at admission and not time-dependent.

**Table 3.** Categories for vital signs

|  |  |
| --- | --- |
| Vitals | Categories |
| BMI | <18.5, [18.5–24.9], [25.0–29.9], >30.0, Unknown |
| Diastolic BP | <80, [80–89], [90–99], >100, Unknown |
| Systolic BP | <120, [120–139], [140–159], >160, Unknown |
| Pulse | <50, [50–65], [66–80], [81–100], >100, Unknown |
| Temperature | <95.0, [95.0–97.6], [97.7–99.5], [99.5–104.0], >104.0, Unknown |

**References**

1. Waitman LR, Warren JJ, Manos EL, Connolly DW. Expressing observations from electronic medical record flowsheets in an i2b2 based clinical data repository to support research and quality improvement. AMIA Annu Symp Proc. 2011:1454-63, 2011.
2. KDIGO: Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO clinical practice guideline for acute kidney injury. Kidney Int Suppl 2: 1-138, 2012.