```
┌─────────────────────────────────────┐
│        Cases with 10+ visits,       │
│         304.* ICD-9 code,           │
│            20+ years old            │
│             n=11,573                │
└─────────────────────────────────────┘
                  ⇩
┌─────────────────────────────────────┐
│       Controls with 10+ visits,     │
│      No addiction ICD-9 codes,      │
│            20+ years old            │
│            n=828,062                │
└─────────────────────────────────────┘
                  ⇩
┌─────────────────────────────────────┐
│     Labs/vitals for cases + controls,│
│            outlier removal          │
│    Cases: 9,520; Controls: 707,015  │
└─────────────────────────────────────┘
                  ⇩
┌─────────────────────────────────────┐
│  Optional removal of sparse patients, labs, │
│              and vitals             │
│       Non-sparse Cases: 7,797       │
└─────────────────────────────────────┘
                  ⇩
┌─────────────────────────────────────┐
│    Case-control matching (1:10) with │
│  replacement by age, gender, race, HIV, │
│  hepatitis C, sickle-cell status using kNN │
└─────────────────────────────────────┘
```