Antibiotic treatment windows were created to identify high risk patients to prevent recurrent Clostridium difficile infection (rCDI). Time Zero (T0) was defined as day 0 for antibiotic treatment. These datasets and their associated methodology permit a more flexible modeling structure that permits rapid variation of datasets than previously available.

**Methods**

**Retrospective cohort study dataset using Kaiser Permanente Northern California data from 2007-2014.**

**Inclusion criteria:** adults ≥18 years of age with at least one positive enzyme linked immunosorbent assay or polymerase chain reaction CDI test associated with a hospitalization.

**Exclusion criteria:** patients who had a positive CDI test or CDI code for CDI in the 48 days prior to study entry.

**Data collection:** CDI-2 clinical data review, demographic and administrative data, inpatient beds, hospital encounters, microbiologic results, medications (e.g., antibiotics, proton pump inhibitors), laboratory tests, and vital signs.

**Two predictive models were developed using a derivation dataset from 01/01/2007-12/31/2012 to predict rCDI within 30 days and 48 days after the index antibiotic treatment (see figure).** Data from 2013 – 2014 will be used as a validation dataset.

**Time Zero (T0) was defined at the date/time when a CDI was ordered by the physician.**

**RESULTS**

- Within the derivation dataset, there were 7,694 adults with CDI
- 30 (6.1%) met criteria for rCDI within 30 days
- 470 (6.9%) met criteria for rCDI within 48 days
- 30-day model has a c-statistic of 0.662 and the 84-day model has a c-statistic of 0.629

**Conclusions**

- Availability of robust, detailed EMR data permits development of larger, generalizable, and more comprehensive CDI predictive models than previously available.

- This is the largest dataset to date that examines predictors for rCDI.

- The datasets and their associated methodology permit a more flexible modeling structure that permits rapid variation of specific eligibility criteria (e.g., time frame for recurrence vs. number of study sites), which enhances overall model utility.

- Models can help identify which CDI patients are at greatest risk for recurrence. This can help clinicians determine the most appropriate therapeutic strategy.

- Since we are in an era of gradual transition to comprehensive EMRs, one of the greatest challenges will be to define models that can be used in most settings, including those without comprehensive EMR data.

- Next steps are to conduct additional exercises to simplify and improve the model and validate it with additional sources of data, including clinical trial data.