Validation of a predictive model for Legionella culture-positivity (cx+) within a healthcare system water distribution system (WDS)

Brooke Decker1, Monique Bourdoux Kelly1, Jospeh Mikolok3, Jon Walker1, Cornelius Clancy*2

1VA Pittsburgh, 2University of Pittsburgh, 3VA Pittsburgh

Background: Hospitals devote considerable resources to WDS surveillance and remediation for Legionella. Rates of Legionnaires disease (LD) are highest in warm months, and have been linked to relative humidity (RH) and precipitation (ppt). Our goal was to build and validate a model based on weather and water parameters that predicts Legionella cx+ in our hospital WDS.

Materials/methods: 1 L water samples from fixtures at 2 campuses were cultured for Legionella on BCYE plates with cysteine as part of infection prevention protocols. Logistic regression (LR) and random forest (RF) models included daily hospital WDS measurements (14 d prior to cultures) and U.S. NOAA weather data (7 d). Training and validation used 2014-15 and 2016-17 data, respectively. Models predicted (within 14 d windows): 1) first +cx; 2) subsequent +cx after first +cx.

Results: >2000 cxs were performed each year. Data from 1,046 loop-specific d were collected. Overall, 1.8% of cx were + for Legionella. 14.6% (2014-15) and 13.6% (2016-17) of d had a first +cx. 53.6% (2014-15) and 81.2% (2016-17) of subsequent d were cx+. Significant predictors of first +cx were minimum/maximum WDS chlorine (Cl) (p=0.04, Wilcoxon) and RH (p=0.015). Significant predictors of subsequent +cx were minimum/maximum Cl (p≤0.003), RH (p=0.0003), and average 24-h ppt (p=0.006). The LR model misclassified 31% and 23% of first +cx and subsequent +cx, respectively. RF models built upon 9 parameters had AUCs=1 for predicting first +cx and subsequent +cx using training data. In validation studies, the RF model for first +cx had sensitivity/specificity of 100%/97% and PPV/NPV of 71%/100%. The RF model for subsequent +cx had sensitivity/specificity of 100%/99.7% and PPV/NPV of 94%/100%.

Conclusions: Models using water and weather data were validated as accurate predictors of new Legionella cx+ and subsequent Legionella cx+ within a hospital WDS. Most importantly, NPVs for the models were 100%, meaning that no positive Legionella cxs were recovered during periods identified as low-risk. The models are powerful tools for most efficiently directing resources to Legionella surveillance and LD prevention.