VALIDATING THE USE OF ELECTRONIC HEALTH RECORD DATA FOR POPULATION HEALTH SURVEILLANCE

Katharine H. McVeigh, PhD, MPH1; Remle Newton-Dame, MPH1; Lorna E. Thorpe, PhD2; Sharon E. Perlman, MPH2; Lauren Schreibstein, MA2; Elisabeth Snell, MPH3; Claudia Chernov, MPH1
1) Division of Epidemiology, 2) Primary Care Information Project, New York City Department of Health & Mental Hygiene, and 3) City University of New York School of Public Health

Introduction:
Networked electronic health records (EHRs) could play an important role in population health surveillance by providing data about:
• disease prevalence, management and control
• risk factor prevalence
• uptake of preventive services

The NYC Macroscope project is evaluating the validity of EHR-derived prevalence estimates for 13 outcomes:
• prevalence, management and control of hypertension, high cholesterol, and diabetes
• the prevalence of obesity, smoking and depression
• the use of preventive services (flu shots)

Analysis:
NYC Macroscope estimates have been stratified into 6 groups defined by age group (3 levels) and sex (2 levels). For each stratum, we compared:
1) The NYC Macroscope estimate against those of reference survey estimate
2) The prevalence ratio and the absolute and standardized differences between the two estimates

Summary measures of goodness of fit include the proportion of strata with significantly different estimates, the average differences between estimates across all 6 strata, and the Spearman correlation coefficient

Research Objective:
To evaluate the construct validity of an EHR-based indicator of smoking prevalence relative to an established survey measure

Data Sets and Sources:
NYC Macroscope estimates of smoking prevalence for 2013 are generated from EHR data aggregated across 467,983 patients at 384 practices throughout New York City. Reference survey estimates are obtained from the:
• 2013-14 New York City Health and Nutrition Examination Survey (NYC HANES; N = 1,543; 1,135 in care)
• 2013 NYC Community Health Survey (CHS; N = 8,698; 6,531 in care)

Study Design:
This analysis compares NYC Macroscope prevalence estimates, weighted to the age group, sex and neighborhood poverty distribution of the NYC adult population ages 20 and older that has seen a doctor in the past year (population in care), with estimates from the NYC HANES and CHS in-care populations

Principal Findings:
Benchmarking comparisons for the same self-reported measure of smoking behavior between CHS and HANES showed excellent alignment.

Compared to NYC HANES, Macroscope estimates were
• Not significantly different from NYC HANES estimates overall
• Were consistently lower than reference estimates in women but not men, with some significant differences

The fit between NYC Macroscope and CHS was extremely good and all fit criteria were met

Implications for Public Health Practice and Policy
• Estimates of smoking prevalence derived from EHR data are comparable to those from telephone and examination surveys, but may be somewhat lower than survey estimates for women
• These results are promising, but future studies assessing the sensitivity and specificity of the EHR-based indicator, both within the Macroscope context and across EHR vendors, are needed before validity can be established
• If we are able to demonstrate similar results across a variety of NYC Macroscope indicators, we have the potential to revolutionize local health surveillance