12-15-2016

Can Electronic Health Records Be Used for Population Health Surveillance? Validating Population Health Metrics Against Established Survey Data

Katharine H. McVeigh PhD, MPH
NYC Department of Health and Mental Hygiene, tmcveigh@health.nyc.gov

Remle Newton-Dame MPH
NYC Department of Health and Mental Hygiene (at the time of the study), remle.newton.dame@gmail.com

Pui Ying Chan MPH
NYC Department of Health and Mental Hygiene, pchan7@health.nyc.gov

Lorna E. Thorpe PhD
NYU School of Medicine, Department of Population Health, lorna.thorpe@sph.cuny.edu

See next pages for additional authors

The Electronic Data Methods (EDM) Forum is supported by the Agency for Healthcare Research and Quality (AHRQ), Grant 1U18HS022789-01. eGEMs publications do not reflect the official views of AHRQ or the United States Department of Health and Human Services.

Can Electronic Health Records Be Used for Population Health Surveillance? Validating Population Health Metrics Against Established Survey Data

Abstract

Introduction: Electronic health records (EHRs) offer potential for population health surveillance but EHR-based surveillance measures require validation prior to use. We assessed the validity of obesity, smoking, depression, and influenza vaccination indicators from a new EHR surveillance system, the New York City (NYC) Macroscope. This report is the second in a 3-part series describing the development and validation of the NYC Macroscope. The first report describes in detail the infrastructure underlying the NYC Macroscope; design decisions that were made to maximize data quality; characteristics of the population sampled; completeness of data collected; and lessons learned from doing this work. This second report, which addresses concerns related to sampling bias and data quality, describes the methods used to evaluate the validity and robustness of NYC Macroscope prevalence estimates; presents validation results for estimates of obesity, smoking, depression and influenza vaccination; and discusses the implications of our findings for NYC and for other jurisdictions embarking on similar work. The third report applies the same validation methods described in this report to metabolic outcomes, including the prevalence, treatment and control of diabetes, hypertension and hyperlipidemia.

Methods: NYC Macroscope prevalence estimates, overall and stratified by sex and age group, were compared to reference survey estimates for adult New Yorkers who reported visiting a doctor in the past year. Agreement was evaluated against 5 *a priori* criteria. Sensitivity and specificity were assessed by examining individual EHR records in a subsample of 48 survey participants.

Results: Among adult New Yorkers in care, the NYC Macroscope prevalence estimate for smoking (15.2%) fell between estimates from NYC HANES (17.7 %) and CHS (14.9%) and met all 5 *a priori* criteria. The NYC Macroscope obesity prevalence estimate (27.8%) also fell between the NYC HANES (31.3%) and CHS (24.7%) estimates, but met only 3 *a priori* criteria. Sensitivity and specificity exceeded 0.90 for both the smoking and obesity indicators. The NYC Macroscope estimates of depression and influenza vaccination prevalence were more than 10 percentage points lower than the estimates from either reference survey. While specificity was > 0.90 for both of these indicators, sensitivity was < 0.70.

Discussion: Through this work we have demonstrated that EHR data from a convenience sample of providers can produce acceptable estimates of smoking and obesity prevalence among adult New Yorkers in care; gained a better understanding of the challenges involved in estimating depression prevalence from EHRs; and identified areas for additional research regarding estimation of influenza vaccination prevalence. We have also shared lessons learned about how EHR indicators should be constructed and offer methodologic suggestions for validating them.

Conclusions: This work adds to a rapidly emerging body of literature about how to define, collect and interpret EHR-based surveillance measures and may help guide other jurisdictions.

Acknowledgements

The authors would like to thank Thomas Farley, Carolyn Greene, Jesse Singer, Elisabeth Snell, Laura Jacobson, Amy Freeman, Jesica Rodriguez-Lopez, Rhoda Schlamm, Kevin Konty, Stephen Immerwahr, Shadi Chamany, Sarah Shih, Tiffany Harris, James Hadler, and Charon Gwynn for their contributions to this work. This work has been made possible by the financial support of the de Beaumont Foundation, the Robert Wood Johnson Foundation including its National Coordinating Center for Public Health Services and Systems Research, the Robin Hood Foundation, the NY State Health Foundation, the Doris Duke Charitable Foundation, and the U.S. Centers for Disease Control and Prevention (U28EH000939). The contents of this paper are solely the responsibility of the authors and do not represent the official views of the funders.

Keywords
Population Health, Electronic Health Records, Surveillance, Validity, Chronic Disease

Disciplines
Epidemiology | Medicine and Health Sciences | Public Health

Authors
Katharine H McVeigh, *NYC Department of Health and Mental Hygiene*; Remle Newton-Dame, *NYC Department of Health and Mental Hygiene (at the time of the study)*; Pui Ying Chan, *NYC Department of Health and Mental Hygiene*; Lorna E Thorpe, *NYU School of Medicine, Department of Population Health*; Lauren Schreibstein, *NYC Department of Health and Mental Hygiene*; Kathleen S Tatem, *NYC Department of Health and Mental Hygiene*; Claudia Chernov, *NYC Department of Health and Mental Hygiene*; Elizabeth Lurie-Moroni, *NYC Department of Health and Mental Hygiene*; Sharon E Perlman, *NYC Department of Health and Mental Hygiene*.

eGEMs
Generating Evidence & Methods
to improve patient outcomes

Can Electronic Health Records Be Used for Population Health Surveillance? Validating Population Health Metrics Against Established Survey Data

Katharine H. McVeigh, PhD, MPH;[i] Remle Newton-Dame, MPH;[ii] Pui Ying Chan, MPH;[iii] Lorna E. Thorpe, PhD;[iv] Lauren Schreibstein, MS;[ii] Kathleen S. Tatem, MPH;[ii] Claudia Chernov, MPH;[i] Elizabeth Lurie-Moroni, MPH;[ii] Sharon E. Perlman, MPH[i]

ABSTRACT

Introduction: Electronic health records (EHRs) offer potential for population health surveillance but EHR-based surveillance measures require validation prior to use. We assessed the validity of obesity, smoking, depression, and influenza vaccination indicators from a new EHR surveillance system, the New York City (NYC) Macroscope. This report is the second in a 3-part series describing the development and validation of the NYC Macroscope. The first report describes in detail the infrastructure underlying the NYC Macroscope; design decisions that were made to maximize data quality; characteristics of the population sampled; completeness of data collected; and lessons learned from doing this work. This second report, which addresses concerns related to sampling bias and data quality, describes the methods used to evaluate the validity and robustness of NYC Macroscope prevalence estimates; presents validation results for estimates of obesity, smoking, depression and influenza vaccination; and discusses the implications of our findings for NYC and for other jurisdictions embarking on similar work. The third report applies the same validation methods described in this report to metabolic outcomes, including the prevalence, treatment and control of diabetes, hypertension and hyperlipidemia.

Methods: NYC Macroscope prevalence estimates, overall and stratified by sex and age group, were compared to reference survey estimates for adult New Yorkers who reported visiting a doctor in the past year. Agreement was evaluated against 5 a priori criteria. Sensitivity and specificity were assessed by examining individual EHR records in a subsample of 48 survey participants.

[i]New York City Department of Health and Mental Hygiene, [ii]Formerly New York City Department of Health and Mental Hygiene, [iii]New York City Department of Health and Mental Hygiene, [iv]NYU School of Medicine Department of Population Health

CONTINUED

Results: Among adult New Yorkers in care, the NYC Macroscope prevalence estimate for smoking (15.2%) fell between estimates from NYC HANES (17.7 %) and CHS (14.9%) and met all 5 a priori criteria. The NYC Macroscope obesity prevalence estimate (27.8%) also fell between the NYC HANES (31.3%) and CHS (24.7%) estimates, but met only 3 a priori criteria. Sensitivity and specificity exceeded 0.90 for both the smoking and obesity indicators. The NYC Macroscope estimates of depression and influenza vaccination prevalence were more than 10 percentage points lower than the estimates from either reference survey. While specificity was > 0.90 for both of these indicators, sensitivity was < 0.70.

Discussion: Through this work we have demonstrated that EHR data from a convenience sample of providers can produce acceptable estimates of smoking and obesity prevalence among adult New Yorkers in care; gained a better understanding of the challenges involved in estimating depression prevalence from EHRs; and identified areas for additional research regarding estimation of influenza vaccination prevalence. We have also shared lessons learned about how EHR indicators should be constructed and offer methodologic suggestions for validating them.

Conclusions: This work adds to a rapidly emerging body of literature about how to define, collect and interpret EHR-based surveillance measures and may help guide other jurisdictions.

Introduction

Across the United States, robust local population health surveillance systems are needed to guide and support policies and programs aimed at improving health outcomes. Local data are needed because disease burden, and risk and protective factors, can vary widely across communities within a single county or state. Timely and accurate community data provide the evidence base necessary to support locally relevant program and policy interventions and to measure their impact.[1] Data from geographically defined electronic health record (EHR) networks offer the promise of being timely and population specific. However, this promise is tempered by concerns about system governance, data quality, sampling bias, and a host of technical extraction-related issues, all of which influence both the ability to produce local prevalence estimates from EHR data and the data's accuracy.[2,3] In 2012, with support from external funders and in partnership with the City University of New York School of Public Health (CUNY), the New York City (NYC) Department of Health and Mental Hygiene (DOHMH) sought to test whether EHR data obtained from a convenience sample of more than 700 outpatient practices could be used to produce accurate estimates of population prevalence for NYC.

This novel EHR-based surveillance system, named the NYC Macroscope, was designed to measure health outcomes among the NYC adult population actively seeking medical care, defined as having visited a doctor in the reporting year

Volume 4 (2016) Issue Number 1

of interest. Health outcomes included prevalence, treatment, and control of diabetes, hypertension, and hyperlipidemia; prevalence of smoking, obesity, and depression; and uptake of vaccination against influenza. This report is the second in a three-part series describing the development and validation of the NYC Macroscope. The first report describes in detail the infrastructure underlying the NYC Macroscope, design decisions that were made to maximize data quality, characteristics of the population sampled, completeness of data collected, and lessons learned from doing this work.[4] This second report, which addresses concerns related to sampling bias and data quality, describes the methods used to evaluate the validity and robustness of NYC Macroscope prevalence estimates; presents validation results for estimates of obesity, smoking, depression and influenza vaccination; and discusses the implications of our findings for NYC and for other jurisdictions embarking on similar work. The third report applies the same validation methods described in this report to metabolic outcomes, including the prevalence, treatment and control of diabetes, hypertension and hyperlipidemia.[5]

Methods

We assessed the validity of the NYC Macroscope in two ways: (1) by comparing population-level NYC Macroscope estimates to in-care population estimates from two reference surveys—the gold standard 2013–2014 NYC Health and Nutrition Examination Survey (NYC HANES) and the 2013 NYC Community Health Survey (CHS); and (2) through a review of EHRs belonging to 48 NYC HANES participants who received primary care from a practice that contributed data to the NYC Macroscope estimates.

Design of the NYC Macroscope

The NYC Macroscope uses data from the Hub Population Health System (the Hub),[6] which is one of the largest ambulatory care data networks in the country. Contributing practices are located throughout NYC and are concentrated in low-income neighborhoods. Participating practices use eClinicalWorks EHR software and have signed agreements to share data with DOHMH. Data are collected using a distributed model. SQL language queries are sent to participating practices and aggregate counts are returned automatically to a secure database without transmitting patient-identifiable data. Providers who share data on the Hub regularly engage with DOHMH on improving documentation quality and on using EHRs to increase the delivery of needed preventive care, track chronic disease, and improve disease management.[7-10]

Hub data are transformed into NYC Macroscope data through filtering and weighting.[4] Filtering, which is intended to reduce double counting and improve data quality, limits records to primary care providers (internal medicine without a subspecialty, pediatrics, geriatrics, or family medicine) with at least 10 patients ages 20 years and older. Obstetricians, gynecologists, and specialists were excluded to minimize double counting of patients. Specialists were excluded also because they are less likely to address and document general health issues, including obesity, diabetes hyperlipidemia, smoking, depression, and influenza vaccination. Additional filtering restricts NYC Macroscope providers to those who meet documentation quality criteria that are largely aligned with the Centers for Medicare and Medicaid stage 1 Meaningful Use requirements for reimbursement.[11] Filtering for documentation quality reduced the number of contributing providers by 7.6 percent and the number of patient records by 5.5 percent.[4]

Patients of these providers were included in the 2013 NYC Macroscope sample if they were ages 20–100, had their sex recorded as male or female, resided in an NYC ZIP code, and had visited a provider in 2013 (were in care). The sample included 716,076 patients, representing 15.2 percent of the estimated 4.7 million NYC adults ages 20 and older who received primary care in the past year.[12] In most of the city, 10.0 percent to 19.9 percent of adults in care visited a NYC Macroscope provider in 2013, and coverage was 30.7 percent and 47.9 percent in the two neighborhoods with the deepest penetration. The unweighted NYC Macroscope population distribution, stratified by age group, sex, and neighborhood poverty, was similar to that of all NYC adults in care, though NYC Macroscope patients were slightly more likely to be younger and to have lower income. Compared with other NYC primary care providers, NYC Macroscope providers were less likely to be pediatricians, more likely to practice family medicine, and more likely to work in small sites of 1–5 providers.[4] To reduce the impact of patient and practice selection bias, each indicator was weighted to the sex (male, female), age group (20–39, 40–59, 60–100), and neighborhood poverty distribution of the adult NYC population in care. Neighborhood poverty was defined as the percent of the population in the patient's home ZIP code with an annual income below the federal poverty threshold (<10.0 percent, 10.0–19.9 percent, 20.0–29.9 percent, 30.0–100.0 percent).[13] Combined 2008–2012 ACS ZIP code approximations were used to identify the percent living in poverty.[14]

NYC Macroscope indicator definitions were developed based on three criteria: (1) information about EHR data element documentation quality from a previous Hub chart review study,[15] (2) indicator definitions used in the gold-standard NYC HANES survey,[16] and (3) consistency with national EHR-based measure sets such as Meaningful Use.[11] NYC Macroscope indicators were selected to capture potentially modifiable risk factors and conditions that contribute to a high burden of disease.[4]

Reference Data Sources and Analytic Sample

Other data sources for this study included two cross-sectional surveys—the gold standard NYC Health and Nutrition Examination Survey (NYC HANES) and the NYC Community Health Survey (CHS)—and primary care EHR data from a subsample of NYC HANES participants. In order to understand whether the EHR-based surveillance estimates were comparable to traditional survey estimates, survey data served as the reference data source against which the EHR data were evaluated in both population- and individual-levels analyses.

The primary reference survey, the 2013–2014 NYC HANES, was a population-based household examination survey of noninstitutionalized NYC residents ages 20 and older, modeled on the gold standard National Health and Nutrition Examination Survey (HANES).[16] National and local HANES are considered the "gold standard" in survey-based surveillance initiatives because blood pressure, height, and weight are measured in well-validated and standardized ways, and laboratory testing is conducted in research laboratories to ensure high-quality testing results. The NYC HANES sample consisted of 1,524 adults, of whom 1,135 reported having seen a health care provider in the past year (in care). NYC HANES data were statistically weighted to the 2013 American Community Survey (ACS) population of NYC adults ages 20 and older, and each outcome was adjusted for nonresponse by dropping all observations with missing values on that outcome before weighting. All estimates were limited to the in-care population and age adjusted to the U.S. 2000 Standard Population. See Thorpe et al. (2015) for more information about NYC HANES.[16]

Volume 4 (2016) Issue Number 1

The 2013 Community Health Survey (CHS) was a supplemental reference survey for this analysis. CHS is an annual, population-based, random-digit dialed telephone survey of adult New Yorkers, modeled on the Behavioral Risk Factor Surveillance System.[17] The 2013 CHS had a sample size of 8,698, of whom 6,166 were ages 20 and older and reported being in care. CHS data were weighted to the NYC population based on the 2010 U.S. Census, the 2012 ACS, and the 2011 NYC Housing and Vacancy Survey to represent NYC adults ages 18 and older in 2013. All estimates were then limited to the in-care population ages 20 and older and age adjusted to the U.S. 2000 Standard Population. More information about CHS can be found online.[17]

The 48 record chart-review sample was drawn from NYC HANES. Of 1,524 NYC HANES participants, 1,089 met eligibility criteria because they had reported visiting a health care provider within the previous year (i.e., "in care"), and did not have a proxy interview. Of these participants, 491 individuals signed a consent form and completed a Health Insurance Portability and Accountability Act wavier granting access to their medical records (45 percent consent rate). We were able to obtain printed copies of EHRs for 277 participants, of which 190 contained primary care data recorded within a year prior to the participant's NYC HANES interview. Of these 190 records, 48 were obtained from a NYC Macroscope provider.

Measures

All NYC Macroscope data were extracted from structured fields within the EHR. In NYC Macroscope, obesity was classified based on body mass index (BMI) calculated within the EHR from height and weight data. Height and weight were self-reported in CHS and measured in NYC HANES.

The NYC Macroscope smoking indicator was extracted from a dedicated field in the EHR that documented whether the patient was a current smoker. This field is tied to a prevention-oriented feature of the eClinicalWorks EHR software that reminds providers to assess patient smoking status annually. In NYC HANES and CHS, respondents were classified as smokers if they reported they had smoked at least 100 cigarettes in their lifetime and currently smoke.

Depression was captured in NYC Macroscope either by a Patient Health Questionnaire (PHQ 9)[18] screening with a score of 10 or higher (moderate depression) recorded in a dedicated field in the EHR, or by an ICD-9 code diagnosis of depression in the assessment or problem list sections of the EHR. Participants in NYC HANES were classified as depressed if they had a self-reported depression diagnosis (reported ever being told they had depression by a health care professional) or if they scored 10 or higher on the PHQ-9. Since CHS did not include the PHQ 9, the NYC Macroscope depression indicator was also evaluated against NYC HANES and CHS measures of self-reported depression diagnosis alone. We did not formally evaluate a depression measure that included medication for depression because those medications are often prescribed to treat other conditions.[19,20]

Receipt of influenza vaccination in the past year was captured by NYC Macroscope as the presence of an appropriate ICD-9, CPT, or CVX code. Vaccinations recorded in the unstructured portions of the EHR could not be captured by NYC Macroscope. NYC HANES and CHS used the same self-reported measure of having received an influenza vaccination in the past 12 months.

Statistical Analysis

NYC Macroscope data were weighted to generalize the findings from the Hub convenience sample of patients from practices that exchange data with the DOHMH to the target population of all adult New

Yorkers in care. For each indicator, the stratified, provider-level aggregate count data were pulled by the Hub, filtered using NYC Macroscope inclusion criteria, and converted to line-level data using Proc Freq in SAS software version 9.4 (SAS Institute Inc., Cary, N.C.). Records with missing outcomes were dropped (for smoking and obesity only), and the line-level data were weighted, separately for each outcome, to the age group (20–39, 40–59, 60–100), sex (male, female) and neighborhood poverty distributions (< 10 percent, 10–19 percent, 20–29 percent, >= 30 percent) of the NYC HANES and CHS populations in care. Patients from the same practice are not independent observations. To control for this, the NYC Macroscope population-based estimates were computed using SAS-callable (meaning that SUDAAN runs from within the SAS session) SUDAAN software, version 11.0 (Research Triangle Institute, Research Triangle Park, N.C.) using a sampling with replacement design and nested within practice. NYC HANES and CHS estimates were also computed using SAS-callable SUDAAN software to account for their complex survey designs. All estimates were age adjusted to the U.S. 2000 Standard Population to facilitate comparison with prevalence estimates across data sources.[21]

We compared population-level NYC Macroscope estimates with reference survey estimates for the population in care overall and stratified by sex and age group (20–39, 40–59, 60+). We assessed agreement based on five criteria: statistical equivalence,[22-25] statistical difference,[26-28] absolute prevalence difference,[29-31] prevalence ratio,[29] and internal consistency.[32,33,34] These criteria captured agreement across a variety of dimensions for prevalence estimates ranging in size from 12.6 percent to 47.6 percent. Statistical equivalence, which quantifies the probability that two estimates are equivalent within a predefined margin, and is not sensitive to sample size, was used to directly measure exchangeability. Statistical equivalence was evaluated using the two one-sided test of equivalence (TOST)[22-25] with a +/- 5 percentage point equivalence margin. Equivalence testing, which is required by the U.S. Food and Drug Administration for assessing bioequivalence of new drug formulations, has rarely been used in epidemiologic research.[22,25] Traditional epidemiologic assessment using the Student's t-test[26-28] was also carried out. Difference testing quantifies the probability that two estimates are different, but does not assess whether they are the same. A chief concern with difference testing in the context of evaluating estimates from EHR data is the method's sensitivity to sample size. With the NYC Macroscope sample exceeding 700,000 records, we were concerned that statistically significant differences might not be meaningful. For this reason, absolute and relative differences in estimate magnitude of 5 percentage points and 15 percent, respectively, were also assessed.[29] Finally, Spearman correlation coefficients with a threshold of 0.80 and scatterplots (not shown) were used to evaluate differential agreement across the six strata defined by age group and sex, and to identify poorly performing strata for further investigation.[32,33,34] The impact of adjustment for nonresponse on NYC Macroscope obesity and smoking estimates was also evaluated.

Measures of criterion-related validity, including percent agreement, Cohen's Kappa, sensitivity, and specificity were assessed relative to NYC HANES using the EHR data obtained from 48 NYC HANES participants who had received care from a NYC Macroscope provider in the year prior to their NYC HANES interview. Kappa was evaluated against criteria established by Landis and Koch[35] that characterize agreement as slight (Kappa: 0.0–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80) and almost perfect (0.81v1.0). Sensitivity was characterized as high (0.90–1.00), moderate

Volume 4 (2016) Issue Number 1

(0.70–0.89) and low (< 0.70), and specificity was characterized as high (0.90–1.00), moderate (0.80–0.89) and low (< 0.80). Data abstracted from unstructured fields in the EHR were used to assess whether sensitivity would have been higher if Hub queries had been able to access those data.

Results

Obesity

Completeness of NYC Macroscope Obesity Data. Obesity data were returned by 384 practices. Among the 703,978 patients in these practices, 7.8 percent were missing BMI. There was little difference in the percentage of missing BMI data by sex or by age group, and adjustment for nonresponse had no impact on the NYC Macroscope prevalence estimate.

Assessment of Validity. As seen in Table 1, the NYC Macroscope obesity-prevalence estimate for NYC adults in care of 27.8 percent fell between the objectively measured NYC HANES (31.3 percent) and self-reported CHS (24.7 percent) estimates. Although the comparison with NYC HANES failed both the TOST and t-test, it met all other a priori criteria for agreement. TOST and t-test comparisons with CHS gave mixed results, with poorest agreement among men ages 20–39 and women ages 40–59. The Spearman correlation across strata between NYC Macroscope and the reference surveys was 1.0 and 0.83 for NYC HANES and CHS, respectively (Table 2). Among the 44 chart-review participants with valid BMI data from both sources, the sensitivity of the NYC Macroscope BMI indicator relative to NYC HANES was 0.92 and the specificity was 0.97 (Table 3).

Smoking

Completeness of NYC Macroscope Smoking Data. Smoking status was documented for 468 219 patients at 382 practices. The percentage of patients with missing smoking status was 32.1 percent overall and ranged from 30.7 among New Yorkers from the poorest neighborhoods to 35.0 among New Yorkers ages 60 and older. The impact of adjustment for nonresponse on the overall prevalence estimate was less than 0.1 percentage points.

Table 1. Prevalence of Obesity, Smoking, Depression and Influenza Vaccination among Adults in Care, New York City, 2013

OUTCOME	2013 NYC MACROSCOPE[a] % (95% CI)	2013–2014 NYC HANES[b] % (95% CI)	2013 NYC CHS[c] % (95% CI)
Obesity	27.8 (27.7–27.9)	31.3 (28.5–34.2)	24.7 (23.2–26.3)
Smoking	15.2 (15.1–15.3)	17.7 (15.1–20.8)	14.9 (13.6–16.3)
Depression Self-Report (SR)[d]	8.2 (8.1–8.2) n/a	19.0 (16.4–21.9) 15.2 (13.0–17.7)	n/a 16.4 (15.1–7.9)
Influenza Vaccination	20.9 (20.8–21.0)	47.6 (44.0–51.3)	47.3 (45.5–49.0)

Notes: [a] Weighted to the NYC HANES distribution of the population in care.
[b] New York City Health and Nutrition Examination Survey.
[c] New York City Community Health Survey.
[d] Alternate definition: Self-reported diagnosis.

Table 2. Comparability of Prevalence Estimates of Obesity, Smoking, Depression, and Influenza Vaccination Across 2013 NYC Macroscope, 2013–2014 NYC HANES and 2013 CHS

EVALUATION CRITERIA	STATISTICALLY EQUIVALENT (TOST) $P < 0.05$	STATISTICALLY DIFFERENT (T-TEST) $P < 0.05$	PREVALENCE RATIO 0.85-1.15	PREVALENCE DIFFERENCE ± 5.0	INTERNAL CONSISTENCY $r \geq 0.80$
OBESITY					
NYC Macroscope vs. NYC HANES	0.14	0.02	**0.89**	**-3.5**	**1.0**
NYC Macroscope vs. CHS	**0.01**	<0.01	**1.13**	**3.2**	**0.83**
SMOKING					
NYC Macroscope vs. NYC HANES	**0.04**	**0.08**	**0.86**	**-2.6**	**0.83**
NYC Macroscope vs. CHS	**<0.01**	**0.85**	**1.01**	**0.1**	**0.94**
DEPRESSION					
NYC Macroscope vs. NYC HANES Self-Report (SR)*	>0.99 0.96	<0.01 <0.01	0.43 0.54	-10.8 -7.1	0.71 0.66
NYC Macroscope vs. CHS (SR)*	>0.99	<0.01	0.50	-8.2	**0.94**
INFLUENZA VACCINATION					
NYC Macroscope vs. NYC HANES	>0.99	<0.01	0.44	-26.7	**1.00**
NYC Macroscope vs. CHS	>0.99	<0.01	0.44	-26.3	**0.94**

Notes: **BOLD** entries meet a priori criteria for agreement; TOST = two one-sided test for statistical equivalence.
*Alternate definition: Self-reported diagnosis.

Assessment of Validity. The NYC Macroscope prevalence estimate for smoking among NYC adults in care (15.2 percent) fell between estimates from NYC HANES (17.7 percent) and CHS (14.9 percent) and met all a priori criteria (Tables 1 and 2). Among women ages 20–39 and 40–59, the NYC Macroscope estimate was lower than the NYC HANES estimate. The Spearman correlation across strata between NYC Macroscope and the reference surveys was 0.83 and 0.94 for NYC HANES and CHS, respectively. Among the 43 chart-review participants with valid EHR and NYC HANES smoking data, sensitivity and specificity relative to NYC HANES were both 1.0 (Table 3).

Volume 4 (2016) Issue Number 1

Table 3. Measures of Criterion-Related Validity of 2013 NYC Macroscope Indicator Definitions Relative to 2013–2014 NYC HANES from a Review of Individual EHRs

INDICATOR	% AGREEMENT	KAPPA	SENSITIVITY (95% CI)	SPECIFICITY (95% CI)
Obesity (n = 44)	95	0.89	0.92 (0.64, 1.00)	0.97 (0.83, 1.00)
Smoking (n = 43)	100	1.00	1.00 (0.54, 1.00)	1.00 (0.91, 1.00)
Depression (n = 48)	81	0.39	0.31 (0.09, 0.61)	1.00 (.90, 1.00)
Influenza Vaccination (n = 48)	81	0.61	0.64 (0.41, 0.83)	0.96 (0.80, 1.00)

Note: CI = Confidence Interval

Depression

Completeness of NYC Macroscope Depression Data. Depression data were available for 384 practices and 700,260 patients. The depression measure itself (consisting of either a diagnosis or a PHQ-9 score ≥10) had no missing data because the absence of a positive diagnosis was interpreted as "not depressed." However, 272 (70.8 percent) of these practices completed a PHQ 9 screening for less than 50 percent of their patients. Patients with no diagnosis and no screening may have been misclassified as not having depression. The percent of missing PHQ-9 data was higher in men, and increased with age and with higher neighborhood income. We were unable to adjust for nonresponse at the patient level because we had not used a nested approach to construct this compound indicator.

Assessment of Validity. NYC Macroscope estimates of depression prevalence among adult New Yorkers in care were 10.8 percentage points (57 percent) lower than NYC HANES estimates (Table 1) and showed relatively low internal consistency across strata (Spearman r = 0.71) (Table 2). When NYC HANES depression was defined only as self-report of diagnosis, prevalence dropped from 19.0 percent to 15.2 percent but remained higher than the NYC Macroscope prevalence of 8.2 percent based on both diagnosis and (inconsistent) PHQ-9 screening. The CHS self-reported depression prevalence among NYC adults in care was 16.4 percent, and the Spearman correlation between NYC Macroscope and CHS was 0.94. In the 48 charts reviewed, NYC Macroscope sensitivity relative to NYC HANES was 0.31 and specificity was 1.0 (Table 3). Incorporating unstructured data increased sensitivity to 0.38.

Influenza Vaccination

Completeness of NYC Macroscope Influenza Vaccination Data. Influenza vaccination data were returned on 712,043 patients from 391 practices. No patients were dropped from the denominator because the indicator could not differentiate between negative and missing vaccination status.

Assessment of Validity. NYC Macroscope estimates of influenza vaccination prevalence among adult New Yorkers in care were 26.7 percentage points (56 percent) lower than NYC HANES estimates (Table

1), but were perfectly correlated (Spearman r = 1.0), indicating high internal consistency across strata (Table 2). The relationship between NYC Macroscope and CHS mirrored the comparison with NYC HANES. In the 48 charts reviewed, NYC Macroscope influenza vaccination indicator sensitivity was 0.64 and specificity was 0.96 (Table 3). None of the eight NYC Macroscope false-negative influenza vaccination cases were reclassified based on findings in unstructured data. Only two patients were false negative for both influenza vaccination and depression.

Discussion

The NYC Macroscope indicators presented here demonstrated a wide range of strengths and weaknesses. We had hypothesized that obesity prevalence among adults in care would be well measured in the NYC Macroscope,[15] but found the NYC Macroscope prevalence to be 3.5 percentage points (11 percent) lower than the directly measured NYC HANES survey estimate. Nevertheless, the NYC Macroscope estimate was closer to the NYC HANES estimate than the estimate produced by the widely used CHS indicator, demonstrating that the NYC Macroscope indicator is acceptable for use in NYC. The difference between NYC HANES and CHS estimates was expected, as studies have demonstrated that people often overreport height and underreport weight when surveyed by telephone.[36] The reasons for the difference in prevalence between NYC Macroscope and NYC HANES are less clear. The sensitivity and specificity of the NYC Macroscope obesity indicator were 0.92 and 0.97, respectively, indicating there was little measurement error in this sample—a finding consistent with data from a 2015 anthropometric study that found little measurement error in weight and height data recorded by general practitioners.[37] Additionally, two previous studies in pediatric populations have found differences of less than 0.1 percentage point between EHR-derived, weighted obesity-prevalence estimates and National HANES estimates.[38,39] For these reasons, we suspect that the difference we observed in obesity prevalence between NYC Macroscope and NYC HANES was primarily attributable to differences in sample composition along dimensions other than age group, sex, and neighborhood poverty level. It may be that adult New Yorkers in care who are found at home and respond to a household survey are less active and thus have higher BMI than other adult New Yorkers in care.

Contrary to our original expectations,[15] smoking was well measured in NYC Macroscope, closely mirroring results obtained from the reference surveys and achieving perfect criterion-related validity in the EHR chart review. The amount of missing smoking data was substantial (32 percent) but essentially nondifferential (by age group, sex, and neighborhood poverty) in this sample from providers using a prevention-oriented EHR platform designed to support annual assessment of smoking status. We will be interested to learn from analysis of the other 142 medical records we have obtained whether our results are generalizable to providers who do not contribute data to the NYC Macroscope and who may not have provider alerts around smoking embedded in their EHR system. However, our results are consistent with findings from British[40,41] and U.S.[32] studies comparing EHR data with survey estimates, and with evidence of improvement in smoking history documentation over time.[42] The improvement in smoking history documentation in the U.S. was likely attributable to federal incentive payments to providers who met specific Meaningful Use criteria for EHRs, including structured documentation of smoking status in the EHR for at least 50 percent of their patients.[43]

These two indicators—obesity and smoking—demonstrated sufficient validity to be included in future iterations of the NYC Macroscope. In the majority of local and state jurisdictions, BMI and smoking data are limited to state- or county-level survey estimates only. Once validated locally, EHR-based indicators could be especially useful in providing geographic or population subgroup estimates, as Tabano et al. have done with obesity in Denver[44] and Linder et al. have done with smoking in Boston.[32] In jurisdictions that are able to monitor obesity and smoking prevalence through both local surveys *and* EHR systems, relative strengths and weaknesses of each data source can continue to be evaluated.

The poor performance of the NYC Macroscope depression indicator relative to NYC HANES is consistent with research findings that depression is underdiagnosed in the United States.[45-48] EHR-based depression indicators may perform better with widespread depression screening.[48] However, patients may answer the PHQ-9 differently when interviewed at home during a household survey than when screened in a primary care office. Furthermore, in the context of universal screening, the simple depression indicator definition we tested may not have been sufficient. While one Spanish study found that a simple depression indicator based only on diagnosis produced acceptable prevalence estimates,[49] other studies have demonstrated that achieving sufficient indicator sensitivity may require a more complicated definition that incorporates medications[19,50,51] as well as diagnoses recorded within unstructured fields.[50] In the 48 EHR charts reviewed, we found one instance (2 percent of records) of a depression diagnosis that had not been recorded in the structured field, and incorporation of that record into the indicator definition only marginally improved sensitivity. We chose not to include medications in our standard depression definition because medications used to treat depression are also prescribed to treat a number of other conditions. Other jurisdictions may prefer a different balance of sensitivity and specificity.

The NYC Macroscope influenza-vaccination prevalence estimate is less than half the survey estimates. The absence of vaccination documentation in structured fields of the EHR may be because vaccination was received in nontraditional settings, such as pharmacies and workplaces.[52,53] Other studies have also found low rates of influenza vaccination in EHR data relative to self-report, but attribute at least some of the difference to survey overreporting.[54,55] Further work is needed to determine whether the influenza vaccination indicator could be used to monitor trends in vaccination coverage over time, or how it could be used in conjunction with data from other sources, such as pharmacy vaccination sales data. EHR surveillance systems with the ability to incorporate data from unstructured fields may have better success in measuring influenza vaccination than we did, but results from our small chart-review sample are not promising. We expect that influenza vaccination prevalence will be better assessed using data from state or local Immunization Information Systems (IIS) or registries rather than from EHRs.

We learned a number of important lessons from our experience. First, after careful consideration, we selected the in-care population as the target population to which NYC Macroscope prevalence estimates were generalized. According to NYC HANES, 75 percent of the adult NYC population is in care. As we recently demonstrated elsewhere,[12] the population not in care is heterogeneous, and health profiles differ from in-care population profiles and differ as well among those not in care. The in-care adult population in NYC is more likely to be older, female, non-Hispanic, and insured compared with the not-in-care population. For this reason, we do not believe it is appropriate to generalize findings from the NYC Macroscope to the total population including persons not in care. As the proportion of uninsured New Yorkers declines pursuant to implementation of the Affordable Care Act, we anticipate that the proportion in care and represented by the NYC Macroscope will increase. Second, evaluating validity at both the individual- and population levels was important in assessing measurement error and, especially when little error was found, in quantifying sampling bias. For example, chart review demonstrated that obesity and smoking indicators had very little measurement error in NYC Macroscope, with sensitivities of 0.92 and 1.00, and specificities of 0.97 and 1.00, respectively. We were therefore able to attribute the differences between NYC Macroscope and NYC HANES estimates primarily to sampling bias. Third, we were able to use differences between NYC HANES and CHS estimates of obesity and smoking prevalence to inform our interpretation of differences between NYC Macroscope and NYC HANES. Fourth, while we adjusted our obesity and smoking estimates for nonresponse to reduce the impact of data missing differentially across strata, doing so did not change the obesity estimate and only changed the smoking estimate by 0.01 percentage points. Fifth, contrary to our expectation and findings from Wu et al.,[56] our chart review of 48 EHRs found that scanning unstructured fields only minimally improved indicator sensitivity. This finding is important and reassuring because natural language processing to extract unstructured data is not possible within NYC Macroscope and is complicated and burdensome in any setting.

Our study had a number of limitations. First, the sample of providers contributing to NYC Macroscope was not random, and the sample of patients excludes those who did not visit a NYC Macroscope provider. NYC Macroscope providers are unique in that they use a particular EHR platform and participate with DOHMH in data exchange and clinical quality improvement. Evaluating the impact of this limitation on NYC Macroscope prevalence estimates was the primary goal of this study. We were able to demonstrate that indicators with minimal measurement error produced prevalence estimates that were comparable to survey estimates. We are currently evaluating whether the criterion-related validity of NYC Macroscope indicators is generalizable beyond our unique provider sample through a review of 142 medical charts provided by 133 non-Macroscope providers and recorded on more than 20 EHR platforms.

Second, while NYC HANES served as our primary reference data source, it is not without its limitations. The sample size was small for some strata, which reduced the reliability of estimates for some groups. We should also point out that in our chart review we designated as the reference NYC HANES instead of the complete medical record. We did this to assess the utility of the NYC Macroscope estimates as potential replacements for survey data. We must acknowledge, however, that in some cases the medical record may better represent the true outcome.

Third, in NYC Macroscope we constructed several compound indicators based on information in the diagnosis fields as well as on objective measurement—i.e., depression (PHQ-9), blood pressure, A1C, and total cholesterol. These compound indicators were challenged by the lack of an explicit negative finding in the diagnosis field as well as by differential completion rates of the measurement. Nested approaches should potentially be taken to indicator construction so that diagnosis and measurement components can be evaluated both together and separately.

Last, the distributed data model upon which NYC Macroscope is built limits our ability to stratify NYC Macroscope estimates by factors not used

Volume 4 (2016) Issue Number 1

in weighting, including neighborhood. To directly estimate prevalence of a single outcome across NYC's 59 Community Districts, for example, currently requires 2,784 queries in addition to the standard 48. Recent system upgrades will soon allow us to stratify query results by residential neighborhood, but we will need to carefully assess sampling bias within each neighborhood to determine the most accurate approach for generating neighborhood prevalence estimates. These same upgrades will also make it possible to obtain estimates by race and ethnicity.

This robust validation study has many strengths. The well-established and temporally aligned reference data sources, NYC HANES and CHS, provided state-of-the-art surveillance estimates as comparisons with NYC Macroscope and, when compared to each other, provided empirical benchmarks of agreement. The assessment of validity at both the population- and individual levels, with a unique chart-review sample drawn from a population-based survey, provided insight into both measurement error and sampling bias. And, the set of metrics used to evaluate agreement against a priori criteria, including tests of equivalence, absolute and relative difference, and internal consistency, provided a multidimensional assessment of validity that enabled the evaluation of outcomes with different prevalence magnitudes within a single analytic framework.

Conclusions

Through this work we have developed evidence for the validity of the obesity and smoking prevalence estimates produced by the NYC Macroscope, gained a better understanding of the challenges involved in estimating depression prevalence from EHRs, and documented that EHR data alone are insufficient to measure influenza vaccination prevalence. We have also demonstrated approaches that other researchers may find useful for evaluating the validity of EHR-based surveillance indicators and shared lessons learned about how EHR indicators should be constructed. This work adds to a rapidly emerging body of literature about how to define, collect, and interpret EHR-based surveillance measures and may help guide other jurisdictions.

Acknowledgements

The authors would like to thank Thomas Farley, Carolyn Greene, Jesse Singer, Elisabeth Snell, Laura Jacobson, Amy Freeman, Jesica Rodriguez-Lopez, Rhoda Schlamm, Kevin Konty, Stephen Immerwahr, Shadi Chamany, Sarah Shih, Tiffany Harris, James Hadler, and Charon Gwynn for their contributions to this work. This work has been made possible by the financial support of the de Beaumont Foundation, the Robert Wood Johnson Foundation including its National Coordinating Center for Public Health Services and Systems Research, the Robin Hood Foundation, the NY State Health Foundation, the Doris Duke Charitable Foundation, and the U.S. Centers for Disease Control and Prevention (U28EH000939). The contents of this paper are solely the responsibility of the authors and do not represent the official views of the funders.

References

1. Castrucci BC, Rhoades EK, Leider JP, Hearne S. What Gets Measured Gets Done: An Assessment of Local Data Uses and Needs in Large Urban Health Departments. Joural of public health management and practice. 2015;21(1 Supp): S38–S48.
2. Paul MM, Greene CM, Newton-Dame R, Thorpe LE, Perlman SE, McVeigh KH, et al. The state of population health surveillance using electronic health records: a narrative review. Population health management. 2015;18(3):209-16.
3. McVeigh KH, Newton-Dame R, Perlman S, Chernov C, Thorpe L, Singer J, Greene C. Developing an Electronic Health Record-Based Population Health Surveillance System. New York: New York City Department of Health and Mental Hygiene, 2013.
4. Newton-Dame R, McVeigh KH, Schreibstein L, Perlman S, Lurie E, Greene CM, et al. Design of the New York City Macroscope: Innovations in Population Health Surveillance using Electronic Health Records. http://repository.edm-forum.org/egems/vol4/iss1/26/ 2016.

5. Thorpe LE, McVeigh KH, Perlman SE, Chan PY, Bartley K, Scheribstein L, et al. Monitoring Prevalence, Treatment, and Control of Metabolic Conditions in New York City Adults Using 2013 Primary Care Electronic Health Records: A Surveillance Validation Study. http://repository.edm-forum.org/egems/vol4/iss1/28/ 2016.
6. Buck MD, Anane S, Taverna J, Amirfar S, Stubbs-Dame R, Singer J. The Hub Population Health System: distributed ad hoc queries and alerts. Journal of the American Medical Informatics Association : JAMIA. 2012;19(e1):e46-50.
7. Ryan AM, McCullough CM, Shih SC, Wang JJ, Ryan MS, Casalino LP. The intended and unintended consequences of quality improvement interventions for small practices in a community-based electronic health record implementation project. Medical care. 2014;52(9):826-32.
8. Sebek KM, Virkud A, Singer J, Pulgarin CP, Schreibstein L, Wang JJ. Preliminary evaluation of a comprehensive provider feedback report. Journal of medical practice management. 2014;29(6):397-405.
9. Mostashari F, Tripathi M, Kendall M. A tale of two large community electronic health record extension projects. Health affairs (Project Hope). 2009;28(2):345-56.
10. Summer L. Using Health Information Technology to Improve Health and Health Care in Underserved Communities: The Primary Care Information Project [Internet]. Washington: AcademyHealth [n.d.; cited 2015 Dec 2]. Available from: https://www.academyhealth.org/files/publications/HIT4AKPCIP.pdf.
11. Centers for Medicare and Medicaid Services. Medicare & Medicaid EHR Incentive Program Meaningful Use Stage 1 Requirements Overview [Internet]. Maryland: Centers for Medicare and Medicaid Services; 2010 [cited 2016 Mar 31]. Available from: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/MU_Stage1_ReqOverview.pdf.
12. Romo ML, Chan PY, Lurie-Moroni E, Perlman SE, Newton-Dame R, Thorpe LE, et al. Characterizing Adults Receiving Primary Medical Care in New York City: Implications for Using Electronic Health Records for Chronic Disease Surveillance. Preventing chronic disease. 2016;13:E56.
13. Toprani A, Hadler JL. Selecting and applying a standard area-based socioeconomic status measure for public health data: analysis for New York City. New York City Department of Health and Mental Hygiene: Epi Research Report, May 2013; 1-11.
14. United States Census Bureau. What is the American Community Survey? [Internet]. Maryland: United States Census Bureau [updated Jun; cited 2015 Nov 2]. Available from: http://www.census.gov/programs-surveys/acs/about.html.
15. Parsons A, McCullough C, Wang J, Shih S. Validity of electronic health record-derived quality measurement for performance monitoring. Journal of the American Medical Informatics Association : JAMIA. 2012;19(4):604-9.
16. Thorpe LE, Greene C, Freeman A, Snell E, Rodriguez-Lopez JS, Frankel M, et al. Rationale, design and respondent characteristics of the 2013–2014 New York City Health and Nutrition Examination Survey (NYC HANES 2013–2014). Preventive Medicine Reports. 2015;2:580-5.
17. New York City Department of Health and Mental Hygiene. Community Health Survey: Public Use Data [Internet]. New York: New York City Department of Health and Mental Hygiene [n.d.; cited 2015 Dec 2]. Available from: http://www.nyc.gov/html/doh/html/data/chs-data.shtml.
18. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. Journal of general internal medicine. 2001;16(9):606-13.
19. Coleman N, Halas G, Peeler W, Casaclang N, Williamson T, Katz A. From patient care to research: a validation study examining the factors contributing to data quality in a primary care electronic medical record database. BMC family practice. 2015;16:11.
20. Gardarsdottir H, Egberts AC, van Dijk L, Sturkenboom MC, Heerdink ER. An algorithm to identify antidepressant users with a diagnosis of depression from prescription data. Pharmacoepidemiology and drug safety. 2009;18(1):7-15.
21. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People 2010 Statistical Notes No. 20. January 2001.
22. Barker LE, Luman ET, McCauley MM, et al. Assessing equivalence: an alternative to the use of difference tests for measuring disparities in vaccination coverage. Am J Epidemiol 2002;156:1056–61.
23. Walker E, Nowacki AS. Understanding equivalence and noninferiority testing. Journal of general internal medicine. 2011;26(2):192-6.
24. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. Journal of pharmacokinetics and biopharmaceutics. 1987;15(6):657-80.
25. Liu H, Cella D, Gershon R, et al. Representativeness of the Patient-Reported Outcomes Measurement Information System internet panel. J Clin Epidemiol 2010;63:1169–78.
26. Student. The probable error of the mean. Biometrika. 1908;6(1):1-25.
27. Lumley T, Diehr P, Emerson S, Chen L. The importance of the normality assumption in large public health data sets. Annual reviw of public health. 2002; 23:151–69.
28. Yun S, Zhu BP, Black W, Brownson RC. A comparison of national estimates of obesity prevalence from the behavioral risk factor surveillance system and the National Health and Nutrition Examination Survey. International journal of obesity (2005). 2006;30(1):164-70.
29. Yi SS, Johns M, Lim S. Use of regional data to validate and recalibrate self-reported hypertension: highlighting differences in immigrant groups in New York City. Journal of immigrant and minority health / Center for Minority Public Health. 2015.
30. Agaku IT, Awopegba AJ, Filippidis FT. The impact of inter-survey differences in the definition of current smokeless tobacco use on comparability of US national and state-specific prevalence estimates, 2009-2011. Preventive medicine. 2015;74:86-92.

31. Li C, Balluz LS, Ford ES, Okoro CA, Zhao G, Pierannunzi C. A comparison of prevalence estimates for selected health indicators and chronic diseases or conditions from the Behavioral Risk Factor Surveillance System, the National Health Interview Survey, and the National Health and Nutrition Examination Survey, 2007-2008. Preventive medicine. 2012;54(6):381-7.
32. Linder JA, Rigotti NA, Brawarsky P, Kontos EZ, Park ER, Klinger EV, et al. Use of practice-based research network data to measure neighborhood smoking prevalence. Preventing chronic disease. 2013;10:E84.
33. Le A, Judd SE, Allison DB, Oza-Frank R, Affuso O, Safford MM, et al. The geographic distribution of obesity in the US and the potential regional differences in misreporting of obesity. Obesity (Silver Spring, Md). 2014;22(1):300-6.
34. Lee Dc, Long JA, Wall SP, Carr BG, Satchell SN, Braithwaite RS, Elbel B. Determining chronic disease prevalence in local populations using emergency department surveillance. American Journal of Public Health: September 2015; 105(9):e67-e74.
35. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics. 1977;33(2):363-74.
36. Connor Gorber S, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. Obesity reviews : an official journal of the International Association for the Study of Obesity. 2007;8(4):307-26.
37. Sebo P, Haller D, Pechère-Bertschi A, Bovier P, Herrmann F. Accuracy of doctors' anthropometric measurements in general practice. Swiss Med Wkly. 2015 Feb 21;145:w14115
38. Flood TL, Zhao YQ, Tomayko EJ, Tandias A, Carrel AL, Hanrahan LP. Electronic health records and community health surveillance of childhood obesity. American journal of preventive medicine. 2015;48(2):234-40.
39. Bailey LC, Milov DE, Kelleher K, Kahn MG, Del Beccaro M, Yu F, et al. Multi-Institutional Sharing of Electronic Health Record Data to Assess Childhood Obesity. PloS one. 2013;8(6):e66192.
40. Szatkowski L, Lewis S, McNeill A, Huang Y, Coleman T. Can data from primary care medical records be used to monitor national smoking prevalence? Journal of epidemiology and community health. 2012;66(9):791-5.
41. Booth HP, Prevost AT, Gulliford MC. Validity of smoking prevalence estimates from primary care electronic health records compared with national population survey data for England, 2007 to 2011. Pharmacoepidemiology and drug safety. 2013;22(12):1357-61.
42. Chen LH, Quinn V, Xu L, Gould MK, Jacobsen SJ, Koebnick C, et al. The accuracy and trends of smoking history documentation in electronic medical records in a large managed care organization. Substance use & misuse. 2013;48(9):731-42.
43. Centers for Medicare and Medicaid Services. Eligible Professional Meaningful Use Core Measures Measure 9 of 13 [Internet]. Maryland: Centers for Medicare and Medicaid Services [updated 2014 May; cited 2015 Dec 4]. Available from: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf.
44. Tabano D, Barrow J, McCormick E, Bol K, Anthamatten P, Thomas D, et al. PS2-20: Obesity Mapping in Colorado: A Novel System for Monitoring and Tracking BMI. Clinical Medicine & Research. 2014;12(1-2):83.
45. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). Jama. 2003;289(23):3095-105.
46. Nichols GA, Brown JB. Following depression in primary care: do family practice physicians ask about depression at different rates than internal medicine physicians? Archives of family medicine. 2000;9(5):478-82.
47. Gwynn RC, McQuistion HL, McVeigh KH, Garg RK, Frieden TR, Thorpe LE. Prevalence, diagnosis, and treatment of depression and generalized anxiety disorder in a diverse urban community. Psychiatric services (Washington, DC). 2008;59(6):641-7.
48. U.S. Preventive Services Task Force. Draft Recommendation Statement: Depression in Adults: Screening [Internet]. Maryland: U.S. Preventive Services Task Force [updated 2015 Jul; cited 2015 Nov 25]. Available from: http://www.uspreventiveservicestaskforce.org/Page/Document/draft-recommendation-statement115/depression-in-adults-screening1.
49. Violan C, Foguet-Boreu Q, Hermosilla-Perez E, Valderas JM, Bolibar B, Fabregas-Escurriola M, et al. Comparison of the information provided by electronic health records data and a population health survey to estimate prevalence of selected health conditions and multimorbidity. BMC public health. 2013;13:251.
50. Williamson T, Green ME, Birtwhistle R, Khan S, Garies S, Wong ST, et al. Validating the 8 CPCSSN case definitions for chronic disease surveillance in a primary care database of electronic health records. Annals of family medicine. 2014;12(4):367-72.
51. Orueta JF, Nuno-Solinis R, Mateos M, Vergara I, Grandes G, Esnaola S. Monitoring the prevalence of chronic conditions: which data should we use? BMC health services research. 2012;12:365.
52. Sy LS, Liu IL, Solano Z, Cheetham TC, Lugg MM, Greene SK, et al. Accuracy of influenza vaccination status in a computer-based immunization tracking system of a managed care organization. Vaccine. 2010;28(32):5254-9.
53. Kwong JC, Manuel DG. Using OHIP physician billing claims to ascertain individual influenza vaccination status. Vaccine. 2007;25(7):1270-4.
54. Jimenez-Garcia R, Hernandez-Barrera V, Rodriguez-Rieiro C, Carrasco Garrido P, Lopez de Andres A, Jimenez-Trujillo I, et al. Comparison of self-report influenza vaccination coverage with data from a population based computerized vaccination registry and factors associated with discordance. Vaccine. 2014;32(35):4386-92.
55. Rolnick SJ, Parker ED, Nordin JD, Hedblom BD, Wei F, Kerby T, et al. Self-report compared to electronic medical record across eight adult vaccines: do results vary by demographic factors? Vaccine. 2013;31(37):3928-35.
56. Wu CY, Chang CK, Robson D, Jackson R, Chen SJ, Hayes RD, et al. Evaluation of smoking status identification using electronic health records and open-text information in a large mental health case register. PloS one. 2013;8(9):e74262.