Boston University

OpenBU http://open.bu.edu

Theses & Dissertations Boston University Theses & Dissertations

2019

Comparative evaluation of methods that adjust for reporting biases in participatory surveillance systems

https://hdl.handle.net/2144/38799
Boston University

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

COMPARATIVE EVALUATION OF METHODS THAT ADJUST FOR REPORTING BIASES IN PARTICIPATORY SURVEILLANCE SYSTEMS

by

KRISTIN BALTRUSAITIS

BS, Lehigh University, 2006
MS, Brooklyn College, 2008
MA, Boston University, 2014

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2019

© Copyright by
KRISTIN BALTRUSAITIS
2019

Approved by

First Reader

Laura Forsberg White, PhD
Associate Professor, Biostatistics

Second Reader

Helen Jenkins, PhD
Assistant Professor, Biostatistics

Third Reader

Paola Sebastiani, PhD
Professor, Biostatistics

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Laura White for the continuous support of my Ph.D. study and research. Thank you to the rest of my dissertation committee: Mauricio Santillana, Paola Sebastiani, Helen Jenkins, and Elaine Nsoesie for their insightful comments, questions, and advice. I would also like to thank my research assistantship advisor, Joseph Massaro, and fellow RAs: Jeremiah Perez, Ellie Gurary, and Taylor Mahoney for making the unit a wonderful place to learn and laugh. Thank you to all of my friends and classmates, especially Samantha Lent, Dan Posner, Aya Mitani, Darlene Lu, and Kendra Plourde for making qual studying bearable and post-qual partying amazing. Most of all, I want to thank my family (Joe, Peg, Michael, Mark, and all the little ones), Christopher Reeder, Shannon Sorenson, and Rita, for all of their unwavering support, encouragement, and love that enabled me to have a worthwhile life outside of BU.

COMPARATIVE EVALUATION OF METHODS THAT ADJUST FOR REPORTING BIASES IN PARTICIPATORY SURVEILLANCE SYSTEMS

KRISTIN BALTRUSAITIS

Boston University, Graduate School of Arts and Sciences, 2019

Major Professor: Laura Forsberg White, PhD

ABSTRACT

Over the past decade the widespread proliferation of mobile devices and wearable technology has significantly changed the landscape of epidemiological data gathering and evolved into a field known as Digital Epidemiology. One source of active digital data collection is online participatory syndromic surveillance systems. These systems actively engage the general public in reporting health-related information and provide timely information about disease trends within the community. This dissertation comprehensively addresses how researchers can effectively use this type of data to answer questions about Influenza-like Illness (ILI) disease burden in the general population. We assess the representativeness and reporting habits of volunteers for these systems and use this information to develop statistically rigorous methods that adjust for potential biases. Specifically, we evaluate how different missing data methods, such as complete case and multiple imputation models, affect estimates of ILI disease burden using both simulated data as well as data from the Australian system, Flutracking.net. We then extend these methods to data from the American system, Flu Near You, which has different patterns of participant reporting, and evaluate additional methods of bias adjustment.

Finally, we provide examples of how this data has been used to answer questions about ILI in the general community and promote better understanding of disease surveillance and data literacy among volunteers.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AR	Attack Rate
ARI	Acute Respiratory Infection
AU	Australia
BRFSS ...	Behavioral Risk Factor Surveillance Survey
BPHC ...	Boston Public Health Commission
CDC	Centers for Disease Control and Prevention
CI	Confidence Interval
DE	Digital Epidemiology
ED	Emergency Department
EHR	Electronic Health Records
ER	Emergency Room
FNY	Flu Near You
HDI	Human Development Index
HHS	Health and Human Services
ILI	Influenza-Like Illness
ILINet ...	Outpatient Influenza-like Illness Surveillance Network
IP	Incidence Proportion
IQR	Interquartile Range
IR	Incidence Rate
FluSurv-NET	Influenza Hospital Surveillance Network
MAR	Missing at Random
MCAR ...	Missing Completely at Random
MI	Multiple Imputation

MICE . . .	Multivariate Imputation by Chained Equations
MMWR . .	Morbidity and Mortality Weekly Report
MNAR . . .	Missing Not at Random
MoAR . . .	Motivated at Random
MoCAR . .	Motivated Completely at Random
MoNAR . .	Motivated Not at Random
NCHS . . .	National Center for Health Statistics
NNDSS . .	National Notifiable Diseases Surveillance System
NRMSE . .	Normalized Root Mean Square Error
NREVSS . .	National Respiratory and Enteric Virus Surveillance System
OR	Odds Ratio
SciFri . . .	Science Friday
SES	Socioeconomic status
SIR	Susceptible-Infected-Recovered
US	United States of America
WHO . . .	World Health Organization
WP	Weekly Prevalence

CHAPTER 1

Introduction

1.1 INFLUENZA SURVEILLANCE

1.1.1 Global

Every year influenza epidemics are responsible for substantial clinical and economic burdens that include an estimated 290 000 to 650 000 deaths worldwide.(Lee et al., 2018c; Putri et al., 2018) National estimates of disease burden in the population are essential to track the potential impact of influenza on hospitalizations and death, aid in clinical resource allocation decisions, assess vaccine effectiveness, and understand the overall global burden of influenza disease.(Lee et al., 2018c; Lipsitch et al., 2011) In 2015, the World Health Organization (WHO) released a manual for estimating seasonal influenza burden. The primary goal of this manual is to provide a guideline for countries who do not have established burden of disease studies because building accurate global burden of disease estimates for influenza requires better data from all regions of the world.(World Health Organization, 2015)

1.1.2 United States of America

In the United States of America (U.S.), the Centers for Disease Control and Prevention (CDC) has a national influenza surveillance system that collects and reports weekly data from five different categories of surveillance: virology, outpatient illness, mortality, hospitalization, and geographical spread. These systems provide a picture of national influenza activity that determines the location and timing of influenza activity, defines the types and subtypes of circulating influenza

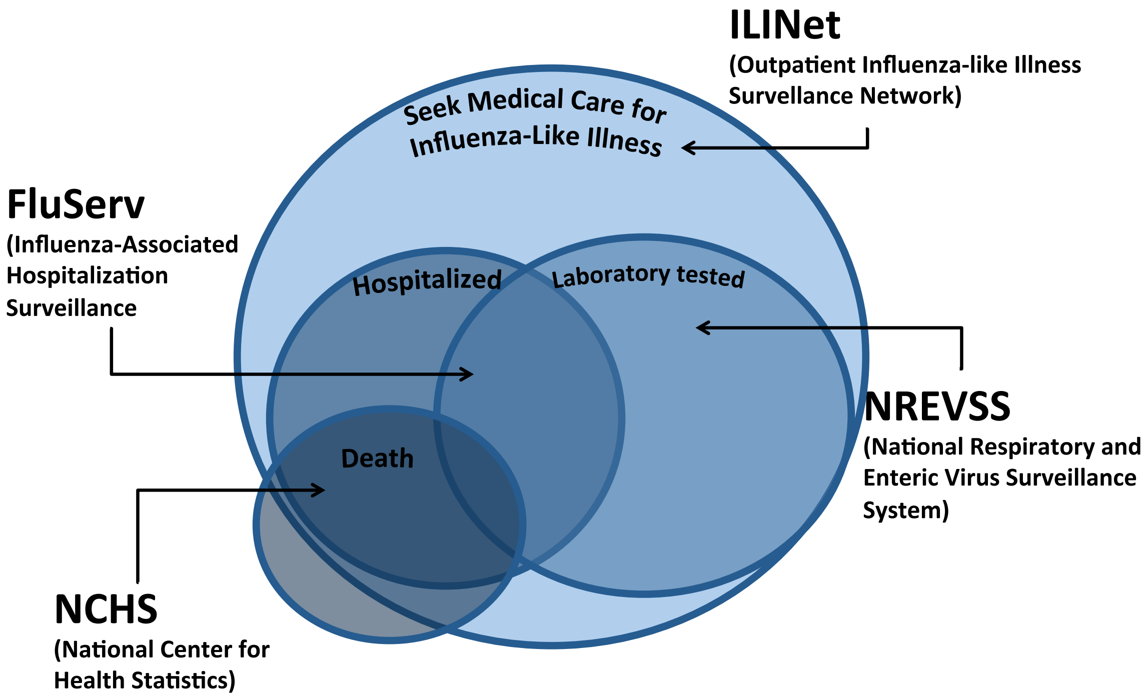

Figure 1.1: Schematic of CDC influenza surveillance.

viruses, detects changes in circulating viruses, tracks Influenza-Like Illness (ILI), and measures the impact of influenza-related hospitalizations and deaths (Figure 1.1). (Thompson et al., 2006) Although influenza surveillance occurs throughout the calendar year, the influenza season is defined by the Morbidity and Mortality Weekly Report (MMWR) week 40 through week 20, which usually corresponds with months October through May. The peak of the influenza season typically occurs between December and March.

Influenza surveillance in the U.S. is robust and comprehensive, however, there are several limitations to this system. Because state and health care provider reporting is voluntary and each state is responsible for the recruitment of healthcare providers, the composition of provider-types, coverage of geographical regions, and provider reporting vary from state to state. This convenience sample-driven model of surveillance results in certain parts of the population being over or under represented in the reported influenza activity.(Polgreen et al., 2009a; Lee et al., 2018a; Scarpino et al., 2012) Furthermore, all estimates are based only on those who seek medical care and there is an inherent delay of approximately 1-2 weeks between the day of visit and published date of estimates. Together, these systems can provide an indication of where, when, and what influenza viruses are circulating, but they do not provide the actual number of influenza infections during a season.(Thompson et al., 2006)

1.2 DIGITAL EPIDEMIOLOGY

1.2.1 Overview

Over the past decade the widespread proliferation of mobile devices and wearable technology has significantly changed the landscape of epidemiological data gathering and evolved into a field known as Digital Epidemiology (DE).(Salathé, 2018; Choi et al., 2016) DE provides an informal, complementary approach to traditional sentinel surveillance methods by leveraging data generated outside of the public health system through digital data sources, such as Google (Ginsberg et al., 2009), Yahoo (Polgreen et al., 2008), and Baidu (Yuan et al., 2013) Internet searches; Twitter posts (Signorini et al., 2011; Dredze et al., 2014; Chen et al., 2015); Wikipedia article views (McIver & Brownstein, 2014; Generous et al., 2014); clinicians database queries (Santillana et al., 2014); and cloud-based Electronic Health Records (EHR) (Santillana et al., 2016). These systems have the ability to reach a wider population and provide real-time access to information about influenza activity. Ensemble methods that use machine learning techniques to combine multiple Internet data sources have led to influenza tracking systems that accurately monitor and forecast CDC estimates of ILI activity at multiple geographical scales.(Santillana et al., 2015; Yang et al., 2015b; Santillana et al., 2014; Lu et al., 2018) While these alternative data sources provide real-time information about trends and general patterns of disease activity, defining the underlying population at risk is often challenging.(Chunara et al., 2017)

1.2.2 Participatory syndromic surveillance systems

One source of active digital data collection is online participatory syndromic surveillance systems.(Smolinski et al., 2017) Through these systems participants volunteer to report health information via online or mobile communication technologies on a weekly basis. The first of these systems, de Grote Griepmeting, or the Great Influenza Survey, started in 2003 in the Netherlands and Belgium.(Marquet et al., 2006) Since that time multiple systems have been established throughout Europe, Influenzanet; Australia (AU), Flutracking.net; U.S., Flu Near You (FNY); and Japan, Flu-Report.(Paolotti et al., 2014; Carlson et al., 2013; Smolinski et al., 2015; Fujibayashi et al., 2018) These systems actively engage the general public in reporting and provide timely information about disease trends within the community, thereby providing a mechanism for members of the community to become "citizen-scientists".(Smolinski et al., 2017; Kullenberg & Kasperowski, 2016)
Participatory syndromic surveillance systems complement traditional health care-based surveillance systems because they reduce the time-delay associated with visiting a health care provider and capture individuals who do not seek medical care. Studies in the U.S. have shown that approximately 45% of adults and 57% of children with ILI seek healthcare.(Biggerstaff et al., 2014) Both FNY and the Italian crowd-sourced counterpart to Influenzanet, INFLUWEB, have reported that approximately one third of their participants seek medical assistance.(Baltrusaitis et al., 2017; Guerrisi et al., 2016) Furthermore, these systems allow for a longitudi-

nal view of illness burden and have a well-defined population at risk, individuals who reported whether or not they have symptoms. However, because not every participant reports every week the population at risk can be inconsistent and include systematic biases.(Chunara et al., 2017)

1.2.2.1 Flu Near You

FNY is a U.S. based participatory syndromic surveillance system that was created in 2011 through collaboration between HealthMap of Boston Childrens Hospital and the Skoll Global Threats Fund.(Smolinski et al., 2015) Any resident of the U.S. or Canada can register as a user through the FNY website, mobile app, or Facebook. Upon registration, users provide information on their sex, month and year of birth, residential zip code, and email address. Although individuals must be at least 13 years of age to register, users can also add household members of any age and submit reports on their behalf. Following registration, FNY users are asked to submit brief weekly reports (Figure 1.2) where they can report any symptoms that they or any registered household members had during the previous week (Monday through Sunday). The symptoms in the report include fever, cough, headache, sore throat, diarrhea, body aches, fatigue, shortness of breath, chills or night sweats, nausea, rash, or runny nose. If a user did not have any of these symptoms, he, she, or they can also choose "I did not have any of the listed symptoms." However, if a user reports any of these symptoms, he, she, or they is asked to provide the date of symptom onset and whether or not they received medical care for the symptom(s). Starting in early October, users are also asked if they received an influenza vaccination for the current influenza season. Users are sent a reminder to complete the symptom report every Monday through either an email with a survey link or a push notification on their mobile phone. In

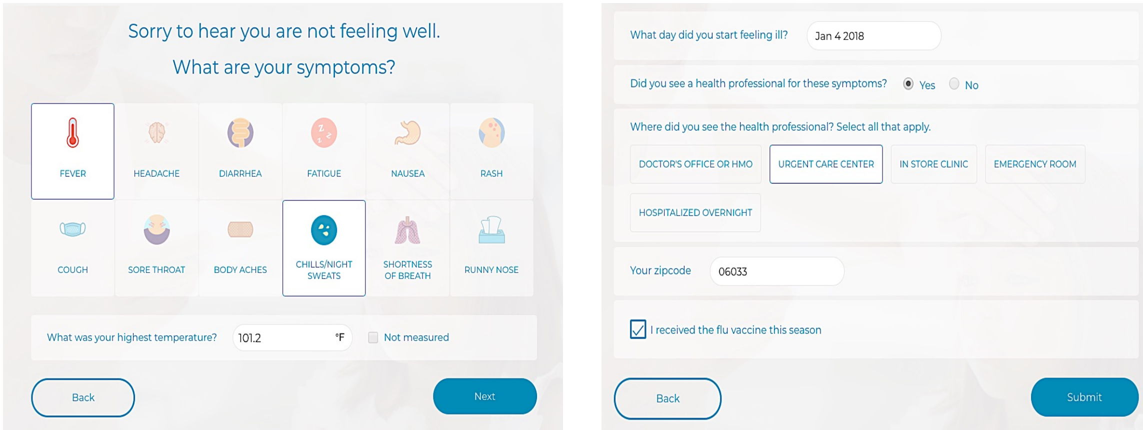

Figure 1.2: Screenshot of a weekly FNY report.

exchange for participating in FNY, users can visualize local ILI activity on maps, connect with local public health organizations, and find nearby locations offering influenza vaccines.

1.2.2.2 Flutracking.net

Flutracking.net is an online health surveillance system of influenza in AU and, as of 2018, New Zealand. Launched in 2006, the Flutracking system has grown to include over 30 000 participants.(Flutracking.net, 2018) Upon registration, Flutracking users provide basic demographic information, including month and year of birth, gender, postcode of residence, indigenous status, highest level of education, and whether or not they work directly with patients in a health care setting. Users then complete weekly surveys about the presence of ILI symptoms including fever, cough, and/or sore throat (Figure 1.3). Users who experience any of these symptoms are asked follow-up questions about absenteeism from work or normal duties, visits to health care providers, results of laboratory tests, and vaccination. Symptom surveys are sent every Monday, however, unlike other participatory surveillance systems, Flutracking participants have the option to complete

Figure 1.3: Screenshot of a weekly Flutracking.net report.

missed surveys up to five weeks previous.

1.3 OBJECTIVES

In this dissertation, we will address how researchers can effectively use data from online participatory syndromic surveillance systems to answer questions about ILI disease burden in the general population. We assess the representativeness and reporting habits of volunteers for these systems and use this information to develop statistically rigorous methods that adjust for potential biases. Specifically, we evaluate how different missing data methods, such as complete case and multiple imputation models, affect estimates of ILI disease burden using both simulated data as well as data from Flutracking.net. We then extend these methods to data from FNY, which has different patterns of participant reporting, and evaluate additional methods of bias adjustment. Finally, we provide examples of how this data has been used to answer questions about ILI in the general community and promote better understanding of disease surveillance and data literacy among volunteers.

CHAPTER 2

Identifying the biases and limitations of participatory syndromic surveillance systems

2.1 CHARACTERIZATION OF PARTICIPANT REPRESENTATIVENESS AND DETERMINANTS OF PARTICIPANTS' FOLLOW-UP IN FLU NEAR YOU

2.1.1 Overview

Crowd-sourced participatory syndromic surveillance programs, such as Influenzanet, FluTracking.net, and FNY, correlate well with traditional, sentinel ILI activity surveillance tools, and other platforms, such as GoViral, have validated the use of participatory information for disease surveillance by comparing volunteers' self-reported symptoms to specimens.(van Noort et al., 2015; Dalton et al., 2013; Smolinski et al., 2015; Goff et al., 2015) Although participatory syndromic surveillance systems track influenza activity in a timely fashion, a large, diverse, cohort of users who participate regularly and are representative of the population is essential for these systems to work effectively. Current participatory surveillance systems in Europe have assessed the representativeness of their participant population compared to the general populations and investigated factors that influence participant follow-up.(Cantarelli et al., 2014; Blanchon et al., 2013; Bajardi et al., 2014a,b) Here, we evaluate the representativeness of the FNY participant population compared to the general population of the U.S. and explore the demographic and behavioral characteristics that are associated with FNY's "good" users.

2.1.2 Representativeness

2.1.2.1 Methods

Data For this analysis, the FNY participant population includes all registered users and household members residing within the 50 U.S. states with complete sex, month and year of birth, and zip code information, and who submitted at least one symptom report during the 2014-2015 influenza season, defined by MMWR week 40 (week ending October 4, 2014) through week 20 (week ending May 23, 2015). National estimates of sex and age are from the U.S. Census Bureau's 2014 annual estimates of resident population.(U.S. Census Bureau, 2015) Socioeconomic status (SES) is estimated at the county level using the Human Development Index (HDI) as a proxy.(Lewis & Burd-Sharps, 2015) The HDI represents the county-level average of the educational index and income index and is measured on scale from 0-10, where 0 represents the lowest HDI and 10 represents the highest HDI. The educational index is a weighted average of the educational attainment index (i.e., the measure of overall level of educational attainment achieved by the adult population) and the enrollment index (i.e., total number of students enrolled in school divided by the total school-aged population of 3 to 24 year olds). The income index is calculated from county-level median income. The use of county-level HDI as an SES proxy is further assessed by estimating user-specific HDI from the cohort survey results and comparing these estimates to the corresponding county-level HDI estimate. Consistent with the method established by the Measure of America, the income index of the user-specific HDI is estimated by dividing the difference between the log of the zip-code level median income of the user and the log of the minimum U.S. median income by the difference between the log of the maximum and the log of the minimum U.S. median incomes. This ratio is multiplied by 10 to scale the index between 0 and 10. The educational index is estimated from the level of education response of the user survey, where lower educational attainment, such as "did not graduate High School", are assigned smaller values compared to higher educational attainment, such as "Ph.D., law, or medical degree."

Statistical methods The representativeness of sex (male and female) and age groups (<5, 5-14, 15-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80+) of FNY participants compared to the general U.S. population are assessed using a two-sided chi-square goodness of fit test. The county-level distribution of HDI of FNY participants is compared to the general U.S. population using a two-sample Kolmogorov-Smirnov test. We also assess the representativeness during the 2012-2013 and 2013-2014 influenza season as a sensitivity analysis.

2.1.2.2 Results and discussion

Although all 50 states are represented, FNY participants tend to cluster around major cities and along the coasts (Figure 2.1). During the 2014-2015 influenza season, California had the most number of participants (n=6595), while Wyoming had the fewest (n=89). When we adjusted for state population size, Rhode Island had the greatest per capita representation (0.04%) and Mississippi had the smallest per capita representation (0.008%). The 2012-2013 and 2013-2014 influenza seasons display a similar geographic distribution.(Baltrusaitis et al., 2017)

Over the course of the 2014-2015 influenza season, 47 234 unique participants had at least one symptom report that was either self-reported or submitted on their behalf. Of these participants, 28 906 (61.20%) are female and 18 328 (38.80%) are male. The proportion of female FNY participants is significantly over-represented

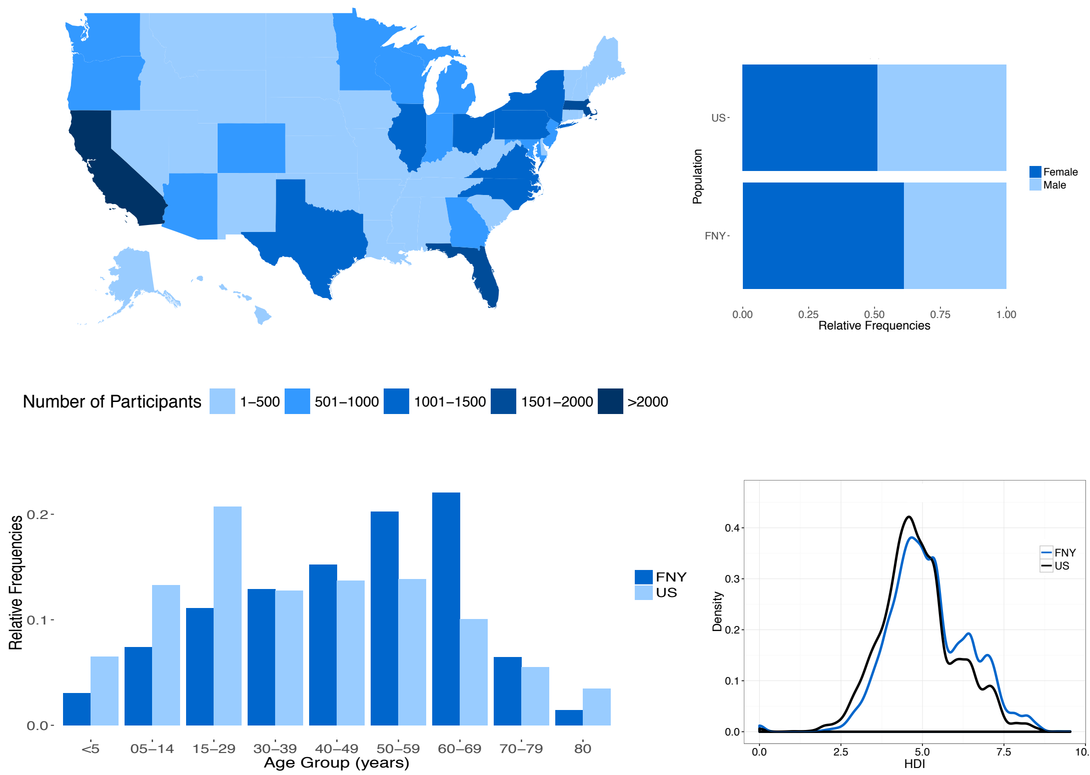

Figure 2.1: Representativeness of FNY participatants compared to the general U.S. population.

compared to the general U.S. population (51.1% female, P<.001) (Figure 2.1). Other participatory surveillance systems have reported an over representation of female participants. During the 2011-2012 influenza season, Influenzanet participants were more likely to be female than in the general population (56.8% vs 50.9%, P<.001), and among FluTracking participants who completed at least one survey, 66% and 64% were female in 2011 and 2012, respectively.(Cantarelli et al., 2014; Carlson et al., 2013) This over-representation of female participants is reflective of other studies showing that women are more likely than men to seek online health information.(Nölke et al., 2015; Fox et al., 2000)

Although each age group is represented in the FNY population, the distribution of age is significantly different from the U.S. population (P<.001). Overall, adult populations are over-represented (ages 40-79 years), while both younger populations (ages <30 years) and older populations (ages 80+ years) are underrepresented (Figure 2.1). As with sex, patterns of age representations are similar to both Influenzanet and FluTracking participants.(Cantarelli et al., 2014; Carlson et al., 2013)

The HDI range in the FNY population is 0-9.54 with a median of 5.03. As shown in Figure 2.1, the distribution of HDI scores is significantly greater for the FNY population compared to the U.S. population (P<.001). When comparing the FNY user-specific HDI estimates to the county-level HDI estimates in the population of FNY users who completed the 2016 survey, we found that, in general, the county-level HDI underestimates the user-specific HDI. This finding further supports our conclusion that FNY participants have a higher HDI than the U.S. population. The relatively high levels of HDI in the FNY population may be in part due to patterns in Internet penetration. Studies from the Pew Research Center have shown that Americans with high education levels and those in relatively affluent households have high Internet penetration.(Perrin & Duggan, 2015) The characteristics of participants are similar for the 2012-2013 and 2013-2014 influenza seasons.(Baltrusaitis et al., 2017)

2.1.3 Reporting behaviors

2.1.3.1 Methods

Data For this analysis only users who reported their own information, completed at least one symptom report during the 2014-2015 influenza season before MMWR week 18, and provided sex information at registration are included. In addition, only residents of the U.S. between ages 13 and 80 at registration date are selected because users must be at least 13 years of age to register. A limit of 80 years of age is used to account for possible errors in year of birth input at user registration. Users who meet these criteria are classified as either a "good user" or not based upon the number of symptom reports they submitted during the season. Users who complete more than three symptom reports during the season are classified as "good users."

Statistical methods The demographic factors used in this analysis are sex (male or female), age group (13-29, 30-39, 40-49, 50-59, 60-69, and 70-79), and HDI as a continuous variable. In addition, whether or not a report of ILI, as defined by the CDC, was reported at first entry is included. Although information from individual household members is not examined in the analysis, whether or not primary participants reported for other household members is also included.
Associations between these demographic and behavioral factors and "good users" are analyzed using multivariable logistic regression. For odds ratio (OR) comparisons among age groups, 50-59 is used as the reference group because it has the largest number of users. The demographic and behavioral factors are independent

variables, while level of follow-up is a dichotomous outcome ("good user" versus not). The outcome is dichotomized because the distribution of number of reports is not normally distributed, and the cut-off value of three is determined empirically by assessing a histogram of number of reports. Sensitivity analyses are conducted using more and less stringent definitions of a "good user", specifically, more than ten entries and more than one entry, respectively, for the 2012-2013, 2013-2014, and 2014-2015 influenza seasons to confirm the robustness of our findings.

2.1.3.2 Results and discussion

Figure 2.2 summarizes the ORs and 95% Confidence Intervals (CI) across all characteristics assessed. Overall, being a "good user" is associated with sex (male), reporting for household members, higher HDI score, not reporting an ILI at the first survey, and older age. These findings are consistent using both more-stringent (>10 entries) and less-stringent (>1 entry) definitions of a "good user", and the results are consistent across all three seasons, except for sex.(Baltrusaitis et al., 2017) While females were less likely to be a "good user" during the 2014-2015 season, this was not consistent with the 2012-2013 and 2013-2014 seasons. Given the differences in reporting patterns by sex across years, an underlying factor, such as method of member recruitment, may be a confounder of this association. In addition, the OR comparing participation habits between males and females are close to 1, and a previously published study from Influenzanet found that there are no significant differences between reporting for males and females.(Bajardi et al., 2014a)

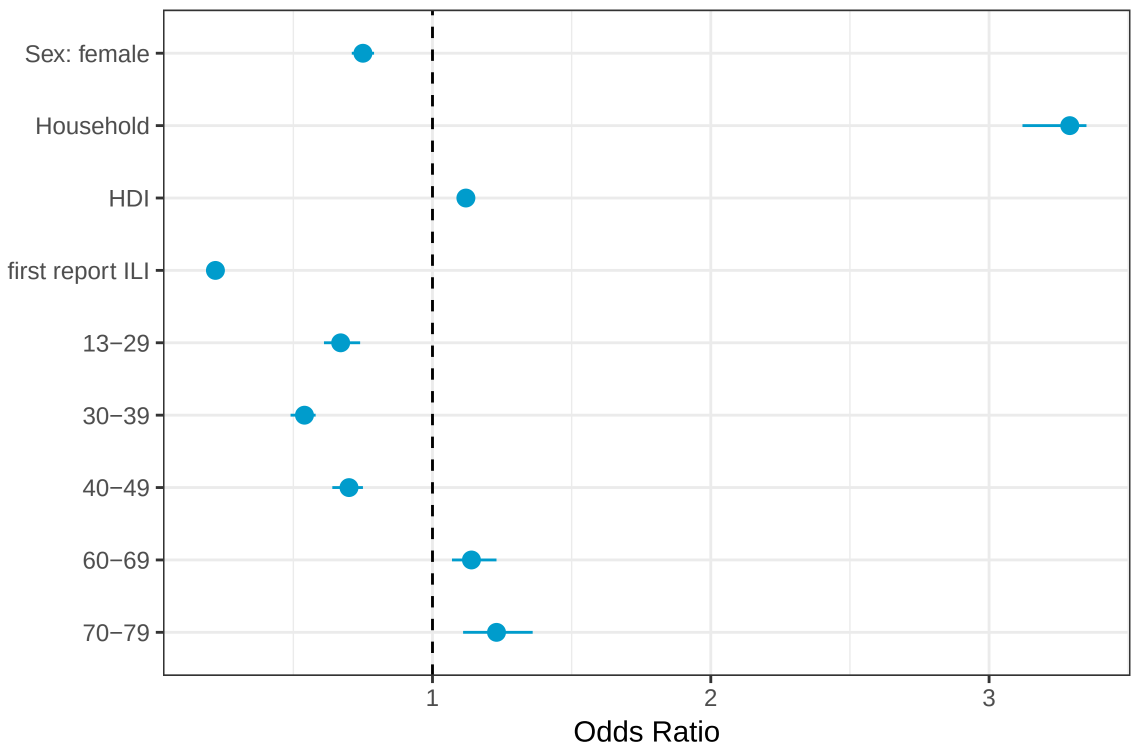

Figure 2.2: ORs and 95% CIs for different characteristics of a "good user" during the 2014-2015 influenza season.

2.2 COMPARISON OF FLU NEAR YOU TO TRADITIONAL HEALTH CARE-BASED INFLUENZA TRACKING SYSTEMS AT MULTIPLE SPATIAL RESOLUTIONS IN THE UNITED STATES OF AMERICA

2.2.1 Overview

The ability of FNY to complement, track, and forecast traditional provider-based influenza surveillance systems has been established at the national and regional levels in the U.S.(Smolinski et al., 2015; Santillana et al., 2015) However, because characteristics of activity may differ across states and sub-populations, further investigation of these novel systems is essential at finer spatial resolutions.(Lipsitch et al., 2011; Althouse et al., 2015; Lee et al., 2018a) The objectives of this project are to assess whether FNY correlates with traditional influenza surveillance systems across multiple spatial resolutions with different sample sizes and to determine the minimum number of reports necessary to produce influenza activity estimates that resemble the historical trends recorded by traditional sentinel surveillance systems for a given spatial resolution.

2.2.2 Multiple geographical scales

2.2.2.1 Methods

Data FNY percent ILI is calculated by dividing the number of participants reporting ILI, defined by a symptom report of fever plus cough and/or sore throat, in a given week, by the total number of FNY participant reports in that same week at each spatial resolution. Participants are aggregated at each spatial resolution using the zip code provided at registration for the time-period of 2012-2016. Information on patient visits to health care providers for ILI is collected

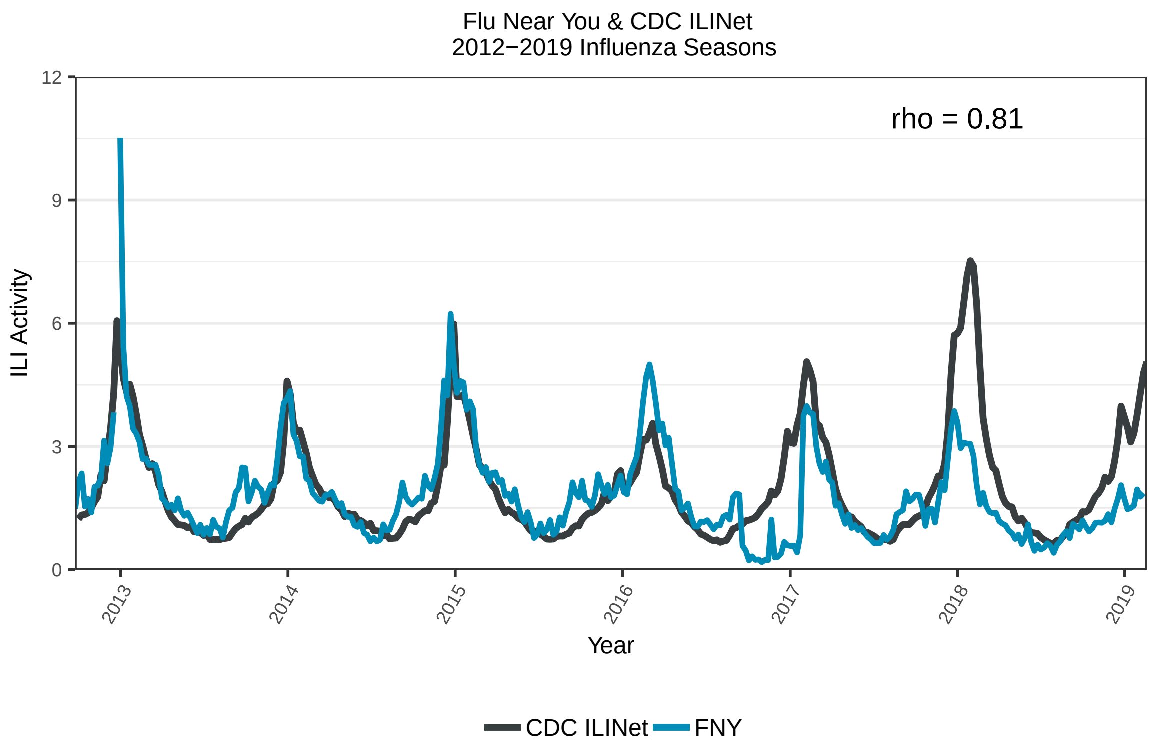

Figure 2.3: National time series of FNY percent ILI (blue) with CDC ILINet ILI activity (black) from October 2012 through February 2019.

through the U.S. Outpatient Influenza-like Illness Surveillance Network (ILINet, https://www.cdc.gov/flu/weekly/overview.htm).(Centers for Disease Control , CDC) For this system, ILI was defined as fever (temperature of $37.8\,^{\circ}\mathrm{C}$ [$100\,^{\circ}\mathrm{F}$] or greater) plus cough and/or sore throat without a known cause other than influenza. Weighted percent ILI, calculated by weighting the percentage of patient visits to healthcare providers for ILI reported each week on the basis of state population, is used as the ILI activity measure. For regional analyses, we use the ten Health and Human Services (HHS) defined regions (Appendix Table A.1). Data from Boston is collected through the Boston Public Health Commission (BPHC), which has operated a syndromic surveillance system since 2004. All nine acute care Boston hospitals electronically send limited data for all emergency department (ED) visits every 24 hours. Data sent includes visit date, chief complaint, zip code of residence, age, gender, and race/ethnicity. ILI visits are defined as fever and a cough or sore throat using chief complaints. Greater Boston was defined as zipcodes associated with Suffolk, Norfolk, Middlesex, Essex and Plymouth counties. These zipcodes are associated with over 90% of Boston ED visits. Percent ILI for Greater Boston is calculated by dividing the number of ILI visits by the total number of ED visits.

Statistical methods We use Pearson correlations to compare FNY percent ILI to CDC ILINet ILI activity. Correlations are calculated at the national and HHS-defined regional resolutions during the time period of October 1, 2012 through May 21, 2016, and for each of the four individual influenza seasons within this time period (MMWR weeks 40 to 20) separately. We also present comparisons of FNY percent ILI to CDC ILINet ILI activity for 49 states that voluntarily provided historic data across all seasons. Finally, FNY percent ILI is compared to perecent

ILI estimated from ED visits in the Greater Boston area. Boston was chosen as a pilot city because of the large FNY user base and availability of data. Descriptive statistics of the mean weekly reports are displayed as median (Interquartile Range IQR) for each geographical resolution.

2.2.2.2 *Results and discussion*

Pearson correlations of FNY percent ILI versus CDC ILINet and BPHC as well as mean weekly reports at all spatial resolutions are shown in Figure 2.4 and Appendix Table A.2. The national mean weekly reports across all seasons is 9699, and the correlation is 0.81. At the regional level, the median of the mean weekly reports is 889 (707, 1157). Region 7 has the smallest mean weekly reports (415), and Region 4 has the largest mean weekly reports (1798). The median correlation is 0.74 (0.73, 0.76). The median of the mean weekly reports at the state level is 128 (57, 263), and the median correlation with CDC ILINet is 0.55 (0.43, 0.63). For Boston, the mean weekly reports is 304 and the correlation with BPHC ILI activity estimates is 0.69. In general, the correlation with CDC ILINet ILI activity decreases as the geographical scale and corresponding mean weekly reports decreases (Figure 2.4). The geographic distribution of FNY mean weekly reports shows large gaps of information especially in the middle and southern areas of the US, and participants tend to cluster around large urban areas, with especially large user bases in the greater metropolitan areas surrounding Boston, New York City, and San Francisco (Figure 2.4). Our findings suggest that FNY percent ILI estimates correlate with ILI estimates from traditional influenza surveillance systems in various spatial resolutions if there is a sufficient number of reports.

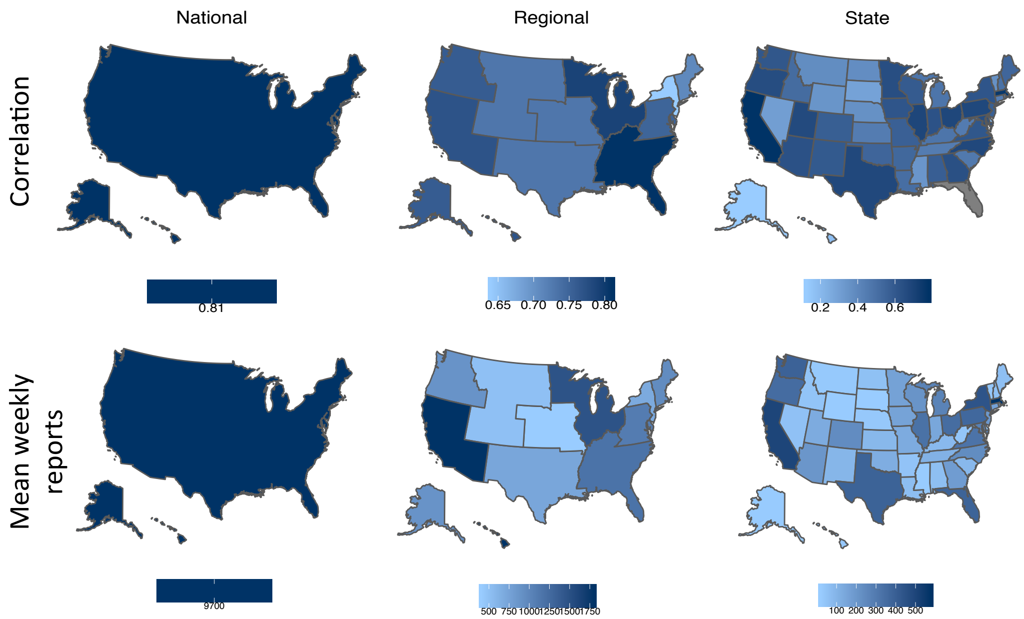

Figure 2.4: Correlation of FNY percent ILI with CDC ILINet ILI activity and mean weekly reports at the national, regional, and state levels.

2.2.3 Sample size estimation

2.2.3.1 Methods

We plot the Pearson correlations of the weekly proportion of FNY participants reporting ILI with the proportion of individuals visiting healthcare providers in CDC ILINet with ILI as function of the mean weekly FNY reports at the national, regional, state, and city resolutions during time-period of 2012-2016 to visually assess the relationship between the number of reports and correlation with established sentinel surveillance systems. A bootstrap sampling approach is also used to estimate the minimum number of FNY reports necessary to produce estimates that resemble the historical government-lead surveillance system trends. For this approach, Pearson correlations are calculated for subsets of the FNY from 0.1% to 15% of the full dataset in increments of 0.1% and compared to national weekly

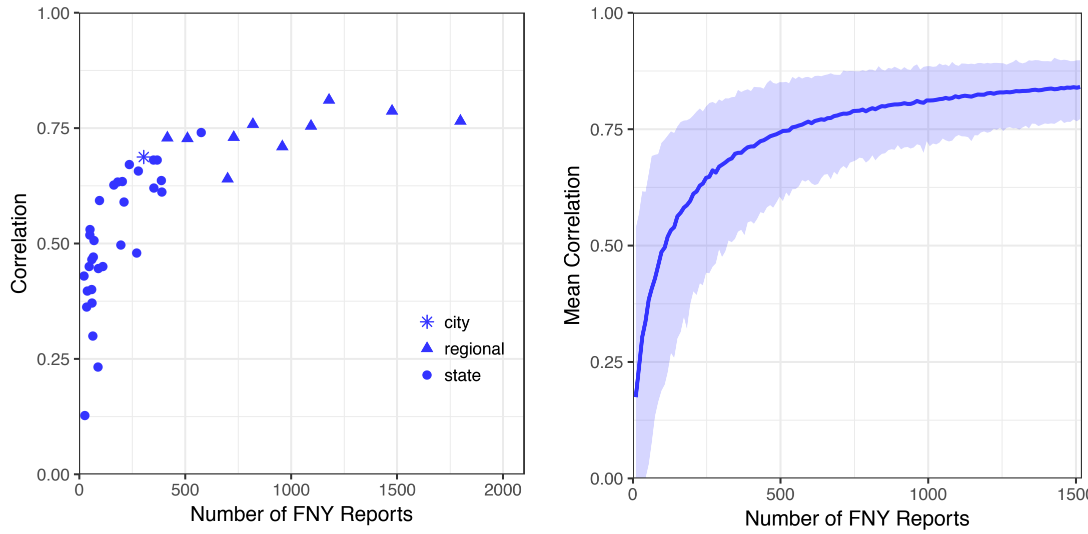

Figure 2.5: (A) Correlation of FNY ILI with CDC ILINet at different geographical scales (B) Correlation between bootstrapped samples of FNY ILI and CDC ILINet ILI with 95% CIs.

estimates of ILI from CDC ILINet. This process is repeated 1000 times using sampling with replacement, and the 95% CIs were calculated by ordering the Pearson correlation coefficients and selecting the 2.5th and 97.5th percentiles.

2.2.3.2 Results and discussion

As shown in Figure 2.5A, in general, correlation values increase as the mean weekly FNY reports increase at all geographic resolutions. Spatial resolutions with at least 2.5% (approximately 250/ 9699) of total weekly FNY reports have correlations greater than 0.5. As shown in Figure 2.5B, the correlation coefficient increases as the number of weekly reports increases, but the rate of growth slows around 250 weekly reports, similar to the results shown in Figure 2.5A.

Correlations between FNY ILI and CDC ILINet ILI activity never reach perfect correlation. Instead, they converge to approximately 0.8-0.9, as shown using both empirical and theoretical approaches. A similar observation was observed when comparing methods of provider recruitment in Texas.(Scarpino et al., 2012) This difference in correlation saturation may be a result of differences in the activity being measured (e.g. ILI reports out of all persons enrolled vs. visits with ILI out of the total number of patient visits) and the population under surveillance, as the crowd-sourced population includes individuals who may not seek medical attention. As mentioned in section 2.1.2.2, FNY also differs by demographics. Specifically, females and middle-aged individuals are over-represented in FNY.

With a sufficient number of weekly reports, approximately 250, data from FNY can complement traditional healthcare-based systems, especially in populations who do not access health care systems, areas with limited surveillance data, and community based populations.

CHAPTER 3

Developing and comparing appropriate methods to adjust for biases in participatory syndromic surveillance systems

3.1 COMPARATIVE EVALUATION OF MISSING DATA METHODS FOR P-ARTICIPATORY SYNDROMIC SURVEILLANCE DATA

3.1.1 Overview

In addition to tracking weekly prevalence of ILI, participatory syndromic surveillance systems have been used to produce age-specific attack rates (AR), (Patterson-Lomba et al., 2014; Chunara et al., 2015; Reed et al., 2016), determine risk factors of ILI, (van Noort et al., 2015) estimate influenza vaccine effectiveness, (Carlson et al., 2010; Debin et al., 2014; van Noort et al., 2015), and assess health care seeking behavior. (Tilston et al., 2010; Peppa et al., 2017) However, as discussed in section 1.2.2, because not every participant reports every week, the population at risk can be temporally inconsistent and include systematic biases.(Chunara et al., 2017) The most common approach to address the inconsistencies in user reporting habits has been to select a cohort of "active users", where the definition of "active user" varies by system and study, and assume that all missing reports were asymptomatic.(van Noort et al., 2015; Patterson-Lomba et al., 2014; Chunara et al., 2015; Reed et al., 2016) Unfortunately, no study has yet assessed how this deterministic assumption affects estimates of disease burden. In this study, we assess how different missing data methods affect estimates of ILI disease burden using both simulated data as well as data from Flutracking.net.

3.1.2 Methods

3.1.2.1 Data

Data collection for Flutracking.net is described in section 1.2.2.2. For this study, ILI is defined as report of both fever and cough, with or without sore throat. Surveys submitted more than one week after the initial reminder are referred to as "retrospective reports". Although Flutracking.net collects data from the beginning of May through mid-October, we use reports submitted during the influenza season in the southern hemisphere, defined as MMWR weeks 25 through 41, or approximately mid-June through the beginning of October. Descriptive statistics, including age, sex, household status, and vaccination status are displayed as median (IQR) for continuous variables and n (%) for categorical variables for of all participants who submitted at least one symptom report during the 2016, 2017, and 2018 influenza seasons.

3.1.2.2 Outcomes

As recommended by the WHO, we use the incidence rate (IR) as our measure of influenza disease burden.(World Health Organization, 2015) The IR is equal to the number of incident ILI reports, defined as a report of ILI in which ILI was not reported the previous week, divided by the total person-time reported by participants:

$$\text{Incidence Rate} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} Y_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T}(1-R_{it})} \times 10000 \tag{3.1}$$

where,

$$Y_{it} = \begin{cases} 1, & \text{if ILI} \\ 0, & \text{otherwise} \end{cases}$$

$$R_{it} = \begin{cases} 0, & \text{if } Y_{it} \text{ observed} \\ 1, & \text{if } Y_{it} \text{ missing} \end{cases}$$

$$t = \{25, .., 41\}.$$

The rate is expressed as per 10 000 person weeks. Because person-time at risk is unavailable for most routine influenza surveillance data, we also present the incidence proportion (IP) for comparability across systems. The IP is equal to the number of participants who reported ILI at least once during the influenza season divided by the total number of participants:

$$\text{Incidence Proportion} = \frac{\sum_{i=1}^{N} Q_i}{N} \tag{3.2}$$

where,

$$Q_i = \begin{cases} 1, & \sum_{t=1}^{T} Y_{it} \geq 1 \\ 0, & \text{otherwise.} \end{cases}$$

The 95% CIs for these estimates are given by:

$$\text{95\% Confidence Interval} = \left(\frac{\text{IR or IP}}{e^{1.96/\sqrt{d}}}, \text{ IR or IP} \times e^{1.96/\sqrt{d}} \right), \tag{3.3}$$

where d is the number of cases.(Kirkwood & Sterne, 2003; Giesecke, 2002) We calculate these measures for the overall population as well as by age group (<5, 5-17, 18-49, 50+). Finally, we present the weekly prevalence (WP) of ILI at each week, which is calculated by dividing the number of ILI reports by the total number of reports observed,

$$\text{Weekly Proportion}_t = \frac{\sum_{i=1}^{N} Y_{it}}{\sum_{i=1}^{N} (1 - R_{it})}. \tag{3.4}$$

Because weekly prevalence is a measure that is estimated in near-real time, all retrospective reports are assumed missing when calculating estimates. We assess and compare these measures across three influenza seasons (2016, 2017, and 2018) in Australia.

3.1.2.3 Missing data methods

We assess five different methods that account for missing data:

1. Ignore all missing data
2. Complete Case
3. Assume all missing reports are non-ILI reports
4. Multiple Imputation (MI)
5. MI with delta (δ) adjustment.

The first method ignores all missing data (i.e. select all Y_{it} for $R_{it} = 0$), whereas the second method includes only complete cases (i.e. select all individuals, i, where $R_{it} = 0$ for all t). The third method assumes that all missing reports are non-ILI reports (i.e.: $P(Y_{it} = 1|R_{it} = 1) = 0$), similar to past studies. The next two methods use MI methods to produce 10 point estimates, which are aggregated using Rubin's rules with a log transformation to account for non-normality.(Rubin, 2004) For the first MI method (method 4), we fit a model assuming Missing at Random (MAR),

$$logitP(Y_t = 1|\mathbf{X_t}) = \beta_0 + \beta_1 X_1 + \beta_2 X_1 + \beta_3 X_{3t} + .. + \beta_{2+t} Y_{t-1} \tag{3.5}$$

where

$$\mathbf{X_t} = \{\text{age group, gender, vaccination status at week t}\}.$$

Because the MAR assumption may not be valid for this type of data, we also perform MI using a δ adjustment, which is a flexible and transparent method to impute missing data under Missing Not at Random (MNAR) assumptions.(Leacy et al., 2017) The δ MI method uses 3.5, however, prior to imputing the missing data a fixed quantity, δ, is added to the linear predictor of the regression model,

$$logitP(Y_t = 1|\mathbf{X_t}, R_t) = \beta_0 + \beta_1 X_1 + \beta_2 X_1 + \beta_3 X_{3t} + .. + \beta_{2+t} Y_{t-1} + \delta R_t. \tag{3.6}$$

In this case, δ represents the difference in log-odds of ILI for participants who did not report compared to participants who did report. Both MI methods are fit using the Multivariate Imputation by Chained Equations (MICE) package for R.(R Core Team (R Foundation for Statistical Computing), 2016; van Buuren & Groothuis-Oudshoorn, 2011)

3.1.2.4 Estimation of δ

Because the log-odds of ILI for participants who did not report are unknown, we use the retrospective reports to estimate this value and create season-specific $\tilde{\delta}$ estimates. In other words,

$$\tilde{\delta} = log\left[\frac{odds\left(P\left(Y = 1|\tilde{R} = 1\right)\right)}{odds\left(P\left(Y = 1|\tilde{R} = 0\right)\right)}\right] \tag{3.7}$$

where,

$$\tilde{R}_{it} = \begin{cases} 0, & \text{if } Y_{it} \text{ report for same week} \\ 1, & \text{if } Y_{it} \text{ report for previous week} \end{cases}.$$

Negative values of δ indicate that participants were less likely to report ILI for retrospective reports compared to reports submitted during the same week.

3.1.2.5 Simulations

We also evaluate the missingness data methods under three missingness models: missing completely at random (MCAR), MAR, and MNAR, using simulated data. The data is simulated using a three-step process. First, 1000 Flutracking.net populations (n=30 000 each) are simulated using the characteristics, including age group, sex, and vaccination status, of the 2016 influenza season participant population. Simulated participants are assigned an age group, sex, vaccination status, and 17 weeks of symptom reports, Y_{it}. These weekly symptom reports are simulated using a multinomial distribution, where n, which is Poisson distributed with an age-group specific mean, represents the total number of ILI reports for the participant and p is the vector of weekly percent of sentinel general practitioner ILI consultations as reported by AU's Department of Health.(Australian Government Department of Health, 2016a) Next, 17 missingness indicators, R_{it}, are simulated to reflect distribution of Flutracking.net participant reports (Figure 3.1). As shown in this figure, approximately 80% of Flutracking.net participants submitted 15-17 reports, whereas the remaining 20% of participants are approximately uniformally distributed between 1 and 14 reports. We simulate these missingness indicators using three separate missingness models:

1. MCAR

$$\mathbf{R}_i \sim \text{Binomial}\left(n = 17, p_i\right) \tag{3.8}$$

where,

$$p_i = \begin{cases} 0.05, & \text{with probability } 0.8 \\ \text{Uniform}\left(0.1, 0.95\right), & \text{with probability } 0.2 \end{cases}$$

2. MAR

$$R_{it} \sim \text{Bernoulli}\left(p_{it}\right) \tag{3.9}$$

where,

$$p_{it} = \frac{e^{Z_i + \gamma R_{it-1}}}{1 + e^{Z_i + \gamma R_{it-1}}}$$

and

$$Z_i = \begin{cases} -2.944, & \text{with probability } W_i \\ \text{Uniform}\left(-2.197, 2.197\right), & \text{with probability } 1 - W_i \end{cases}$$

and

$$W_{it} = \frac{e^{\beta_0 + \beta_1 X_1 + .. + \beta_7 X_7}}{1 + e^{\beta_0 + \beta_1 X_1 + .. + \beta_7 X_7}}$$

for

$$\mathbf{X} = \{\text{age group, gender, vaccination status}\}.$$

3. MNAR

$$R_{it} \sim \text{Bernoulli}\left(p_{it}\right) \tag{3.10}$$

where,

$$p_{it} = \frac{e^{Z_i + \gamma R_{it-1} + \delta Y_t}}{1 + e^{Z_i + \gamma R_{it-1} + \delta Y_t}}.$$

The values of β_0 through β_7 are estimated for Flutracking.net data using the methods described in section 2.1.3.1. We define δ equal to the 2016 Flutracking $\tilde{\delta}$ estimate, however, we also present sensitivity analysis that assesses how varying the MNAR assumption affects IR estimates. Finally, each of the five methods described in the previous section is applied to produce overall and age-specific IR and IP estimates, as well as the overall WP estimates for each simulated dataset. We compare adjusted-estimates to the original simulation parameters through violin plots and Normalized Root Mean Square Errors (NRMSE) normalized by the original parameter,

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{\sum_i^T \left(\hat{B}-B\right)^2}{T}}}{B} \tag{3.11}$$

for $T = 1000$. In this equation, B represents either IR, IP, or WP.

3.1.3 Results

3.1.3.1 Flutracking.net

During the 2016, 2017, and 2018 influenza seasons, 29 671, 32 778, and 43 389 unique participants submitted at least one symptom report between week 25 and week 41, respectively. Across all seasons, approximately 60% of the participants identify as female, and the median age of participants range from 47 to 49. The largest age group is 50+, followed by 18 to 49, 5 to 17, and finally <5. More than half of the participants are primary users who submitted reports on their own behalf, and 59%, 61%, and 67% of participants reported that they received the influenza vaccination during the 2016, 2017, and 2018 influenza seasons, respectively (Table 3.1).

Table 3.1: Descriptive statistics of the Flutracking.net cohort during the 2016, 2017, and 2018 influenza seasons. Continuous variables are displayed as median (IQR) and categorical variables are displayed as n (%).

Variable		2016	2017	2018
Participants	n	29,671	32,778	43,389
Sex	male	11,153 (37.59)	12,665 (38.64)	17,086 (39.38)
	female	17,267 (58.19)	19,277 (58.81)	25,561 (58.91)
	other	1 (0)	5 (0.02)	19 (0.04)
	unknown	1250 (4.22)	831 (2.54)	723 (1.67)
Age	median (IQR)	47 (31, 58)	48 (31, 59)	49 (32, 61)
Age Group	<5	963 (3.25)	1062 (3.24)	1506 (3.47)
	5 to 17	3387 (11.42)	3813 (11.63)	4961 (11.43)
	18 to 49	11,812 (39.81)	12,306 (37.54)	15,667 (36.11)
	50+	13,509 (45.53)	15,597 (47.58)	21,255 (48.99)
Household	Primary user	17,525 (59.06)	19,170 (58.48)	24,945 (57.49)
	Household member	12,146 (40.94)	13,608 (41.52)	18,444 (42.51)
Vaccinated	yes	17,526 (59.07)	19,966 (60.91)	29,081 (67.02)
	no	12,145 (40.93)	12,812 (39.09)	14,308 (32.98)

The descriptive statistics of the reporting habits of participants during the 2016, 2017, and 2018 influenza seasons are shown in Table 3.2. The total number of symptom reports submitted during the influenza season increased from approximately 450 000 in 2016, to almost 500 000 in 2017, and finally to more than 650 000 in 2018.

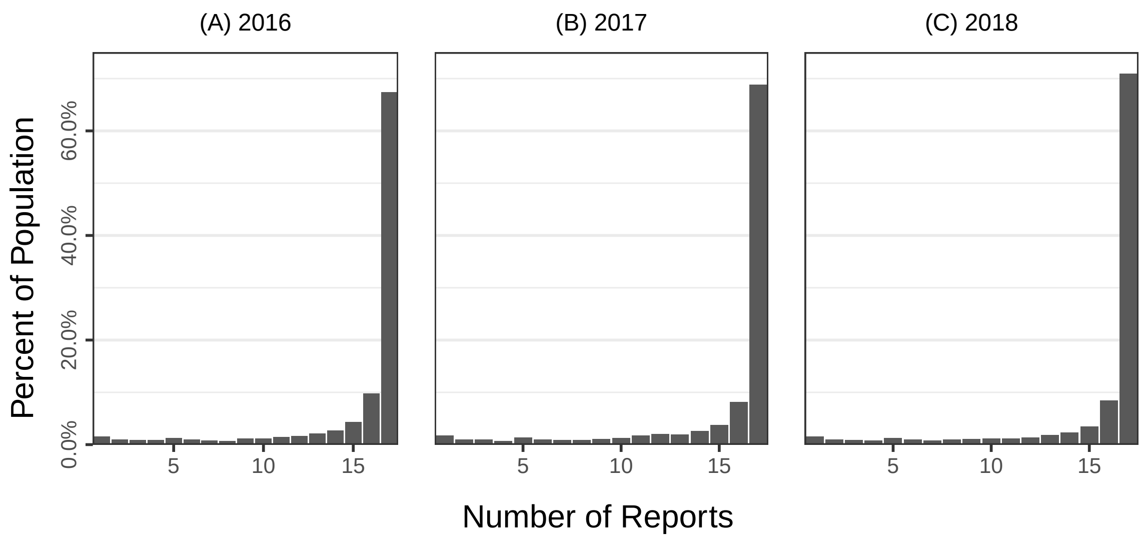

Figure 3.1: Distributions of the number of Flutracking.net participant reports during the (A) 2016, (B) 2017, and (C) 2018 influenza seasons.

The median number of weekly reports also increased from approximately 26 000 in 2016 to 39 000 in 2018. The distribution of the number of participant reports is shown in Figure (3.1). During each influenza season, the median number of reports per participant was 17 (16, 17), indicating that more than half of the participants submitted a symptom report for every week during the influenza season. As shown in Figure 3.2, most Flutracking.net participants register before week 25, however, the percentage of participants who are lost to follow-up increases as the season progresses, as shown by the dark gray bars. Most reports are submitted within one week of the symptom report date, however, a larger fraction of these reports are ILI compared to the retrospective reports, resulting in a negative value of $\tilde{\delta}$. The exact value of $\tilde{\delta}$ varies by season, and the corresponding ORs of reporting ILI for a retrospective report compared to reporting ILI for a report submitted the same week range from 0.80 in 2018 to 0.92 in 2017. The percentage of retrospective reports (dark blue) is fairly consistent through the season, but the proportion of missing reports (light grey) appears to increase until mid-season, at which point it slowly decreases as more participants are lost to follow-up (dark grey). These patterns are consistent across all influenza seasons.

Table 3.2: Descriptive statistics of Flutracking.net participant reporting habits during the 2016, 2017, and 2018 influenza seasons. Continuous variables are displayed as median (IQR) and categorical variables are displayed as n (%).

Variable		2016	2017	2018
Participants	n	29,671	32,778	43,389
Reports	Total	452,627 (100)	498,465 (100)	665,935 (100)
	Non-ILI	442,817 (97.83)	486,380 (97.58)	655,187 (98.39)
	ILI	9810 (2.17)	12,085 (2.42)	10,748 (1.61)
Reports Submitted Within One Week	Total	398,496 (88.04)	432,733 (86.81)	585,295 (87.89)
	Non-ILI	389,706 (97.79)	422,139 (97.55)	575,611 (98.35)
	ILI	8790 (2.21)	10,594 (2.45)	9684 (1.65)
Retrospective Reports	Total	54,131 (11.96)	66,096 (13.26)	80,640 (12.11)
	Non-ILI	53,111 (98.12)	64,605 (97.74)	79,576 (98.68)
	ILI	1020 (1.88)	1491 (2.26)	1064 (1.32)
Reports per Week	median (IQR)	26,570 (26,470, 26,860)	29,370 (29,200, 29,520)	39,150 (38,550, 39,800)
Reports per Participant	median (IQR)	17 (16, 17)	17 (16, 17)	17 (16, 17)
delta ($\tilde{\delta}$)		-0.158	-0.082	-0.226
OR ($e^{\tilde{\delta}}$)		0.85	0.92	0.80

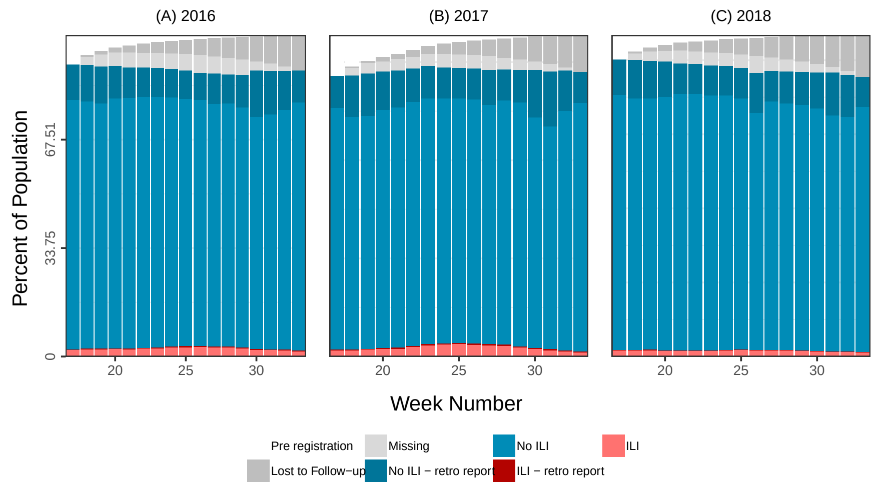

Figure 3.2: Histograms of Flutracking.net participant reporting habits during the (A) 2016, (B) 2017, and (C) 2018 reporting periods.

3.1.3.2 Outcomes

Incidence Rate Appendix Table A.3 and Figure 3.3 display overall and age group specific IRs and 95% CIs, expressed as number of ILI reports per 10 000 person weeks, by influenza season. Although the 2017 influenza season had higher IRs compared to the 2016 and 2018 seasons, the general patterns in estimates are consistent across all seasons and age groups. The method that assumes that all missing reports are non-ILI has the lowest IR estimates, whereas ignoring the missing value and MI methods have the highest IR estimates. As expected, IR estimates from the δ MI method are slightly less than estimates from the MI method without the δ adjustment indicating that retrospective reports are less likely to be ILI. In most age groups, IR estimates from the complete case method are similar or slightly greater than estimates from the method that assumes all missing reports are non-ILI.

Incidence Proportion The overall and age group specific IPs and 95% CIs for each influenza season are shown in Appendix Table A.4 and Figure 3.4. Estimates

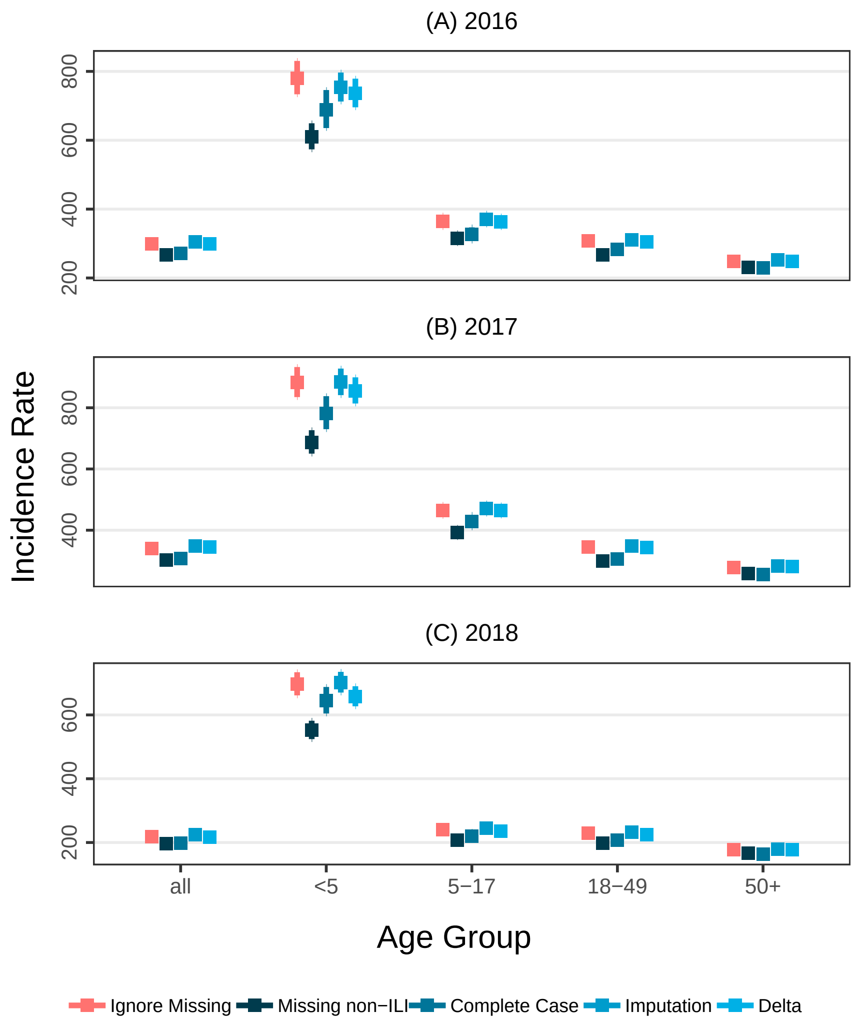

Figure 3.3: Overall and age group specific IRs and 95% CIs, expressed as number of ILI reports per 10 000 person weeks, for the (A) 2016, (B) 2017, and (C) 2018 influenza seasons.

for the method that ignores missing data are not shown because this method is equivalent to assuming that all missing reports are non-ILI. Similar to IR estimates, IPs estimates from the method that assumes all missing reports are non-ILI and complete case method are less than IP estimates from the MI and δ MI methods. However, the differences in IP estimates appear to be less pronounced compared to IR estimates.

Weekly Proportion Near-real time WP estimates are shown in Figure 3.5. Because the complete case population is unknown during the season, this method is not applied. We also present the complete data with the retrospective reports for comparison. The method that assumes all missing reports are non-ILI results in WP estimates that are less than the other methods. WP estimates from the MI method are slightly higher than WP estimates from the method that ignores missing values and the δ MI method. The estimates from these two methods are similar to the complete data.

3.1.3.3 Simulations

Violin plots are shown in Appendix Figure A.1 and Appendix Figure A.2, for IR and IP respectively, and WP time series results are shown in Appendix Figure A.3. NRMSEs are displayed in Table 3.3. Under each model, assuming that all missing reports are non-ILI underestimates the IR, IP, and WP, and results in the largest NRMSE. When estimating the IR, the ignoring missing reports and MI methods have smaller NRMSEs compared to the complete case and δ MI methods under MCAR and MAR assumptions. However, under MNAR, IR estimates using the δ MI have the smallest NRMSEs compared to the other methods. For IP, the complete case and MI methods outperform ignoring missing reports, which is equivalent

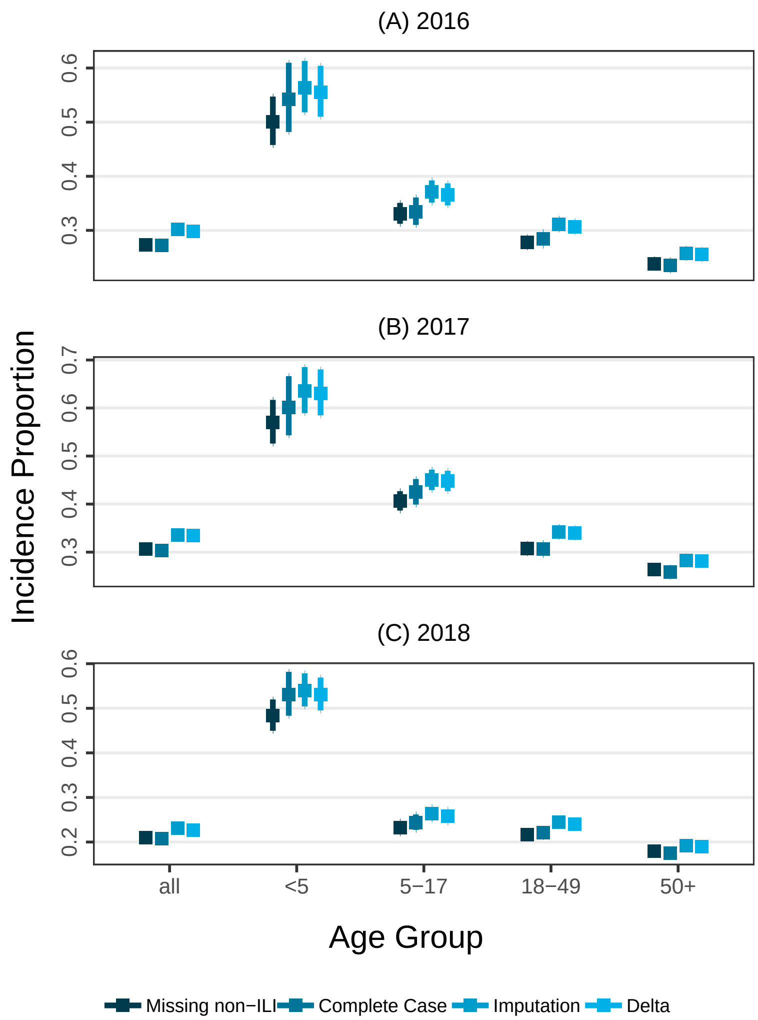

Figure 3.4: Overall and age group specific IPs and 95% CIs, expressed as percent of population reporting ILI at least once, for the (A) 2016, (B) 2017, and (C) 2018 influenza seasons.

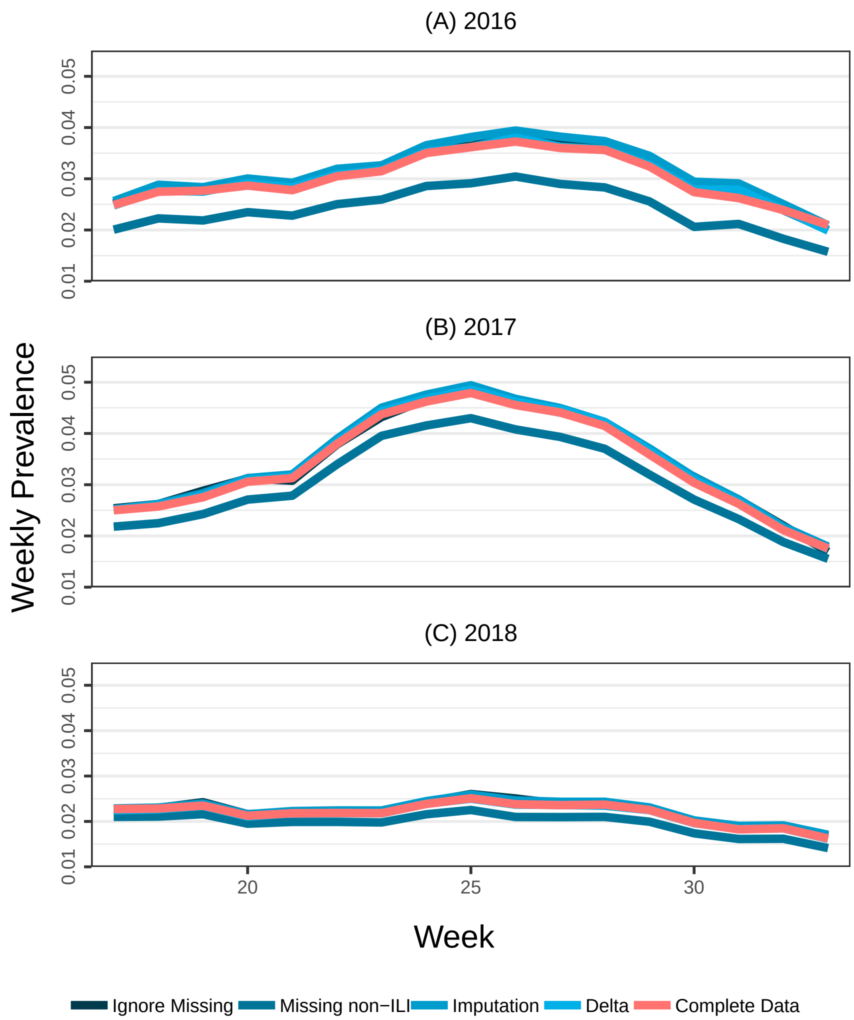

Figure 3.5: WPs for the (A) 2016, (B) 2017, and (C) 2018 influenza seasons.

to assuming that all missing reports are non-ILI in this scenario. Under MCAR and MAR models, the MI method has the smallest NRMSEs, and even though the δ MI method underestimates the IP, the NRMSEs are similar to those from the complete case method. Similar to IR, the δ MI method is the best method when data are MNAR. When estimating the WP, the method that ignores missing reports outperforms the other methods under each missingness model.

Table 3.3: NRMSE, expressed as percentage, by age-group for IR and IPs and overall for WP under MCAR, MAR, and MNAR missingness models.

Model	Age Group	Ignore missing data	Assume missing are non-ILI	Complete Case	MI	δ MI
IR						
MCAR	Overall	0.44	14.02	1.52	0.73	2.51
	<5	1.94	14.05	6.82	2.8	3.31
	5-17	1.17	14.07	4	1.69	2.87
	18-49	0.65	14.01	2.45	1.02	2.57
	50+	0.67	14.06	2.43	1.05	2.68
MAR	Overall	0.56	14.59	1.73	0.83	2.57
	<5	2.23	19.47	7.41	3.45	4.36
	5-17	1.26	16.72	4.13	1.98	3.24
	18-49	0.68	15.71	2.45	1.18	2.84
	50+	0.64	12.08	2.34	1.05	2.28
MNAR	Overall	2.08	12.5	1.61	2.95	0.81
	<5	3.65	16.95	7.57	5.02	3.24
	5-17	2.91	14.45	4.25	3.71	1.77
	18-49	2.64	13.52	2.78	3.19	1.11
	50+	2.18	10.22	2.68	2.65	0.91
IP						
MCAR	Overall	-	12.98	1.44	0.64	2.55
	<5	-	12.46	6.05	2.53	3.14
	5-17	-	12.84	3.63	1.58	2.83
	18-49	-	12.95	2.25	0.96	2.59
	50+	-	13.16	2.32	0.96	2.74
MAR	Overall	-	13.54	1.61	0.69	2.62
	<5	-	17.64	6.69	2.98	4.19
	5-17	-	15.46	3.79	1.76	3.2
	18-49	-	14.65	2.3	1.04	2.91
	50+	-	11.31	2.2	0.96	2.33
MNAR	Overall	-	11.58	1.51	2.42	0.87
	<5	-	15.29	6.78	4.08	3
	5-17	-	13.32	3.89	3.02	1.75
	18-49	-	12.59	2.54	2.66	1.17
	50+	-	9.56	2.48	2.26	0.94
WP						
MCAR	Overall	2.17	14.14	-	3.01	4.04
MAR	Overall	2.17	14.37	-	3.1	3.97
MNAR	Overall	2.88	12.21	-	4.13	2.98

3.1.4 Discussion

National estimates of disease burden in the population are essential to determine the health and economic impact of influenza.(Lee et al., 2018c) However, most sentinel surveillance includes only individuals who visit a medical care facility and there is typically a delay from onset of patient symptoms to final publication of reports. Alternative data sources, such as Flutracking.net, have the potential to complement these traditional systems by capturing a population not routinely included among the other healthcare-based systems and minimizing delays in reporting.(Smolinski et al., 2017)

Although global and AU estimates of IRs and IPs for ILI are not currently available, laboratory-confirmed influenza is a nationally notifiable disease in AU.(Sullivan et al., 2016) The National Notifiable Diseases Surveillance System (NNDSS) provides estimates of the notification rate of laboratory confirmed influenza per 100 000 population. These age-specific rates estimates range from 246.7 to 1237.2 per 100 000 population during the 2016 influenza season.(Australian Government Department of Health, 2016b) As expected, these estimates are much lower than the estimates from Flutracking.net because NNDSS estimates laboratory confirmed influenza, whereas Flutracking.net estimates ILI. Furthermore, the number of cases represent only a proportion of the total cases occurring in the community, that is, only those cases for which health care was sought, a test conducted and a diagnosis made, followed by a notification to health authorities. However, the patterns in age group specific estimates are similar.

There are several limitations of this study. Because the estimates of ILI are based on syndromic data, the specificity with respect to the actual circulation of influenza viruses among the population is relatively low. Furthermore, the demographics of the Flutracking.net population differ from AUs national population. Females and middle-aged individuals are over-represented in the Flutracking.net population. Finally, in our MNAR simulation model the value of δ in equation 3.10 is known, and the resulting δ MI model can be properly parameterized. In reality, this value is unknown, and we were only able to estimate it from retrospective reports. Sensitivity analysis show that as the probability of missing given a report is ILI decreases, the method that assumes missing reports are non-ILI becomes a better method of adjustment, as shown in Appendix FigureA.4. However, under modest changes in the δ, for example increasing δ from .3 to .5 or 1.3, which corresponds an increase in OR from 1.35 to 1.65 or 3.67, respectively, the δ MI method still outperforms the method that assumes missing are non-ILI.

When using participatory surveillance data to estimate ILI disease burden in the general population, the final estimates depend on the method used to account for missing data. Under each simulation scenario, assuming that all missing reports are non-ILI underestimates all estimates. Although the optimal method depends on the estimate of interest and the missingness model, when properly parameterized, the δ MI method provides estimates of disease burden that are similar to the true parameter under MNAR models.

Based on this study, we recommend following Flutracking.net's lead by providing users with the opportunity to complete missing surveys. This system accommodation not only adds approximately 10% more weekly reports, but also provides valuable insight into reporting behaviors. Furthermore, the δ MI method accurately predicted end of season WP estimates from real-time data. In the future, the value of δ can be easily updated and adapted over the course of an influenza season.

3.2 COMPARATIVE EVALUATION OF METHODS THAT ADJUST FOR REPORTING BIASES WHEN ESTIMATING INFLUENZA-LIKE ILLNESS BURDEN USING DATA FROM FLU NEAR YOU

3.2.1 Overview

Because the dynamics and severity of influenza in the U.S. varies each season, yearly population estimates of influenza burden are essential to determine the risk of morbidity and mortality in different segments of the population, guide vaccination programs, evaluate the use of diagnostic tests and antiviral drugs, and plan for seasonal epidemics and future pandemics.(Thompson et al., 2006) Since the influenza A (H1N1) Pandemic in 2009, the CDC has used a probabilistic multiplier model to estimate the seasonal influenza burden for the entire U.S. using laboratory-confirmed influenza-associated hospital rates from Influenza Hospital Surveillance Network (FluSurv-NET) and results from health-seeking behavior studies.(Reed et al., 2015; Shrestha et al., 2017; Rolfes et al., 2018)

To obtain these age group specific estimates of influenza burden, influenza associated hospital rates from FluSurv-NET are first adjusted to correct for the under-detection of influenza hospitalizations by multiplying the reported rate by both the probability that a person hospitalized with an influenza infection would be tested and the probability that a person who is positive for an influenza would have a positive test. These adjusted hospitalization rates are then extrapolated to the U.S. population. Next, the age-specific number of influenza illnesses who sought health care are estimated based using the ratio of the estimated number of ill persons per hospitalization. Finally, the estimated number symptomatic illnesses for each age group are calculated using estimates of the percentage of persons with a respiratory illness who sought medical care. These percentages are obtained from

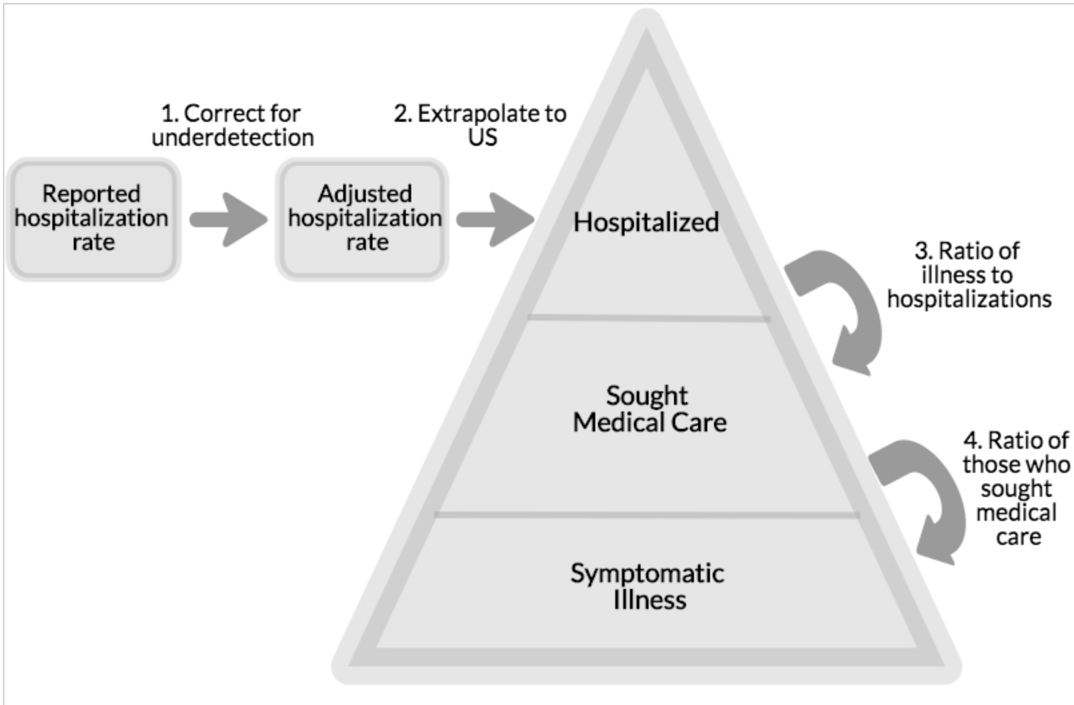

Figure 3.6: Schematic of the steps to estimate influenza disease burden in the U.S. population from FluSurv-NET laboratory-confirmed influenza-associated hospital rates (adapted from Reed et al. (2015)).

the 2010 Behavioral Risk Factor Surveillance Survey (BRFSS). A schematic of these steps is presented in Figure 3.6.(Biggerstaff et al., 2012)

There are several limitations to this method. First, the number of hospital-based influenza-confirmed cases come from FluServ-NET, which includes data from only 13 geographical areas representing only 9% of the U.S. population. As a result, the data may not be representative of the entire U.S., especially among those who do not seek medical care. In addition, estimates of health care-seeking behavior from the 2010 BRFSS are based on trends in visits to health care facilities and influenza diagnosis and treatment during the 2009 influenza pandemic. These trends may not be applicable to post-pandemic years.

FNY has the potential to complement the CDC's estimation methods because it captures individuals who do not seek medical care. However, as discussed in previous chapters, because not all participants report every week and participants are more likely to report when ill, the estimates of disease burden may be biased.(Baltrusaitis et al., 2017) Past studies have used various methods to address these issues including restricting analyses to cohorts of users that report regularly,(Chunara et al., 2015; Cantarelli et al., 2014; Patterson-Lomba et al., 2014) dropping the first report of all users, and using a spike detector.(Smolinski et al., 2015) In this study, we apply these approaches that adjust for common reporting biases to estimate ILI burden in the general population using data from FNY and compare these estimates to the CDC's estimates of influenza burden. We also present results from a simulation study that compares these approaches under different missingness and motivation to report assumptions.

3.2.2 Methods

3.2.2.1 Data

Data collection for FNY is described in Section 1.2.2.1. This analysis uses FNY data from the 2015-2016, 2016-2017, and 2017-2018 influenza seasons, defined as MMWR week 40 through week 20. These weeks usually corresponds with months October through May.

3.2.2.2 Outcomes

Age group specific estimates of ILI burden are calculated by dividing the sum of the weekly incident cases of ILI, defined as a report of fever with cough and/or sore throat, by the population at risk at the beginning of the period. These proportions are then extrapolated to the general population using the U.S. Census

Bureau's 2013-2017 American Community Survey 5-Year Estimates(U.S. Census Bureau, 2015):

$$\text{ILI Burden}_{agegroup} = \frac{\sum_{i=1}^{N_{agegroup}} Q_i}{N_{agegroup}} \times N_{US_{agegroup}} \tag{3.12}$$

where,

$$Q_i = \begin{cases} 1, & \sum_{t=1}^{T} Y_{it} \geq 1 \\ 0, & \text{otherwise}. \end{cases}$$

$$Y_{it} = \begin{cases} 1, & \text{if ILI} \\ 0, & \text{otherwise} \end{cases}$$

$$t = \{40, .., 20\}.$$

The 95% CIs for these estimates are given by:

$$\text{95\% CI} = \left(\frac{\frac{\sum_{i=1}^{N_{agegroup}} Q_i}{N}}{e^{1.96/\sqrt{\sum Q_i}}} \times N_{US_{agegroup}}, \frac{\sum_{i=1}^{N_{agegroup}} Q_i}{N} \times e^{1.96/\sqrt{\sum Q_i}} \times N_{US_{agegroup}} \right). \tag{3.13}$$

The age groups correspond to the CDC's defined age groups: <5, 5-17, 18-49, 50-64, 65+.

3.2.2.3 *Bias adjustment methods*

Five different methods of bias adjustment are assessed:

1. All reports

2. Drop first report

3. Cohort

4. Cohort with MI

5. Cohort with δ MI.

The first method includes all FNY participants, defined as users and household members, who submitted at least one symptom report, whereas the second method includes all FNY participants who submitted at least two symptom reports and drops the first symptom report for all participants. All missing reports are assumed to be non-ILI for both of these methods. The final methods include a cohort of FNY participants who meet a specific reporting entry criteria, at least 10 symptom reports, and applies the missing data methods explored in Section 3.1.2.3. Specifically, the third method assumes that all missing reports are non-ILI reports, the fourth method uses MI under the MAR assumption

$$logitP(Y_t = 1|\mathbf{X_t}) = \beta_0 + \beta_1 X_1 + \beta_2 X_1 + .. + \beta_{1+t} Y_{t-1} \tag{3.14}$$

where

$$\mathbf{X_t} = \{\text{age group, sex}\},$$

and the fifth method uses MI with a δ adjustment

$$logitP(Y_t = 1|\mathbf{X_t}, R_t) = \beta_0 + \beta_1 X_1 + \beta_2 X_1 + .. + \beta_{1+t} Y_{t-1} + \delta R_t. \tag{3.15}$$

Unlike Flutracking.net, FNY does not collect retrospective reports. Instead, we use information from end of season surveys and expert opinion to define δ equal to -0.5, which corresponds to an OR of 0.6. In other words, reports that are missing have 0.6 times the odds of being ILI compared to reports that are not missing. For both MI methods, the MICE package in R is used to produce 10 point estimates.(R Core Team (R Foundation for Statistical Computing), 2016; van Buuren & Groothuis-Oudshoorn, 2011) These point estimates are aggregated using Rubin's rules with a log transformation to account for non-normality.(Rubin, 2004)

3.2.2.4 Simulations

We also evaluate the missingness data methods using simulated data. The data is simulated through a four-step process. First, 1000 populations (n=100 000 each) of potential FNY participants are simulated using the characteristics of the 2016-2017 influenza season participant population. Simulated potential participants are assigned an age group, sex, and 33 weeks of symptom reports, Y_{it}. These weekly symptom reports are simulated using a multinomial distribution, where n, which is Poisson distributed with an age-group specific mean, represents the total number of ILI reports for the participant and p is the vector of weekly percent of sentinel general practitioner ILI consultations as reported by CDC ILINet. Next, a sub-population of approximately 30 000 individuals are chosen to be FNY participants. We assess the motivation (M_i) for being a FNY participant under three different models:

1. Motivated Completely At Random (MoCAR)

$$M_i \sim \text{Bernoulli}\left(p_i\right) \tag{3.16}$$

where,

$$p_i = 0.3$$

2. Motivated At Random (MoAR)

$$M_i \sim \text{Bernoulli}\,(p_i) \tag{3.17}$$

where,

$$p_i = \frac{e^{\beta_0+\beta_1 X_{i1}+..+\beta_7 X_{i7}}}{1+e^{\beta_0+\beta_1 X_{i1}+..+\beta_7 X_{i7}}}$$

for

$$\mathbf{X} = \{\text{age group, sex}\}.$$

3. Motivated Not At Random (MoNAR)

$$M_i \sim \text{Bernoulli}\,(p_i) \tag{3.18}$$

where,

$$p_i = \frac{e^{\beta_0+\beta_1 X_{i1}+..+\beta_7 X_{i7}+\delta\sum Y_{it}}}{1+e^{\beta_0+\beta_1 X_{i1}+..+\beta_7 X_{i7}+\delta\sum Y_{it}}}$$

for

$$\mathbf{X} = \{\text{age group, sex}\}.$$

Next, 33 missingness indicators, R_{it}, are simulated to reflect distribution of FNY participant reports (Figure 3.7). All three missingness models start with a Beta-binomial distribution with α=0.5 and β=1, to reflect the distribution of FNY reports:

1. MCAR

$$\mathbf{R}_i \sim \text{Binomial}\,(n=33, p_i) \tag{3.19}$$

where,

$$p_i = \frac{Z_i+1}{33}$$

and

$$Z_i \sim \text{Beta-binomial}\,(n = 32, \alpha = 0.5, \beta = 1)$$

2. MAR

$$R_{it} \sim \text{Bernoulli}\,(p_{it}) \tag{3.20}$$

where,

$$p_{it} = \frac{e^{Z_i' + \beta_0 + \beta_1 X_{i1} + .. + \beta_7 X_{i7} + \gamma R_{it-1}}}{1 + e^{Z_i' + \beta_0 + \beta_1 X_{i1} + .. + \beta_7 X_{i7} + \gamma R_{it-1}}}$$

and

$$Z_i' = log\left(\frac{Z_i}{1.000001 - Z_i}\right)$$

for

$$\mathbf{X}_t = \{\text{age group, gender}\}.$$

3. MNAR

$$R_{it} \sim \text{Bernoulli}\,(p_{it}) \tag{3.21}$$

where,

$$p_{it} = \frac{e^{Z_i' + \beta_0 + \beta_1 X_{i1} + .. + \beta_7 X_{7i} + \gamma R_{it-1} + \delta Y_{it}}}{1 + e^{Z_i' + \beta_0 + \beta_1 X_{i1} + .. + \beta_7 X_{i7} + \gamma R_{it-1} + \delta Y_{it}}}$$

The values of β_0 through β_7 are estimated for FNY data using the methods described in section 2.1.3.1. We define γ equal to -0.2 so that subsequent reports are less likely to be observed if the previous report is also missing and δ equal to 0.5 so that reports are less likely to be missing if they are ILI. Although there are nine potential combinations of motivation and missingness models, for this simulation study, we assume that the motivation and missingness models are concordant. For example, under MoCAR the missingness model is also MCAR. Finally, each of the five methods described in the previous section is applied to produce age-specific

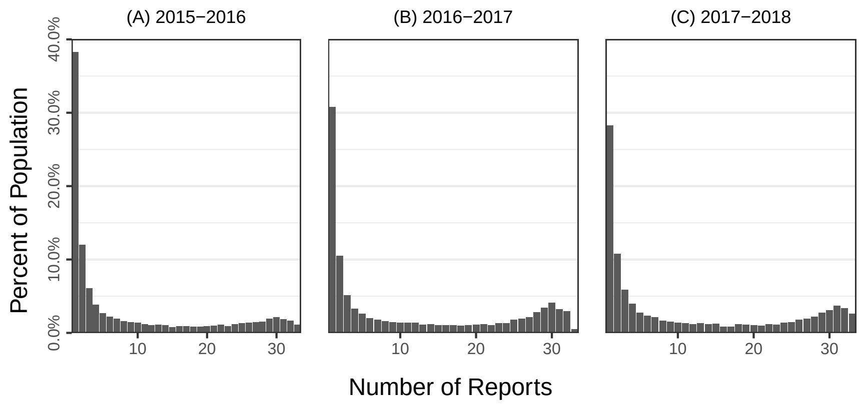

Figure 3.7: Distributions of the number of FNY participant reports during the (A) 2015-2016, (B) 2016-2017, and (C) 2017-2018 influenza seasons.

IP estimates (Equation 3.2). We compare adjusted-estimates to the original simulation parameters through violin plots and NRMSE (Equation 3.11).

3.2.3 Results

3.2.3.1 Flu Near You population characteristics

The descriptive statistics for each population: total, drop first report, and at least 10 report cohort for the 2015-2016, 2016-2017, and 2017-2018 influenza seasons are shown in Table 3.4. Although the 2015-2016 influenza season had more total participants (n=43 944) than the 2016-2017 (n=28 526) and 2017-2018 (n=30 531) influenza seasons, a greater percentage of participants were included in the drop first report and at least 10 report cohort populations during these seasons. For the 2015-2016 influenza season, 63% of the total population was included in the drop first report population and less than 30% of the total population was included in the at least 10 report cohort, whereas over 80% and 40% of the total population were included in the drop first report population and at least 10 report cohort, respectively, during the 2016-2017 and 2017-2018 influenza seasons. The 2016-2017 and 2017-2018 influenza seasons also had a greater median number of reports per participant for the total population compared to the 2015-2016 influenza season, 5 (1, 23) vs. 2 (1, 13). This change in reporting habits is captured in Figure 3.8. As indicated by the blue bars, approximately 20% of the potential reports are submitted each week during the 2015-2016 influenza season, on the other hand, approximately 30% of the potential reports are submitted each week during the 2016-2017 and 2017-2018 influenza seasons. Across all seasons, the percent of females decreased and the median age increased as the entry criteria became more restrictive.

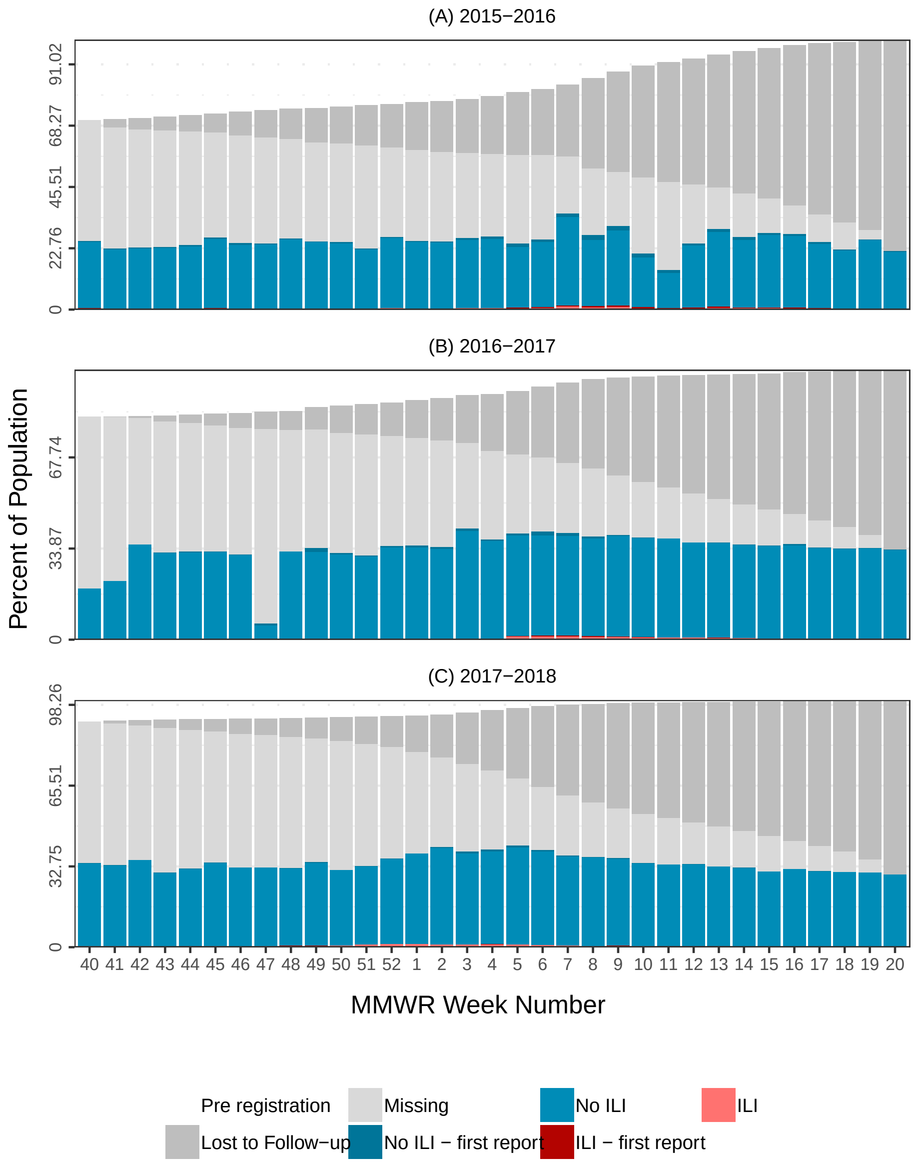

Figure 3.8: Histograms of FNY participant reporting habits during the (A) 2015-2016, (B) 2016-2017, and (C) 2017-2018 influenza seasons.

Table 3.4: Descriptive statistics of different FNY populations for the 2015-2016, 2016-2017, and 2017-2018 influenza season. Continuous variables are displayed as median (IQR) and categorical variables are displayed as n (%).

Variable	Total	Drop 1st report	At least 10 cohort
2015-2016			
N	43944 (100%)	27715 (63.07%)	13068 (29.74%)
Number of reports	369235 (100%)	353006 (95.6%)	297603 (80.6%)
Reports per week	11157 (10509, 11827)	10490 (9503, 11052)	9209 (8546, 9653)
Reports per participant	2 (1 , 13)	8 (3, 23)	24 (16, 29)
Male	15585 (36.24%)	10552 (38.9%)	5594 (43.66%)
Female	27304 (63.5%)	16537 (60.96%)	7218 (56.33%)
Unknown	112 (0.26%)	39 (0.14%)	2 (0.02%)
Age	50.9 (34.7, 62.3)	52.5 (35.2, 63.5)	56.5 (38.2, 65.4)
2016-2017			
N	29526 (100%)	24218 (82.02%)	12036 (40.76%)
Number of reports	328455 (100%)	323147 (98.38%)	285844 (87.03%)
Reports per week	10285 (9694, 10985)	9983 (9498, 10659)	9227 (8435, 9570)
Reports per participant	5 (1, 23)	9 (2, 26)	26 (18, 30)
Male	11041 (38.17%)	9486 (40.01%)	5102 (43.04%)
Female	17818 (61.6%)	14191 (59.86%)	6750 (56.94%)
Unknown	66 (0.23%)	30 (0.13%)	2 (0.02%)
Age	53.6 (36.4, 64.1)	54.45 (37.2, 64.8)	58.2 (41.2, 66.5)
2017-2018			
N	30531 (100%)	26546 (86.95%)	12354 (40.46%)
Number of reports	344948 (100%)	340963 (98.84%)	297815 (86.34%)
Reports per week	10255 (9721, 11018)	10155 (9644, 10860)	9047 (8702, 9366)
Reports per participant	5 (1, 23)	8 (2, 25)	26 (18, 30)
Male	11505 (38.62%)	10242 (39.48%)	5187 (42.69%)
Female	18188 (61.06%)	15640 (60.29%)	6953 (57.23%)
Unknown	96 (0.32%)	60 (0.23%)	9 (0.07%)
Age	52.8 (35.8, 64.3)	54.1 (36.9, 64.9)	58.9 (42, 67.3)

3.2.3.2 *Disease Burden*

Appendix Table A.5 and Figure 3.9 display age group specific estimates of ILI disease burden and 95% CIs, expressed as per 1 000 000 persons, by influenza season. Overall, the 2017-2018 influenza season has the largest estimates of disease burden for both influenza and ILI, and the 2016-2017 influenza season has the smallest estimates of disease burden for influenza and ILI. Although the general age group patterns are consistent across all seasons, there are some slight differences in estimates within each age group and season. The method that drops the first report has smaller ILI burden estimates compared to the method that includes all reports. This difference indicates that a large percentage of individuals report ILI for their first report, particularly in the 18-49 and 50-64 age groups, where this difference is more pronounced. For most age groups and seasons, the cohort methods have larger estimates compared to the method that includes all reports, except for older age groups during the 2015-2016 influenza season. As expected, both MI methods have larger estimates of ILI burden compared to all methods that assume all missing reports are non-ILI. The estimates from the δ MI method are less than the estimates from the MI, except for the <5, 50-64, and 65+ age groups during the 2016-2017 influenza season.

3.2.3.3 *Simulations*

Violin plots are shown in Figure 3.10, and NRMSEs are displayed in Table 3.5. Under each model, using all reports and dropping the first report underestimate the IP and have the largest NRMSEs. Under MoCAR/MCAR and MoAR/MAR assumptions, the cohort MI method best estimates the original parameter and has the smallest NRMSEs. However, under MoNAR/MNAR assumptions, the cohort

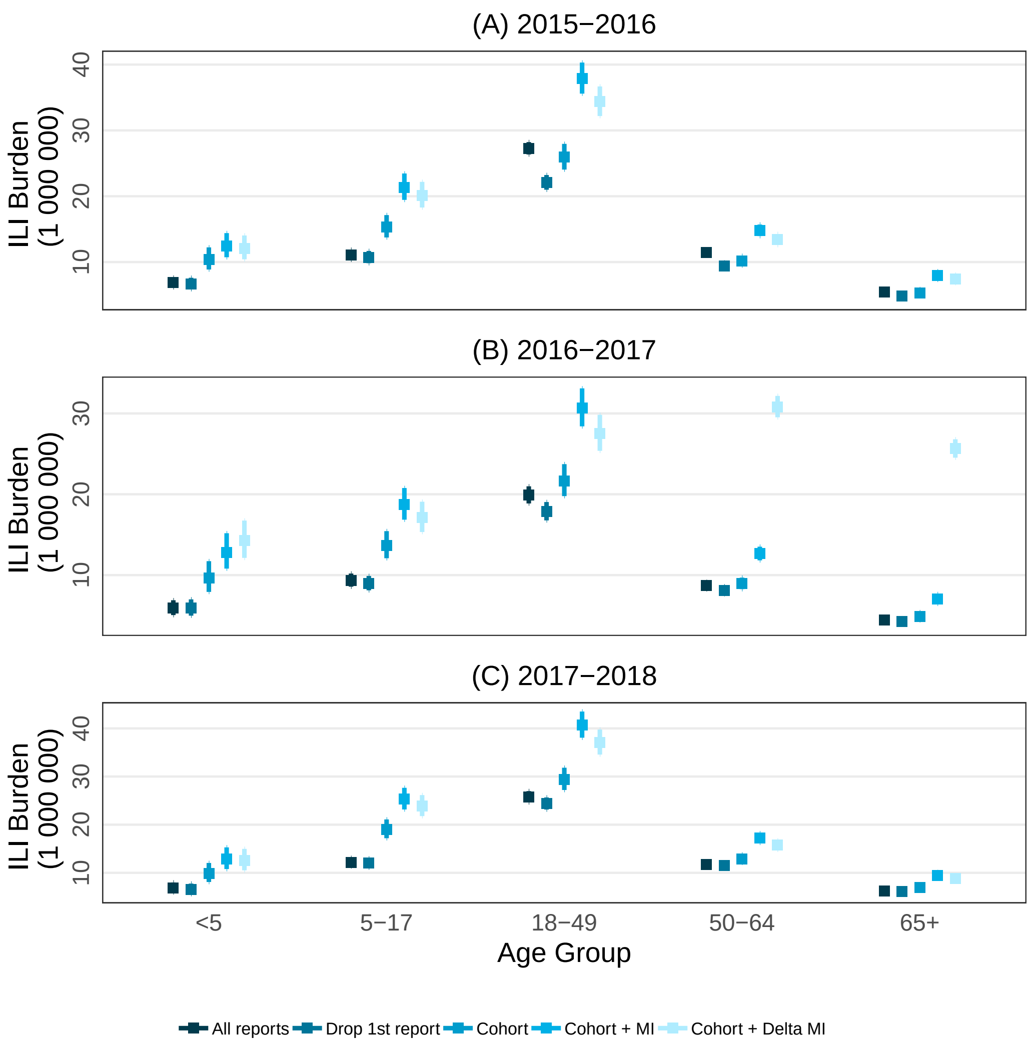

Figure 3.9: Age group specific ILI burden, expressed as per 1 000 000 persons, for the (A) 2015-2016 (B) 2016-2017 and (C) 2017-2018 influenza seasons.

method that assumes all missing reports are non-ILI has the smallest NRMSEs. Although the cohort δ MI method has the second smallest NRMSE, there is a large range of estimates in the <5 age group.

Table 3.5: NRMSE, expressed as percentage, by age-group for IR and IPs and overall for WP under MCAR, MAR, and MNAR missingness models.

Age Group	All reports	Drop 1st report	Cohort	Cohort + MI	Cohort + δ MI
MCAR					
<5	61.42	56.32	32.34	18.50	89.33
5-17	61.90	56.88	32.53	7.23	15.70
18-49	62.30	57.32	32.96	5.20	12.25
50-64	62.61	57.67	33.38	5.91	12.59
65+	62.95	58.04	33.78	8.03	12.60
MAR					
<5	59.84	54.05	30.28	12.29	50.72
5-17	60.33	54.55	30.50	4.64	11.67
18-49	60.68	55.00	30.94	3.22	11.48
50-64	61.01	55.36	31.37	3.52	11.70
65+	46.95	45.65	26.45	3.04	9.98
MNAR					
<5	47.91	40.01	10.60	32.90	37.53
5-17	48.26	40.28	9.46	32.32	16.95
18-49	48.35	40.37	9.08	33.38	17.13
50-64	48.45	40.48	8.98	35.15	18.06
65+	36.15	34.79	12.08	22.63	8.71

3.2.4 Discussion

The CDC has a well-established, robust influenza surveillance system. However, this system primarily uses data from sentinel systems and relies on mathematical models to estimate the overall influenza burden in the U.S. population. FNY has the potential to compliment CDC by reaching populations who do not seek health care for symptoms.

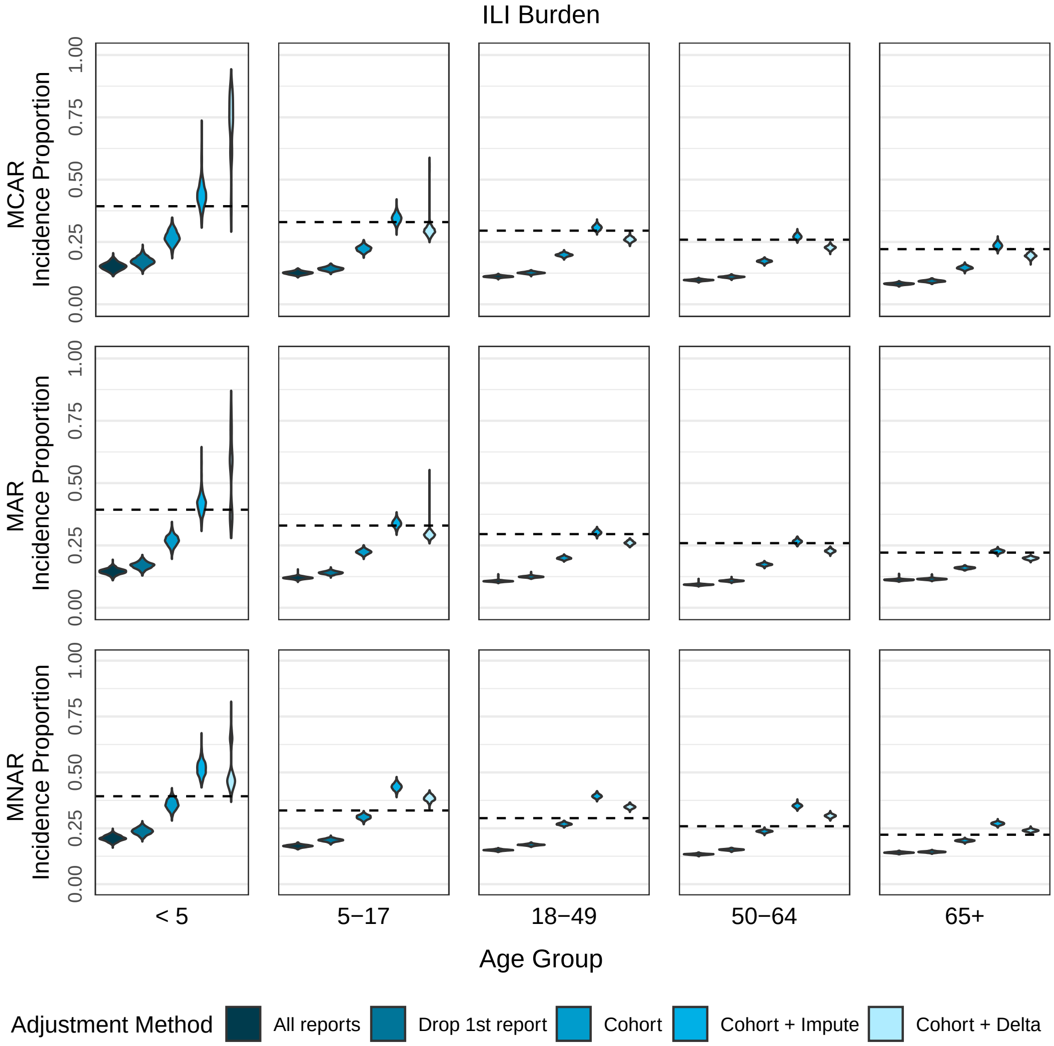

Figure 3.10: Violin plots of simulated age group specific IPs under three models of motivation and missingness: MoCAR/MCAR, MoAR/MAR, and MoNAR/MNAR. Dotted line represents the original simulated parameter.

Although studies have proposed different methods to address reporting inconsistencies in participatory syndromic surviellance data, no study has evaluated these approaches using both real data from FNY as well as simulated data. We found that when estimating ILI burden, the drop first report method produces ILI burden estimates that are less than using all reports. This difference indicates that a large fraction of participants report ILI on their first report. In addition, both the MI and δ MI models have the largest estimates. Based on the simulations, the cohort approach with MI is most appropriate under MoCAR/MCAR and MoAR/MAR models, whereas the cohort method that assumes that all missing data is non-ILI is most appropriate under MoNar/ MNAR models. Past research has shown that MoCAR/MCAR and MoAR/MAR models may not reflect reality. For example, 20-30% of FNY users who completed a post-season survey reported that they are more likely to report when ill (L. Goodwin, personal communication, March, 2019). Additionally, as discussed in Section 2.1.3.1, FNY users who reported ILI symptoms are less likely to continue reporting. As a result, when using FNY data to retrospectively estimate ILI burden, the cohort method that assumes that all missing data is non-ILI may be most appropriate.

3.2.4.1 Choice of reporting entry criteria

A reporting entry criteria of at least 10 reports is used for the cohort methods to be consistent with past studies.(Chunara et al., 2015; Reed et al., 2016) As a sensitivity analysis, we assess the relationship between the reporting entry criteria with both the percent of participants remaining and the percent of total data observed as well as with the IP for all three seasons. As discussed in the Section 3.2.3.1, participant reporting habits during the 2015-2016 influenza season are different than the 2016-2017 and 2017-2018 influenza seasons, however, the overall patterns between the reporting entry and percent of subjects remaining and percent of total observed are similar (Figure 3.11). If all participants are included in the analysis (100% participants remaining) then only 25-30% of the total possible reports for a season are observed, which means that 70-75% of the data would need to be imputed. If a reporting entry criteria of at least 3 reports is used, then 50-60% of the total participants remain and 45-55% of the total possible reports are observed. A reporting entry criteria of at least 10 reports includes 30-40% of the total population and 65-75% of the total possible reports are observed. There is a distinct trade off between keeping participants and percent of observed reports when selecting the reporting entry criteria, and this criteria should be carefully considered when applying MI methods under MNAR. Despite this trade-off, there is little variation in IP estimates (i.e.: the percent of participants who reported an incidence of ILI) between reporting entry criteria of 5 through 25 across most age groups and seasons (Figure 3.12). Based on these results, a reporting entry criteria of 5 may be an appropriate choice, unless a MI method is applied.

3.2.4.2 Next Steps

As mentioned in Section 3.2.2.4, we assume that the motivation and missingness models are concordant. However, because there are nine possible motivation and missingness combinations, we will assess how these methods work under different combinations. In addition, for the MoNAR and MNAR models, we only assessed one value of δ. In the future, we will also assess how well the methods perform as the value of δ changes.

Unlike Flutracking.net, FNY does not collect missed reports from users. As a result, there is no clear way to estimate the probability that a report is ILI given the report is missing for FNY. For this analysis, we use information from end of sea-

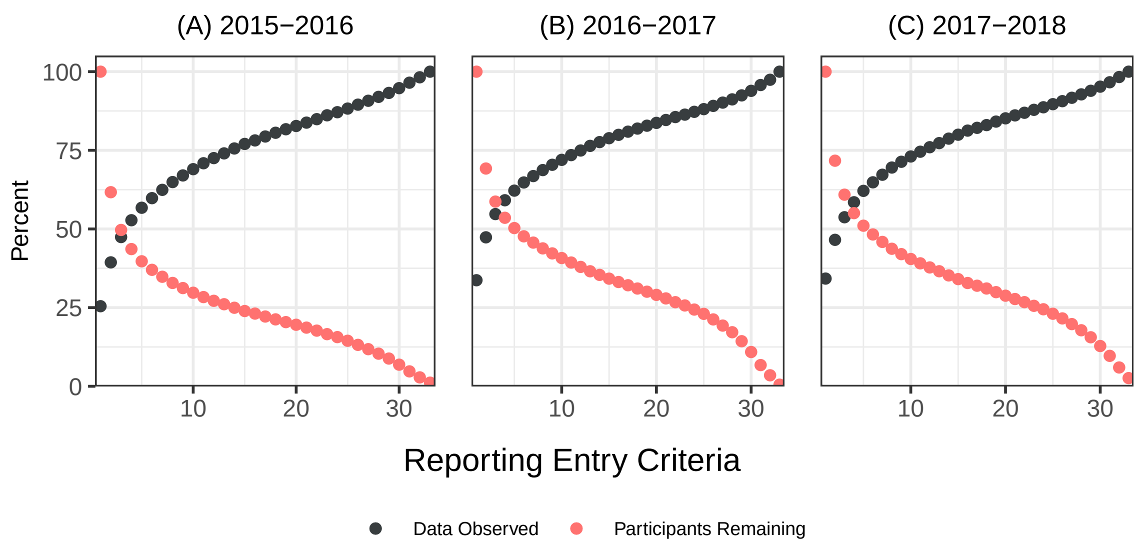

Figure 3.11: Percent of data observed (black) and percent of FNY population remaining (red) by reporting entry criteria for the (A) 2015-2016, (B) 2016-2017, and (C) 2017-2018 influenza seasons.

son surveys and expert opinion to estimate δ. We are currently investigating better ways to estimate this value. Furthermore, the estimates of ILI burden using the δ MI method varied widely during the 2016-2017 influenza season. There were errors in data collection from October 2016 through December 2016, and these errors may have contributed to the issues in convergence for the δ MI method. Both the MI and δ MI models will be re-run without data from this time period to assess whether more consistent estimates of ILI burden can be obtained.

Since January, the CDC has been releasing weekly estimates of cumulative influenza burden. Our study has shown that a cohort method is the best approach to estimate ILI burden. However, this method can only be applied for retrospective analyses because we do not know which participants will reach the reporting entry criteria at the beginning of the influenza season. As a result, other methods, such as adjustment through a scaling factor, will be assessed to determine if near-real time estimates of cumulative ILI burden can be produced.

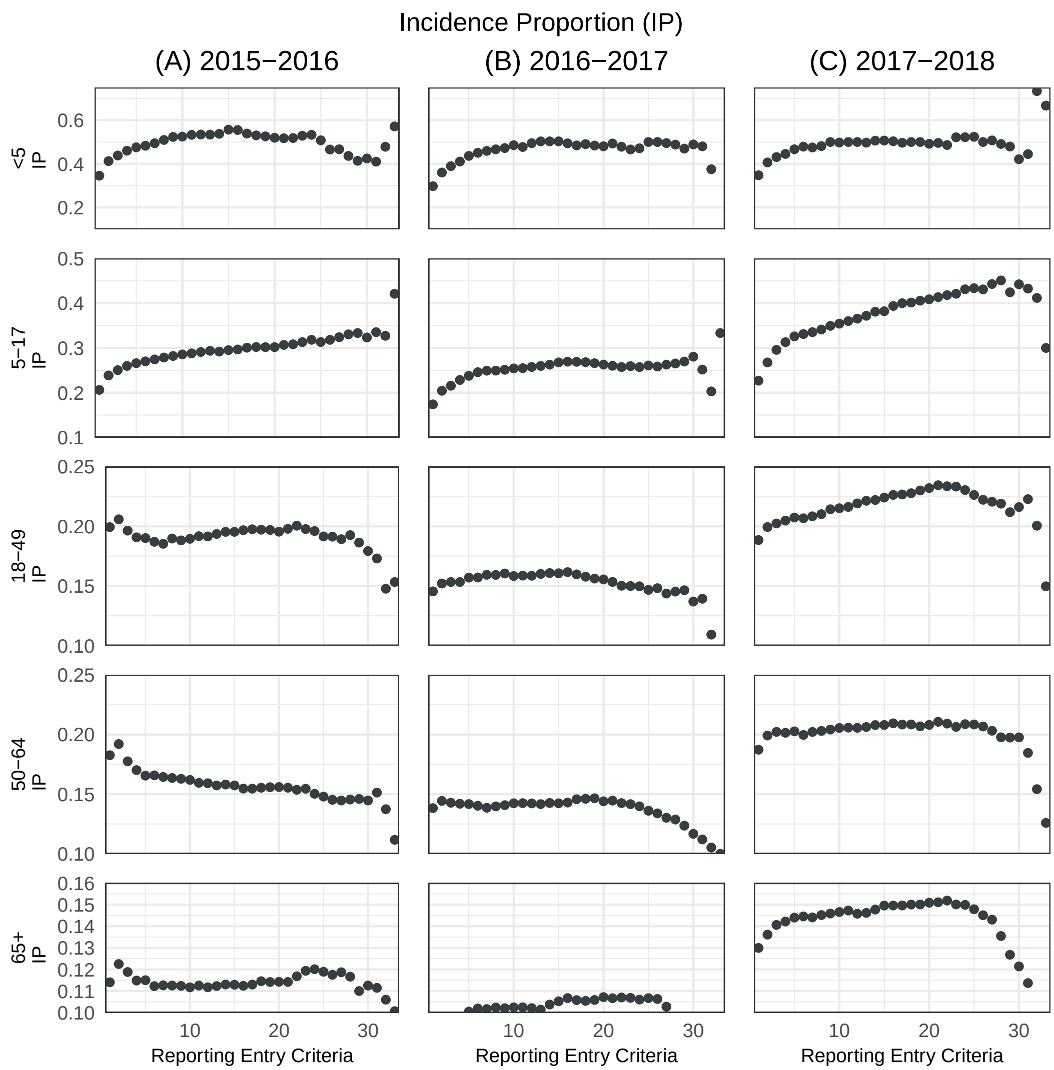

Figure 3.12: Estimated age-group specific IPs for different reporting entry criteria for the (A) 2015-2016, (B) 2016-2017, and (C) 2017-2018 influenza seasons.

CHAPTER 4

Applying these methods to answer questions about influenza in the community

4.1 IS THERE REALLY MORE FLU IN THE SOUTH? SURVEILLANCE SYSTEMS SHOW DIFFERENCES IN INFLUENZA ACTIVITY ACROSS REGIONS

4.1.1 Overview

As discussed in Section 1.1.2, the CDC tracks patients who seek medical attention with ILI symptoms through ILINet. This surveillance system includes thousands of volunteer health care specialists, including individual providers, group practices, and hospital-based clinics located throughout all 50 states, Puerto Rico, the District of Columbia, and the US Virgin Islands. Because participation in ILINet is voluntary and each state is responsible for their own recruitment of healthcare providers, the composition of provider-types, coverage of geographical regions, and consistency of provider reporting varies from state to state. This convenience sample-driven model of surveillance results in certain parts of the population being over- or under-represented in the reported influenza activity.(Polgreen et al., 2009b; Lee et al., 2018b; Scarpino et al., 2012) At both national and HHS-defined regional levels, the CDC routinely reports the weekly percentages of patients presenting with ILI to healthcare providers. In addition, each season the CDC calculates and reports region-specific baselines, using influenza activity data from previous seasons, to identify the beginning and end of the influenza season and contextualize the severity of a given region-specific outbreak. These baselines vary widely across regions, and the degree to which the differences in baselines, as well as percent ILI visits during an influenza season, reflect actual differences in influenza activity or systematic differences in the methods used to collect the data is unclear. Recent models suggest that the spatial patterns in U.S. sentinel ILI surveillance may be the result of socio-environmental factors, state-specific health policies, and sampling.(Lee et al., 2018b) Identifying and characterizing the presence of potential methodological measurement biases in ILINet is important, as it is frequently used as an indicator of influenza activity for decision-making purposes, as well as the ground truth in mechanistic and statistical predictive modeling efforts aimed at understanding disease transmission dynamics and monitoring and forecasting influenza activity.(Zhang et al., 2017; Biggerstaff et al., 2016, 2018; Brooks et al., 2015; Tizzoni et al., 2012; Shaman et al., 2010; Yang et al., 2015a, 2017; Santillana et al., 2015, 2014) Here, we qualitatively and quantitatively compare national and region-specific baselines and ILI activity during three influenza seasons across four surveillance data sources:

- CDC ILINet
- FNY
- athenahealth, a provider of cloud-based electronic health record (EHR) services
- HealthTweets.org, a research platform that shares health trends data from Twitter

to determine whether these data sources, commonly used as input in influenza modeling efforts, show regional structural patterns that are similar to those observed in CDC ILINets data. We also compare yearly self-reported health care-seeking behavior of FNY participants to determine if this factor can better characterize the differences in ILI activity across regions.

4.1.2 Methods

4.1.2.1 Data

CDC ILINet The CDC reports the weighted percentage of patient visits to healthcare providers presenting ILI symptoms on a weekly basis at the national and regional levels. These values are weighted on the basis of state population and represent the percentage of patient visits to healthcare providers that present as ILI, defined as fever (temperature of $37.8\,^{\circ}\mathrm{C}$ [$100\,^{\circ}\mathrm{F}$] or greater) plus a cough and/or a sore throat without a known cause other than influenza.

FNY Data collection for FNY is described in section 1.2.2.1. National and regional percent of ILI symptoms reported is calculated by dividing the number of participants reporting ILI, defined as a report of fever plus cough and/or sore throat, in a given week by the total number of FNY participant reports in that same week. FNY participants are assigned to a region based on the zip code provided at registration. Unweighted FNY percent of ILI symptoms is used to maintain consistency across previous studies and the FNY website.

athenahealth A provider of cloud-based services and mobile applications for medical groups and health systems originates this data set. National and regional percent of visits for ILI is calculated by dividing the Unspecified Viral or ILI Visit Count, which is equal to the number of visits where the patient had an unspecified viral diagnosis, an influenza diagnosis, or a fever diagnosis with an accompanying sore throat or cough diagnosis, by the total number of provider visits each week.

HealthTweets.org This dataset is generated by an online research platform (HealthTweets.org) that shares the output of Twitter data mining algorithms with researchers and public officials.(Paul et al., 2014) We use weekly aggregated trends data from each state to calculate the influenza prevalence measure for each region. Weekly national and regional influenza prevalence measures are calculated by normalizing the number of influenza infection tweets in the health stream by the total number of tweets in the general stream during the same week.(Broniatowski et al., 2013)

4.1.2.2 Statistical methods

Baseline comparison CDC ILINet national and regional baselines for the 2017-2018 influenza season are available on the CDC website.(Centers for Disease Control , CDC) National and regional baselines for FNY, athenahealth, and HealthTweets.org are estimated following CDCs baseline definition. A baseline is defined as the mean percentage of ILI activity during non-influenza weeks, for the previous three seasons, plus two standard deviations. Non-influenza weeks during these seasons are the same for all three systems and are delineated, by the CDC, as periods of two or more consecutive weeks in which each week accounted for less than 2% of the seasons total number of specimens that tested positive for influenza in public health laboratories. Descriptive statistics of baselines are presented as median (IQR).

ILI activity comparison Differences in weekly ILI activity across geographical areas within each surveillance data source are assessed using data from the start of the 2015-2016 influenza season (week of October 5, 2015) through the end of the 2017-2018 influenza season (week of October 1, 2018). Weekly ILI activity across geographical areas within each data source is quantitatively compared by dividing the difference in ILI activity between two areas by the maximum within each week, defined by

$$\text{Mean Relative Difference} = \frac{1}{K}\sum_{week_k}\frac{ILI_{ik} - ILI_{jk}}{max\left(ILI_{ik}, ILI_{jk}\right)} \tag{4.1}$$

for

$$i, j \in \{\text{Region 1:10, National}\}.$$

Mean relative differences within each data source are summarized using heatmaps, where the geographical areas along the rows are represented by i in the equation above and the geographical areas along the columns are represented by j. Time series heatmaps are also presented to qualitatively compare weekly ILI activity across geographical areas for each data source.

Health care-seeking behavior National and regional health care-seeking percents are calculated for each influenza season by dividing the number of FNY participants who sought medical care for ILI symptoms, as defined above, by the total number of ILI reports within an influenza season. Because health care-seeking behavior varies by age,(Biggerstaff et al., 2014) health care-seeking rates are also adjusted by age-group (<18, 18-49, 50-64, 65+) using population data from the 2010 U.S. census.(U.S. Census Bureau, 2015)

4.1.3 Results

4.1.3.1 CDC ILINet

Table 4.1 and Figure 4.1 provide the ILI activity baselines for each data source across national and regional levels. The national baseline for CDC ILINet during the 2018-2019 influenza season is 2.2, and the median CDC ILINet regional baseline is 2.1 (1.8-2.3). Region 10 has the smallest baseline, 1.1, whereas Region 6 has the largest baseline, 4.0. As shown in Figure 4.2A, Regions 2 and 6 have consistently higher weekly percent of ILI visits compared to other regions, indicated by the red shades across the row, whereas Regions 1, 8, and 10 have consistently lower weekly percent of ILI visits, indicated by the blue shades across the row. This pattern is also shown qualitatively in Figure 4.3A and Appendix Figure A.5, where darker shades of blue, as seen for Regions 2, 6, and 9, correspond to higher percent of ILI visits.

Table 4.1: Regional and national ILI activity baselines for the 2018-2019 influenza season for CDC ILINet, FNY, athenahealth, and HealthTweet-s.org

Geographical Area	CDC ILINet	FNY	athenahealth	HealthTweets.org
Region 1	1.8	2.1	1.3	0.8
Region 2	3.1	2.4	1.7	0.4
Region 3	2.0	2.4	1.5	0.5
Region 4	2.2	2.7	1.4	0.6
Region 5	1.8	2.6	1.1	0.5
Region 6	4.0	2.6	1.9	0.7
Region 7	1.6	2.5	1.0	0.7
Region 8	2.2	2.9	1.0	0.8
Region 9	2.3	2.5	1.7	0.6
Region 10	1.1	2.5	0.6	0.7
National	2.2	2.3	1.4	0.5

4.1.3.2 FNY

For FNY, the national baseline is 2.3, and the median regional baseline is 2.5 (2.4-2.6). The minimum baseline is 2.1, Region 1, and the maximum baseline is 2.9, Region 8. Compared to other data sources, the mean relative differences for FNY in Figure 4.2B show less heterogeneity and no consistent patterns in percent ILI across regions. Although the timing of peaks in percent ILI varies between regions,

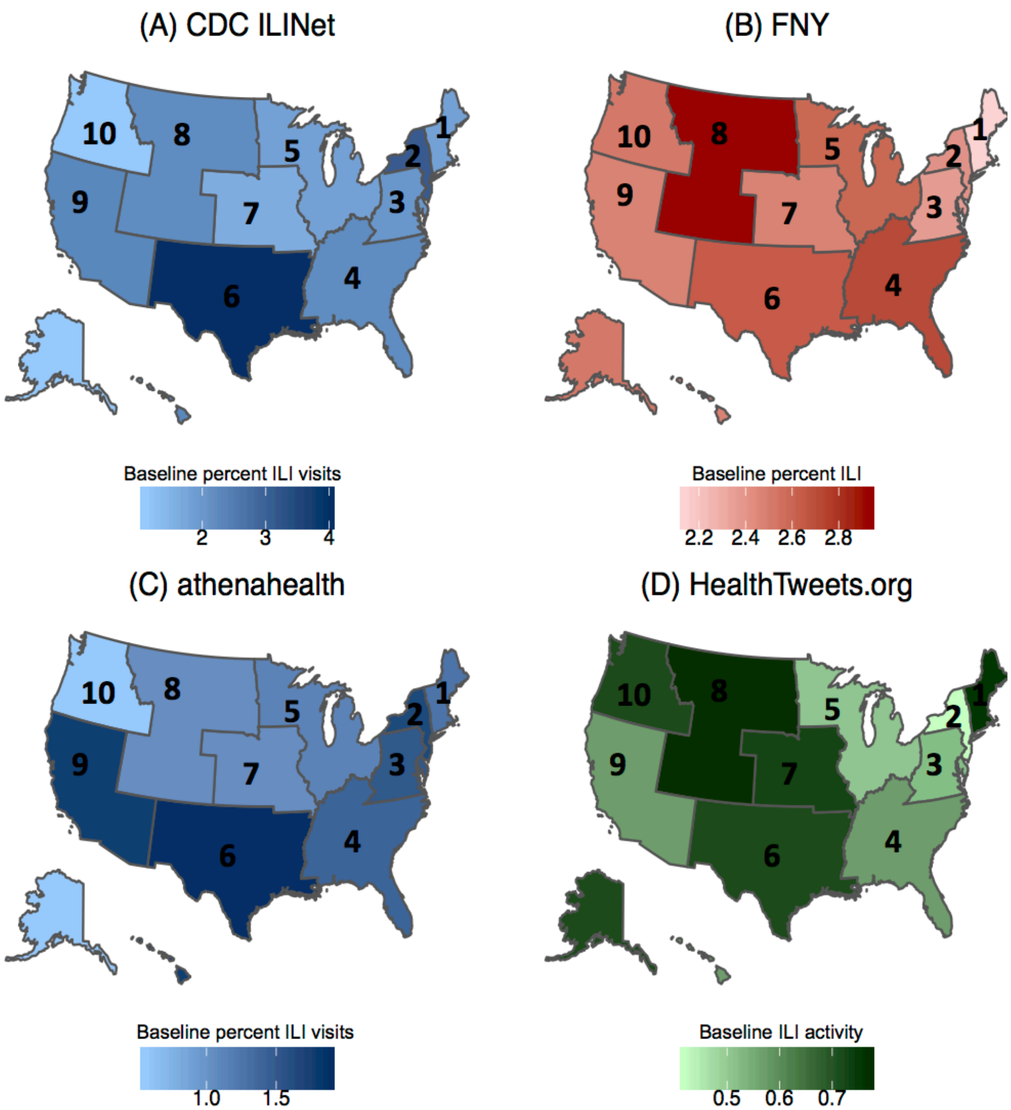

Figure 4.1: Spatial heatmaps of U.S. regional baseline ILI activity for the 2017-2018 influenza season for (A) CDC ILINet, (B) FNY, (C) athenahealth, and (D) HealthTweets.org

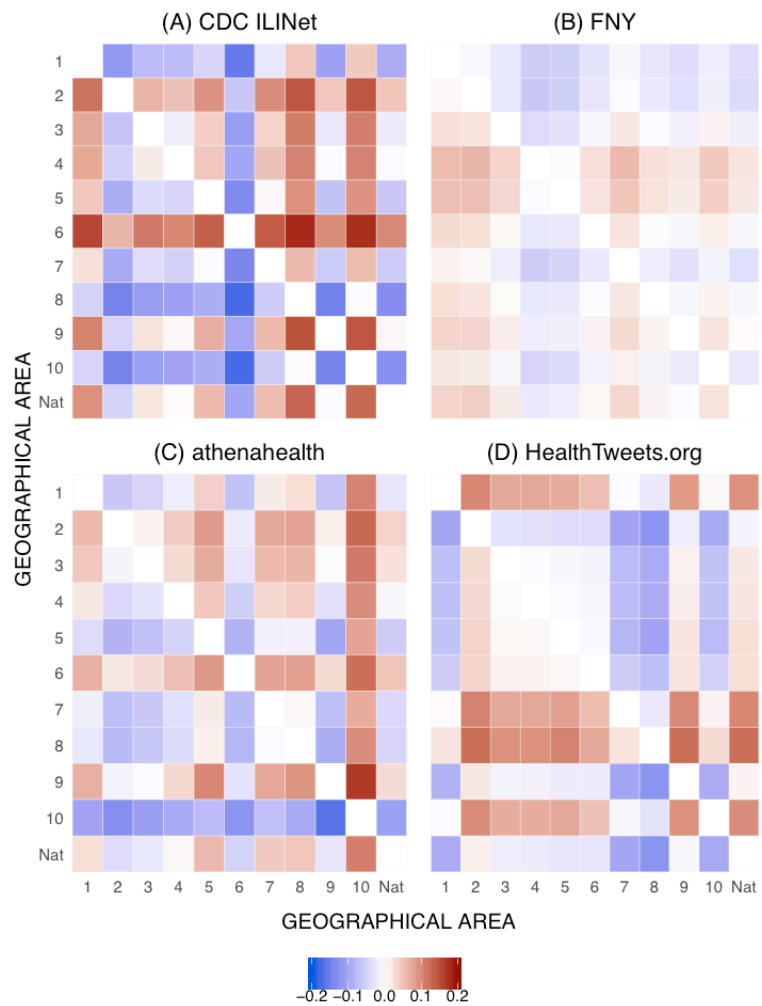

Figure 4.2: Heatmaps of the mean relative difference of ILI activity across geographical areas for (A) CDC ILINet, (B) FNY, (C) athenahealth, and (D) HealthTweets.org

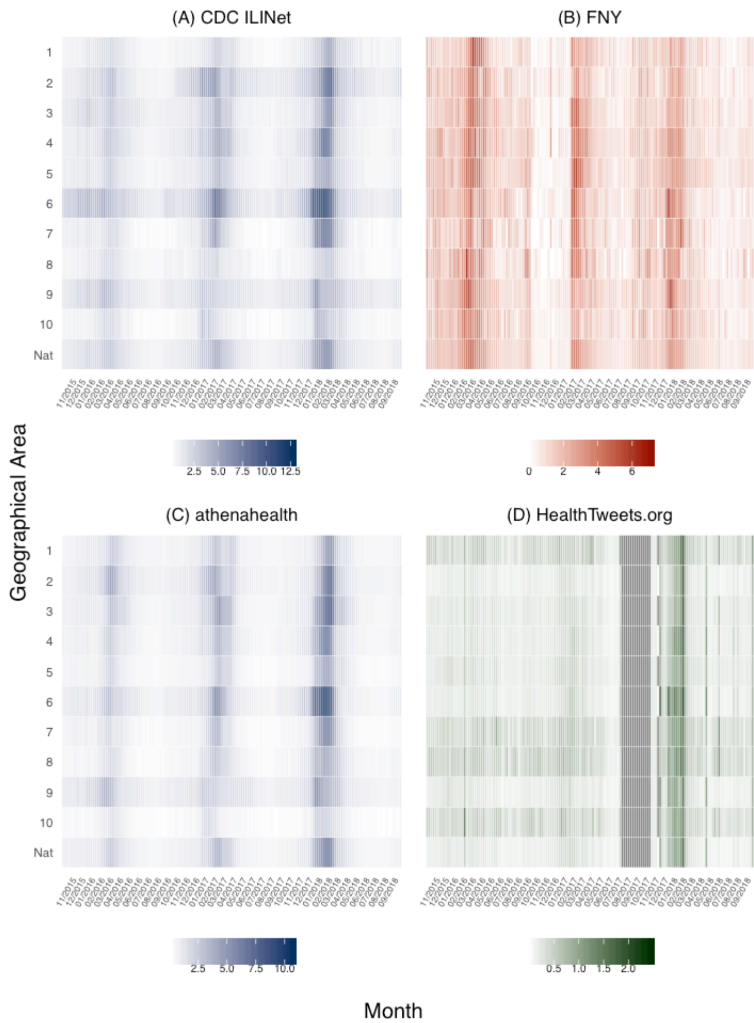

Figure 4.3: Time series heatmaps of ILI activity across geographical areas for (A) CDC ILINet, (B) FNY, (C) athenahealth, and (D) HealthTweets.org

the relative percent ILI is consistent across regions and seasons (Figure 4.3B and Appendix Figure A.6).

4.1.3.3 athenahealth

The national baseline for athenahealth is 1.4, and the median regional baseline is 1.3 (1.0-1.6). Region 10 has the minimum baseline of 0.6, and Region 6 has the maximum baseline of 1.9. Similar to CDC ILINet, Regions 2, 6 and 9 have consistently higher weekly percent of ILI visits compared to other regions, and Regions 7, 8, and 10 have consistently lower weekly percent of ILI visits (Figure 4.2C). This pattern is reflected in Figure 4.3C and Appendix Figure A.7 as Regions 2, 6, and 9 have consistently higher percent of ILI visits across all seasons.

4.1.3.4 HealthTweets.org

The national baseline is 0.5, the median baseline is 0.6 (0.5-0.7), the minimum baseline is 0.4, Region 2, and the maximum baseline is 0.8, Region 8. Unlike CDC ILINet and athenahealth, HealthTweets.org show higher ILI activity in Regions 1, 7, 8, and 10 (Figures 4.2D). These regions have mean normalizing constants that are less than half the mean normalizing constants of other regions (Table 4.2). As shown in Figure 4.3D and Appendix Figure A.8, this pattern is consistent across seasons.

4.1.3.5 Health care-seeking behaviors

The age-adjusted percent of FNY participants who sought health care for ILI symptoms are shown by season and across all seasons in Table 4.3 and Figure 4.4. At the national level, a higher percent of participants sought health care for ILI symptoms

Table 4.2: Descriptive statistics of the HealthTweets.org normalizing constant at the national and regional level

Geographical Area	Normalizing Constant [Mean $\pm$ standard deviation]
Region 1	210.82 $\pm$ 114.917
Region 2	627.69 $\pm$ 330.270
Region 3	599.53 $\pm$ 293.320
Region 4	1103.78 $\pm$ 553.374
Region 5	798.25 $\pm$ 387.266
Region 6	845.30 $\pm$ 414.785
Region 7	171.05 $\pm$ 82.077
Region 8	121.96 $\pm$ 63.936
Region 9	5848.54 $\pm$ 3775.923
Region 10	181.33 $\pm$ 97.756
National	6352.25 $\pm$ 3351.390

during the 2016-2017 season, 35.1%, compared to the 2015-2016 and 2017-2018 seasons, 21.7% and 29.2%, respectively. Within each season, Regions 2, 4, and 6 have the highest percent of participants who sought health-care, whereas Regions 1, 5, 9, and 10 have the lowest percent of participants who sought health care.

Table 4.3: Age-adjusted regional and national percent of FNY participants who sought health care for ILI symptoms

Geographical Area	All Seasons	2015-2016	2016-2017	2017-2018
Region 1	25.98	20.82	33.29	27.77
Region 2	29.97	26.05	36.03	31.79
Region 3	28.66	22.07	37.03	31.73
Region 4	32.61	25.47	43.23	34.77
Region 5	26.43	21.53	34.59	26.73
Region 6	35.17	28.58	44.83	37.47
Region 7	30.93	23.79	41.95	32.09
Region 8	25.50	22.74	30.86	26.16
Region 9	22.49	19.06	27.77	24.69
Region 10	20.03	17.03	23.39	22.33
National	27.12	21.73	35.06	29.23

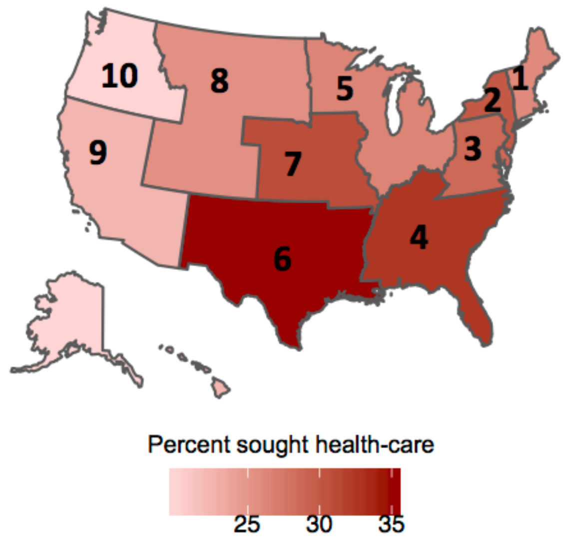

Figure 4.4: Spatial heatmap of age-adjusted regional and national percent of FNY participants who sought health care for ILI symptoms

4.1.4 Discussion

Our findings show that differences in ILI activity across regions, as reported by a given surveillance system, are not consistent across surveillance platforms. In other words, regions that show larger baselines (and thus higher overall historical ILI activity) in one surveillance system appear to be different to their counterparts in other surveillance systems. The heterogeneity of recruitment practices of health-care providers for each system, the composition of provider types, and the variability and consistency of coverage of geographical regions have the potential to contribute substantially to these systematic differences in baselines.(Lee et al., 2018a) Our findings suggest that these structural differences reflect methodological collection practices rather than actual differences in influenza activity across regions. The observed structural patterns within each surveillance system were consistent across individual influenza seasons (Appendix Figure A.9). This pattern implies that the differences in ILI active do reflect a specific time period heterogeneity. Specifically, baselines from CDC ILINet vary across different geographical areas, and the geographical areas with the largest baseline values also have consistently larger percent of ILI visits during the influenza season. On the other hand, FNYs baselines and percent ILI were similar across geographical areas. This similarity is captured by the homogeneity in the mean relative differences. One potential contributing factor to the observed differences in patterns between these surveillance systems is the activity being measured. CDC ILINet measures the number of patient visits with ILI symptoms out of the total number of patient visits, whereas FNY measures the number of ILI reports out of enrolled persons who submitted a report. Furthermore, the population under surveillance also differs, as FNY includes individuals who may not seek medical attention and FNY has a different demographic profile compared to CDC ILINet. For example, females and middle-aged participants are over-represented in FNY.(Baltrusaitis et al., 2017)

Although not identical, athenahealth showed similar patterns in both baseline measures as well as percent of ILI visits to CDC ILINet across geographical areas. Both CDC ILINet and athenahealth use data from individuals seeking medical care. However, athenahealth has only a partially overlapping coverage of healthcare providers, and the proportion of visit settings differs slightly between the two systems. Most of athenaheaths providers see patients in office-based settings. Other settings, such as Emergency Rooms (ER) and nursing facilities, are under-represented compared to CDC ILINet.(Santillana et al., 2016)

Unlike FNY, patterns across geographical areas within Twitter ILI activity appear to be the opposite of the patterns shown by CDC ILINet and athenahealth, as areas with consistently lower Twitter ILI activity had consistently higher percent of ILI visits for CDC ILINet and athenahealth, and vice versa. One potential reason for the differences in patterns in ILI activity across data sources is the difference in the activity being measured. As mentioned above, both CDC ILINet and athenahealth measure the number of ILI visits out of total visits, whereas HealthTweets.org normalizes the number of influenza infection tweets by the total number of tweets in the general stream. Also, the groups most susceptible to influenza illness, young children and the elderly, may be underrepresented on Twitter. Furthermore, we found that smaller normalizing constants correspond to higher values of ILI activity.

Despite these differences in patterns of ILI activity within systems, current research shows that these alternative data sources track CDC ILINet at both the national and regional levels. At the national level, the correlation between CDC ILINet and athenahealth is 0.97, and regional correlations range from 0.90 to 0.97.(Baltrusaitis et al., 2018) The correlation between CDC ILINet and FNY at the national level is 0.81, and regional correlations range from 0.64 to 0.81.(Baltrusaitis et al., 2018) Twitter-based influenza prevalence measures show a correlation of 0.93 with CDC ILINet at the national level, and a correlation of 0.88 with New York Citys weekly ED visits for ILI.(Broniatowski et al., 2013)

Compared to other recent publications, the percent of FNY participants who sought medical care for ILI was less than reported estimates. A recent meta-analysis that used estimates from multiple countries across different influenza seasons estimated an overall pooled health care-seeking rate of 0.52 (95% CI 0.46-0.59).(Ma et al., 2018) In the U.S., national reported health care-seeking percents for children were 56% and 57% during the 2009-2010 and 2010-2011 influenza seasons, respectively. Among adults, 40% reported that they sought health during the 2009-2010 influenza season, and 45% reported that they sought health during 2009-2010 influenza season.(Biggerstaff et al., 2012, 2014) Interestingly, the percent of FNY participants seeking health care for ILI symptoms differs slightly across regions. These differences may contribute to the differences in CDC ILINet and athenahealth baseline activity, as regions with higher health care-seeking percents correspond to regions with higher CDC ILINet and athenahealth baselines.

From a predictive modeling perspective, our findings may explain why certain approaches designed to predict CDC ILINet values for the "Predict the Influenza season challenge", weeks ahead of the publication of official CDC reports, may work better than others. As discussed in the two existing reports that document the performance of different methodologies to predict influenza activity, models that rely on local statistical approaches that exploit region-specific autoregressive information and historically observed ILI activity from previous seasons, as well as external predictors (such as humidity data, Google searches, Wikipedia),(Brooks et al., 2015; Shaman et al., 2010) outperform mechanistic agent-based stochastic Susceptible-Infected-Recovered (SIR) models that aim at modeling individual humans behavior, to infer epidemic activity across spatial resolutions.(Biggerstaff et al., 2016, 2018; Tizzoni et al., 2012) The former modeling approaches are "trained" to track ILI levels in a region-specific fashion (frequently ignoring inconsistency across spatial resolutions), whereas the latter agent-based stochastic SIR models aim to predict the whole national epidemic outbreak across geographic areas. In other words, if the ILI activity report varies depending on how data is aggregated, then even a very accurate agent-based model may not be able to capture influenza activity correctly for every geographic area.

Our study has several limitations. There were errors in FNY data collection from October 2016 through December 2016 resulting in an underestimation in the weekly percent of ILI reports. We did not adjust the ILI estimates for these weeks. There was also an issue in data collection during week of August 28, 2017. We adjusted the estimates for this week by taking the average percent of ILI reports of the previous and subsequent weeks. In addition, there were a few weeks during the summer of 2017 during which there were no reports of ILI activity for HealthTweets.org. We did not input or estimate these missing weeks. Because the overall patterns of ILI activity were similar across seasons (Appendix Figure A.9), we do not suspect that these data issues affected our overall conclusions. Another limitation is that because each system has a different measure of ILI activity we cannot directly compare measures across systems.

Although ILI activity differs across geographical areas and data sources, the general region-specific seasonal trends are similar and provide valuable information about changes in influenza activity. Together, these platforms offer a more comprehensive view of influenza surveillance that helps public health offices monitor and respond to seasonal influenza epidemics.

4.2 HEALTH CARE-SEEKING BEHAVIOR AMONG FLU NEAR YOU PARTICIPANTS

4.2.1 Overview

As discussed in 3.2.1, the CDC estimates seasonal influenza burden in the community through a probabilistic multiplier model that uses laboratory-confirmed influenza-associated hospital rates from FluSurv-NET and results from health care-seeking behavior surveys.(Shrestha et al., 2017; Rolfes et al., 2018; Reed et al., 2009) The health care-seeking behavior surveys used in these models were conducted through a module in BFRSS in select states during the 2009 pandemic and 2010-2011 influenza season.(Reed et al., 2011; Biggerstaff et al., 2012, 2014) The goal of these surveys is to collect information on the incidence and risk factors of ILI and health care-seeking behaviors. Although telephone-based surveys are a well-established method of determining this information and have the ability to reach a large number of individuals, information is often collected sporadically because of the expense and selection bias due to exclusion of households without landlines is possible.

In 2015 FNY implemented follow-up survey questions that asked if and where participants received medical attention for reported symptoms. Here, we summarize the health care-seeking behavior of FNY participants who reported symptoms consistent with either ILI or acute respiratory infection (ARI) during three influenza seasons (2015-2016, 2016-2017, and 2017-2018) to see if information collected through this system is consistent with past trends and can be used to update information on care seeking behaviors, which is 10 years old. Specifically, we assess and compare monthly and seasonal trends in the percent of FNY participants who sought health care across different demographic and census region subgroups.

4.2.2 Methods

4.2.2.1 Data

As described in Section 1.2.2.1, FNY users report any symptoms that they or any registered household members experienced on a weekly basis. If FNY users report any symptom, they are asked to provide the date of symptom(s) onset, whether or not they received medical care for the symptom(s), and, if so, whether they received this care from a doctor's office, urgent care, clinic, ER, hospital, or other facility. Although FNY collects data throughout the entire year, for the purposes of this study, we examine FNY participant data reported only during the influenza season, defined as MMWR weeks 20 through 40.

For this study, ARI is defined as a symptom report that included a combination of at least two of the following symptoms: fever, cough, sorethroat, running nose, or breathlessness, and ILI is defined as a symptom report of fever plus cough and/or sorethroat. Only registered FNY participants who submitted at least three symptom reports during the study time period, October 2015 through September 2018, and provided valid date and month of birth and zip code information are included in the analyses. Appendix Figure A.10 summarizes a senstivity analysis that assesses changes in the percent of FNY participants who sough health care for symptoms over time for different inclusion criteria.

4.2.2.2 Statistical methods

We compare the monthly trends in the percent of FNY participants who sought health care for both ARI and ILI symptoms across three influenza seasons by aggregating the weekly symptom report data by month (based on the calendar month of the start of the reporting week). Trends in the percent of FNY participants who sought health care are evaluated by age group (<18, 18-49, 50-64, 65+), sex (male, female), census region (Northeast, South Midwest, and West), and place of care (doctor's office, urgent care, or other facility, which includes clinic, ER, hospital, or other).

We also calculate seasonal estimates for the percent of FNY participants who sought health care for ARI and ILI symptoms in the overall FNY population, as well as by age group, sex, census region, and place of care. Because the age distribution of the FNY population differs from the U.S. population,(Baltrusaitis et al., 2017) we calculate age-adjusted estimates for the overall, sex, and census region populations using the U.S. Census Bureaus 2013-2017 American Community Survey 5-Year Estimates.(U.S. Census Bureau, 2015) We compare statistical significance across across groups using chi-square tests.

4.2.3 Results

4.2.3.1 Overtime

Acute respiratory illness Time series of the monthly percent of FNY participants who sought health care for ARI symptoms is shown for the overall population as well as by census region, age group, and sex in Figure 4.5. Between September 2015 and October 2018, the monthly percent of FNY participants who sought health care for ARI symptoms ranged from 8.9 to 25.0%. Although no seasonal trends in health care-seeking behavior for ARI are evident, a smaller percentage of participants from the West census region sought health care for ARI symptoms compared to participants from the Northeast, South, and Midwest. There are also differences across age groups. The youngest, <18, and oldest, 65+, age groups have the highest percentage of participants who sought health care for ARI symptoms, followed by the 50-64 age group, and finally the 18-49 age group. There are no differences in health care-seeking behavior for ARI symptoms over time between males and females.

As shown in Figure 4.6, where participants sought health care for ARI symptoms was consistent over time. Across all seasons, most participants sought health care for ARI symptoms at doctor's offices, followed by urgent care facilities, then other facilities, which includes ERs, clinics, and hospitals.

Influenza-like illness Figure 4.7 shows the monthly time series of the percentage of FNY participants who sought health care for ILI symptoms for the overall population as well as by census region, age group, and sex. The timing of the peak percentage of FNY participants who sought health care varies each season. During the 2015-2016 influenza season, the percentage of FNY participants who sought health care for ILI symptoms peaked between April and July. Whereas, the percentage of FNY participants who sought health care for ILI symptoms peaked between February and May during the 2016-2017 influenza season and between January and June during the 2017-2018 influenza season. In general, the percentage of FNY participants from the South census region who sought health care for ILI symptoms is greater than other regions, and percentage of FNY participants from the West census region who sought health care for ILI symptoms is less than other regions. The age group-specific trends in health care-seeking behavior varies by season. During the 2016-2017 and 2017-2018 influenza seasons, a greater percentage of FNY participants in the <18 and 65+ age groups sought health care for ILI symptoms compared to the 18-49 and 50-64 age groups. However, there are no

Figure 4.5: Time series of the monthly percent of FNY participants who sought health care for ARI symptoms by (A) Overall (B) Census Region (C) Age Group (D) Sex.

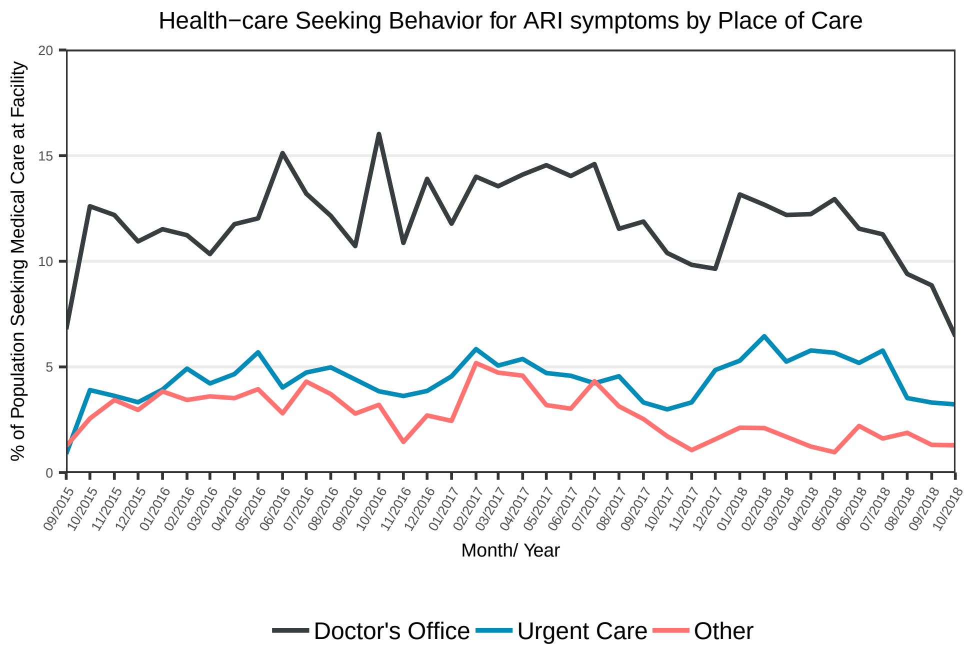

Figure 4.6: Time series of the monthly percent of FNY participants who sought health care for ARI symptoms by place of care.

clear differences in the percentage of FNY participants who sought health care for ILI symptoms between age groups during the 2017-2018 influenza season. Across all seasons, no distict differences in health care-seeking behavior for ILI symptoms are evident for males and females.

Similar to ARI, most FNY participants sought health care for ILI symptoms at doctor's offices, followed by urgent care, then other facilities (Figure 4.8). The patterns of place of care do not change over time, until the 2017-2018 influenza season. During this season, the percentage of FNY participants who sought care for ILI symptoms at urgent care facilities increased, while the percentage of FNY participants who sought care for ILI symptoms at other facilities decreased. Reporting of "other" decreased during this season because "other" was removed as an option for web-based submissions at the start of the season.

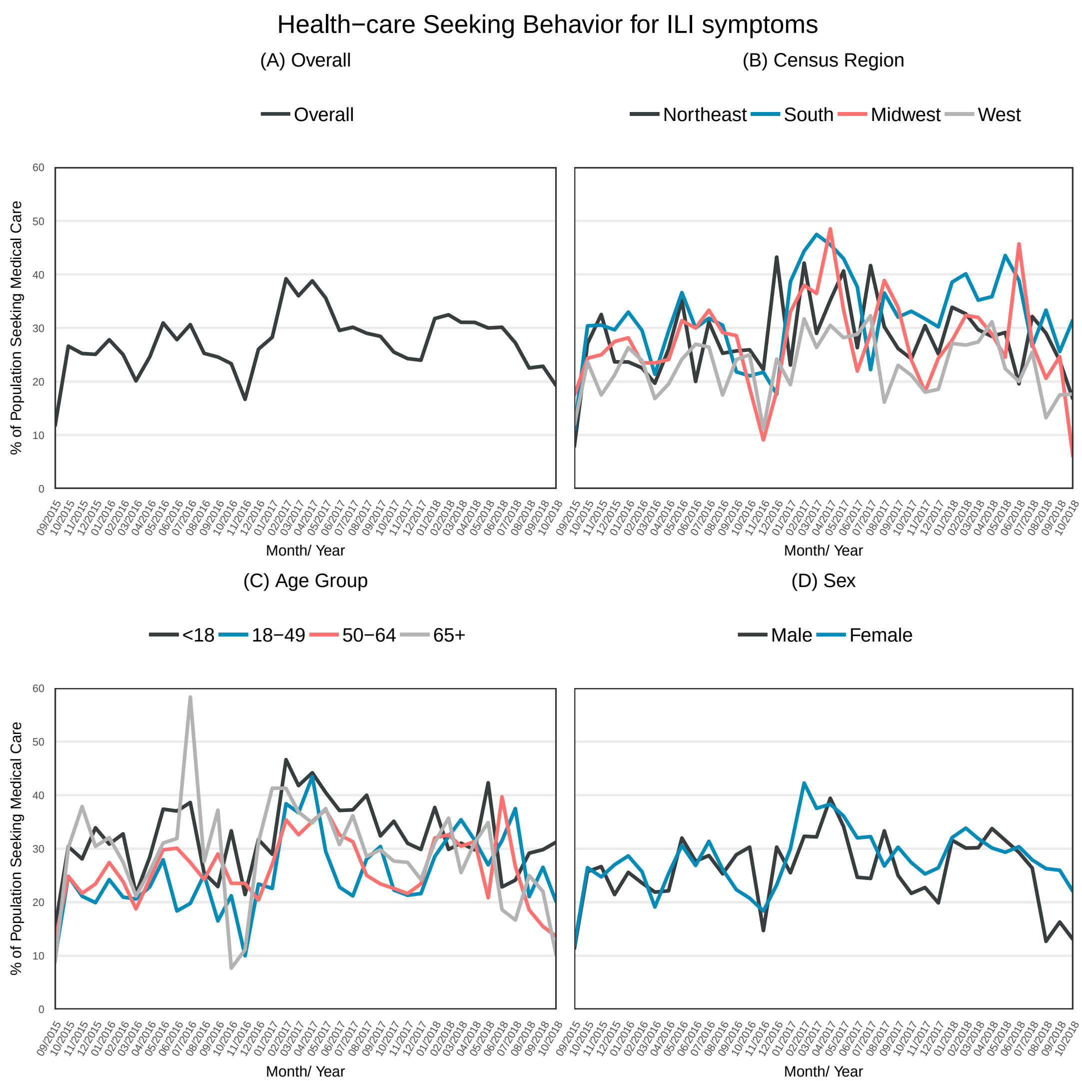

Figure 4.7: Time series of the monthly percent of FNY participants who sought health care for ILI symptoms by (A) Overall (B) Census Region (C) Age Group (D) Sex.

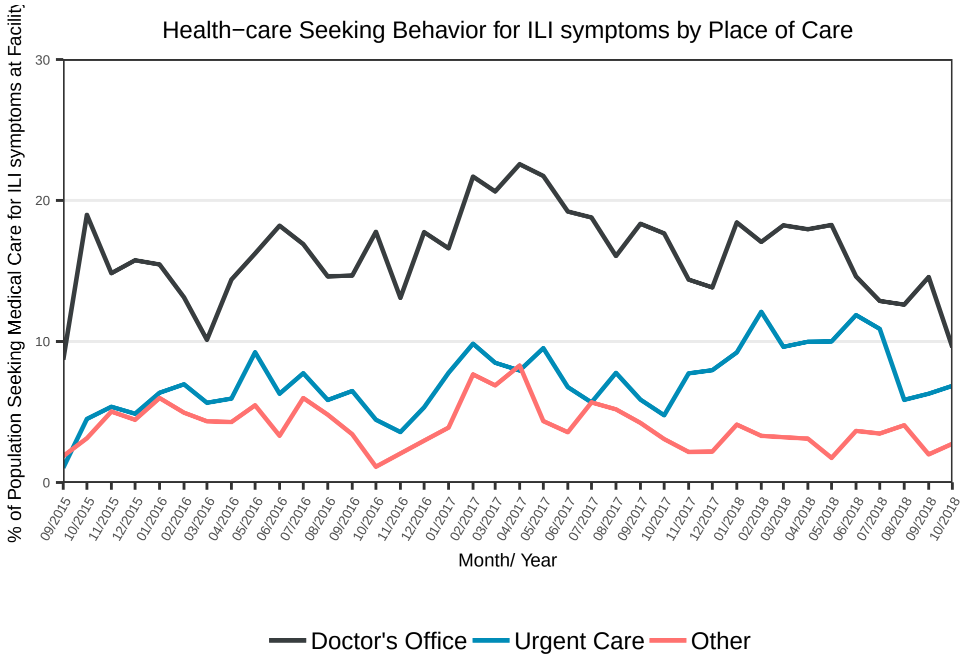

Figure 4.8: Time series of the monthly percent of FNY participants who sought health care for ILI symptoms by place of care.

4.2.3.2 By season

Table 4.4 displays the number of participants and total number of reports submitted each influenza season along with the number of symptomatic reports and age-adjusted health care-seeking percentages for ARI and ILI. For ARI, the overall age-adjusted percent of FNY participants who sought health care are 19.06%, 22.88%, and 17.72% for the 2015-2016, 2016-2017, and 2017-2018 influenza seasons, respectively. Across all seasons, the health care-seeking behaviors are significantly different across age groups. The <18 and 65+ age groups have greater ARI health care-seeking percentages compared to the 18-49 and 50-64 age groups. Although females have greater age-adjusted ARI health care-seeking percentages compared to males, this difference was statistically different only during the 2016-2017 and 2017-2018 influenza seasons (p<0.001). Health care-seeking behaviors across census regions are statistically significant during each season (p<0.001). Participants from the South census region have greater age-adjusted ARI health care-seeking percentages compared to the other regions each season.

The overall age-adjusted percent of FNY participants who sought health care for ILI symptoms for the 2015-2016, 2016-2017, and 2017-2018 influenza seasons are 25.35%, 35.59%, and 28.81%, respectively. Similar to ARI, the health care-seeking behaviors are significantly different across age groups. The <18 and 65+ age groups have greater age-adjusted ILI health care-seeking percentages compared to the 18-49 and 65+ age groups. Females also have greater age-adjusted ILI heath care-seeking percentages compared to males, however, this difference is statistically significant for only the 2016-2017 influenza season (p=0.001). Health care-seeking behaviors across census regions are statistically significant during each season (p<0.001). Across all seasons, FNY participants from the South census region have the largest age-adjusted ILI heath care-seeking percentages, while FNY participants from the West census region have the smallest age-adjusted ILI heath care-seeking percentages.

Table 4.4: Number of FNY participants, total symptom reports, symptomatic reports, and estimates of the percent of FNY participants who sought health care for ILI and ARI symptoms by selected demographics and census regions for the 2015-2016, 2016-2017, and 2017-2018 influenza seasons.*Indicates age-adjusted percent.

variable	value	No. part.	No. of reports	No. (%) ILI	% ILI care	No. (%) ARI	% ARI care
2015-2016							
overall*		27368	348903	6491 (1.86%)	25.35%	16262 (4.66%)	19.06%
age group	<18	3167	36827	1174 (3.19%)	30.07%	2056 (5.58%)	22.42%
	18 to 49	8850	96528	2198 (2.28%)	22.20%	5579 (5.78%)	15.88%
	50 to 64	9121	121596	2174 (1.79%)	23.74%	5541 (4.56%)	18.88%
	65 +	6230	93952	945 (1.01%)	29.21%	3086 (3.28%)	23.23%
	p-value				<0.001		<0.001
sex*	male	10557	147829	2150 (1.45%)	23.16%	5434 (3.68%)	17.48%
	female	16171	194223	4240 (2.18%)	26.21%	10610 (5.46%)	19.61%
	p-value				0.267		0.096
census region*	northeast	5422	71565	1194 (1.67%)	24.96%	3125 (4.37%)	18.61%
	midwest	7550	92756	1660 (1.79%)	30.06%	3979 (4.29%)	22.42%
	south	5565	70363	1315 (1.87%)	25.02%	3294 (4.68%)	20.05%
	west	8708	112440	2291 (2.04%)	22.04%	5774 (5.14%)	16.35%
	p-value				<0.001		<0.001

variable	value	No. part.	No. of reports	No. (%) ILI	% ILI care	No. (%) ARI	% ARI care
2016-2017							
overall*		23432	321058	3897 (1.21%)	35.59%	11906 (3.71%)	22.88%
age group	<18	2647	31703	680 (2.14%)	40%	1493 (4.71%)	26.05%
	18 to 49	6809	78016	1168 (1.5%)	33.73%	3361 (4.31%)	20.38%
	50 to 64	7858	112571	1359 (1.21%)	32.89%	4327 (3.84%)	22.65%
	65 +	6118	98768	690 (0.7%)	37.68%	2725 (2.76%)	25.43%
	p-value				0.004		<0.001
sex*	male	9256	136758	1278 (0.93%)	29.74%	4019 (2.94%)	18.72%
	female	13656	179109	2552 (1.42%)	37.59%	7736 (4.32%)	24.70%
	p-value				0.001		<0.001
census region*	northeast	4632	66362	722 (1.09%)	34.58%	2294 (3.46%)	22.15%
	midwest	6546	86840	1147 (1.32%)	42.45%	3150 (3.63%)	27.03%
	south	4889	65845	890 (1.35%)	36.52%	2822 (4.29%)	25.36%
	west	7281	100582	1121 (1.11%)	28.81%	3575 (3.55%)	17.82%
	p-value				<0.001		<0.001

variable	value	No. part.	No. of reports	No. (%) ILI	% ILI care	No. (%) ARI	% ARI care
2017-2018							
overall*		24204	344515	5531 (1.61%)	28.81%	18623 (5.41%)	17.72%
age group	<18	2725	32750	976 (2.98%)	33.30%	2406 (7.35%)	20.41%
	18 to 49	7078	82355	1665 (2.02%)	26.79%	5467 (6.64%)	15.53%
	50 to 64	7806	115537	1885 (1.63%)	27.80%	6260 (5.42%)	18.21%
	65 +	6595	113873	1005 (0.88%)	29.05%	4490 (3.94%)	19.22%
	p-value				0.003		<0.001
sex*	male	9468	144070	1880 (1.3%)	26.33%	6265 (4.35%)	18.28%
	female	14166	194441	3547 (1.82%)	29.86%	12096 (6.22%)	19.69%
	p-value				0.106		<0.001
census region*	northeast	4868	72383	987 (1.36%)	28.96%	3770 (5.21%)	17.11%
	midwest	6805	94580	1496 (1.58%)	35.85%	4739 (5.01%)	22.28%
	south	5017	71444	1239 (1.73%)	26.98%	4210 (5.89%)	16.92%
	west	7456	105132	1790 (1.7%)	23.88%	5859 (5.57%)	14.98%
	p-value				<0.001		<0.001

The number and percent of FNY participants who sought health care for ARI and ILI symptoms by place of care are shown in Table 4.5. The patterns in the place of care are similar for ARI and ILI. During each season, approximately 60% of FNY participants sought care at doctors offices, approximately 25% of FNY participants sought care at urgent care facilities, approximately 3-4% of FNY participants sought care at clinics, approximately 5-6% of FNY participants seek care at

EDs, and approximately 1-2% of FNY participants seek care at hospitals. During the 2015-2016 and 2016-2017 influenza seasons, approximately 8% of FNY participants sought care at an "other" facility, however, less than 1% of participants sought care at an "other" facility during the 2017-2019 influenza season. Again, reporting of "other" decreased during this season because "other" was removed as an option for web-based submissions at the start of the season.

Table 4.5: Number and percent of FNY participants who sought health care for ARI and ILI symptoms by place of care

season	Place of care	No. (%) ILI visit	No. (%) ARI visit
2015- 2016	total	1633	3110
	dr office	933 (57.13%)	1882 (60.51%)
	urgent care	391 (23.94%)	677 (21.77%)
	clinic	57 (3.49%)	107 (3.44%)
	ed	93 (5.7%)	152 (4.89%)
	hospital	26 (1.59%)	42 (1.35%)
	other	133 (8.14%)	250 (8.04%)
2016- 2017	total	1373	2747
	dr office	799 (58.19%)	1627 (59.23%)
	urgent care	326 (23.74%)	613 (22.32%)
	clinic	51 (3.71%)	107 (3.9%)
	ed	71 (5.17%)	107 (3.9%)
	hospital	13 (0.95%)	20 (0.73%)
	other	113 (8.23%)	273 (9.94%)
2017- 2018	total	1587	3343
	dr office	928 (58.48%)	2134 (63.83%)
	urgent care	488 (30.75%)	896 (26.8%)
	clinic	48 (3.02%)	107 (3.2%)
	ed	101 (6.36%)	161 (4.82%)
	hospital	10 (0.63%)	18 (0.54%)
	other	12 (0.76%)	27 (0.81%)

4.2.4 Discussion

Because traditional ILI surveillance relies on reports of medically attended ILI from health care providers, understanding the health care-seeking behavior of individuals can provide a more comprehensive picture of ILI disease burden in the community. A recent meta-analysis reported an overall pooled health care-seeking rate of 0.52 (95% CI: 0.46-0.59).(Ma et al., 2018) However, this meta-analysis used health care-seeking estimates from different countries and seasons and fails to capture the potential seasonal, regional, and age group dynamics in health care-seeking behavior. In the U.S., the 2010-2011 BFRSS survey estimated that 45% of adults and 57% of children sought health care for ILI during this season. Although our overall estimates range from 25% to 36%, the age-group and sex specific patterns are similar.

FNY has several benefits compared to the BFRSS surveys. Because FNY collects data from participants throughout the influenza season, we can assess how health care-seeking behaviors change over time. We can also compare health care-seeking behavior across seasons and for different combinations of symptoms. In addition, all states and age groups are represented in FNY.

There are several limitations to this study. As discussed in Section 2, FNY is not representative of the U.S. population. While we adjusted for age differences for the seasonal estimates, we did not age-adjust the time series analysis. We also did not adjust for age group while comparing health care-seeking percentages across different demographics and census regions. In the future, we will use a log-linear model to assess if there are any significant interactions between the demographics and census regions and age groups. In addition, FNY relies on self-reported data that is subject to recall and social desirability bias. FNY participant reporting is also not consistent throughout the influenza season (Refer to Section 3.2). For this study, we used a cohort method to address these inconsistencies, however, we did not apply any additional missing data methods, as these methods were developed to estimate ILI burden. Finally, there were errors in FNY data collection from October 2016 through December 2016 resulting in an underestimation in the weekly number of symptom reports. This error may contribute the observed troughs in the time series during this period.

The percent of FNY participants who sought health care for ARI and ILI symptoms varies by season, geographical region, age-group, and sex. FNY can compliment existing sentinal surveillance systems by adding important insight into community-level disease trends and health care-seeking behavior.

4.3 SCIENCE FRIDAY

4.3.1 Overview

This past fall Science Friday (SciFri) partnered with FNY to understand and enhance community disease surveillance by enrolling SciFri listeners as FNY participants (Figure 4.9). The goals of this project were to encourage weekly participation in the FNY campaign, provide factual information about influenza virus, symptoms, spread, and vaccination, address misconceptions about the flu and vaccines, and promote better understanding of disease surveillance and data literacy. The project launched on November 16, 2018 with an on-air interview with Dr. John Brownstein and wrapped on March 29, 2019 with an on-air interview with myself. As part of this project, SciFri listeners signed up for FNY using a special landing page, which allowed us to differentiate SciFri participants from all other FNY participants. SciFri listeners also had the opportunity to sign up for weekly text reminders that included facts and tips related to influenza. All participants of FNY were able to track the weekly percent ILI for the SciFri cohort, FNY cohort, and CDC ILINet on through a time series visualization that was embedded on the FNY website.

4.3.2 Summary

During the study period, 3150 SciFri listeners registered for FNY, and 2905 users submitted at least one symptom report. An additional 543 household members were registered by users, bringing the total SciFri cohort to 3448 participants. Compared to registered FNY users who submitted at least one symptom report during the 2018-2019 influenza season, SciFri users are younger, 55.3 (IQR: 35.3, 66.4) vs. 58.1 (IQR: 45.3, 66.7) and have a higher percentage of males (38.5% vs. 29.6%).

Figure 4.9: Screenshot of the SciFri home page.

Overall, 28 056 symptom reports were submited, and the median number of symptom reports per week was 1659 (IQR: 1627, 1705). As shown in Figure 4.10, symptom reports submitted by SciFri participants make up about 10-15% of the total weekly FNY symptom reports.

Overall, SciFri participants have a smaller weekly percentage of ILI symptoms compared to both all reports submitted to FNY and registered FNY users (Figure 4.11). The overall ILI burden for SciFri participants (8.1%) is also less than the overall ILI burden for all registered FNY users (12.4%). These estimates are not age-adjusted and no missing data methods were applied.

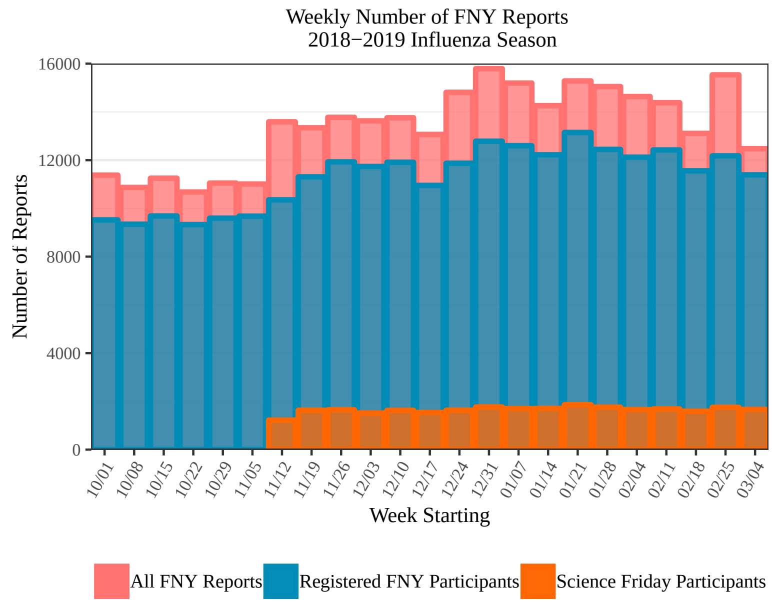

Figure 4.10: Weekly number of reports for all FNY reports, registered FNY participants, and SciFri participants during the 2018-2019 influenza season.

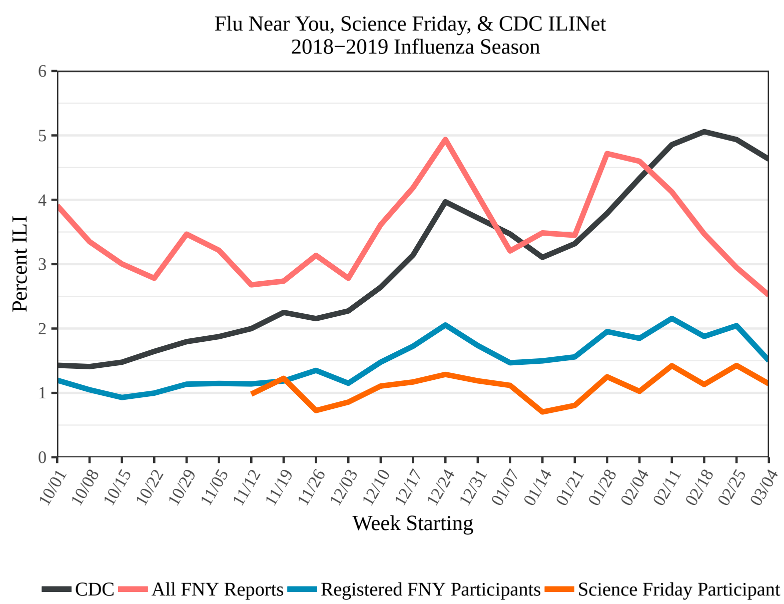

Figure 4.11: Time series of percent ILI for CDC ILINet, all FNY reports, registered FNY participants, SciFri participants, and CDC ILINet for the 2018-2019 influenza season.

CHAPTER 5

Conclusion

Participatory syndromic surveillance systems complement traditional health-care based influenza surveillance systems by monitoring disease activity in populations who do not seek health care, areas with limited surveillance data, and community based populations. They also provide an opportunity for the public to engage directly in community-level disease surveillance. However, the potential public health benefits of participatory surveillance data are maximized only by the development of statistically rigorous methods that address potential biases.

Based on our research, we have outlined our conclusions and included recommendations for establishing, maintaining, and analyzing data from partcipatory syndromic surveillance systems:

1. Because females and older populations are over-represented in these populations, we recommend providing sex or age group specific estimates or weighting overall estimates by sex and age group.

2. We find that approximately 300-500 weekly symptom reports are necessary to accurately track ILI in a given geographical area.

3. We recommend following Flutracking.net's lead by providing users with the opportunity to complete missing surveys. This system accommodation not only adds approximately 10% more weekly reports, but also provides valuable insight into reporting behaviors.

4. For Flutracking.net, the δ MI method accurately predicted end of season WP estimates from real-time data. In the future, the value of δ can be easily updated and adapted over the course of an influenza season.

5. When estimating ILI burden using a participatory syndromic surveillance system with inconsistent reporting habits, such as FNY, we recommend using a cohort method that assumes all missing reports are non-ILI. Unless imputation methods are applied, an entry criteria of at least 5 is appropriate.

6. Differences in ILI activity across regions, as reported by a given surveillance system, are not consistent across surveillance platforms. However, the general region-specific seasonal trends are similar and provide valuable information about changes in influenza activity.

7. The health care-seeking behaviors of individuals vary by season, age, and geographical area. These dynamics should be taken into consideration when estimating the overall disease burden in the general population from sentinel surveillance data.

8. Finally, based on our collaboration with SciFri, we found that providing users with an interactive environment that includes information about influenza and ILI in their community improves user activity.

With the increase in the emergence of infectious diseases over the last few decades, the application of this method of disease surveillance may prove useful in broader early disease detection for other emergent diseases such as Zika virus and dengue by providing actionable insights for public health stakeholders.

APPENDIX A

Appendix

A.1 APPENDIX TABLES

Table A.1: HHS-defined regions and corresponding states.

Region	States
Region 1	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Region 2	New Jersey, New York, Puerto Rico, US Virgin Islands
Region 3	Delaware, District of Columbia, Maryland, Pennsylvania, Virginia, West Virginia
Region 4	Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee
Region 5	Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin
Region 6	Arkansas, Louisiana, New Mexico, Oklahoma, Texas
Region 7	Iowa, Kansas, Missouri, Nebraska
Region 8	Colorado, Montana, North Dakota, South Dakota, Utah, Wyoming
Region 9	Arizona, California, Guam, Hawaii, Nevada
Region 10	Alaska, Idaho, Oregon, Washington

Table A.2: Pearson correlations between FNY and CDC ILINet/ BPHC and mean weekly reports at the national, regional, state, and city resolutions.

Geographical Resolution	All ρ	Seasons $\bar{n}$
National	0.81	9699
Region 1	0.71	958
Region 2	0.64	700
Region 3	0.75	1093
Region 4	0.81	1178
Region 5	0.79	1476
Region 6	0.73	729
Region 7	0.73	415
Region 8	0.73	510
Region 9	0.77	1798
Region 10	0.76	819
Alaska	0.13	27
Alabama	0.58	66
Arkansas	0.53	51
Arizona	0.63	204
California	0.78	1503
Colorado	0.57	237
Connecticut	0.44	163
Delaware	0.19	26
Florida	-	384
Georgia	0.64	182
Hawaii	0.18	34
Iowa	0.62	160
Idaho	0.52	50
Illinois	0.69	333
Indiana	0.63	147
Kansas	0.45	90

Geographical Resolution	All ρ	Seasons $\bar{n}$
Kentucky	0.48	84
Louisiana	0.51	70
Massachusetts	0.75	576
Maryland	0.65	279
Maine	0.31	64
Michigan	0.48	271
Minnesota	0.65	163
Missouri	0.57	131
Mississippi	0.34	26
Montana	0.38	34
North Carolina	0.68	237
North Dakota	0.38	60
Nebraska	0.36	35
New Hampshire	0.49	59
New Jersey	0.57	196
New Mexico	0.6	96
Nevada	0.31	57
New York	0.62	442
Ohio	0.68	351
Oklahoma	0.55	125
Oregon	0.65	352
Pennsylvania	0.68	368
Rhode Island	0.4	58
South Carolina	0.46	88
South Dakota	0.43	23
Tennessee	0.45	112
Texas	0.68	388
Utah	0.67	131
Virginia	0.61	327
Vermont	0.36	38
Washington	0.61	390
Wisconsin	0.59	211
West Virginia	0.45	47
Wyoming	0.35	26
Boston	0.69	304

Table A.3: Overall and age group specific IRs and 95% CIs, expressed as number of ILI reports per 10 000 person weeks, by influenza season for Flutracking.net.

Age Group	Ignore missing data	Assume missing non-ILI	Complete Case	MI	δ MI
2016					
Overall	299.4 (294.4, 304.5)	267.1 (262.6, 271.7)	271.9 (266.3, 277.6)	305.3 (300.6, 310.2)	299.6 (294.9, 304.4)
<5	780.4 (733.5, 830.3)	610.2 (573.5, 649.3)	688.4 (635.4, 745.8)	753.4 (712.5, 796.6)	736.1 (695.6, 778.8)
5-17	364.2 (347.8, 381.3)	315.2 (301, 330.1)	327.2 (308.5, 346.9)	370.6 (355.2, 386.7)	362.8 (347.6, 378.7)
18-49	308.4 (300.3, 316.8)	267.7 (260.7, 275)	283.5 (274.3, 293.1)	311 (303.4, 318.8)	304.6 (297.1, 312.4)
50+	248 (241.4, 254.7)	230 (223.9, 236.3)	228.7 (221.4, 236.2)	252.1 (245.7, 258.7)	248.3 (242, 254.8)
2017					
Overall	340.8 (335.7, 346)	303.1 (298.6, 307.7)	307.1 (301.5, 312.8)	348.7 (343.8, 353.6)	344.3 (339.5, 349.3)
<5	882.6 (834.8, 933.1)	687.4 (650.2, 726.7)	782.1 (730, 837.8)	883.6 (841.3, 928)	855.5 (813.9, 899.3)
5-17	464.5 (446.8, 482.9)	392.3 (377.4, 407.9)	429.1 (409.1, 450.1)	470.5 (454.1, 487.5)	464.2 (448, 481.1)
18-49	345.7 (337.2, 354.4)	298.6 (291.2, 306.1)	306.1 (296.8, 315.6)	347.9 (339.9, 355.9)	343.6 (335.7, 351.6)
50+	278.7 (272.2, 285.4)	258.8 (252.7, 265)	255.3 (248.1, 262.7)	283.2 (276.8, 289.6)	280.8 (274.5, 287.3)

Age Group	Ignore Missing data	Assume missing non-ILI	Complete Case	MI	δ MI
2018					
Overall	218.5 (214.9, 222)	196.2 (193, 199.4)	197.6 (193.8, 201.6)	224.1 (220.7, 227.5)	217.5 (214.1, 220.8)
<5	696.7 (661.3, 734)	552.3 (524.3, 581.9)	644.8 (604.5, 687.7)	701.9 (670.2, 735.2)	657.8 (627.2, 690)
5-17	240.6 (229.6, 252.2)	207.3 (197.8, 217.2)	220.6 (208.2, 233.6)	244.5 (234.2, 255.3)	236.1 (225.9, 246.7)
18-49	229.6 (223.5, 235.9)	198.4 (193.1, 203.8)	206.7 (200, 213.7)	231.8 (226.1, 237.6)	224.1 (218.5, 229.9)
50+	177.5 (173.1, 182.1)	166.7 (162.6, 171)	163 (158.2, 167.9)	179.8 (175.5, 184.2)	177 (172.7, 181.4)

Table A.4: Overall and age group specific Incidence Proportions and 95% CIs, expressed as percent of population reporting ILI at least once, by influenza season for Flutracking.net.

Age Group	Ignore missing data	Complete Case	MI	δ MI
2016				
Overall	27.33 (26.74, 27.93)	27.22 (26.49, 27.97)	30.2 (29.58, 30.83)	29.82 (29.21, 30.45)
<5	50.05 (45.78, 54.73)	54.21 (48.19, 60.98)	56.37 (51.82, 61.31)	55.5 (50.99, 60.41)
5-17	33.1 (31.22, 35.09)	33.45 (31.01, 36.08)	37.12 (35.13, 39.23)	36.6 (34.62, 38.69)
18-49	27.79 (26.86, 28.76)	28.4 (27.19, 29.66)	31.16 (30.17, 32.18)	30.69 (29.71, 31.71)
50+	23.85 (23.04, 24.69)	23.53 (22.57, 24.53)	25.77 (24.92, 26.64)	25.53 (24.69, 26.4)
2017				
Overall	30.63 (30.04, 31.24)	30.31 (29.58, 31.05)	33.6 (32.98, 34.24)	33.42 (32.8, 34.06)
<5	56.97 (52.6, 61.69)	60.16 (54.31, 66.65)	63.56 (58.94, 68.54)	63.08 (58.48, 68.04)
5-17	40.62 (38.65, 42.7)	42.48 (39.9, 45.22)	45 (42.92, 47.18)	44.77 (42.69, 46.94)
18-49	30.72 (29.75, 31.71)	30.63 (29.43, 31.89)	34.22 (33.2, 35.27)	34.02 (33, 35.06)
50+	26.33 (25.53, 27.14)	25.82 (24.88, 26.79)	28.29 (27.47, 29.14)	28.16 (27.34, 29.01)

Age Group	Ignore missing data	Complete Case	MI	δ MI
2018				
Overall	20.95 (20.53, 21.39)	20.81 (20.29, 21.33)	23.13 (22.68, 23.58)	22.71 (22.27, 23.17)
<5	48.34 (44.95, 51.98)	53.02 (48.33, 58.18)	54 (50.41, 57.84)	53.08 (49.53, 56.89)
5-17	23.22 (21.92, 24.6)	24.39 (22.71, 26.2)	26.4 (25.01, 27.87)	25.78 (24.4, 27.23)
18-49	21.67 (20.95, 22.41)	22.05 (21.15, 23)	24.46 (23.7, 25.25)	23.96 (23.2, 24.73)
50+	17.95 (17.39, 18.53)	17.55 (16.9, 18.22)	19.19 (18.61, 19.79)	18.93 (18.35, 19.52)

Table A.5: Age-group specific estimates of disease burden and 95% CIs expressed as per 1 000 000 persons by influenza seasons.

	<5	5-17	18-49	50-64	65+
2015-2016					
All reports	6.86 (6.17, 7.63)	11.09 (10.35, 11.87)	27.27 (26.34, 28.22)	11.48 (11.05, 11.93)	5.44 (5.11, 5.8)
Drop 1st Report	6.7 (5.89, 7.62)	10.72 (9.83, 11.69)	22.08 (20.99, 23.22)	9.43 (8.95, 9.94)	4.86 (4.49, 5.27)
Cohort	10.42 (8.87, 12.23)	15.34 (13.74, 17.14)	25.93 (24.05, 27.96)	10.18 (9.48, 10.94)	5.33 (4.83, 5.89)
Cohort + MI	12.43 (10.73, 14.4)	21.36 (19.45, 23.46)	37.88 (35.59, 40.31)	14.8 (13.95, 15.71)	7.93 (7.31, 8.6)
Cohort + δ MI	12.09 (10.42, 14.04)	20.14 (18.29, 22.18)	34.36 (32.18, 36.68)	13.4 (12.59, 14.26)	7.42 (6.82, 8.07)
2016-2017					
All reports	5.9 (5.06, 6.89)	9.35 (8.57, 10.2)	19.9 (18.86, 20.99)	8.7 (8.25, 9.17)	4.43 (4.1, 4.8)
Drop 1st Report	5.91 (4.98, 7)	8.97 (8.12, 9.9)	17.87 (16.77, 19.04)	8.12 (7.64, 8.62)	4.29 (3.94, 4.67)
Cohort	9.64 (7.92, 11.72)	13.66 (12.09, 15.44)	21.66 (19.77, 23.73)	8.95 (8.27, 9.69)	4.89 (4.42, 5.41)
Cohort + MI	12.81 (10.81, 15.18)	18.71 (16.86, 20.77)	30.66 (28.4, 33.1)	12.64 (11.83, 13.52)	7.03 (6.46, 7.65)
Cohort + δ MI	14.25 (12.13, 16.75)	17.1 (15.33, 19.08)	27.51 (25.38, 29.83)	30.82 (29.52, 32.17)	25.64 (24.53, 26.8)

	<5	5-17	18-49	50-64	65+
2017-2018					
All reports	6.9 (5.94, 8.01)	12.19 (11.35, 13.1)	25.78 (24.64, 26.98)	11.77 (11.25, 12.33)	6.2 (5.82, 6.62)
Drop 1st Report	6.59 (5.58, 7.77)	12.01 (11.09, 13)	24.37 (23.14, 25.66)	11.53 (10.97, 12.11)	6.11 (5.71, 6.54)
Cohort	9.88 (8.09, 12.05)	19.04 (17.19, 21.09)	29.43 (27.21, 31.83)	12.92 (12.09, 13.81)	7 (6.45, 7.58)
Cohort + MI	12.84 (10.78, 15.29)	25.31 (23.16, 27.66)	40.7 (38.07, 43.51)	17.2 (16.23, 18.22)	9.48 (8.85, 10.16)
Cohort + δ MI	12.55 (10.52, 14.98)	23.86 (21.78, 26.15)	37.08 (34.57, 39.76)	15.73 (14.8, 16.71)	8.84 (8.23, 9.5)

A.2 APPENDIX FIGURES

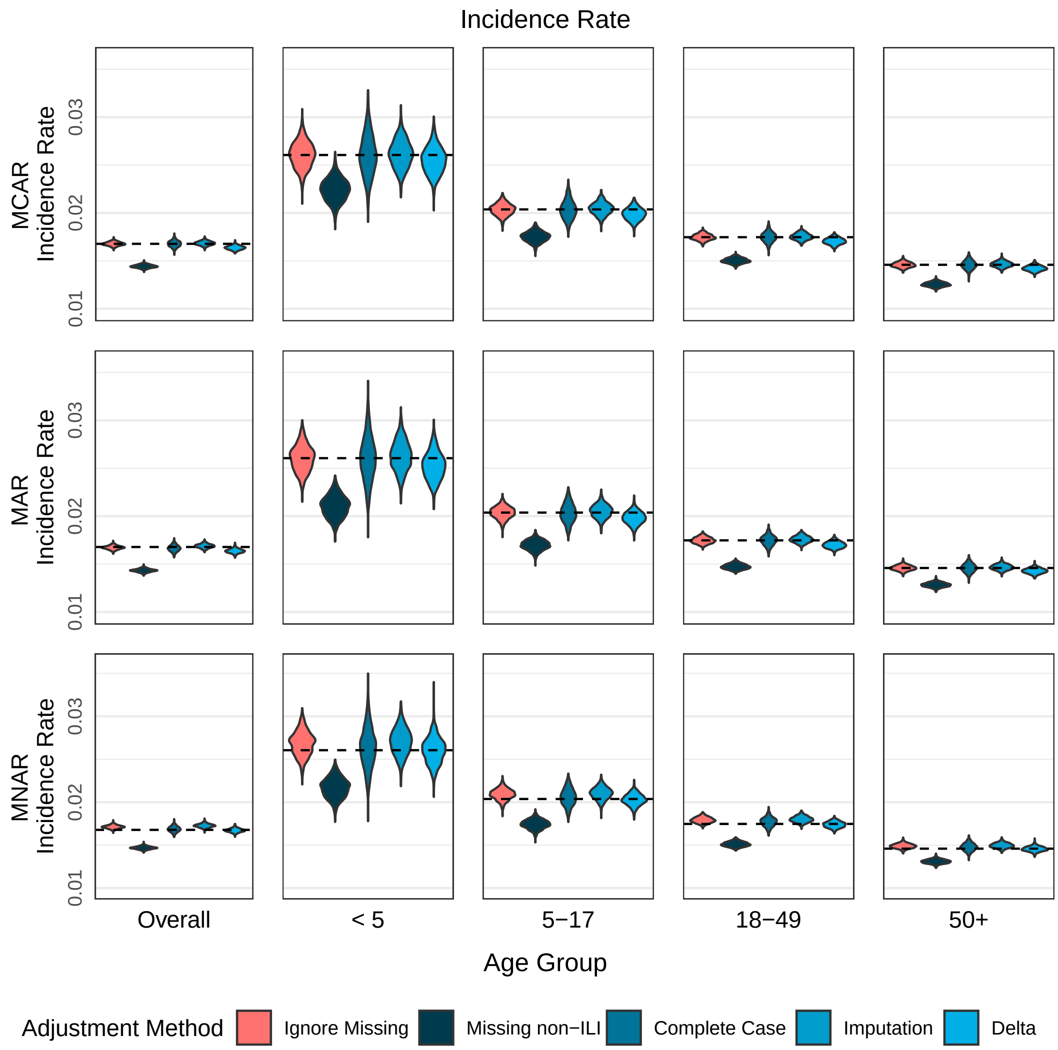

Figure A.1: Violin plots of simulated overall and age group specific IRs under three models of missingness: MCAR, MAR, and MNAR. Dotted line represents the original simulated parameter.

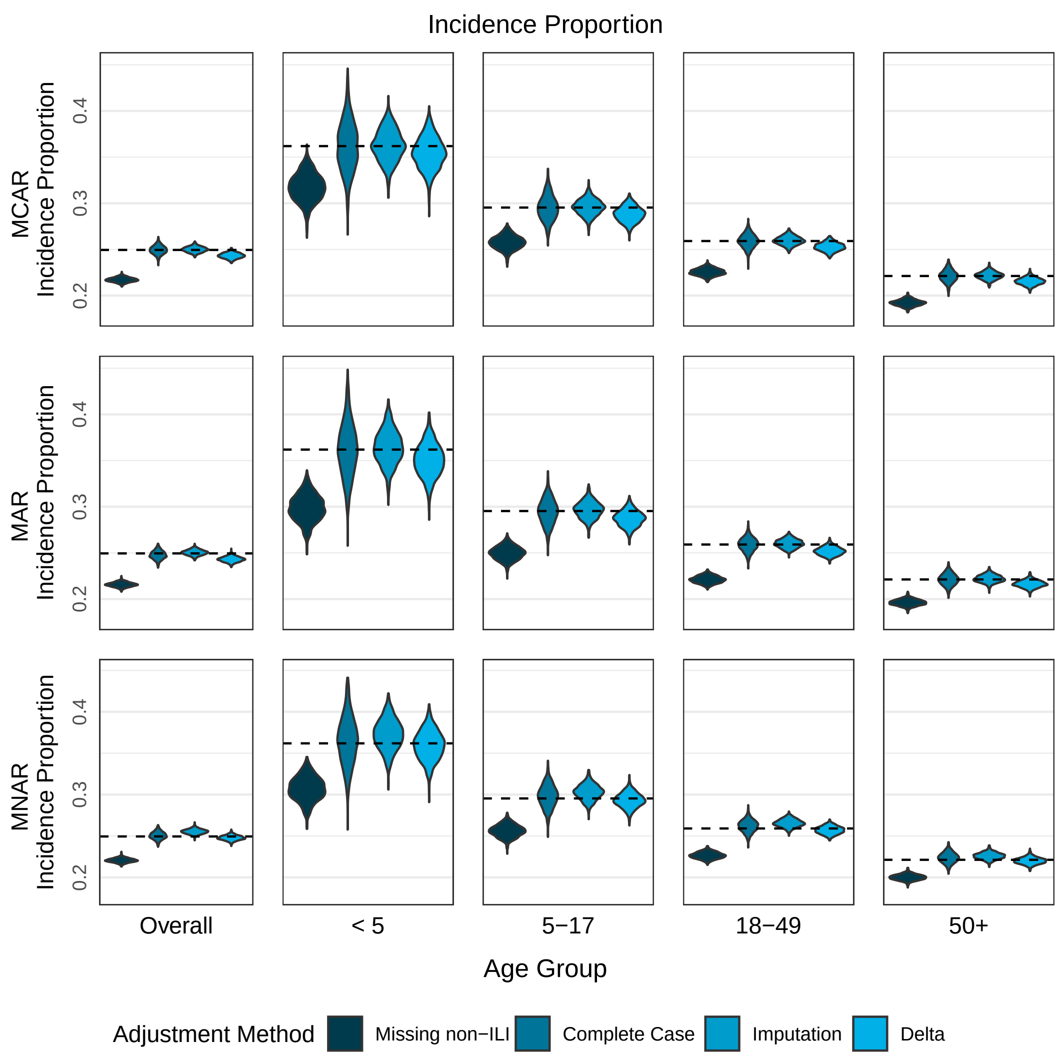

Figure A.2: Violin plots of simulated overall and age group specific IPs under three models of missingness: MCAR, MAR, and MNAR. Dotted line represents the original simulated parameter.

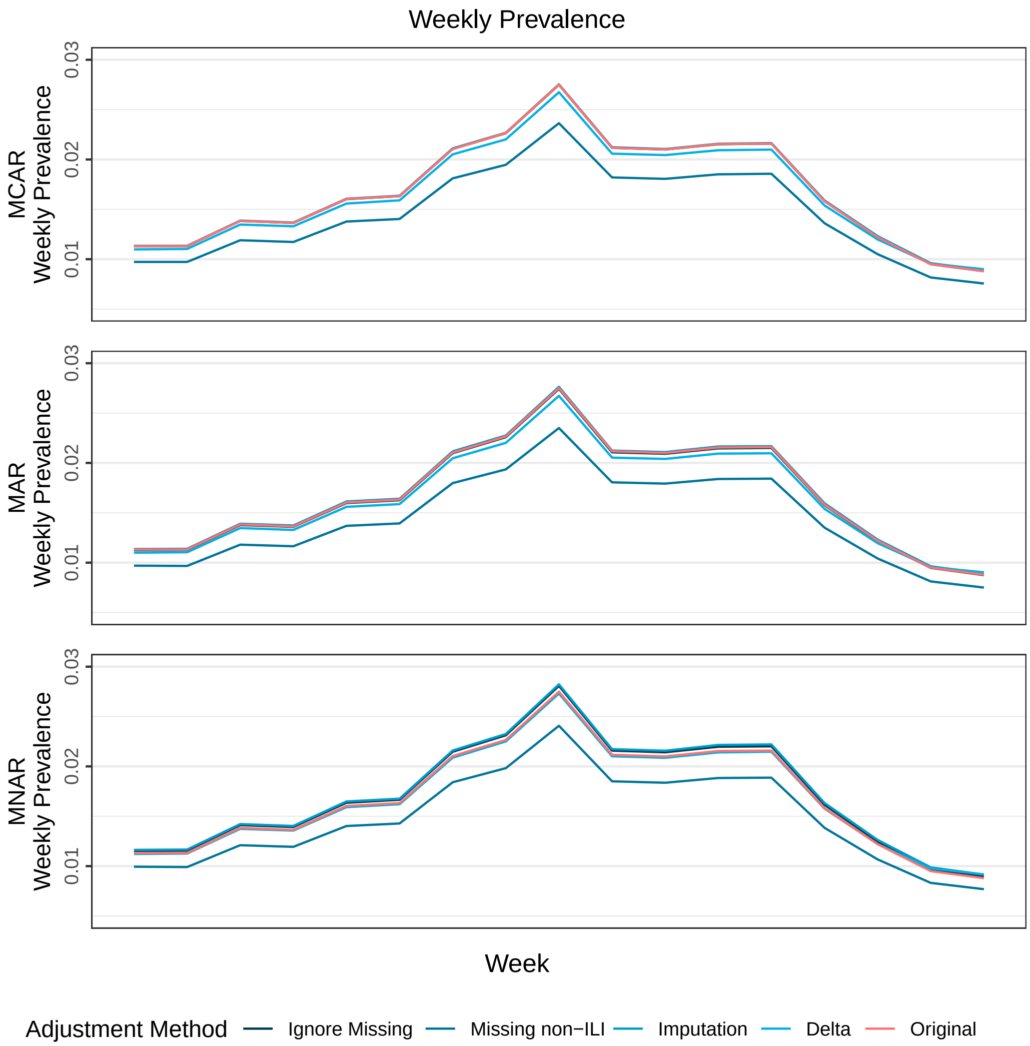

Figure A.3: Times series of mean simulated overall and age group specific IRs under three models of missingness: MCAR, MAR, and MNAR. Red line represents the original simulated parameter.

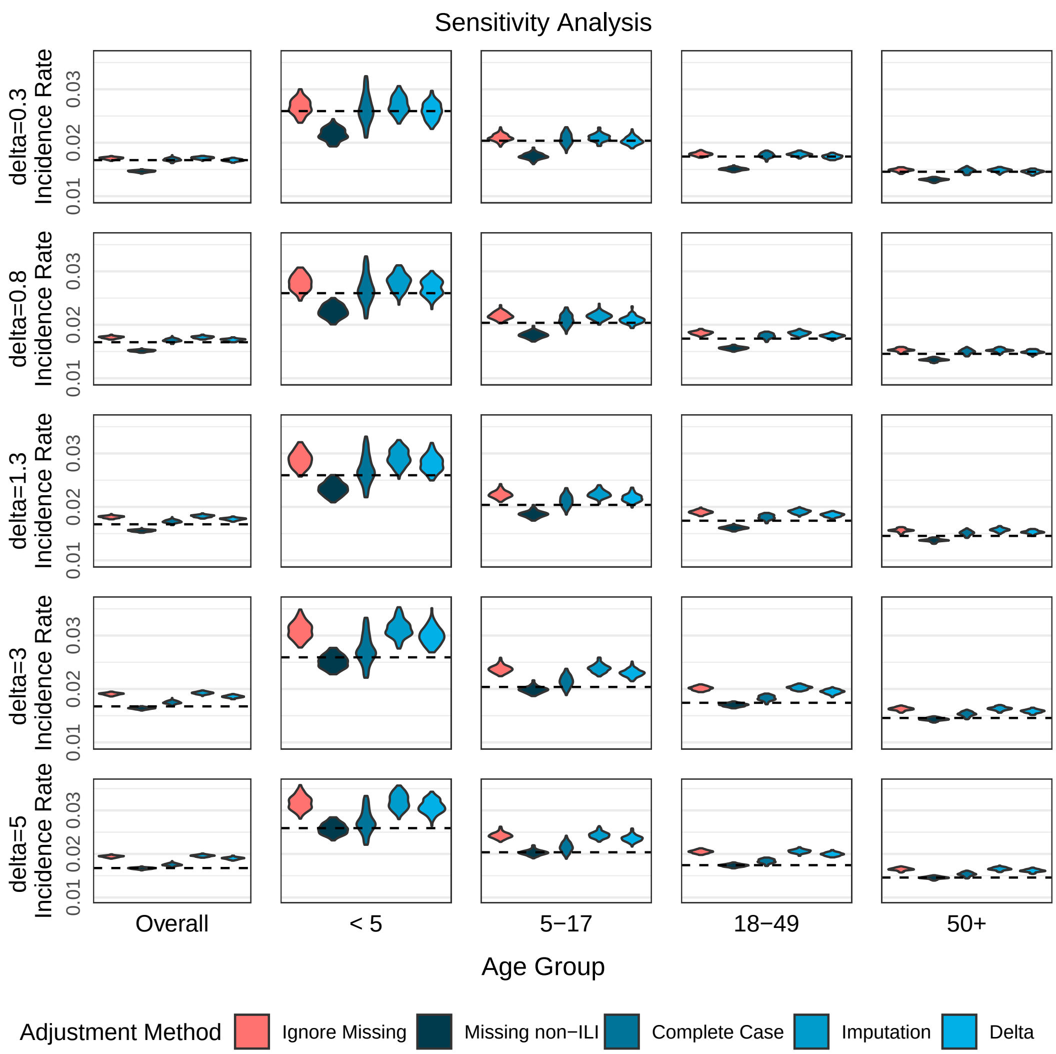

Figure A.4: Violin plots of simulated overall and age group specific IRs under five MNAR models. Dotted line represents the original simulated parameter.

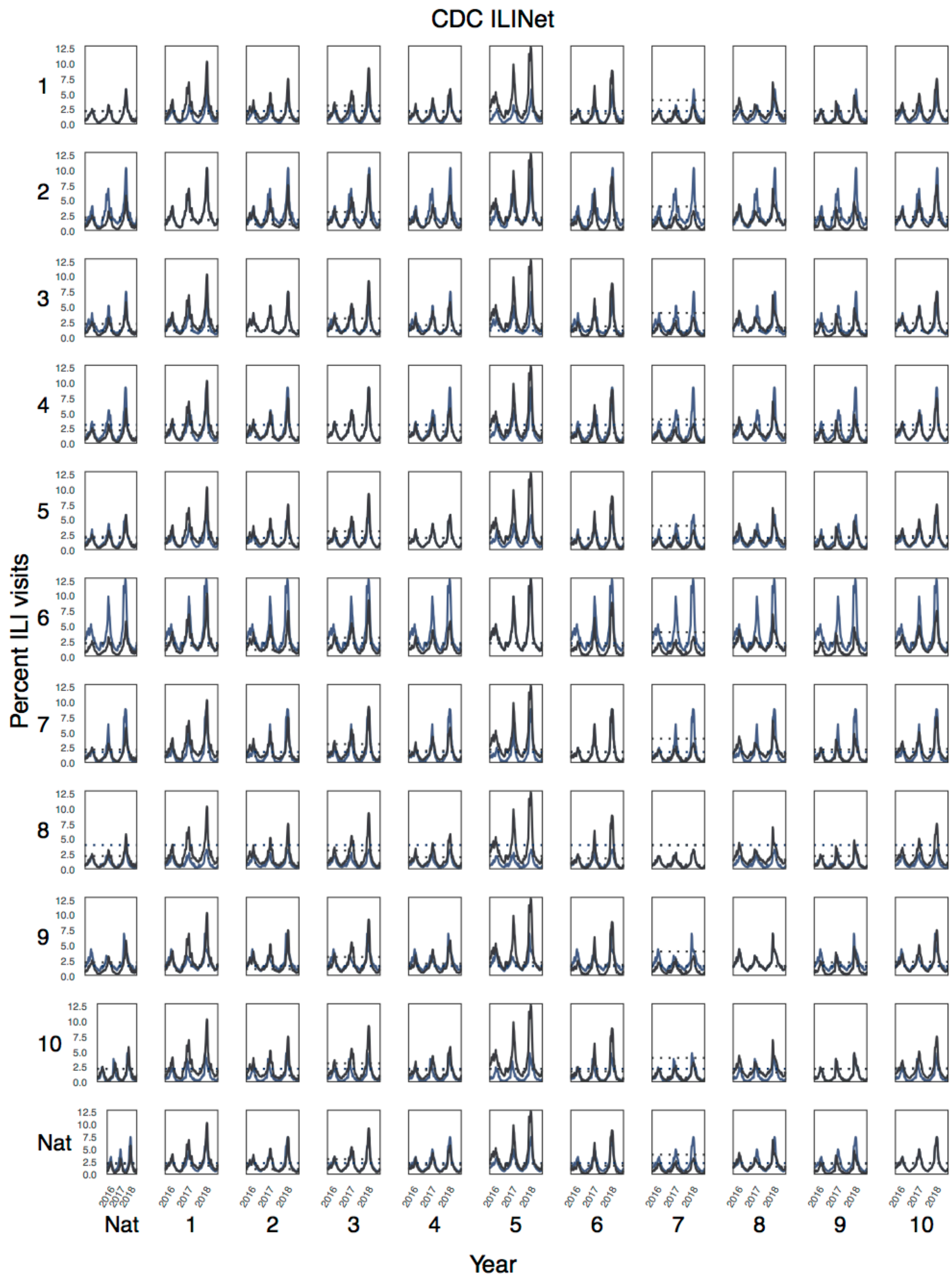

Figure A.5: Time series plots of weekly percent of ILI visits from CDC ILINet across three influenza seasons (2015-2016, 2016-2017, and 2017-2018) with baselines. Geographical areas on the columns are represented by black and geographical areas on the rows are represented by blue.

Figure A.6: Time series plots of weekly percent ILI from FNY across three influenza seasons (2015-2016, 2016-2017, and 2017-2018) with baselines. Geographical areas on the columns are represented by black and geographical areas on the rows are represented by red.

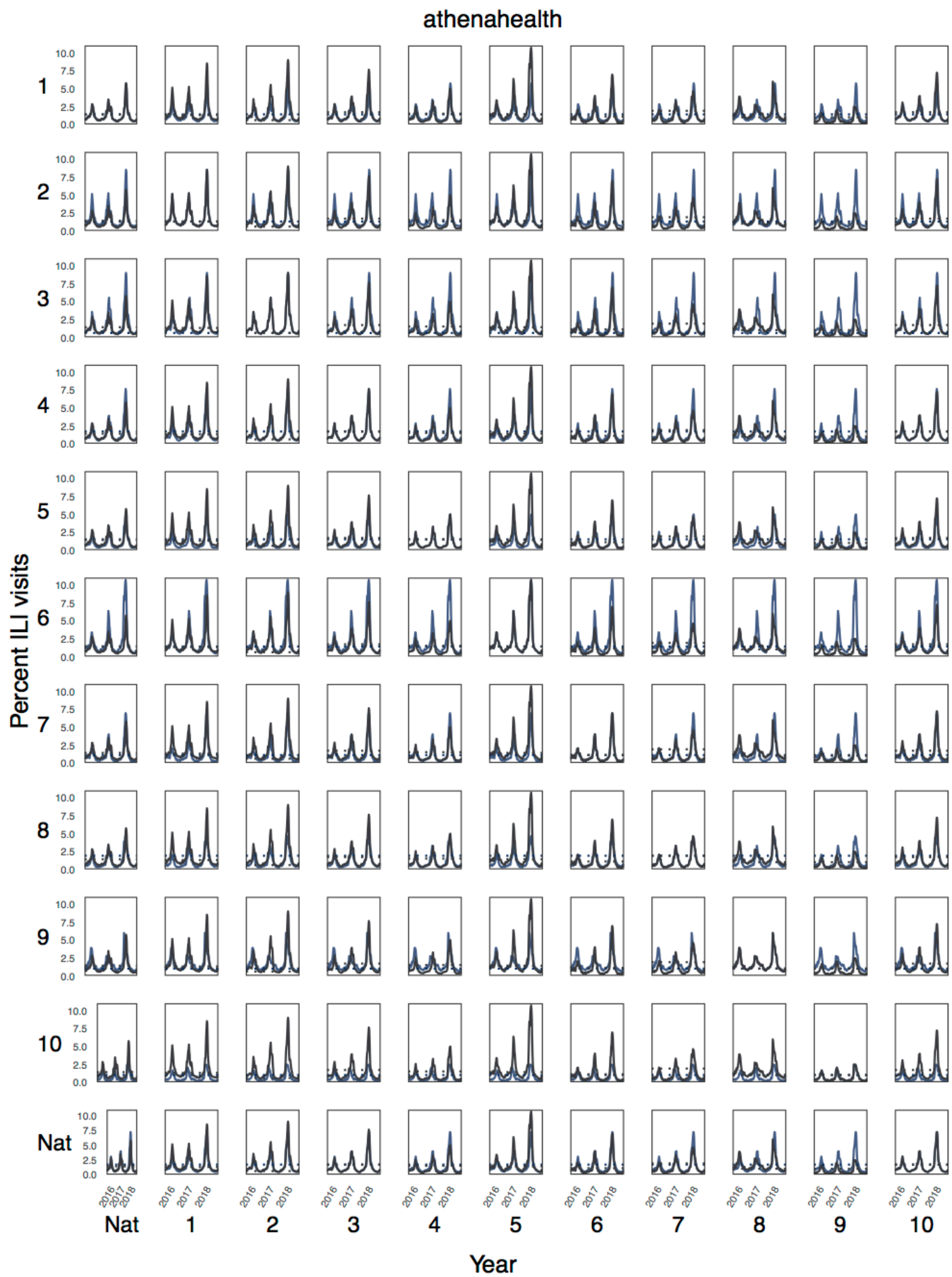

Figure A.7: Time series plots of weekly percent of ILI visits from athenahealth across three influenza seasons (2015-2016, 2016-2017, and 2017-2018) with baselines. Geographical areas on the columns are represented by black and geographical areas on the rows are represented by blue.

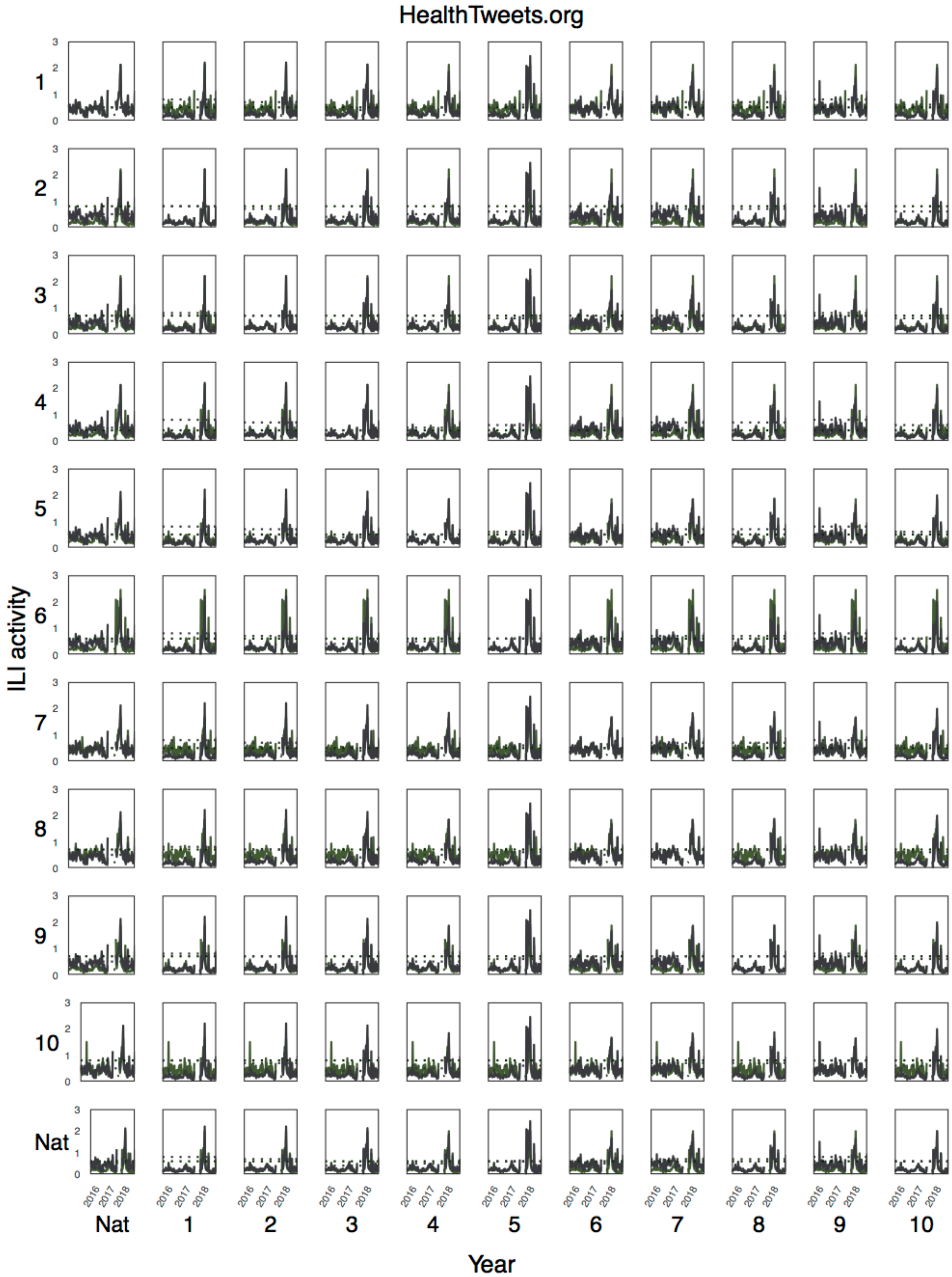

Figure A.8: Time series plots of weekly ILI activity from HealthTweets.org across three influenza seasons (2015-2016, 2016-2017, and 2017-2018) with baselines. Geographical areas on the columns are represented by black and geographical areas on the rows are represented by green.

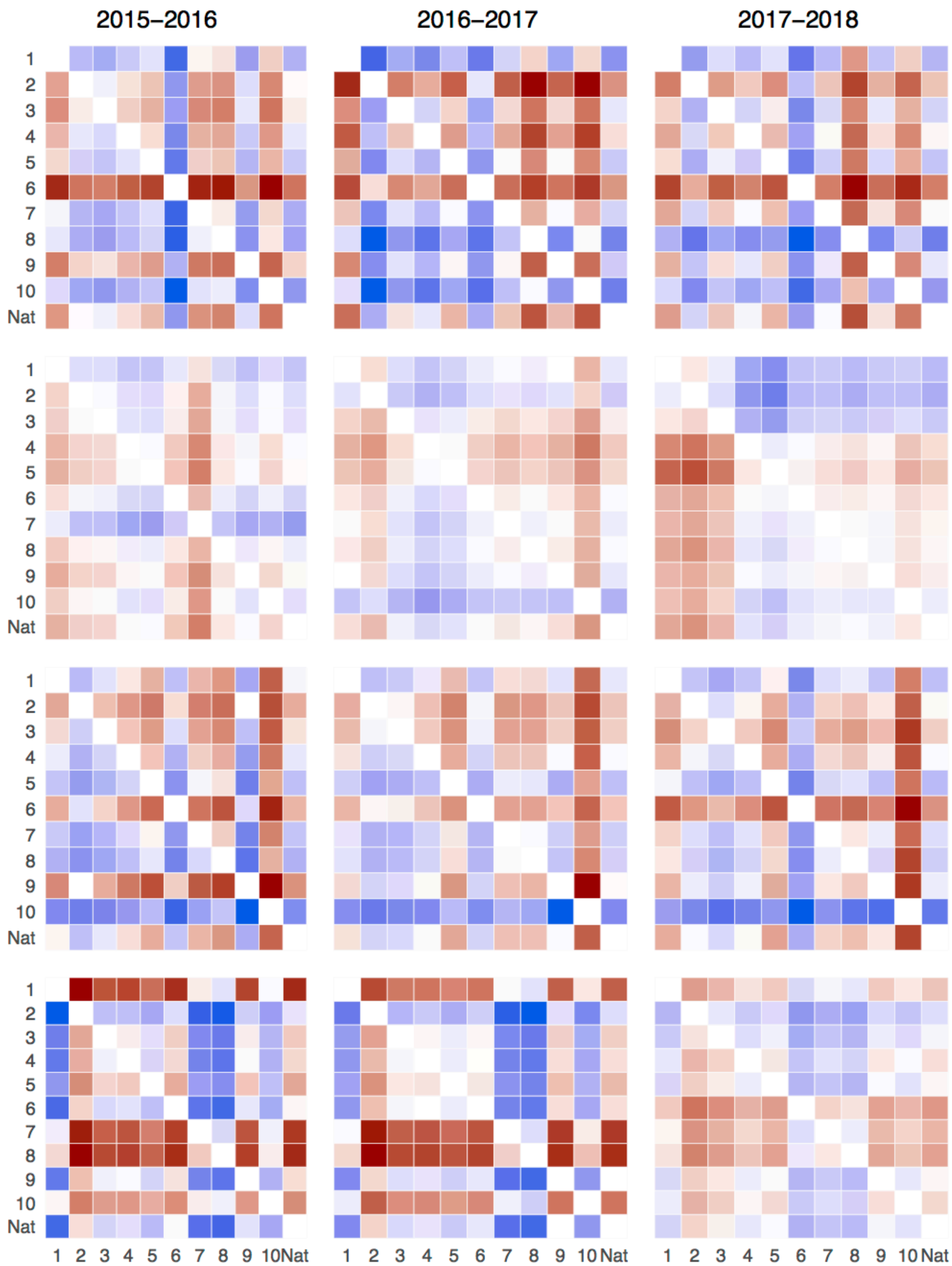

Figure A.9: Heatmaps of the mean relative difference of ILI activity across geographical areas for (A) CDC ILINet, (B) FNY, (C) athenahealth, and (D) HealthTweets.org for each influenza season.

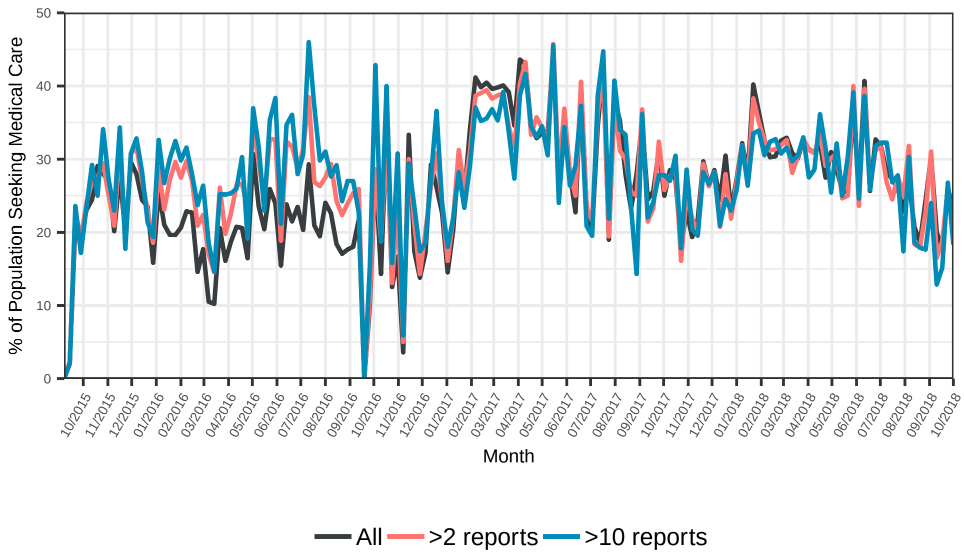

Figure A.10: Comparison of weekly percent of FNY participants who sought health care for ILI symptoms by entry criteria.

LIST OF JOURNAL ABBREVIATIONS

Commun Dis Intell	Communicable Diseases Intelligence
BMC	BioMed Center
JMIR	Journal of Medical Internet Research
PloS	Public Library of Science

BIBLIOGRAPHY

Althouse, B. M., Scarpino, S. V., Meyers, L. A., Ayers, J. W., Bargsten, M., Baumbach, J., Brownstein, J. S., Castro, L., Clapham, H., Cummings, D. A. T., Del Valle, S., Eubank, S., Fairchild, G., Finelli, L., Generous, N., George, D., Harper, D. R., Hébert-Dufresne, L., Johansson, M. A., Konty, K., Lipsitch, M., Milinovich, G., Miller, J. D., Nsoesie, E. O., Olson, D. R., Paul, M., Polgreen, P. M., Priedhorsky, R., Read, J. M., Rodríguez-Barraquer, I., Smith, D. J., Stefansen, C., Swerdlow, D. L., Thompson, D., Vespignani, A., & Wesolowski, A. (2015). Enhancing disease surveillance with novel data streams: challenges and opportunities. *EPJ Data Science*, *4*(1), 1–8.

Australian Government Department of Health (2016a). AUSTRALIAN INFLUENZA Laboratory Confirmed Influenza Activity. Tech. Rep. 11, Australian Government Department of Health. http://www.health.gov.au.

Australian Government Department of Health (2016b). NATIONAL NOTIFIABLE DISEASES SURVEILLANCE SYSTEM. Tech. rep., Australian Government Department of Health. http://www.health.gov.au/nndssdata.

Bajardi, P., Paolotti, D., Vespignani, A., Eames, K., Funk, S., Edmunds, W. J., Turbelin, C., Debin, M., Colizza, V., Smallenburg, R., Koppeschaar, C., Franco, A. O., Faustino, V., Carnahan, A., Rehn, M., Merletti, F., Douwes, J., Firestone, R., & Richiardi, L. (2014a). Association between recruitment methods and attrition in Internet-based studies. *PloS one*, *9*(12), e114925.

Bajardi, P., Vespignani, A., Funk, S., Eames, K. T. D., Edmunds, W. J., Turbelin, C., Debin, M., Colizza, V., Smallenburg, R., Koppeschaar, C. E., Franco, A. O., Faustino, V., Carnahan, A., Rehn, M., & Paolotti, D. (2014b). Determinants of follow-up participation in the internet-based european influenza surveillance platform influenzanet. *Journal of Medical Internet Research*, *16*(3).

Baltrusaitis, K., Brownstein, J. S., Scarpino, S. V., Bakota, E., Crawley, A. W., Conidi, G., Gunn, J., Gray, J., Zink, A., & Santillana, M. (2018). Comparison of crowd-sourced , electronic health records based , and traditional health- care based influenza-tracking systems at multiple spatial resolutions in the United States of America. *BMC infectious diseases*, (pp. 1–8).

Baltrusaitis, K., Santillana, M., Crawley, A. W., Chunara, R., Smolinski, M., & Brownstein, J. S. (2017). Determinants of Participants' Follow-Up and Characterization of Representativeness in Flu Near You, A Participatory Disease Surveillance System. *JMIR public health and surveillance*, *3*(2), e18.

Biggerstaff, M., Alper, D., Dredze, M., Fox, S., Fung, I. C. H., Hickmann, K. S., Lewis, B., Rosenfeld, R., Shaman, J., Tsou, M. H., Velardi, P., Vespignani, A., Finelli, L., Chandra, P., Kaup, H., Krishnan, R., Madhavan, S., Markar, A., Pashley, B., Paul, M., Meyers, L. A., Eggo, R., Henderson, J., Ramakrishnan, A., Scott, J., Singh, B., Srinivasan, R., Bakach, I., Hao, Y., Schaible, B. J., Sexton, J. K., Del Valle, S. Y., Deshpande, A., Fairchild, G., Generous, N., Priedhorsky, R., Hickman, K. S., Hyman, J. M., Brooks, L., Farrow, D., Hyun, S., Tibshirani, R. J., Yang, W., Allen, C., Aslam, A., Nagel, A., Stilo, G., Basagni, S., Zhang, Q., Perra, N., Chakraborty, P., Butler, P., Khadivi, P., Ramakrishnan, N., Chen, J., Barrett, C., Bisset, K., Eubank, S., Anil Kumar, V. S., Laskowski, K., Lum, K., Marathe, M., Aman, S., Brownstein, J. S., Goldstein, E., Lipsitch, M., Mekaru, S. R., Nsoesie, E. O., Gesualdo, F., Tozzi, A. E., Broniatowski, D., Karspeck, A., Tse, Z. T. H., Ying, Y., Gambhir, M., & Scarpino, S. (2016). Results from the centers for disease control and prevention's predict the 2013-2014 Influenza Season Challenge. *BMC Infectious Diseases*, *16*(1), 1–10.

Biggerstaff, M., Jhung, M., Kamimoto, L., Balluz, L., & Finelli, L. (2012). Self-Reported Influenza-Like Illness and Receipt of Influenza Antiviral Drugs During the 2009 Pandemic , United States , 2009 2010. *American Journal of Public Health*, *102*(10), 2009–2010.

Biggerstaff, M., Jhung, M. A., Reed, C., Fry, A. M., Balluz, L., & Finelli, L. (2014). Influenza-like illness, the time to seek healthcare, and influenza antiviral receipt during the 2010-2011 influenza season - United States. *Journal of Infectious Diseases*, *210*(4), 535–544.

Biggerstaff, M., Johansson, M., Alper, D., Brooks, L. C., Chakraborty, P., Farrow, D. C., Hyun, S., Kandula, S., McGowan, C., Ramakrishnan, N., Rosenfeld, R., Shaman, J., Tibshirani, R., Tibshirani, R. J., Vespignani, A., Yang, W., Zhang, Q., & Reed, C. (2018). Results from the second year of a collaborative effort to forecast influenza seasons in the United States. *Epidemics*, (May 2017).

Blanchon, T., Bonmarin, I., Falchi, A., Debin, M., Hanslik, T., Levy-bruhl, D., Poletto, C., & Colizza, V. (2013). Evaluating the Feasibility and Participants Representativeness of an Online Nationwide Surveillance System for Influenza in France. *PloS one*, *8*(9).

Broniatowski, D. A., Paul, M. J., & Dredze, M. (2013). National and local influenza surveillance through twitter: An analysis of the 2012-2013 influenza epidemic. *PLoS ONE*, *8*(12), 1–8.

Brooks, L. C., Farrow, D. C., Hyun, S., Tibshirani, R. J., & Rosenfeld, R. (2015). Flexible Modeling of Epidemics with an Empirical Bayes Framework. *PLoS Computational Biology*, *11*(8), 1–18.

Cantarelli, P., Debin, M., Turbelin, C., Poletto, C., Blanchon, T., Falchi, A., Hanslik, T., Bonmarin, I., Levy-Bruhl, D., Micheletti, A., Paolotti, D., Vespignani, A., Edmunds, J., Eames, K. T. D., Smallenburg, R., Koppeschaar, C., Franco, A. O., Faustino, V., Carnahan, A., Rehn, M., & Colizza, V. (2014). The representativeness of a European multi-center network for influenza-like-illness participatory surveillance. *BMC public health*, *14*, 984.

Carlson, S. J., Dalton, C. B., Butler, M. T., Fejsa, J., Elvidge, E., & Durrheim, D. N. (2013). weekly online community survey and 2012. *Commun Dis Intell*, *37*(4), 398–406.

Carlson, S. J., Dalton, C. B., Durrheim, D. N., & Fejsa, J. (2010). Online Flutracking survey of influenza-like illness during pandemic (H1N1) 2009, Australia. *Emerging Infectious Diseases*, *16*(12), 1960–1962.

Centers for Disease Control (CDC), C. f. D. C. (1991). Nosocomial transmission of multidrug-resistant tuberculosis among HIV-infected persons–Florida and New York, 1988-1991. *MMWR. Morbidity and mortality weekly report*, *40*(34), 585–91.

Chen, L., Tozammel Hossain, K. S. M., Butler, P., Ramakrishnan, N., & Prakash, B. A. (2015). Syndromic surveillance of Flu on Twitter using weakly supervised temporal topic models. *Data Mining and Knowledge Discovery*.

Choi, J., Cho, Y., Shim, E., & Woo, H. (2016). Web-based infectious disease surveillance systems and public health perspectives: a systematic review. *BMC Public Health*, *16*(1), 1238.

Chunara, R., Goldstein, E., Patterson-lomba, O., & Brownstein, J. S. (2015). Estimating influenza attack rates in the United States using a participatory cohort. *Scientific Reports*, (pp. 1–5).

Chunara, R., Wisk, L. E., & Weitzman, E. R. (2017). Denominator Issues for Personally Generated Data in Population Health Monitoring. *American Journal of Preventive Medicine*, *52*(4), 549–553.

Dalton, C. B., Carlson, S. J., Butler, M. T., Elvidge, E., & Durrheim, D. N. (2013). Building Influenza Surveillance Pyramids in Near Real Time, Australia. *Emerging Infectious Diseases*, *19*(11), 19–21.

Debin, M., Colizza, V., Blanchon, T., Hanslik, T., Turbelin, C., & Falchi, A. (2014). Effectiveness of 2012 2013 influenza vaccine against influenza-like illness in general population Estimation in a French web-based cohort. *10*(3), 536–543.

Dredze, M., Cheng, R., Paul, M. J., & Broniatowski, D. (2014). HealthTweets.org: A platform for public health surveillance using Twitter. *AAAI Workshop - Technical Report*, *WS-14-14*, 2–3.

Flutracking.net (2018). FluTracking.net. https://info.flutracking.net/.

Fox, S., Rainie, L., Horrigan, J., Lenhart, A., Spooner, T., Burke, M., Lewis, O., & Carter, C. (2000). The online health care revolution : How the Web helps Americans take better care of themselves. *Pew Internet & American Life Project.*, (pp. 1–23).

Fujibayashi, K., Takahashi, H., Tanei, M., Uehara, Y., Yokokawa, H., & Naito, T. (2018). A new influenza-tracking smartphone app (flu-report) based on a self-administered questionnaire: Cross-sectional study. *Journal of Medical Internet Research*, *20*(6), 1–8.

Generous, N., Fairchild, G., Deshpande, A., Del Valle, S. Y., & Priedhorsky, R. (2014). Global disease monitoring and forecasting with Wikipedia. *PLoS computational biology*, *10*(11), e1003892.

Giesecke, J. (2002). *Modern Infectious Disease Epidemiology*. London, United Kingdom: Taylor & Francis Ltd, 2nd ed.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012–1014.

Goff, J., Rowe, A., Brownstein, J. S., & Chunara, R. (2015). Surveillance of Acute Respiratory Infections Using Community ■ Submitted Symptoms and Specimens for Molecular Diagnostic Testing. *PLoS currents*.

Guerrisi, C., Turbelin, C., Blanchon, T., Hanslik, T., Bonmarin, I., Levy-Bruhl, D., Perrotta, D., Paolotti, D., Smallenburg, R., Koppeschaar, C., Franco, A. O., Mexia, R., Edmunds, W. J., Sile, B., Pebody, R., van Straten, E., Meloni, S., Moreno, Y., Duggan, J., Kjelsø, C., & Colizza, V. (2016). Participatory Syndromic Surveillance of Influenza in Europe. *Journal of Infectious Diseases*, *214*(suppl 4), S386–S392.

Kirkwood, B. R., & Sterne, J. (2003). *Essentials of Medical Statistics*. Blackwell Science Ltd, 2nd ed.

Kullenberg, C., & Kasperowski, D. (2016). What is citizen science? - A scientometric meta-analysis. *PLoS ONE*, *11*(1), 1–16.

Leacy, F. P., Floyd, S., Yates, T. A., & White, I. R. (2017). Analyses of sensitivity to the missing-at-random assumption using multiple imputation with delta adjustment: Application to a tuberculosis/HIV prevalence survey with incomplete HIV-status data. *American Journal of Epidemiology*, *185*(4), 304–315.

Lee, E. C., Arab, A., Goldlust, S. M., & Grenfell, B. (2018a). (Gen) Deploying digital health data to optimize influenza surveillance at national and local scales. (Ili).

Lee, E. C., Arab, A., Goldlust, S. M., Viboud, C., Grenfell, B. T., & Bansal, S. (2018b). Deploying digital health data to optimize influenza surveillance at national and local scales. *Plos Computational Biology*, *14*(3), 1–23.

Lee, V. J., Ho, Z. J. M., Goh, E. H., Campbell, H., Cohen, C., Cozza, V., Fitzner, J., Jara, J., Krishnan, A., & Bresee, J. (2018c). Advances in measuring influenza burden of disease. *Influenza and other Respiratory Viruses*, *12*(1), 3–9.

Lewis, K., & Burd-Sharps, S. (2015). The Measure of America 2013-2014. Tech. rep. http://www.measureofamerica.org/.

Lipsitch, M., Finelli, L., Heffernan, R. T., Leung, G. M., Redd, S. C., & Group, f. t. . H. S. (2011). Improving the evidence base for decision making during a pandemic: the example of 2009 influenza A/H1N1. *Biosecurity and bioterrorism : biodefense strategy, practice, and science*, *9*(2), 89–115.

Lu, F. S., Hou, S., Baltrusaitis, K., Shah, M., Leskovec, J., Sosic, R., Hawkins, J., Brownstein, J., Conidi, G., Gunn, J., Gray, J., Zink, A., & Santillana, M. (2018). Accurate influenza monitoring and forecasting using novel internet data streams:a case study in the boston metropolis. *Journal of Medical Internet Research*, *20*(1).

Ma, W., Huo, X., & Zhou, M. (2018). The healthcare seeking rate of individuals with influenza like illness: a meta-analysis. *Infectious Diseases*, *0*(0), 1–8.

Marquet, R. L., Bartelds, A. I. M., van Noort, S. P., Koppeschaar, C. E., Paget, J., Schellevis, F. G., & van der Zee, J. (2006). Internet-based monitoring of influenza-like illness (ILI) in the general population of the Netherlands during the 2003-2004 influenza season. *BMC public health*, *6*(Ili), 242.

McIver, D. J., & Brownstein, J. S. (2014). Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. *PLoS computational biology*, *10*(4), e1003581.

Nölke, L., Mensing, M., Krämer, A., & Hornberg, C. (2015). Sociodemographic and health-(care-)related characteristics of online health information seekers: a cross-sectional German study. *BMC Public Health*, *15*(1), 31.

Paolotti, D., Carnahan, A., Colizza, V., Eames, K., Edmunds, J., Gomes, G., Koppeschaar, C., Rehn, M., Smallenburg, R., Turbelin, C., Van Noort, S., & Vespignani, A. (2014). Web-based participatory surveillance of infectious diseases: The Influenzanet participatory surveillance experience. *Clinical Microbiology and Infection*, *20*(1), 17–21.

Patterson-Lomba, O., Van Noort, S., Cowling, B. J., Wallinga, J., Gomes, M. G. M., Lipsitch, M., & Goldstein, E. (2014). Utilizing syndromic surveillance data for estimating levels of influenza circulation. *American Journal of Epidemiology*, *179*(11), 1394–1401.

Paul, M. J., Dredze, M., & Broniatowski, D. (2014). Twitter improves influenza forecasting. *PLoS currents*, *6*, 1–13.

Peppa, M., John Edmunds, W., & Funk, S. (2017). Disease severity determines health-seeking behaviour amongst individuals with influenza-like illness in an internet-based cohort. *BMC Infectious Diseases*, *17*(1), 1–13.

Perrin, A., & Duggan, M. (2015). Americans' internet access:2000-2015. http://www.pewinternet.org/2015/06/26/americans-internet-access-2000-2015.

Polgreen, P. M., Chen, Y., Pennock, D. M., Nelson, F. D., & Weinstein, R. A. (2008). Using Internet Searches for Influenza Surveillance. *Clinical Infectious Diseases*, *47*(11), 1443–1448.

Polgreen, P. M., Chen, Z., Segre, A. M., Harris, M. L., Pentella, M. A., & Rushton, G. (2009a). Optimizing influenza sentinel surveillance at the state level. *American Journal of Epidemiology*, *170*(10), 1300–1306.

Polgreen, P. M., Chen, Z., Segre, A. M., Harris, M. L., Pentella, M. A., & Rushton, G. (2009b). Optimizing influenza sentinel surveillance at the state level. *American Journal of Epidemiology*, *170*(10), 1300–1306.

Putri, W. C., Muscatello, D. J., Stockwell, M. S., & Newall, A. T. (2018). Economic burden of seasonal influenza in the United States. *Vaccine*.

R Core Team (R Foundation for Statistical Computing) (2016). R: A Language and Environment for Statistical Computing. https://www.r-project.org/.

Reed, C., Angulo, F. J., Biggerstaff, M., Swerdlow, D., & Finelli, L. (2011). Influenza-like Illness in the community during the emergence of 2009 pandemic influenza a(H1N1)-survey of 10 states, April 2009. *Clinical Infectious Diseases*, *52*(SUPPL. 1), 2009–2012.

Reed, C., Angulo, F. J., Swerdlow, D. L., Lipsitch, M., Meltzer, M. I., Jernigan, D., & Finelli, L. (2009). Estimates of the Prevalence of Pandemic (H1N1) 2009, United States, April-July 2009. *15*(12), 10–13.

Reed, C., Biggerstaff, M., Vargas, C., Crawley, A., Larson, E., Saiman, L., & Stockwell, M. (2016). Evaluation of crowd-sourced vs. research surveillance methods for the incidence of acute respiratory infections in the community. *MoSAIC Conference*.

Reed, C., Chaves, S. S., Kirley, P. D., Emerson, R., Aragon, D., Hancock, E. B., Butler, L., Baumbach, J., Hollick, G., Bennett, N. M., Laidler, M. R., Thomas, A., Meltzer, M. I., & Finelli, L. (2015). Estimating influenza disease burden from population-based surveillance data in the United States. *PLoS ONE*, *10*(3).

Rolfes, M. A., Foppa, I. M., Garg, S., Flannery, B., Brammer, L., Singleton, J. A., Burns, E., Jernigan, D., Olsen, S. J., Bresee, J., & Reed, C. (2018). Annual estimates of the burden of seasonal influenza in the United States: A tool for strengthening influenza surveillance and preparedness. *Influenza and other Respiratory Viruses*, *12*(1), 132–137.

Rubin, D. B. (2004). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons.

Salathé, M. (2018). Digital epidemiology: what is it, and where is it going? *Life Science, Society and Policy*, *14*(1), 1–5.

Santillana, M., Nguyen, A. T., Dredze, M., Paul, M. J., Nsoesie, E. O., & Brownstein, J. S. (2015). Combining Search, Social Media, and Traditional Data Sources to Improve Influenza Surveillance. *PLOS Computational Biology*, *11*(10), e1004513.

Santillana, M., Nguyen, A. T., Louie, T., Zink, A., Gray, J., Sung, I., & Brownstein, J. S. (2016). Cloud-based Electronic Health Records for Real-time, Region-specific Influenza Surveillance. *Scientific Reports*, *6*(April), 25732.

Santillana, M., Nsoesie, E. O., Mekaru, S. R., Scales, D., & Brownstein, J. S. (2014). Using clinicians search query data to monitor influenza epidemics.

Scarpino, S. V., Dimitrov, N. B., & Meyers, L. A. (2012). Optimizing provider recruitment for influenza surveillance networks. *PLoS computational biology*, *8*(4), e1002472.

Shaman, J., Pitzer, V. E., Viboud, C., Grenfell, B. T., & Lipsitch, M. (2010). Absolute humidity and the seasonal onset of influenza in the continental United States. *PLoS Biology*, *8*(2).

Shrestha, S. S., Swerdlow, D. L., Borse, R. H., Prabhu, V. S., Finelli, L., Atkins, C. Y., Owusu-edusei, K., Bell, B., Mead, P. S., Biggerstaff, M., Brammer, L., Davidson, H., Jernigan, D., Jhung, M. A., Kamimoto, L. A., Merlin, T. L., Nowell, M., Redd, S. C., Reed, C., Schuchat, A., & Meltzer, M. I. (2017). Estimating the Burden of 2009 Pandemic Influenza A (H1N1) in the United States (April 2009 April 2010). *52*(April 2010), 75–82.

Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PloS one*, *6*(5), e19467.

Smolinski, M. S., Crawley, A. W., Baltrusaitis, K., Chunara, R., Olsen, J. M., Wójcik, O., Santillana, M., Nguyen, A., & Brownstein, J. S. (2015). Flu Near You: Crowd-sourced Symptom Reporting Spanning 2 Influenza Seasons. *American Journal of Public Health*, *105*(10), 2124–2130.

Smolinski, M. S., Crawley, A. W., Olsen, J. M., Jayaraman, T., & Crawley, A. W. (2017). Participatory Disease Surveillance : Engaging Communities Directly in Reporting , Monitoring , and Responding to Health Threats Corresponding Author :. *3*.

Sullivan, S. G., Kate Pennington, J. R., Franklin, L. J., , R. d., Christina Bareja, Kluyver, , the National Influenza Surveillance Committee, f. t. C. D., & Australia, N. (2016). A Summary of Influenza Surveillance Systems in Australia, 2015. *Communicable Disease Intelligence*, (pp. 1–51).

Thompson, W., Comanor, L., & Shay, D. (2006). Epidemiology of Seasonal Influenza: Use of Surveillance Data and Statistical Models to Estimate the Burden of Disease. *The Journal of Infectious Diseases*, *194*(s2), S82–S91.

Tilston, N. L., Paolotti, D., & Ealden, T. (2010). Internet-based surveillance of Influenza-like-illness in the UK during the 2009 H1N1 influenza pandemic. *BMC Public Health*, *10*, 1–9.

Tizzoni, M., Bajardi, P., Poletto, C., Ramasco, J. J., Balcan, D., Gonçalves, B., Perra, N., Colizza, V., & Vespignani, A. (2012). Real-time numerical forecast of global epidemic spreading: Case study of 2009 A/H1N1pdm. *BMC Medicine*, *10*(1), 165.

U.S. Census Bureau (2015). Annual Estimates of the Resident Population for Selected Age Groups by Sex for the United States, States, Counties, and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2014. http://factfinder.census.gov.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3).

van Noort, S. P., Codeço, C. T., Koppeschaar, C. E., van Ranst, M., Paolotti, D., & Gomes, M. G. M. (2015). Ten-year performance of Influenzanet: ILI time series, risks, vaccine effects, and care-seeking behaviour. *Epidemics*, *13*, 28–36.

World Health Organization (2015). A Manual for Estimating Disease Burden Associated with Seasonal Influenza. Tech. rep., World Health Organization. http://www.who.int/influenza/resources/publications/manual_burden_of_disease/en/.

Yang, S., Santillana, M., Brownstein, J. S., Gray, J., Richardson, S., & Kou, S. C. (2017). Using electronic health records and Internet search information for accurate influenza forecasting. *BMC Infectious Diseases*, *17*(1), 332.

Yang, S., Santillana, M., & Kou, S. C. (2015a). Accurate estimation of influenza epidemics using Google search data via ARGO. *Proceedings of the National Academy of Sciences*, *112*(47), 14473–14478.

Yang, W., Lipsitch, M., & Shaman, J. (2015b). Inference of seasonal and pandemic influenza transmission dynamics. *Proceedings of the National Academy of Sciences*, *112*(9), 2723–2728.

Yuan, Q., Nsoesie, E. O., Lv, B., Peng, G., Chunara, R., & Brownstein, J. S. (2013). Monitoring influenza epidemics in china with search query from baidu. *PloS one*, *8*(5), e64323.

Zhang, Q., Perra, N., Perrotta, D., Tizzoni, M., Paolotti, D., & Vespignani, A. (2017). Forecasting Seasonal Influenza Fusing Digital Indicators and a Mechanistic Disease Model. *Proceedings of the 26th International Conference on World Wide Web - WWW '17*, (pp. 311–319).

CV

[redacted]

[redacted]

[redacted]