Evaluation of approaches that adjust for biases in participatory surveillance systems

Kristin Baltrusaitis[1], Kathleen Noddin[2], Colleen Nguyen[2], Adam Crawley*[3], John S. Brownstein[2, 4] and Laura F. White[1]

[1]Department of Biostatistics, Boston University School of Public Health, Boston University, Boston, MA, USA; [2]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, USA; [3]Skoll Global Threats Fund, San Francisco, CA, USA; [4]Harvard Medical School, Boston, MA, USA

Objective

To estimate and compare influenza attack rates (AR) in the United States (US) using different approaches to adjust for reporting biases in participatory syndromic surveillance data.

Introduction

Because the dynamics and severity of influenza in the US vary each season, yearly estimates of disease burden in the population are essential to evaluate interventions and allocate resources. The CDC uses data from a national health-care based surveillance system and mathematical models to estimate the overall burden of disease in the general population. Over the past decade, crowd-sourced syndromic surveillance systems have emerged as a digital data source that collects health-related information in near real-time. These systems complement traditional surveillance systems by capturing individuals who do not seek medical care and allowing for a longitudinal view of illness burden. However, because not all participants report every week and participants are more likely to report when ill, the number of weekly reports is temporally and spatially inconsistent and the estimates of disease burden and incidence may be biased. In this study, we use data from Flu Near You (FNY), a participatory surveillance system based in the US and Canada[1], to estimate and compare Influenza-like Illness (ILI) ARs using different approaches to adjust for reporting biases in participatory surveillance data.

Methods

This analysis uses FNY data from the 2015-16 influenza season. Four different approaches of bias adjustment were assessed. The first approach includes all FNY participants, defined as users and household members, who submitted at least one symptom report, whereas the second approach only includes FNY participants who submitted at least 10 symptom reports. The third approach includes all FNY participants who submitted at least one symptom report, but drops the first symptom report for all participants. For the first three approaches, all missing reports were assumed to be non-ILI when estimating attack rates. Finally, the fourth approach includes FNY participants who submitted at least 10 symptom reports and uses multiple imputation to account for missing reports. Age-stratified and overall estimates of ILI ARs were calculated for each of the four approaches to bias adjustment by dividing the sum of the weekly incident cases of ILI, defined as the first report of fever with cough and/ or sore throat, by the population at risk at the beginning of the period.

Results

During the 2016-2017 influenza season, FNY received an average of 10,723 unique symptom reports per week from 46,390 registered users and their household members. For FNY, the youngest age group assessed, 5-17, had the largest ILI AR, and the ILI ARs decreased as the age group increased for all approaches. Overall, the approach that drops all first reports had the smallest ARs, whereas the approach that selects a cohort of users who submit at least 10 reports during the season and imputes the missing reports had the largest ARs. Although the influenza ARs estimated by the CDC were less than the ILI ARs estimated using FNY data for all age-groups, a similar pattern was observed across age groups, except for the 50-64 age group, which had the largest influenza AR.

Conclusions

As expected, the ARs estimated using FNY data were greater than the CDC's influenza ARs because FNY estimates ARs of ILI and does not adjust for the probability of reporting ILI when experiencing non-flu illness. The approach of dropping the first report had the smallest ARs because during the 2015-16 influenza season the weekly percent of ILI cases that were first time reports ranged from 18-59%. This approach was developed to adjust for the potential correlation between symptom presence and willingness to join the platform. However, important information about the dynamics of disease may be lost when using this approach. The multiple imputation method was used only for individuals who submitted at least 10 reports to maintain a missing data rate below 30%. The imputation model also assumed that data were missing at random, which may not be appropriate in this case, because approximately 30% of FNY users have reported that they are more likely to report when ill. As shown in Table 1, the AR estimate depends on the bias adjustment approach. Simulation-based studies should be performed to further evaluate these methods.

Estimated Attack Rates (%) with 95% CI by age group for the 2015-2016 influenza season

Age Group	CDC [2]	FNY			
		Raw	≥ 10 reports	Drop 1st report	Multiple Imputation
5-17	7.42 (6.26,8.90)	20.72 (19.41,22.03)	28.92 (26.19,31.65)	19.85 (18.45,21.25)	42.00 (38.96,45.04)
18-49	7.10 (6.17,8.34)	19.87 (19.23,20.51)	18.73 (17.41,20.05)	17.27 (16.57,17.97)	29.28 (27.73,30.83)
50-64	11.04 (9.99,12.36)	18.26 (17.61,18.91)	16.24 (15.15,17.33)	16.26 (15.56,16.96)	22.88 (21.63,24.13)
≥ 65	3.53 (3.09,4.24)	11.30 (10.60,12.00)	11.33 (10.26,12.40)	10.55 (9.82,11.28)	16.74 (15.47,18.01)
All ages	7.65 (6.69,8.91)	18.14 (17.77,18.51)	17.41 (16.75,18.07)	16.21 (15.81,16.61)	25.35 (24.58,26.13)

Keywords

Community-based participatory research; digital disease detection; influenza; public health surveillance

Acknowledgments

References

1. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu Near You: Crowdsourced Symptom Reporting Spanning 2 Influenza Seasons. Am J Public Health. 2015
2. Rolfes MA, Foppa IM, Garg S, Flannery B, Brammer L, Singleton JA, et al. Estimated Influenza Illnesses, Medical Visits, Hospitalizations, and Deaths Averted by Vaccination in the United States. 2016 Dec 9 [2017 Sept 25];
https://www.cdc.gov/flu/about/disease/2015-16.htm

*Adam Crawley
E-mail: acrawley@skollglobalthreats.org