State-level demographics and tobacco-control correlates of smoking cessation behavioral change techniques on Twitter

Sherine El-Toukhy, Chris Vargo, Toby Hopp, Kelvin Choi

Abstract

Objective: Identify behavioral change techniques for smoking cessation reflected in discourse on Twitter and explore state-level demographic composition, smoking prevalence, quit rates, smoking-related health conditions, cigarette prices, and tobacco-control expenditures that are associated with smoking cessation tweets volume.

Methods: Using Social Studio Radian 6 application programming interface (API), we retrieved tweets containing smoking-related keywords (eg, smoke) from 1/1/2009 to 12/7/2015. We developed a codebook based on Michie et al. (2011) taxonomy for smoking cessation behavioral change techniques. Two coders manually annotated 5715 random tweets, which were then used to build a machine learning algorithm using LightSide. We used geo-tags or self-report location to geocode tweets to the states using Google Maps Places API. We retrieved state-level data from Centers for Disease Control and Prevention and census data. A least absolute shrinkage and selection operator regression model was used to identify correlates between tweet volume and state-level variables.

Results: A total of 1,431,790 tweets that included behavioral change techniques for smoking cessation were retained for analysis. We identified 10 techniques that fell under four categories: motivation, self-regulatory capacity, adjuvant activities, and information gathering. Variance explained in tweet volume ranged from 3.5% for adjuvant activities to 14.4% for self-regulation. Motivation and information gathering were positively associated with percent smokers under age 35 and negatively associated with racial heterogeneity. Motivation was positively associated with percent population below poverty line and percent smoking adults but negatively associated with price per cigarette pack. Percent of smoking adults was positively associated with self-regulation and negatively associated with adjuvant activity.

Discussion: Geographic clustering of cessation tweets was associated with high smoking prevalence especially among young adults, high poverty levels, and cigarette prices. Social media reflect public interest in smoking cessation and state-level smoking profile and tobacco-control policies.