Big Data for Reproductive Health (bd4rh) - Team RBG

Amy Finnegan, PhD[1, 2, 3], Megan Huchko, MD, MPH [1, 2, 5], Kelly Hunter [1, 2, 4],
Dennis Harrsch, Chelsea Jubitana, Manvi Malhotra, Linda Tang, Bhamini Vellanki

Affiliation: 1. Duke Global Health Institute, 2. Duke Center for Global Reproductive Health,
3. IntraHealth International, 4. Stanford School of Public Policy, 5. Duke OB/GYN

Overview

Reproductive rights and access to long-acting reversible contraception (LARC) have become especially contentious in light of recent significant political events.

DukeBD4RH
@TeamRBG
Following

"The number of annual tweets mentioning contraception has increased nearly three hundred-fold since 2007."

2:46 PM · Apr 21 2021

Goal: understand social sentiments through social media analysis with big data methods is a recent development, posing significant implications for analyzing reproductive behavior and health trends, especially in the context of political events

Question: What are the effects of large-scale **conservative political events** in the U.S. on the **discussion** around Long-acting reversible contraception (LARC) and LARC uptake on Twitter?

Hypothesis: Normally, the discussion on Twitter around LARC and its uptake will be "**negative**". But following **significant conservative political** events, we expect a significant **increase in "positive"tweets**, thus the **overwhelming sentiment** immediately following the event will be positive.

Methods

Search Terms:
"IUD"
"Nexplanon"

Two Events

Trump Election: 11/8/2016

RBG Death: 9/18/2020

Question

What are the effects of large-scale **conservative political events** in the U.S. on the **discussion** around:
1. **long-acting reversible contraception (LARC)**
2. **LARC uptake** on Twitter.

01

02 Methods: Tweet Extraction and Cleaning
1. Identified **timeframes** of political events for tweet extraction
2. Identified **key search terms** for tweet extraction
3. Used **Twitter Advanced Search** to pull tweets to refine search terms and dates
4. Used **Twitter API** to extract tweets
4. **Cleaned** tweets

04 Methods: Training a Model
1. Used random sample (100 tweets) to **train a classifier model**
2. Refining model -- training with random sample of **70% of handcoded tweets**
3. **Test model** with rest of tweets

03 Methods: Handcoding
1. Identified **standardized criteria** for handcoding
2. Handcoded **random sample** of 100 tweets
3. Refined **handcoding criteria**
4. Handcoded entire tweet dataset

Time Period 1
Entire week of the event: 3 days before & after, day of

Time Period 2
Same week: 1 month before & after

Time Period 3
Same week: 6 months before, & after

Methods & Results

Classifying Tweets for sentiments surrounding:
1) LARC
2) LARC uptake

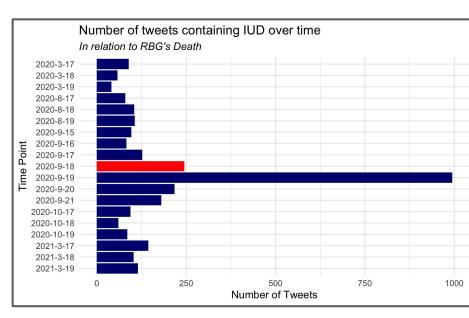**Positive**
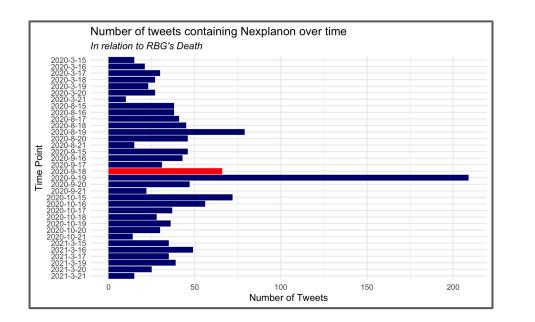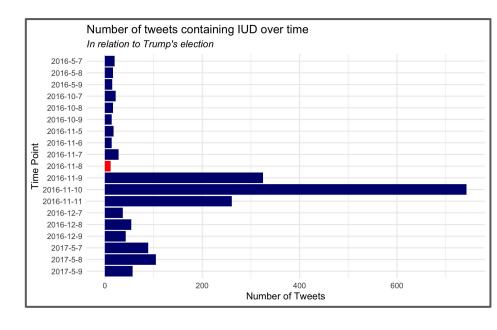1. Net positive sentiment expressed in relation to IUD/Nexplanon
2. Net positive sentiment expressed in relation to IUD/Nexplanon uptake

Negative
1. Negative sentiment expressed in relation to IUD/Nexplanon
2. Negative sentiment expressed in relation to IUD/Nexplanon uptake

Inconclusive
1. Balanced or neutral sentiment expressed in relation to IUD/Nexplanon
2. Balanced or neutral sentiment expressed in relation to IUD/Nexplanon uptake

Not applicable
1. No sentiment expressed in relation to IUD/Nexplanon
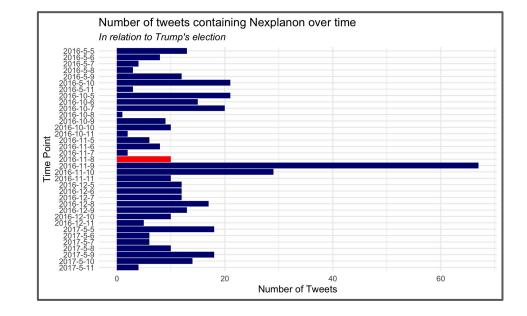2. No sentiment expressed in relation to IUD/Nexplanon uptake

Words that appear most frequently in tweets surrounding IUD

Tweet Volume for Event 1: RBG's Death (~5000)

Number of tweets containing IUD over time — In relation to RBG's Death

Number of tweets containing Nexplanon over time — In relation to RBG's Death

Tweet Volume for Event 2: Trump's Election (~3000)

Number of tweets containing IUD over time — In relation to Trump's election

Number of tweets containing Nexplanon over time — In relation to Trump's election

Topic Modeling of Tweets before and after Trump's Election

Topic Modeling — In relation to Trump's election, November before political event

Topic Modeling — In relation to Trump's election, November after political event

Sentiment Analysis in Handcoded Tweets

Relative frequencies of sentiments towards Nex RBG

Relative frequencies of sentiments towards Nex uptake RBG
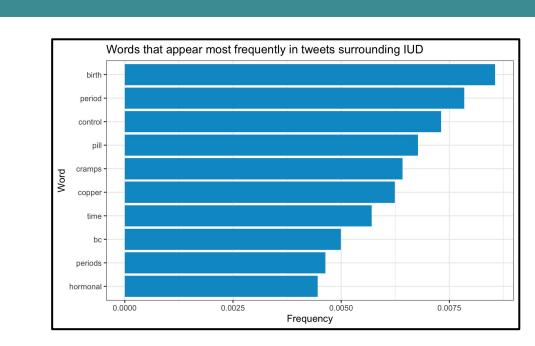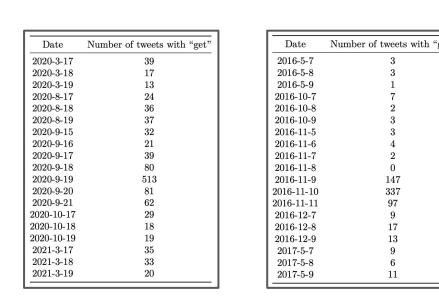
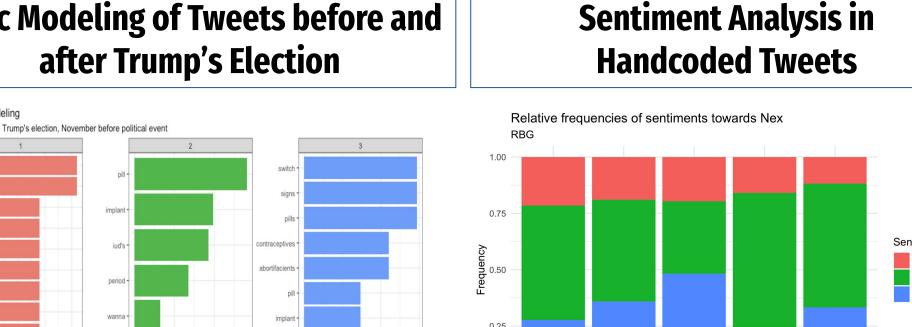## Conclusion

After a conservative political event:
- The total volume of tweets surrounding LARC drastically increase.
  - This increase in discussion tends to persist for around 1-2 weeks.
- Fear emerged as a cluster in topic modeling, more prominent than concerns about the adverse effects of contraceptives.
- The sentiment towards LARC in general shifts slightly positive.
- Sentiment towards LARC uptake becomes highly positive mainly due to concerns around the loss of access to contraceptives.
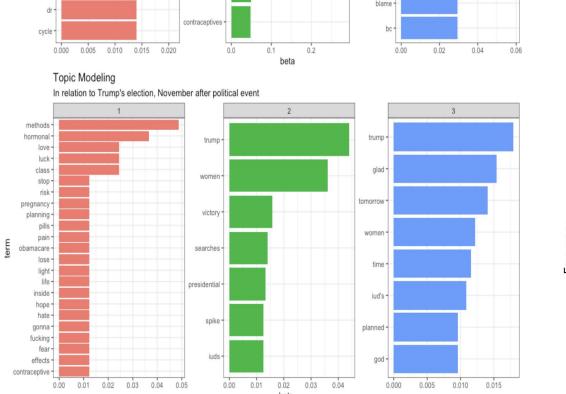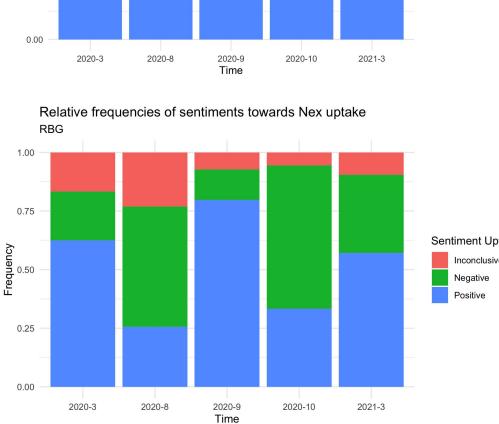
Next steps

Training Model: Train model with randomized set of 70% of hand-coding results to evaluate performance in accurately classifying tweets with remaining 30%

Future Research Question: Does social media discussion following political events translate to "action"? How, to what extent, and in what direction?