Understanding Public Perception of COVID-19 Social Distancing on Twitter







Sameh N Christoph U. Richard J. Saleh, MD^{1,2} Lehmann, MD^{2,3} Medford, MD^{2,4} ¹Department of Internal Medicine, UT Southwestern Medical Center, Dallas, TX, USA ²Clinical Informatics Center, UT Southwestern Medical Center, Dallas, TX, USA ³Department of Pediatrics, UT Southwestern Medical Center, Dallas, TX, USA ⁴Division of Infectious Diseases, UT Southwestern Medical Center, Dallas, TX, USA

Background

- Community non-pharmaceutical interventions (NPIs) such as social distancing are important COVID-19 mitigation strategies.
- Public opinion, behavior and adherence to these NPIs are critical to their success.
- Public perception of social distancing on social media platforms like Twitter has not been described.

Methods

- **Design:** Retrospective cross-sectional
- Data Source: Twitter
- Data Parameters:
 - Non-pharmaceutical intervention related hashtags -#socialdistancing, #stayathome
 - English language
 - Representative subset of tweets using Twitter API
- **Timeline:** March 27th April 10th, 2020
- **Population:** 574,903 tweets from 347,142 unique users
- Approach: natural language processing machine learning
 - 1) Syntax Extraction



- Data cleaning
- Lemmatization
- "went" to "go"
- "distancing" to "distance"
- Word segmentation
- Keyword matching

2) Sentiment Analysis

Polarity or Valence





Twitter is a rich medium that can be leveraged to understand public perception about nonpharmaceutical interventions in real-time and target public health messaging based on user interest and emotion.

Table 1. Characteristics of Tweets and Twitter Users

Twitter s Tv

Hashtags Has link Has medi Mentions Includes Includes Is quoted Verified t User follo User post

Figure 1. Word cloud of top 300 words

#socialdistancing











Results

	All	#socialdistancing	#stayathome
	n = 574,903 (564,886	n = 264,254 (46.0%)	n = 332,075 (57.8%)
	unique)		
ource			
witter for iPhone	178,105 (31.0)	79,739 (30.2)	109,808 (33.1)
vitter for Android	158,071 (27.5)	54,043 (20.5)	104,747 (31.5)
Twitter Web App	108,826 (18.9)	51,482 (19.5)	61,802 (18.6)
Instagram	38,809 (6.8)	24,938 (9.4)	15,147 (4.6)
	3 (2 – 5)	3 (2 – 5)	3 (2 – 5)
	207,465 (36.1)	106,428 (40.3)	108,914 (32.8)
a	272,005 (47.3)	131,758 (49.8)	151,429 (45.5)
user	193,392 (33.6)	89,623 (33.9)	109,197 (32.9)
place	52,020 (9.0)	25,333 (9.6)	28,626 (8.6)
geolocation	12,639 (2.2)	8,337 (3.2)	4,677 (1.4)
text	51,422 (8.9)	20,797 (7.9)	32,630 (9.8)
twitter account	27,560 (5.0)	13,463 (5.1)	16,015 (4.8)
owers	437 (98 – 1971)	499 (113 – 2356)	392 (87 – 1714)
ts to date	3,026 (584 - 13,086)	3,455 (729 – 15,280)	2,668 (485 - 11,378)

Median (IQR) is presented for numerical variables and N (%) is presented for categorical variables.

Figure 2. Tweet emotions



Figure 3. Tweet A) subjectivity and B) sentiment polarity