Social Disparities in Oral Health in America amid the COVID-19 Pandemic

Yangxin Fan, [1] **Hanjia Lyu,** [2] **Jin Xiao,** [3] **Jiebo Luo** [2,*]

[1] Goergen Institute for Data Science, University of Rochester, Rochester, 14627, USA
[2] Department of Computer Science, University of Rochester, Rochester, 14627, USA
[3] Eastman Institute for Oral Health, University of Rochester, Rochester, 14620, USA
[*] Corresponding author: Jiebo Luo (jluo@cs.rochester.edu)

Abstract

We conduct a large-scale social media-based study of oral health during the COVID-19 pandemic based on tweets from 9,104 Twitter users across 26 states (with sufficient samples) in the United States for the period between November 12, 2020 and June 14, 2021. To better understand how discussions on different topics/oral diseases vary across the users, we acquire or infer demographic information of users and other characteristics, including age, gender, income, population density, poverty rate, health insurance coverage rate, community water fluoridation rate, and relative change in the number of daily confirmed COVID-19 cases based on retrieved information from user profiles. Women and younger adults (19-29) are more likely to talk about oral health problems. We use the LDA topic model to extract the major topics/oral diseases in tweets. Overall, 26.70% of the Twitter users talk about wisdom tooth pain/jaw hurt, 23.86% tweet about dental service/cavity, 18.97% discuss chipped tooth/tooth break, 16.23% talk about dental pain, and the rest are about tooth decay/gum bleeding. By conducting logistic regression, we find that discussions vary across user characteristics. More importantly, we find social disparities in oral health during the pandemic. Specifically, we find that health insurance coverage rate is the most significant predictor in logistic regression for topic prediction. People from counties with higher insurance coverage tend to tweet less about all topics of oral diseases. People from counties at a higher risk of COVID-19 talk more about tooth decay/gum bleeding and chipped tooth/tooth break. Older adults ($\geq$ 50), who are vulnerable to COVID-19, are more likely to discuss dental pain. To our best knowledge, this is the first large-scale social media-based study to analyze and understand oral health in America amid the COVID-19 pandemic. We hope the findings of our study through the lens of social media can provide insights for oral health practitioners and policy makers.

Introduction

After the World Health Organization declared the global spread of COVID-19 a pandemic on March 11, 2020,[1] lockdowns were enforced nationwide in the US to reduce the spread of the virus. At the early outbreak of the COVID-19 pandemic, the American Dental Association (ADA) recommended that dental practices postpone elective procedure and provide emergency-only dental services.[2] As a result, patients' access to dental services have greatly decreased. During the week of March 23, 2020, an ADA Health Policy Survey indicated that 19% of dental offices were completely closed and 76% were partly closed but seeing emergency patients only (Association et al. 2020). More importantly, loss of dental insurance by many people has also increased the risk of oral diseases. According to a survey commissioned by the CareQuest Institute for Oral Health,[3] an estimated six million American adults have lost their dental insurance and 28 million American adults delayed getting dental care. Although most dental clinics reopened in June 2020, dental services have not rebounded to the full capacity due to office infection control regulation, lack of PPEs, and reduced patient-initiated dental visits.[4] Additionally, studies have found that there is an association between oral health and severity of COVID-19 complications that makes preventing bad oral health even more challenging (Sampson, Kamona, and Sampson 2020; Botros, Iyer, and Ojcius 2020). An indirect connection has also been suggested that due to the work from home (WFH) policy, people increase consumption of products that are likely to be detrimental to dental health (e.g., alcohol, sweets) but also increase the use of products which benefit oral health.[5]

Given the emergency of dental care caused by lack of access to dental services and loss of dental insurance, we attempt to identify the vulnerable groups of people by analyzing factors like age, gender, population density, income, poverty rate, health insurance coverage, community water fluoridation, as well as relative change in the number of daily confirmed COVID-19 cases. In addition, we intend to explore what kinds of oral diseases or issues to which they are

[1] https://pubmed.ncbi.nlm.nih.gov/32191675/ [Accessed September 8, 2021]

[2] https://www.ada.org/en/publications/ada-news/2020-archive/march/ada-recommending-dentists-postpone-elective-procedures [Accessed September 1, 2021]

[3] https://www.carequest.org/system/files/CareQuest-Institute-Coming-Surge-Oral-Health-Treatment-Needs-Report-1.pdf [Accessed September 3, 2021]

[4] https://www.ada.org/en/science-research/health-policy-institute/covid-19-dentists-economic-impact/survey-results [Accessed September 15, 2021]

[5] https://bridgedental.com/2020/04/26/the-impact-of-working-from-home-during-the-coronavirus-lockdown-on-dental-health/ [Accessed September 15, 2021]

more likely to be exposed during the COVID-19 pandemic. To our best knowledge, this is the first large-scale social media-based study to analyze and understand oral health in America amid the COVID-19 pandemic. We hope our study through the lens of social media, especially the findings of social disparities, can provide insights for oral health practitioners and policy makers.

Related Work

Oral health services in the US face an unprecedented challenge during the COVID-19 outbreak. On one hand, the COVID-19 pandemic increased the risk for oral diseases of whom are vulnerable to COVID-19, including those in rural areas, low socioeconomic groups, older adults, disadvantaged and underprivileged children, and the uninsured (Kalash 2020; Marchini and Ettinger 2020; Weintraub 2020). On the other hand, complex effect from intensified COVID-19 therapies and multi-drug treatment could possibly further exacerbate some oral conditions (Dziedzic and Wojtyczka 2021). COVID-19 also has direct effects on oral health through its official symptom ageusia (Daly and Black 2020). Despite of the fact that COVID-19 testing positivity rates were low among practicing US dentists (Estrich et al. 2020), the fear of contacting the virus may still lead to resistance to dental treatment, which in turn will increase of the level of dental anxiety (González-Olmo et al. 2020).

Although the COVID-19 pandemic has greatly impacted people's oral health over the last year, as an innovative way of disease diagnosis, telemedicine has gained public attention since it has the potential to provide the patients with the clinical care they need while remaining the distance (Hollander and Carr 2020). An example of telemedicine for oral health is to utilize the instant text and image messaging functions from social media platforms to diagnose and counsel for oral diseases in the COVID-19 era (Machado et al. 2020). Many large-scale social media-based studies have investigated different public health topics amid the COVID-19 pandemic, such as acquiring insights about the US mental health during the COVID-19 pandemic from Twitter data (Valdez et al. 2020), studying the natural and diffusion of COVID-19 related oral health information using tweets from Chinese social media Weibo (Tao et al. 2020), monitoring depression trends on Twitter during the COVID-19 pandemic (Zhou et al. 2021; Zhang et al. 2021), tracking mental health (Guntuku et al. 2020; Valdez et al. 2020), and investigating public opinions on COVID-19 vaccines (Lyu et al. 2021; Bonnevie et al. 2021; Wu, Lyu, and Luo 2021).

Study Objectives

Twitter has been a popular social media platform for people in the US to express their views and share their lives with each other. As of July 2021, there are about 73 million Twitter users in the US.[6] In this study, we intend to understand online discussions on oral health during the COVID-19 pandemic. We conduct a large-scale social media-based study of 9,104 Twitter users across 26 states (with sufficient samples) in the US for the period between November 12, 2020 to June 14, 2021. We collect our data using Tweepy[7] and acquire or infer user characteristics based on the information of Twitter users. Particularly, our study aims to answer the following research questions:

- **RQ1:** What are the major topics/oral diseases discussed in oral health-related tweets?
- **RQ2:** How does discussion of each type of topic/oral disease vary across user characteristics including age, gender, population density, income and poverty rate?
- **RQ3:** How does health insurance coverage rate, relative change in the number of daily confirmed COVID-19 cases, and community water fluoridation rate influence users' probability of tweeting about different topics/oral diseases?

To summarize, we make the following three major contributions:

- By applying Latent Dirichlet Allocation (LDA) topic modeling (Blei, Ng, and Jordan 2003), our study discovers five major topics/oral diseases, including dental pain, dental service/cavity, tooth decay/gum bleeding, wisdom tooth pain/jaw hurt, and chipped tooth/tooth break.
- By conducting multiple logistic regression analyses, we find that discussions of topics/diseases vary across user demographics.
- Our analyses show social disparities in oral health that people from the counties with higher health insurance coverage rate tend to tweet less about oral diseases in general and people from counties at a higher risk of COVID-19 tend to tweet less about dental service/cavity but more about oral diseases like tooth decay/gum bleeding and chipped tooth/tooth break. Older people mention dental pain more frequently.

Methods

In this section, we introduce the data collection and the methods we use in our analyses. To address RQ1, we discuss how we extract topics using Topic Modeling. To investigate RQ2 and RQ3, we describe how we infer the user characteristics and conduct the logistic regressions.

Data Collection and Preprocessing

We collect oral health-related tweets through Tweepy using a list of keywords including "tooth decay", "cavity", "black hole", "food stuck on teeth", "gums bleeding", "gums red", "gums inflammation", "face swelling", "cheek swelling", "drain in my mouth", "tongue swelling", "cannot swallow", "tooth chipped", "tooth break", "pain", "throbbing", "radiate to the ear", "jaw hurts", "can't open the mouth", and "wisdom tooth hurts". However, simply using keyword search may collect many false positive tweets. In particular, the tweets with only "pain", "black hole", "cavity", or "throbbing" may not be related to dental health. Therefore,

[6] https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ [Accessed August 28, 2021]

[7] https://www.tweepy.org/ [Accessed August 26, 2021]

we remove this kind of tweets by adding one constraint. If the tweet only contains one of the keywords "pain", "black hole", "cavity", or "throbbing" but does not contain any other keywords from the aforementioned keyword list, this tweet is excluded from our dataset. To validate this method, we randomly sample 1,000 tweets, read them and examine if they are related to our study. Using our method, 1.7% are labeled as relevant and 94.1% of them are indeed related to dental health discussions. Of the 98.3% tweets that are labeled as irrelevant, none of them are actually related to dental health discussions. The validation indicates a good performance of our excluding criteria. In addition, since our study focuses on understanding the online discussions of US Twitter users, we exclude the tweets that were not posted by the users whose profile indicates a US location. After removing the irrelevant tweets, the dataset is composed of 21,677 tweets for the period between November 12, 2020 to June 14, 2021 tweeted by 15,133 unique users.

Feature Inference

Age and Gender Following the methods used in Lyu et al. (2020), we use the Face++ API[8] to infer the age and gender information of Twitter users based on their profile images. There may be multiple faces in one profile image. To achieve the most robust inference of the demographic information of the Twitter users, we only include users with one intelligible face. In addition, the invalid image URLs are removed. Face++ is one of the most robust image-based inference method with respect to age and gender inference (Jung et al. 2018).

Age is binned into five groups: ≤ 18, 19-29, 30-49, 50-64, and ≥ 65. Users who are younger than 18 years old, between 19 to 29, between 30 to 49, and between 50 to 64 years old account for 1.6%, 48.6%, 37.3%, and 9.3%, respectively. The rest are over or equal to 65 years old. According to a report from the Pew Research Center (Wojcik and Hughes 2019), among the US adult Twitter users, 29% are between 18-29 years old, 44% are between 30-49 years old, 19% are between 50-64 years old, and 8% are over or equal to 65 years old. Compared to the age distribution of general Twitter users, there are proportionally more adults between 19 to 29 years old in our dataset. This is *consistent* with the finding of a household survey[9] that younger adults are most likely to report problems regarding the condition of their mouth and teeth. With respect to gender, as of January 2021, the gender distribution of Twitter users in the US is biased towards men who account for 61.6% of total users.[10] However, in our dataset, 57.4% users are women. Women tend to tweet about dental health more actively which echos the previous study that women are more dentally anxious (Doerr et al. 1998).

Population Density We apply a Python package - USZipcode search engine to extract population density of each Twitter user's location based on their profile information that were reported by themselves. The population density is categorized into three levels: urban (greater than 3,000 people per square mile), suburban (1,000-3,000 people per square mile), and rural (less than 1,000 people per square mile). In our study population, 72.0% are urban, 12.1% are suburban, and 15.9% are rural which is similar to the share reported in a previous report of the Pew Research Center[11] that most Twitter users live in urban area.

Income and Poverty Rate Studies have shown that income or poverty rate is strongly associated with oral cancer, dental caries prevalence, caries experience, and traumatic dental injuries (Singh, Peres, and Watt 2019). Therefore, we propose to include the socioeconomic status of the Twitter users into our study. Specifically, we use the Census API[12] to retrieve median per capita income and poverty rate at the county level from the 2019 American Community Survey (ACS).

Health Insurance Coverage Rate To our best knowledge, there is no publicly disclosed detailed information about dental insurance coverage rate at the county or city level. However, Pérez-Núñez et al. (2006) found that having medical insurance is positively correlated with better dental care coverage. Thus, we choose to use the health insurance coverage rate to approximately measure each user's accessibility to dental services. We use the Census API to retrieve health insurance coverage at the county level from the 2019 American Community Survey (ACS).

Fluoridation Rate Studies have shown that drinking fluoridated water can keep teeth strong and reduce cavities by about 25% in children and adults.[13] To better understand how community water fluoridation rate influence the dental health topics that people usually tweet about, we use the latest state-level fluoridation statistics from the Centers of Disease Control and Prevention to approximate the fluoridation rate of water that users use and drink.[14]

Pandemic Severity To measure the pandemic severity, we calculate the county-level 7-day average relative change in the number of daily COVID-19 confirmed cases. The data is collected from the data repository maintained by the Center for Systems Science and Engineering (CSSE) at John Hopkins University (Dong, Du, and Gardner 2020).

In the end, after inferring or extracting gender, age, population density, income, poverty rate, health insurance coverage rate, community fluoridation rate, and relative change of the number of daily confirmed COVID-19 cases and retaining states having at least 100 unique users, our final dataset

[8]https://www.faceplusplus.com/ [Accessed August 28, 2021]

[9]https://www.ada.org/~/media/ADA/Science%20and%20Research/HPI/OralHealthWell-Being-StateFacts/US-Oral-Health-Well-Being.pdf [Accessed September 9, 2021]

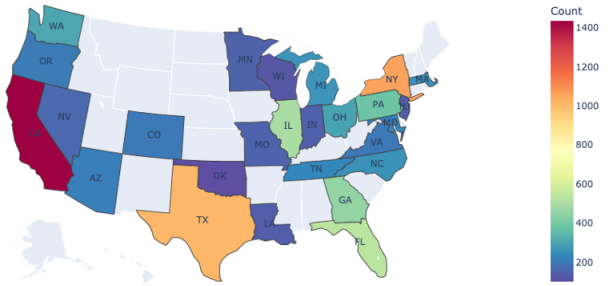
[10]https://www.statista.com/statistics/678794/united-states-twitter-gender-distribution/ [Accessed September 8, 2021]

[11]https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/ [Accessed September 9, 2021]

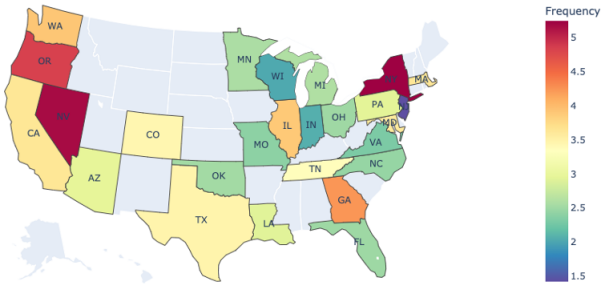
[12]https://www.census.gov/data/developers.html [Accessed September 8, 2021]

[13]https://www.cdc.gov/fluoridation/index.html [Accessed September 6, 2021]

[14]https://www.cdc.gov/fluoridation/statistics/2018stats.htm/ [Accessed September 11, 2021]



(a) Number of Twitter users by state.



(b) Number of Twitter users per 10,000 of total population.

Figure 1: State-level user distributions.

consists of 10,883 tweets posted by 9,104 Twitter users with all inferred features included. Figure 1a shows the geographic distribution of Twitter users in our study. California, Texas, and New York are the top three states with regard to the number of users who tweet about oral health. However, this could be because that these states are most active in general.[15] Figure 1b shows that New York, Nevada, and Oregon tend to have higher relative frequency of users who tweet about oral health. It is noteworthy that Oregon has the highest dental care utilization among adults with private dental benefits.[16] Figure 2 illustrates the trend of daily frequency of unique tweets. Apart from the big "down" and "up" between early January 2021 and mid-March 2021 which roughly correspond to the trend of daily COVID-19 confirmed cases in the US (Dong, Du, and Gardner 2020), the daily tweet activity stays relatively stable and varies mostly between 20 and 60 tweets per day.

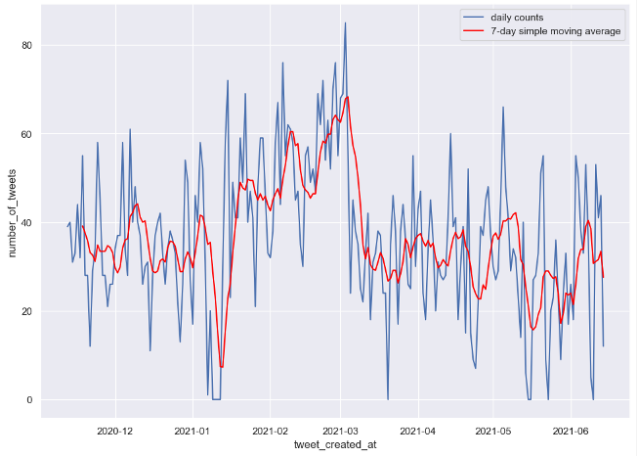


Figure 2: Trend of tweet activity (unique tweets).

[15]https://www.allbusiness.com/twitter-ranking-which-states-twitter-the-most-12329567-1.html [Accessed September 15, 2021]

[16]http://www.ada.org/~/media/ADA/Science%20and%20Research/HPI/OralHealthCare-StateFacts/Oral-Health-Care-System-Full-Report.pdf [Accessed September 15, 2021]

Topic Modeling

We use LDA (Blei, Ng, and Jordan 2003) to extract topics from tweets. We remove Twitter handles, hashtags, links, punctuation, numbers, special characters, and stop words to clean the texts of our tweets. We use the spaCy package to only keep the words whose postag is either "NOUN", "ADJ", "VERB", or "ADV". We apply grid search to find the optimal hyperparameters setting for num_topics, $alpha$, and $beta$. The optimal setting is as follows: num_topics = 5, $alpha$ = 0.01, and $beta$ = 0.41, with a coherence score of 0.53.

Results

Topic Analysis

To address RQ1, we attempt to capture what topics or oral diseases are discussed when Twitter users post about oral health. Table 1 shows the five topics extracted by the LDA topic modeling. We assign the title of each topic based on its top ten keywords. In this section, tweets are the main study objects.

We visualize the relative frequency of keywords in each topic in Figure 3. The differences of the top keywords between topics suggest that LDA has captured the major component of each topic, and there is a good separation. LDA calculates the weights of five topics of each tweet. The topic that has the highest weight is considered as the dominant topic of the tweet. We group the tweets into five classes based on their dominant topic. Figure 4 shows the proportions of the topics of tweets. As indicated by the previous studies, people increase consumption of products that are detrimental to oral health during work from home.[17] To investigate such effects, we construct a list of keywords of snacks and alcohol[18] and perform a keyword search in all

[17]https://www.myfooddata.com/articles/high-sugar-foods.php [Accessed September 15, 2021]

[18]"drunk", "liquor", "beer", "champagne", "wine", "gin", "vodka", "rum", "whisky", "brandy", "tequila", "bourbon", "sweet food", "sugar", "candy", "spaghetti sauce", "sports drinks", "chocolate milk", "granola", "honey", "glucose", "corn sugar", "milkshakes", "juice", "cream soda", "cake", "cereal", "chocolate", "honey", "milk", "yogurt", "ice cream", "cookie", "dried sweetened mango", "candied tamarind".

Table 1: Top 10 keywords of each topic.

Topic	Keywords
(1) Dental pain	pain, dental, work, good, oral, surgery, care, start, lot, people
(2) Dental service/cavity	dentist, cavity, time, make, year, dentistry, ago, fill, month, call
(3) Tooth decay/gum bleeding	tooth, gum, decay, eat, bleed, food, brush, hard, stop, floss
(4) Wisdom tooth pain/jaw hurt	pain, wisdom, bad, tooth, hurt, day, jaw, mouth, week, back
(5) Chipped tooth/tooth break	tooth, chip, break, feel, today, pull, give, front, fix, toothache

Table 2: Topic distribution by user characteristics.

Variable	Dental pain (1)	Dental service /cavity (2)	Tooth decay /gum bleeding (3)	Wisdom tooth pain /jaw hurt (4)	Chipped tooth /tooth break (5)
Total	16.23%	23.86%	14.24%	26.70%	18.97%
Urban	16.01%	24.68%	13.77%	26.66%	18.88%
Suburban	17.00%	24.23%	16.37%	24.50%	17.90%
Rural	16.67%	19.85%	14.73%	28.56%	20.19%
Age $\leq$ 18	10.74%	24.16%	14.09%	34.90%	16.11%
Age 19-29	12.83%	26.99%	11.38%	28.12%	20.67%
Age 30-49	18.12%	21.30%	15.09%	27.21%	18.27%
Age 50-64	23.38%	19.27%	22.68%	19.74%	14.92%
Age $\geq$ 65	28.42%	19.06%	23.38%	14.75%	14.39%
Male	16.34%	20.13%	16.73%	27.42%	19.38%
Female	16.16%	26.63%	12.39%	26.17%	18.66%

the tweets to examine whether or not the tweets mention sweet snacks/drinks or alcohol. Overall, 1.7% tweets mention sweets or alcohol. In particular, the tweets mentioning sweets or alcohol account for 1.0%, 0.4%, 4.2%, 0.7% and 1.8% for the topic of Dental pain, Dental service/cavity, Tooth decay/gum bleeding, Wisdom tooth pain/jaw hurt and Chipped tooth/tooth break, respectively. A higher proportion of tweets mentioning sweets or alcohol is observed in Tooth decay/gum bleeding and Chipped tooth/tooth break. For each user, we assign the topic that the user tweets most frequently as the dominant topic. The proportions of the topics of users are as follows:

Topic 1 (Dental pain) accounts for 16.23% of all Twitter users. In these tweets, people often mention pains caused by dental diseases or infections. An example tweet is:

`I've been dealing with severe dental pain for the past like 10 years. I can handle it but it suuuucks. Compounded that I just had another molar yanked out on Monday morning.`

Topic 2 (Dental service/cavity) represents 23.86% of all Twitter users in our study, where people mainly share their experience with dentists to fix their dental problems and/or talk about specific oral disease like cavity. An example tweet is:

`Shout out to Smile Studio in Zachary. I hate going to the dentist, and I am super-sensitive to pain. 44 years without a cavity and I had to have an`

Figure 3: Word cloud of the five topics extracted by LDA model.

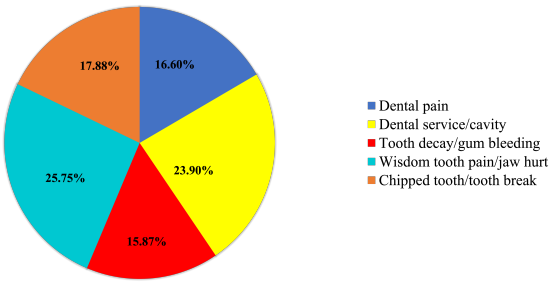


Figure 4: Topics distribution of all tweets (including retweets). Best viewed by zoom-in on screen.

extraction today. They made it almost painless.

Topic 3 (Tooth decay/gum bleeding) constitutes 14.24% of all users, including the keywords "tooth", "gum", "decay", "eat", and "bleed". In this topic, Twitter users mainly talk about two of the most common oral diseases - tooth decay and gum bleeding. An example tweet is:

My fake tooth chipped off on Xmas one month after losing dental so I'm ugly now 29. 62% of people who brush their teeth rinse their mouth out with water, which actually makes tooth decay more likely.

Topic 4 (Wisdom tooth pain/jaw hurt) is the most tweeted topic in our study, which accounts for 26.70% of all users. People mostly post about pains from wisdom tooth or jaw hurt. It contains keywords "chip", "front", "dog", "miss", and "walk". An example tweet is:

Yooo I don't wish the wisdom tooth pain not even on my worse enemy. Shit is wild Painful tooth Can't get straight to your dentist $<$ *hashtag* $>$ $<$ *hashtag* $>$.

Topic 5 (Chipped tooth/tooth break) represents 18.97% of all users, where people mostly talk about chipped tooth/teeth or break their own tooth/teeth. The keywords include "tooth", "chip", "break", "feel", and "today". An example tweet is:

I'm gonna get my chipped front tooth fixed tomorrow. It was broken when my dad smacked me as a kid. His wedding ring. He regretted it. I probably deserved it. I've been getting it fixed for 30 years. God, I miss him.

Logistic Regression Results

To answer RQ2, we conduct multiple logistic regression analyses to examine how each variable including age, gender, population density, income and poverty rate influence whether or not a user will tweet about a specific topic. In this section, users are the main study objects. 9,104 unique users are included in this analysis. Table 3 lists the logistic regression results for different topics against variables of interest. Each column represents a logistic regression model. Figure 5 illustrates the correlations between the variables for logistic regression analyses.

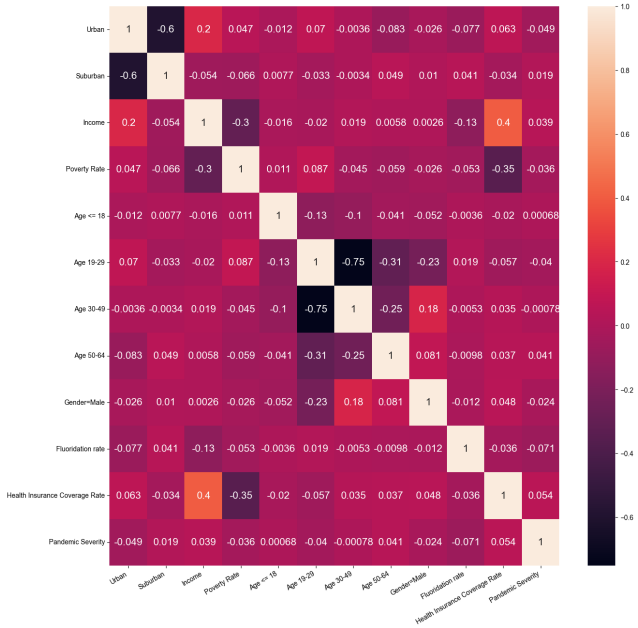


Figure 5: Correlation heat map of the variables for logistic regression (best viewed by zoom-in on screen).

Older adults tend to tweet more about dental pain but less about wisdom tooth pain/jaw hurt. In logistic regression analysis, we divide age into five groups: $\leq$ 18, 19-29, 30-49, 50-64, and $\geq$ 65. As for the topic of tooth pain/dental services, logistic regression coefficients for the first three age groups are all negative, which means adults in their 50s or above are more likely to talk about dental pain in general ($p < .001$). As for the topic of wisdom tooth pain/jaw hurt, logistic regression coefficients for the first three groups are positive, which means adults in their 50s or above are less likely to talk about wisdom tooth pain/jaw hurt ($p < .001$).

Women tend to talk more about dental pain and dental service/cavity while men tweet more tooth decay/gum bleeding and wisdom tooth pain/jaw hurt. We find women are more likely to talk about dental pain ($B = -0.1291, SE = 0.060, p < .05, OR = 0.8789, 95\%CI = [0.7819, 0.9881]$) and dental service/cavity ($B = -0.3082, SE = 0.053, p < .001, OR = 0.7402, 95\%CI = [0.6630, 0.8146]$). Men are more likely to talk about tooth decay/gum bleeding ($B = 0.2700, SE = 0.062, p < .001, OR = 1.3100, 95\%CI = [1.1595, 1.4800]$) and wisdom tooth pain/jaw hurt ($B = 0.1077, SE = 0.050, p < .05, OR = 1.1137, 95\%CI = [1.0101, 1.2275]$).

People from rural areas are less likely to discuss dental service/cavity and people from suburban areas are less likely to talk about wisdom tooth pain/jaw hurt.

Table 3: Logistic regression outputs for all topics against variables of interest.

Variable	Dental pain (1)	Dental service /cavity (2)	Tooth decay /gum bleeding (3)	Wisdom tooth pain /jaw hurt (4)	Chipped tooth /tooth break (5)
Urban (0=No, 1=Yes)	-0.0045	0.1550*	0.0334	-0.1278	-0.0588
	(0.081)	(0.074)	(0.086)	(0.067)	(0.075)
Suburban (0=No, 1=Yes)	-0.0001	0.2044*	0.1419	-0.2290*	-0.1435
	(0.108)	(0.097)	(0.111)	(0.092)	(0.102)
Income	0.0316	0.0979***	-0.0078	-0.0326	-0.0664*
	(0.027)	(0.022)	(0.030)	(0.024)	(0.028)
Poverty Rate	-0.0222**	0.0202***	-0.0138*	-0.0139**	-0.0048
	(0.007)	(0.005)	(0.007)	(0.005)	(0.006)
Age $\leq$ 18 (0=No, 1=Yes)	-1.2473***	0.0010	-0.4958	1.0641***	0.0609
	(0.296)	(0.240)	(0.275)	(0.235)	(0.275)
Age 19-29 (0=No, 1=Yes)	-1.0320***	0.1600	-0.7450***	0.7446***	0.3788*
	(0.139)	(0.149)	(0.149)	(0.163)	(0.165)
Age 30-49 (0=No, 1=Yes)	-0.6085***	-0.0605	-0.4925**	0.6735***	0.2028
	(0.138)	(0.150)	(0.148)	(0.164)	(0.166)
Age 50-64 (0=No, 1=Yes)	-0.3000	-0.1474	-0.0189	0.2632	-0.0519
	(0.154)	(0.168)	(0.162)	(0.181)	(0.187)
Male (0=No, 1=Yes)	-0.1291*	-0.3082***	0.2700***	0.1077*	0.1083
	(0.060)	(0.053)	(0.062)	(0.050)	(0.056)
Fluoridation Rate	-0.0010	-0.0005	-0.0008	-0.0014	0.0005
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
Health Insurance Coverage Rate	-0.0071**	-0.0184***	-0.0147***	-0.0119***	-0.0169***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Pandemic Severity	0.0465	-0.1730***	0.2824***	-0.2760***	0.1718***
	(0.049)	(0.048)	(0.048)	(0.047)	(0.045)
N			9,104		

Note. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$. Table entries are coefficients (standard errors). Income is scaled down by 10,000 while poverty rate, fluoridation rate, pandemic severity, and health insurance coverage rate are scaled up by 100.

By conducting logistic regression, we find that the Twitter users who live in rural areas tweet less about dental service/cavity ($p < .05$) since the coefficients for urban and suburban are both positive in logistic regression analysis of the topic dental service/cavity. We find that suburban people tweet less about wisdom tooth pain/jaw hurt ($B = -0.2290, SE = 0.092, p < .05, OR = 0.7953, 95\%CI = [0.6643, 0.9522]$).

Higher-income people tend to talk more about dental service/cavity but less about chipped tooth/tooth break. If the income increases by 10,000, the odds of tweeting about dental service/cavity is 1.1029 times ($B = 0.0979, SE = 0.022, p < .001, OR = 1.1029, 95\%CI = [1.0555, 1.1526]$) and the odds of tweeting about chipped tooth/tooth break is 0.9358 times ($B = -0.0664, SE = 0.028, p < .05, OR = 0.9358, 95\%CI = [0.8869, 0.9881]$).

People from counties with higher poverty rate talk less about dental pain, tooth decay/gum bleeding, and wisdom tooth pain/jaw hurt and more about dental service/cavity If the poverty rate increase by 1%, the odds of talking about dental pain is 0.9780 times ($B = -0.0222, SE = 0.007, p < .01, OR = 0.9780, 95\%CI = [0.9656, 0.9910]$), the odds of talking about tooth decay/gum bleeding is 0.9863 times ($B = -0.0138, SE = 0.007, p < .05, OR = 0.9863, 95\%CI = [0.9734, 0.9995]$), the odds of talking about wisdom tooth pain/jaw hurt is 0.9862 times ($B = -0.0139, SE = 0.005, p < .01, OR = 0.9862, 95\%CI = [0.9763, 0.9960]$), and the odds of talking about dental service/cavity is 1.0204 times ($B = 0.0202, SE = 0.005, p < .001, OR = 1.0204, 95\%CI = [1.0101, 1.0315]$).

People from counties with higher health insurance coverage rate tend to tweet less about all oral health-related topics. If the health insurance coverage rate increases by 1%, the odds of tweeting about dental pain is 0.9929 times ($B = -0.0071, SE = 0.003, p < .01, OR = 0.9929, 95\%CI = [0.9881, 0.9980]$), dental service/cavity is 0.9817 times ($B = -0.0184, SE = 0.002, p < .001, OR = 0.9817, 95\%CI = [0.9773, 0.9861]$), tooth decay/gum bleeding is 0.9854 times ($B = -0.0147, SE = 0.003, p < .001, OR = 0.9854, 95\%CI = [0.9802, 0.9910]$), wisdom tooth pain/jaw hurt is 0.9882 times ($B = -0.0119, SE = 0.002, p < .001, OR = 0.9882, 95\%CI = [0.9831, 0.9930]$), and chipped tooth/tooth break is 0.9832 times

$(B = -0.0169, SE = 0.003, p < .001, OR = 0.9832, 95\%CI = [0.9782, 0.9881])$.

People from counties at a higher risk of COVID-19 talk less about dental service/cavity, wisdom tooth pain/jaw hurt and more about tooth decay/gum bleeding and chipped tooth/tooth break. If the 7-day average relative change of the number of daily COVID-19 confirmed cases grows by 1%, the odds of tweeting about dental service/cavity is 0.8411 times ($B = -0.1730, SE = 0.048, p < .001, OR = 0.8411, 95\%CI = [0.7657, 0.9240]$), wisdom tooth pain/jaw hurt is 0.7588 times ($B = -0.2760, SE = 0.047, p < .001, OR = 0.7588, 95\%CI = [0.6921, 0.8319]$), tooth decay/gum bleeding is 1.3263 times ($B = 0.2824, SE = 0.048, p < .001, OR = 1.3263, 95\%CI = [1.2080, 1.4564]$), and chipped tooth/tooth break is 1.1874 times ($B = 0.1718, SE = 0.045, p < .001, OR = 1.1874, 95\%CI = [1.0876, 1.2969]$).

Discussions and Conclusions

We have analyzed 10,883 tweets from 9,104 Twitter users across 26 states (with sufficient samples) during the period of November 12, 2020 to June 14, 2021. Among the Twitter users in our study, 26.70% talk about wisdom tooth pain/jaw hurt, 23.86% tweet about dental service/cavity, 18.97% discuss chipped tooth/tooth break, 16.23% talk about dental pain, and 14.24% talk about tooth decay/gum bleeding.

Overall, women are more likely to discuss oral health amid the COVID-19 pandemic. On one hand, this might be because men are more likely to ignore their dental health and visit dentists less frequently for disease prevention (Lipsky et al. 2021). On the other hand, studies (Doerr et al. 1998; Stouthard, Mellenbergh, and Hoogstraten 1993) show that women are more dentally anxious, which might lead to physiological, cognitive, behavioural, health, and social issues (Cohen, Fiske, and Newton 2000). COVID-19 has also changed people's work patterns as many companies encourage or require employees to work from home to prevent the spread of virus,[19] which was found to influence people's oral health by increasing the consumption of products that are detrimental to oral health such as snacks and alcohol and increasing the consumption of oral health products.[20] Another potential reason that women tend to talk about oral health amid the COVID-19 pandemic is that they are more likely to reduce work hours and spend more time on oral health since they can stay home longer (Collins et al. 2020; Xiong et al. 2021). With respect to age, younger adults (19-29) tend to tweet more often about oral health problems. This echoes the finding that younger adults experience a higher prevalence of dental fear and anxiety (DFA), high DFA, and severe DFA (Silveira et al. 2021).

We observe that the topics of interest vary across user characteristics including age, gender, population density, income, poverty rate, and health insurance coverage rate. Older adults, who are identified as the highest risk group for fatal COVID-19 clinical outcomes (Marchini and Ettinger 2020; Sharma 2021) are more likely to talk about dental pain ($p < .001$) and less likely to tweet about wisdom tooth pain/jaw hurt. It is noteworthy that due to the pandemic, older adults are facing lack of access to the oral health care (Wall, Vujicic, and Nasseh 2012). Women tend to focus more on dental pain ($p < .05$) and dental service/cavity ($p < .001$) while men are more interested in discussing tooth decay/gum bleeding ($p < .001$) and wisdom tooth pain/jaw hurt ($p < .05$). Studies (Furuta et al. 2011; Heft et al. 2007) show that women are almost twice as likely to have received a regular dental check-up and are more likely to report general fear of dental pain compared to men. People from rural areas are less likely to discuss dental service/cavity ($p < .05$) which is possibly due to the lack of access to dental care. People living in rural America have about 8% (children) to 10% (adults aged 18–64) less access to dental services compared with their urban counterparts in 2017.[21] Suburban people talk less about wisdom tooth pain/jaw hurt ($p < .05$). Higher-income people talk more about dental service/cavity ($p < .001$) but less about chipped tooth/tooth break ($p < .05$). High cost is the most significant reason for people not visiting dentists in the US. [22] Higher-income people may care less about the high cost and have more frequent dental services which suggests the disparities in oral health. People from counties with a higher poverty rate talk less about dental pain ($p < .01$), tooth decay/gum bleeding ($p < .05$), and wisdom tooth pain/jaw hurt ($p < .01$). Although Kim et al. (2017) suggests that Community Water Fluoridation is associated with lowering the risks of having certain oral diseases like dental caries, our study has shown that state-level fluoridation rate is not associated with the prediction of any oral health-related topics. This may be due to the limitation of not having publicly available more granular fluoridation rate data. Health insurance coverage rate is the most important predictor for the logistic regression for topic prediction. People from counties with a higher insurance coverage rate tend to tweet less about all topics of oral health ($p < .01$) which is *consistent* with the findings of Zivkovic et al. (2020) that health or more specifically dental insurance plays an important role in improving people's oral health conditions. With respect to the pandemic severity, people from counties at a higher risk of COVID-19 talk less about dental service/cavity ($p < .001$) and wisdom tooth pain/jaw hurt ($p < .001$) but more about tooth decay/gum bleeding ($p < .001$) and chipped tooth/tooth break ($p < .001$). On one hand, it is likely that people delay or avoid dental visits because of closure and reduced hours of dental care and the fear of being infected with the virus dur-

[19] https://www.gartner.com/en/newsroom/press-releases/2020-03-19-gartner-hr-survey-reveals-88--of-organizations-have-e [Accessed March 25, 2021]

[20] https://bridgedental.com/2020/04/26/the-impact-of-working-from-home-during-the-coronavirus-lockdown-on-dental-health/ [Accessed September 15, 2021]

[21] https://www.cdc.gov/nchs/data/hus/2018/037.pdf [Accessed September 12, 2021]

[22] https://www.ada.org/~/media/ADA/Science%20and%20Research/HPI/OralHealthWell-Being-StateFacts/US-Oral-Health-Well-Being.pdf [Accessed September 12, 2021]

ing their dental appointments (Brian and Weintraub 2020; WongLaura et al. 2020). On the other hand, COVID-19 has an negative effect on oral health possibly resulting from xerostomia, loss of taste or smell sensation, and mental health breakdown (Farook et al. 2020).

This is the first large-scale social media-based study to understand the public discussions on oral health during the COVID-19 pandemic in the US. We hope our work can promote research on public health issues through the lens of social media, provide insights for oral health practitioners and policy makers, enhance the public awareness of the importance of oral health, and ultimately improve oral health in America amid the COVID-19 pandemic. There are some limitations of our study. Some states with a smaller sample size are not included in the study population. Temporal changes have not been investigated. Future work can be directed to analyzing oral health-related discussions across multiple social media platforms to achieve broader and more comprehensive perspectives.

References

Association, A. D.; et al. 2020. HPI poll examines impact of COVID-19 on dental practices.

Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3: 993–1022.

Bonnevie, E.; Gallegos-Jeffrey, A.; Goldbarg, J.; Byrd, B.; and Smyser, J. 2021. Quantifying the rise of vaccine opposition on Twitter during the COVID-19 pandemic. *Journal of Communication in Healthcare* 14(1): 12–19.

Botros, N.; Iyer, P.; and Ojcius, D. M. 2020. Is there an association between oral health and severity of COVID-19 complications? *Biomedical Journal* 43(4): 325–327.

Brian, Z.; and Weintraub, J. A. 2020. Peer Reviewed: Oral Health and COVID-19: Increasing the Need for Prevention and Access. *Preventing Chronic Disease* 17.

Cohen, S.; Fiske, J.; and Newton, J. 2000. The impact of dental anxiety on daily living. *British Dental Journal* 189(7): 385–390.

Collins, C.; Landivar, L. C.; Ruppanner, L.; and Scarborough, W. J. 2020. COVID-19 and the gender gap in work hours. *Gender, Work & Organization* .

Daly, J.; and Black, E. A. 2020. The impact of COVID-19 on population oral health. *Community Dent Health* 37(4): 236–238.

Doerr, P. A.; Lang, W. P.; Nyquist, L. V.; and Ronis, D. L. 1998. Factors associated with dental anxiety. *The Journal of the American Dental Association* 129(8): 1111–1119.

Dong, E.; Du, H.; and Gardner, L. 2020. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* 20(5): 533–534.

Dziedzic, A.; and Wojtyczka, R. 2021. The impact of coronavirus infectious disease 19 (COVID-19) on oral health. *Oral Diseases* 27: 703–706.

Estrich, C. G.; Mikkelsen, M.; Morrissey, R.; Geisinger, M. L.; Ioannidou, E.; Vujicic, M.; and Araujo, M. W. 2020. Estimating COVID-19 prevalence and infection control practices among US dentists. *The Journal of the American Dental Association* 151(11): 815–824.

Farook, F. F.; Nuzaim, M. N. M.; Ababneh, K. T.; Alshammari, A.; and Alkadi, L. 2020. COVID-19 pandemic: oral health challenges and recommendations. *European Journal of Dentistry* .

Furuta, M.; Ekuni, D.; Irie, K.; Azuma, T.; Tomofuji, T.; Ogura, T.; and Morita, M. 2011. Sex differences in gingivitis relate to interaction of oral health behaviors in young people. *Journal of Periodontology* 82(4): 558–565.

González-Olmo, M. J.; Ortega-Martínez, A. R.; Delgado-Ramos, B.; Romero-Maroto, M.; and Carrillo-Diaz, M. 2020. Perceived vulnerability to Coronavirus infection: impact on dental practice. *Brazilian Oral Research* 34.

Guntuku, S. C.; Sherman, G.; Stokes, D. C.; Agarwal, A. K.; Seltzer, E.; Merchant, R. M.; and Ungar, L. H. 2020. Tracking mental health and symptom mentions on twitter during covid-19. *Journal of General Internal Medicine* 35(9): 2798–2800.

Heft, M. W.; Meng, X.; Bradley, M. M.; and Lang, P. J. 2007. Gender differences in reported dental fear and fear of dental pain. *Community Dentistry and Oral Epidemiology* 35(6): 421–428.

Hollander, J. E.; and Carr, B. G. 2020. Virtually perfect? Telemedicine for COVID-19. *New England Journal of Medicine* 382(18): 1679–1681.

Jung, S.-G.; An, J.; Kwak, H.; Salminen, J.; and Jansen, B. J. 2018. Assessing the accuracy of four popular face recognition tools for inferring gender, age, and race. In *Twelfth International AAAI Conference on Web and Social Media*.

Kalash, D. A. 2020. How COVID-19 deepens child oral health inequities. *Journal of the American Dental Association (1939)* 151(9): 643.

Kim, H.-N.; Kim, J.-H.; Kim, S.-Y.; and Kim, J.-B. 2017. Associations of community water fluoridation with caries prevalence and oral health inequality in children. *International Journal of Environmental Research and Public Health* 14(6): 631.

Lipsky, M. S.; Su, S.; Crespo, C. J.; and Hung, M. 2021. Men and Oral Health: A Review of Sex and Gender Differences. *American Journal of Men's Health* 15(3): 15579883211016361.

Lyu, H.; Chen, L.; Wang, Y.; and Luo, J. 2020. Sense and sensibility: Characterizing social media users regarding the use of controversial terms for covid-19. *IEEE Transactions on Big Data* .

Lyu, H.; Wang, J.; Wu, W.; Duong, V.; Zhang, X.; Dye, T. D.; and Luo, J. 2021. Social media study of public opinions on potential COVID-19 vaccines: informing dissent, disparities, and dissemination. *Intelligent Medicine* .

Machado, R. A.; de Souza, N. L.; Oliveira, R. M.; Júnior, H. M.; and Bonan, P. R. F. 2020. Social media and telemedicine for oral diagnosis and counselling in the COVID-19 era. *Oral Oncology* 105: 104685.

Marchini, L.; and Ettinger, R. L. 2020. COVID-19 pandemics and oral health care for older adults. *Special Care in Dentistry* 40(3): 329.

Pérez-Núñez, R.; Medina-Solis, C. E.; Maupomé, G.; and Vargas-Palacios, A. 2006. Factors associated with dental health care coverage in Mexico: Findings from the National Performance Evaluation Survey 2002–2003. *Community Dentistry and Oral Epidemiology* 34(5): 387–397.

Sampson, V.; Kamona, N.; and Sampson, A. 2020. Could there be a link between oral hygiene and the severity of SARS-CoV-2 infections? *British Dental Journal* 228(12): 971–975.

Sharma, A. 2021. Estimating older adult mortality from COVID-19. *The Journals of Gerontology: Series B* 76(3): e68–e74.

Silveira, E. R.; Cademartori, M. G.; Schuch, H. S.; Armfield, J. A.; and Demarco, F. F. 2021. Estimated prevalence of dental fear in adults: a systematic review and meta-analysis. *Journal of Dentistry* 103632.

Singh, A.; Peres, M.; and Watt, R. 2019. The relationship between income and oral health: a critical review. *Journal of Dental Research* 98(8): 853–860.

Stouthard, M. E.; Mellenbergh, G. J.; and Hoogstraten, J. 1993. Assessment of dental anxiety: a facet approach. *Anxiety, Stress and Coping* 6(2): 89–105.

Tao, Z.-Y.; Chu, G.; McGrath, C.; Hua, F.; Leung, Y. Y.; Yang, W.-F.; Su, Y.-X.; et al. 2020. Nature and diffusion of COVID-19–related oral health information on Chinese social media: analysis of tweets on weibo. *Journal of Medical Internet Research* 22(6): e19981.

Valdez, D.; Ten Thij, M.; Bathina, K.; Rutter, L. A.; and Bollen, J. 2020. Social media insights into US mental health during the COVID-19 pandemic: longitudinal analysis of twitter data. *Journal of Medical Internet Research* 22(12): e21418.

Wall, T. P.; Vujicic, M.; and Nasseh, K. 2012. Recent trends in the utilization of dental care in the United States. *Journal of Dental Education* 76(8): 1020–1027.

Weintraub, J. A. 2020. Oral Health and COVID-19: Increasing the Need for Prevention and Access (vol 17, E93, 2020). *Preventing Chronic Disease* 17.

Wojcik, S.; and Hughes, A. 2019. Sizing up Twitter users. *Pew Research Center* 24.

WongLaura, E.; HawkinsJessica, E.; MurrellKaren, L.; et al. 2020. Where are all the patients? Addressing Covid-19 fear to encourage sick patients to seek emergency care. *NEJM Catalyst Innovations in Care Delivery* .

Wu, W.; Lyu, H.; and Luo, J. 2021. Characterizing Discourse about COVID-19 Vaccines: A Reddit Version of the Pandemic Story. *arXiv preprint arXiv:2101.06321* .

Xiong, Z.; Li, P.; Lyu, H.; and Luo, J. 2021. Social Media Opinions on Working From Home in the United States During the COVID-19 Pandemic: Observational Study. *JMIR Medical Informatics* 9(7): e29195.

Zhang, Y.; Lyu, H.; Liu, Y.; Zhang, X.; Wang, Y.; Luo, J.; et al. 2021. Monitoring Depression Trends on Twitter During the COVID-19 Pandemic: Observational Study. *JMIR Infodemiology* 1(1): e26769.

Zhou, J.; Zogan, H.; Yang, S.; Jameel, S.; Xu, G.; and Chen, F. 2021. Detecting community depression dynamics due to covid-19 pandemic in australia. *IEEE Transactions on Computational Social Systems* .

Zivkovic, N.; Aldossri, M.; Gomaa, N.; Farmer, J. W.; Singhal, S.; Quiñonez, C.; and Ravaghi, V. 2020. Providing dental insurance can positively impact oral health outcomes in Ontario. *BMC Health Services Research* 20(1): 1–9.