

# Analyzing Topics and Sentiments from Twitter to Gain Insights to Refine Interventions for Family Caregivers of Persons with Alzheimer's Disease and Related Dementias (ADRD) During COVID-19 Pandemic

Sunmoo YOON<sup>a,b,1</sup>, Peter BROADWELL<sup>c</sup>, Carmela ALCANTARA<sup>d</sup>,  
Nicole DAVIS<sup>e</sup>, Haeyoung LEE<sup>f</sup>, Amanda Bristol<sup>a</sup>, Dante Tipiani<sup>a</sup>,  
Joo Young Nho<sup>g</sup> and Mary MITTELMAN<sup>g</sup>

<sup>a</sup> General Medicine, Department of Medicine, Columbia University, USA

<sup>b</sup> Data Science Institute, Columbia University, USA

<sup>c</sup> Center for Interdisciplinary Digital Research, Stanford University, USA

<sup>d</sup> School of Social Work, Columbia University, USA

<sup>e</sup> School of Nursing, Clemson University, USA

<sup>f</sup> Department of Nursing, Chung-Ang University, South Korea

<sup>g</sup> Department of Psychiatry, NYU Grossman School of Medicine, USA

**Abstract.** We randomly extracted Tweets mentioning dementia/Alzheimer's caregiving-related terms (n= 58,094) from Aug 23, 2019, to Sep 14, 2020, via an API. We applied a clustering algorithm and natural language processing (NLP) to publicly available English Tweets to detect topics and sentiment. We compared emotional valence scores of Tweets from before (through the end of 2019) and after the beginning of the COVID-19 pandemic (2020-). Prevalence of topics related to caregiver emotional distress (e.g., depression, helplessness, stigma, loneliness, elder abuse) and caregiver coping (e.g., resilience, love, reading books) increased, and topics related to late-stage dementia caregiving (e.g., nursing home placement, hospice, palliative care) decreased during the pandemic. The mean emotional valence score significantly decreased from 1.18 (SD 1.57; range -7.1 to 7.9) to 0.86 (SD 1.57; range -5.5 to 6.85) after the advent of COVID-19 (difference -0.32 CI: -0.35, -0.29). The application of topic modeling and sentiment analysis to streaming social media provides a foundation for research insights regarding mental health needs for family caregivers of a person with ADRD during COVID-19 pandemic.

**Keywords.** dementia caregiving, online intervention, disparities, topic modeling

## 1. Introduction

Alzheimer's disease and related dementias (ADRD) is the sixth leading cause of death in the United States. More than 16 million family members or friends usually provide care for people with ADRD in their homes every year [1]. Family members of persons with

<sup>1</sup> Corresponding Author, Sunmoo Yoon, General Medicine, Department of Medicine, Columbia University Irving Medical Center, 630W 168 Street, PH105, New York, NY, 10032, USA; E-mail: sy2102@cumc.columbia.edu.

ADRD have been physically and psychologically affected by COVID-19 during the pandemic, including loss of support and increased loneliness, as well as likelihood of contracting COVID-19 [1,2]. News media have described patients with ADRD in nursing homes as COVID-19's hidden victims. Although recent qualitative studies from the U.K. and Canada have reported on the burden of family caregivers of persons with ADRD during the COVID-19 pandemic [1,2], it remains unclear which aspects of caregiving gained greater prominence among the general public during the pandemic, and to what degree they were already present in the public consciousness before COVID-19 emerged.

Topic modeling applies statistical analyses to counts of words or word groups that co-occur within text documents to reveal latent word groupings with potential semantic significance across a collection of documents. Compared to manual reading and keyword/keyphrase extraction algorithms (TF-IDF, TextRank), topic modeling offers the ability to find semantically meaningful groupings of multiple words across very large document collections, even in cases when all of the related words do not appear together within any single document or section. Further, sentiment analysis helps quantify the affective state of social media users as calculated via methods such as the AFINN lexicon, proposing a total emotional valence score [2]. Therefore, the purpose of this study was to detect changes in caregiving topics and sentiments to understand emotional distress before and during the pandemic as a foundation for developing future Twitter-based interventions for Hispanic and African American dementia caregivers.

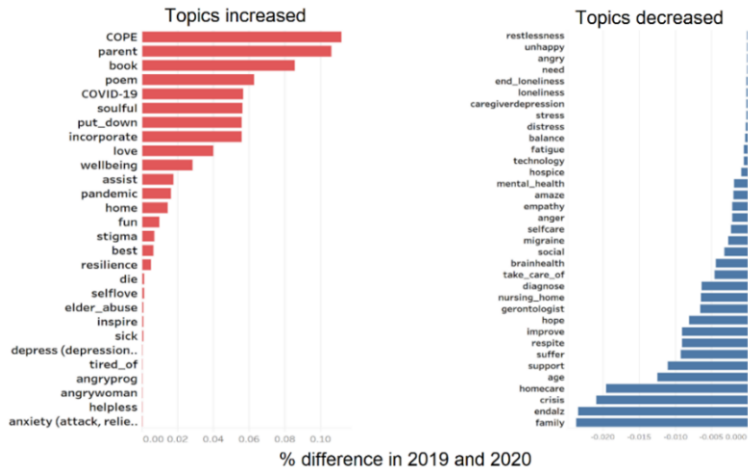
## 2. Methods

We applied topic modeling and sentiment analysis for content mining of a corpus of publicly available English Tweets mentioning dementia/Alzheimer's caregiving-related terms (keywords: Alzheimer's/dementia caregiving/caregiver/care,  $n=58,094$ ) from Aug 23, 2019, to Sep 14, 2020. Any politicians' names were excluded upon data collection to exclude irrelevant Tweets during corpora creation. We used the NCapture and ORA software packages and other scripts run on a High-Performance Computing Cluster (<https://cuit.columbia.edu/shared-research-computing-facility>) to collect and analyze the data. First, 127 XML files were merged using Perl regular expression commands and shell scripting on the Terremoto cluster. Second, we used NLP techniques to remove  $n$ -grams occurring fewer than three times, and applied the Newman clustering algorithm to group associated Tweets in the corpus (Newman modularity: pre-COVID-19: 0.420, during COVID-19: 0.523). Next, we compared emotional valence scores of Tweets from before (through the end of 2019) and after the beginning of the COVID-19 pandemic (2020-). Lastly, experts in dementia caregiving studies reviewed detected topics. The larger study was approved by the Institutional Review Board (IRB). Resources including analytic Python codes and de-identified data are available on GitHub and OSF.io (<https://osf.io/qruf3>).

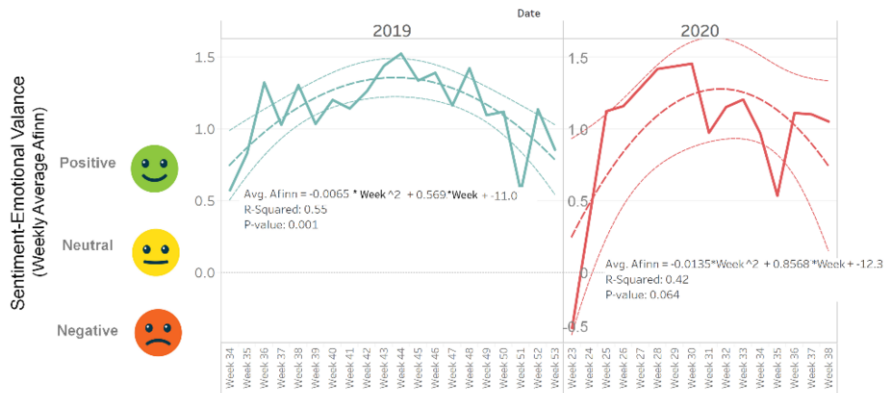
## 3. Results

A total of 58 topics were detected in the English Tweet corpus ( $n=58,094$ ), including 27 topics from 2019, prior to the emergence of COVID-19, and 31 in 2020 during the COVID-19 pandemic. The prevalence of topics related to caregiver emotional distress

(e.g., depression, helplessness, stigma, loneliness, elder abuse) and caregiver coping (e.g., resilience, love, reading books as in bibliotherapy) increased (Figure 1). Conversely, the prevalence of common topics related to late-stage dementia caregiving, such as nursing home placement, hospice, or palliative care, decreased during the pandemic (figure 1). The mean emotional valence score significantly decreased from 1.18 (SD 1.57) to 0.86 (SD 1.57) after the beginning of the COVID-19 pandemic in 2020 (difference -0.32 CI: -0.35, -0.29) (figure 2).



**Figure 1.** Changes in topic prevalence in English Tweets mentioning dementia caregiving in 2019 (before COVID-19) and 2020 (during the COVID-19 pandemic).



**Figure 2.** Weekly average emotional valence score of dementia caregiving English Tweets in 2019 (before COVID-19) and 2020 (during the COVID-19 pandemic).

4. Discussion and Conclusions

This study explored the changes in topics and emotional valence scores of English Tweets mentioning dementia caregiving before and during the COVID-19 pandemic. Consistent with the findings from other qualitative studies [1,2] regarding caregiving burdens among family caregivers in the U.K. and Canada, we found increased prevalence of family caregiver emotional distress such as depression, helplessness, and loneliness in

Tweets mentioning dementia caregiving during COVID-19 pandemic. In the meantime, we also found an increased prevalence of coping strategies and resilience topics in Tweets mentioning dementia caregiving. Historically it is not new to observe mentions of the specific coping strategy of reading books containing caregiving-related stories and poems to alleviate emotional distress, conflicts, and tensions in Tweets that circulate among dementia caregivers. Nevertheless, it was new to observe the actual terms of “resilience” and “cope,” rather than their proxy terms (e.g., reading caregiving stories) in Tweets mentioning dementia caregiving during the pandemic. Consistent with a recent qualitative study emphasizing the role of external encouragement and social support such as from the health system (e.g., a nurse or a social worker in the health system communicates with caregivers and provides support) during the stress of the pandemic [1], this study suggests that communities and organizations on Twitter have been spontaneously taking an active role to promote resilience and coping among dementia caregivers. Moreover, it is remarkable to find new topics on positive coping strategies and resilience (e.g., encouraging love, well-being, and improving spiritual health) in Tweets mentioning dementia caregiving. As considerable evidence has continuously shown the effectiveness of off-line spiritual support for family caregivers for dementia patients since 1985 [4], our finding adds to the knowledge that online social media has the potential to be a platform to promote positive coping strategies and resilience (e.g., love, spiritual support, well-being) at the societal and system level, rather than providing individual-level decision support for late-stage dementia caregiving, such as nursing home placement, hospice, or palliative care, as we found their prevalence decreased over the period considered in this study. In conclusion, the application of topic modeling and sentiment analysis to streaming social media provides caregiving topics and sentiments to understand emotional distress before and during the COVID-19 pandemic as a foundation for developing future Twitter-based interventions for Hispanic and African American dementia caregivers.

## Acknowledgments

U.S. federal grant TweetS2 R01AG060929 (P.I.: Yoon).

## References

- [1] Hanna K, Giebel C, Butchard S, Tetlow H, Ward K, Shenton J, Cannon J, Komuravelli A, Gaughan A, Eley R, Rogers C. Resilience and supporting people living with dementia during the time of COVID-19: A qualitative study. *Dementia*. 2021 Aug 4;14713012211036601.
- [2] Brown EE, Kumar S, Rajji TK, Pollock BG, Mulsant BH. Anticipating and mitigating the impact of the COVID-19 pandemic on Alzheimer's disease and related dementias. *The American Journal of Geriatric Psychiatry*. 2020 Jul 1;28(7):712-21.
- [3] Finn Årup Nielsen A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. Proceedings of the ESWC2011 Workshop on ‘Making Sense of Microposts’: Big things come in small packages 718 in *CEUR Workshop Proceedings* 93-98. 2011 May. <http://arxiv.org/abs/1103.2903>.
- [4] Wright SD, Pratt CC, Schmall VL. Spiritual support for caregivers of dementia patients. *Journal of Religion and Health*. 1985 Mar;24(1):31-8.