

# Applying Social Network Analysis to Compare Dementia Caregiving Networks on Twitter in Hispanic and Black Communities

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

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

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

<sup>c</sup> *School of Nursing, Clemson University, USA*

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

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

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

**Abstract.** We applied social network analysis (SNA) on Tweets to compare Hispanic and Black dementia caregiving networks. We randomly extracted Tweets mentioning dementia caregiving and related terms from corpora collected daily via the Twitter API from September 1 to December 31, 2019 (initial corpus:  $n = 2,742,539$  Tweets, random sample  $n = 549,380$  English Tweets,  $n = 185,684$  Spanish Tweets). After removing bot-generated Tweets, we first applied a lexicon-based demographic inference algorithm to automatically identify Tweets likely authored by Black and Hispanic individuals using Python ( $n = 114,511$  English,  $n = 1,185$  Spanish). Then, using ORA, we computed network measures at macro, meso, and micro levels and applied the Louvain clustering algorithm to detect groups within each Hispanic and Black caregiving network. Both networks contained a similar proportion of dyads and triads (Hispanic 88.2%, Black 88.9%), while the Black caregiving network included a slightly larger proportion of isolates (Hispanic 0.8%, Black 4.0%). This study provides useful baseline information on the composition of existing large groups and small groups. In addition, this work provides useful guidance for future recruitment strategies and the design of social support interventions regarding emotional needs for Hispanic and Black dementia caregivers.

**Keywords.** network analysis, social support, dementia, caregivers, social media

## 1. Introduction

The prevalence of Alzheimer's disease and related dementia (ADRD) is higher among Hispanics and Blacks than non-Hispanic Whites in the United States [1]. Among the proportion of people with dementia who live in the community rather than in assisted living facilities (70%-81%), most are cared for by family or friend caregivers [1]. A systematic review by Petosa and Smith found that peer social support is an effective contributor to meeting one's health goals [2]. Experimental evidence in behavioral science specifically highlights the critical role of dyad or triad friendship relations in

<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.

social support [3]. Although social scientists have reported a strong correlation between beneficial behavior and group cohesion in triad friendships since the mid 19<sup>th</sup> century, triad friendships and their associated benefits were rarely observed in previous social media-based studies of minority dementia caregivers [4].

Social network analysis is commonly used to assess social structures such as the composition of groups (e.g., large groups, small groups, triads, and dyads) and dynamics (e.g., relationships, links, and temporal interactions) using network and graph theory [4]. Social network analysis of online communities, particularly Twitter, may provide insights to aid the design of culturally sensitive and customized social support interventions for Hispanic and Black family caregivers for persons with dementia [4, 5]. This study aims to apply social network analysis to a large corpus of Tweets to gain insights about Hispanic and Black dementia caregiving networks.

## 2. Methods

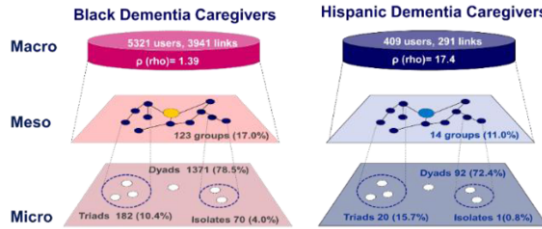
We applied network analysis techniques to publicly available Twitter messages (Tweets) mentioning dementia (keywords/hashtags: #dementia, #demencia, #Alzheimers). We used the Twitter Application Programming Interface (API), Python, and a High-Performance Computing (HPC) Cluster (<https://osf.io/qruf3/>) to extract these messages from a random sample of publicly available Tweets collected daily using via the Twitter API from September 1 to December 31, 2019 (n= 2,742,539 Tweets). Our team's behavioral domain expert and clinician first extracted and categorized U.S.-based Tweets from the initial corpus (n= 549,380 English Tweets, n= 185,684 Spanish Tweets). Second, we applied a Twitter bot detection algorithm to remove bot-generated Tweets using the Botometer Python API (<https://github.com/IUNetSci/botometer-python>). Third, we applied a lexicon-based demographic inference algorithm to automatically identify Tweets likely authored by Black or Hispanic individuals (n= 114,511 English, n = 1,185 Spanish, <https://osf.io/qruf3/>). Fourth, we applied Social Network Analysis to the winnowed Tweet corpus to automatically detect similar communities and to identify stakeholders with a high degree of network centrality. Fifth, to avoid algorithm dependency, we applied and compared the results of five clustering algorithms: Leiden, Principal Component Analysis, Louvain, CONCOR, and Newman. Sixth, network measures at the macro, meso, and micro levels were calculated to compare Black and Hispanic dementia caregiving networks within Twitter using the ORA meta-network analysis software. Seventh, we visualized Black and Hispanic dementia caregiving networks within Twitter at macro, meso, and micro levels. Lastly, behavioral science experts and a clinician evaluated the results according to their clinical meaningfulness for the design and recruitment strategy of a social support intervention for Hispanic and Black dementia caregivers on Twitter. The larger study was approved by the Institutional Review Board (IRB).

## 3. Results

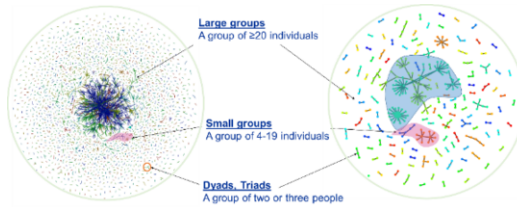
Dementia caregiving-related Tweets potentially created by 5,321 Black users were disseminated to 280,042,783 users, with retweets outnumbering original messages by a factor of 11. In contrast, the dementia caregiving-related Tweets in Spanish created by 409 users were disseminated to 6,141,424 users at a ratio of 11,798 retweets to each

original message. The top 10 stakeholders who disseminated dementia caregiving-related messages in English were identified as media outlets, a lay person, a psychologist, a social worker, and an association; among Spanish-language tweets, the top stakeholders were an individual, a media outlet, a lawyer, a community organization, and an economist. 123 groups (17.0%, Louvain modularity value: 0.802) were found in the Black dementia caregiving network, whereas 14 distinct groups (11.0%, Louvain modularity value: 0.801) were detected in the Hispanic caregiving network.

Both networks contained a similar proportion of dyads or triads (Hispanics: 88.2% dyads or triads, Blacks: 88.9%), while the Black caregiving network included a slightly larger proportion of isolates (Hispanics: 0.8%, Blacks: 4.0%) (Figure 1).



**Figure 1.** Characteristics of macro, meso, and micro level Twitter networks between Hispanic and Black dementia caregivers



**Figure 2.** Social network structure of dyads, triads, and communities of Black and Hispanic dementia caregiving networks from September to December 2021

#### 4. Discussion and Conclusions

This study examined the characteristics of macro, meso, and micro-level social network structures among Hispanic and Black dementia caregiving networks on Twitter. In the macro-level social networking structure, a substantially larger number of Black dementia caregiving network users were engaging in the conversation than in the Hispanic dementia caregiving network in the U.S. (5321 users, 3941 retweet activities vs. 409 users, 291 retweet activities) over three months. According to the U.S census in 2020, the total percentage of the population identifying as Black is smaller than the total number of Hispanics (13.4% vs. 18.5%).

Consistent with the existing literature on the strong presence and cohesiveness among Black communities in Twitter in the U.S., our findings confirms that cohesive communities among Black networks on Twitter are present in the context of dementia caregiving. Indeed, more community groups were observed in the Black dementia caregiving network than in the Hispanic dementia caregiving network in the U.S. (123 groups 17.0% vs. 14 groups, 11.0%) in the meso level social networking structure. Social cohesiveness, a determinant for population health (Healthy People 2030), has been referred to as the way in which individuals within a society are bound together by

attitudes and behaviors. Social cohesion and social support has been associated with decreased depression [6], stress and emotional burden among primary family caregivers of older adults [7].

A novel finding of this study is that the Hispanic dementia caregiving network included a very small proportion (0.8%) of isolates who express themselves without quoting others' messages. Not only is this number smaller than in the Black caregiving network (4.0%), but it is unusual in the context of social networks that develop around diseases such as HIV or COVID-19 on Twitter. Further, the social science-based concept of the optimal number of community members, which in these cases range from two to a hundred people, has not been explored in the context of providing social support for dementia caregivers in the Twitter space. Studies with an interventional design using Twitter are needed to investigate each community's optimal number, which then will enable interventions to deliver social support messages meaningfully and cost-effectively to racial/ethnic minority dementia caregivers.

The Louvain social network clustering algorithm was selected after comparing the utility of the five clustering algorithms listed above. The performance of several of the alternative clustering algorithms on our dataset was suboptimal for the following reasons: 1) Leiden: degree of clustering performance is not available; 2) Principal Component Analysis: results included one large dominant cluster, which is clinically less meaningful; 3) CONCOR: prone to subjective bias; 4) Newman: clustering modularity score was slightly lower than that computed via Louvain (0.801 vs. 0.802).

In conclusion, this study provides useful baseline information on the composition of existing large groups, small groups, and isolates among Hispanic and Black dementia caregivers on Twitter, which will help guide the design and future recruitment strategies of social support interventions involving community engagement and partnership in these minoritized communities.

## Acknowledgments

U.S. federal grant TweetS2 R01AG060929 (PI: Yoon).

## References

- [1] Alzheimer's Ass. Race, Ethnicity, and Alzheimer's in America. *Alzheimer's & Dementia*. 2021;17(3).
- [2] Petosa R, Smith LH. Peer mentoring for health behavior change: A systematic review. *American Journal of Health Education*. 2014;45(6):351-7.
- [3] Simmel G. *The sociology of Georg Simmel*. Wolff KH, editor: The Free Press; 1964.
- [4] Odium M, Hwayoung CH, Broadwell P, Davis, Patrao M, Schauer D, Bales ME, Alcantara C, Yoon S. Application of Topic Modeling to Tweets to Learn Insights on the African American Lived Experience of COVID-19. *Studies in health technology and informatics*. 2020;272:24.
- [5] Yoon S, Lucero R, Mittelman MS, Luchsinger JA, Bakken S. Mining Twitter to Inform the Design of Online Interventions for Hispanic Alzheimer's Disease and Related Dementias Caregivers. *Hispanic Health Care International*. 2020;18(3):138-43.
- [6] Mittelman MS, Roth DL, Coon DW, Haley WE. Sustained benefit of supportive intervention for depressive symptoms in Alzheimer's caregivers. *Am J Psychiatry*. 2004 May;161(5):850-6.PMID: 15121650.
- [7] Naganuma Y, Yamaoka K, Takahashi K. Relationship between social cohesion and the care burden of primary family caregivers in central Tokyo, Japan. *Health Sci Rep*. 2021;4(1):e238.