

Washington University School of Medicine

Digital Commons@Becker

Open Access Publications

12-15-2021

Utility of Facebook's social connectedness index in modeling COVID-19 spread: Exponential random graph modeling study

Beth Prusaczyk

Kathryn Pietka

Joshua M Landman

Douglas A Luke

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Original Paper

Utility of Facebook's Social Connectedness Index in Modeling COVID-19 Spread: Exponential Random Graph Modeling Study

Beth Prusaczyk^{1,2}, MSW, PhD; Kathryn Pietka³; Joshua M Landman^{1,4}, MS; Douglas A Luke⁵, PhD

¹Center for Population Health Informatics, Institute for Informatics, Washington University School of Medicine in St. Louis, Saint Louis, MO, United States

²Department of Medicine, Washington University School of Medicine in St. Louis, Saint Louis, MO, United States

³Truman State University, Kirksville, MO, United States

⁴Division of Computational and Data Sciences, Washington University in St. Louis, Saint Louis, MO, United States

⁵Center for Public Health Systems Science, Brown School, Washington University in St. Louis, Saint Louis, MO, United States

Corresponding Author:

Beth Prusaczyk, MSW, PhD

Center for Population Health Informatics

Institute for Informatics

Washington University School of Medicine in St. Louis

660 S. Euclid Avenue

Saint Louis, MO, 63110

United States

Phone: 1 314 330 0537

Email: beth.prusaczyk@wustl.edu

Abstract

Background: The COVID-19 (the disease caused by the SARS-CoV-2 virus) pandemic has underscored the need for additional data, tools, and methods that can be used to combat emerging and existing public health concerns. Since March 2020, there has been substantial interest in using social media data to both understand and intervene in the pandemic. Researchers from many disciplines have recently found a relationship between COVID-19 and a new data set from Facebook called the Social Connectedness Index (SCI).

Objective: Building off this work, we seek to use the SCI to examine how social similarity of Missouri counties could explain similarities of COVID-19 cases over time. Additionally, we aim to add to the body of literature on the utility of the SCI by using a novel modeling technique.

Methods: In September 2020, we conducted this cross-sectional study using publicly available data to test the association between the SCI and COVID-19 spread in Missouri using exponential random graph models, which model relational data, and the outcome variable must be binary, representing the presence or absence of a relationship. In our model, this was the presence or absence of a highly correlated COVID-19 case count trajectory between two given counties in Missouri. Covariates included each county's total population, percent rurality, and distance between each county pair.

Results: We found that all covariates were significantly associated with two counties having highly correlated COVID-19 case count trajectories. As the log of a county's total population increased, the odds of two counties having highly correlated COVID-19 case count trajectories increased by 66% (odds ratio [OR] 1.66, 95% CI 1.43-1.92). As the percent of a county classified as rural increased, the odds of two counties having highly correlated COVID-19 case count trajectories increased by 1% (OR 1.01, 95% CI 1.00-1.01). As the distance (in miles) between two counties increased, the odds of two counties having highly correlated COVID-19 case count trajectories decreased by 43% (OR 0.57, 95% CI 0.43-0.77). Lastly, as the log of the SCI between two Missouri counties increased, the odds of those two counties having highly correlated COVID-19 case count trajectories significantly increased by 17% (OR 1.17, 95% CI 1.09-1.26).

Conclusions: These results could suggest that two counties with a greater likelihood of sharing Facebook friendships means residents of those counties have a higher likelihood of sharing similar belief systems, in particular as they relate to COVID-19 and public health practices. Another possibility is that the SCI is picking up travel or movement data among county residents. This suggests the SCI is capturing a unique phenomenon relevant to COVID-19 and that it may be worth adding to other COVID-19

models. Additional research is needed to better understand what the SCI is capturing practically and what it means for public health policies and prevention practices.

(*JMIR Public Health Surveill* 2021;7(12):e33617) doi: [10.2196/33617](https://doi.org/10.2196/33617)

KEYWORDS

COVID-19; social media; social networks; network analysis; public health; utility; Facebook; connection; modeling; spread; United States; belief

Introduction

The COVID-19 (the disease caused by the virus SARS-CoV-2) pandemic has underscored the need for additional data, tools, and methods that can be used to combat emerging and existing public health concerns. Since March 2020, there has been substantial interest among researchers, public health professionals, infectious disease experts, and social media companies themselves in using social media data to both understand and intervene in the pandemic [1-7]. This is understandable given that nearly half of the world's population (49% or 3.8 billion people) are social media users, with as many as 7 in 10 Americans reporting using at least one social media site.

One early example of using social media for novel purposes related to the pandemic was done by economists with expertise in modeling geographic and social data, who used a relatively new data set from Facebook called the Social Connectedness Index (SCI) to understand the spread of COVID-19 in the emerging hot spots of Italy and Westchester, New York [8]. The SCI is a measure of the strength of connectedness between two geographic areas as measured by Facebook friendships [9,10]. The researchers found that the SCI was associated with confirmed COVID-19 cases after controlling for geographic distance to the two early hot spots as well as income and population density [8].

Other researchers with backgrounds in economics, engineering, and management have also explored the utility of this data set as it relates to COVID-19. One group of researchers found that households in counties with relatively stronger social connections to early hot spots in China and Italy (as measured by the SCI) were more likely to comply with stay-at-home orders [11]. Others found that public health prevention practices that people in a given region adopt are significantly influenced by the policies and behaviors of people in other regions with whom there is a relatively strong SCI [12]. In other words, even between distant regions, the SCI was associated with people in those two regions having similar COVID-19-related behaviors, suggesting people are influenced by their social connections.

Building off this work, we sought to use the SCI to examine how social similarity of Missouri counties could explain similarities of COVID-19 cases over time. Additionally, we aimed to add to the body of literature on the utility of the SCI by using a novel modeling technique that allows for the modeling of relational data [13]. To our knowledge, this technique has not been used with the SCI, which is a relational data set, thus making it a highly relevant and appropriate method.

Methods

Study Design

In this cross-sectional study, we analyzed publicly available data to test the association between the SCI and COVID-19 spread in Missouri using exponential random graph models (ERGMs). This study was reviewed by the institutional review board and deemed nonhuman participant research.

Data Sources

Social Connectedness Index

The SCI was obtained through the Facebook Data for Good program. The Facebook Data for Good program creates and makes available a variety of tools and data sets that are built from privacy-protected data from the Facebook platform and other publicly available data sources such as satellite imagery. Data sets in the program include the SCI, electrical distribution grid maps, the Inclusive Internet Index (a measure of internet accessibility), the Climate Change Survey, and more.

The SCI measures the relative probability of a Facebook friendship link between a given Facebook user in location A and a user in location B. It is calculated by dividing the number of Facebook friends between two locations divided by the number of Facebook users in location A multiplied by the number of users in location B. The SCI data set includes values for locations from the zip code level up to the country level and is an anonymized snapshot from a single point in time. The locations of Facebook users are assigned based on their information and activity on Facebook, including their public profile information as well as device and connection information.

The SCI is a single data set calculated based on Facebook friendships in March 2020; therefore, additional time points of the SCI could not be included in the model or in sensitivity analyses.

COVID-19 Data

To determine which Missouri counties had similar COVID-19 spread, we used data obtained from the Johns Hopkins University's Coronavirus Resource Center. The data on United States COVID-19 cases and deaths made available through the Center are compiled by the Johns Hopkins Center for Systems Science and Engineering, which is updated daily. They retrieve all state data from their respective state departments of health or other local government reporting agency, and for Missouri, those sources are the Missouri Department of Health, St. Louis City Department of Health, St. Louis County Department of Health, and Nodaway County Health Center. We obtained daily

new case counts for every county in Missouri starting on March 8, 2020 (the day the first case of COVID-19 was recorded in the state), through September 30, 2020, which was the time we conducted the analyses.

Population, Rurality, and Distance Data

Data on each county's population and its rurality were obtained from the United States Census Bureau from the 2010 Census database [14]. Distance between each county pair was obtained from the 2020 TIGER/Line shapefiles, also available from the US Census Bureau [15].

Analysis

All analyses were conducted using R version 4.0.3 (R Foundation for Statistical Computing) with the packages `statnet` and `ergm`. Alpha levels were set at .05.

Data Management

Every county pair has an SCI value, so this variable did not need to be computed, though this variable was log transformed.

To create a measure of two counties' similarity in COVID-19 case counts, we used the daily new case counts as each county's "trajectory" of COVID-19 and conducted a Pearson correlation test between each county's trajectory. We then used a 0.60 correlation coefficient cutoff to classify each county pair as either having highly correlated COVID-19 case count trajectories or not. The 0.60 cutoff was chosen based on established recommendations [16]. This binary variable was our primary outcome.

The total county population was log transformed, and the distance between every county pair was calculated using the distance between the centroids of each county in the shapefiles. The percent of the county that was classified as rural was not computed or transformed before being entered into the analytical model.

We originally intended to include demographic characteristics of residents at the county level, but given the lack of diversity on characteristics such as age, race, and ethnicity across Missouri, including these data in the model caused it to not converge. Therefore, we were unfortunately unable to include them.

Modeling

Our basic modeling approach was to examine the relationship between the social media connections (as measured by the SCI)

and COVID-19 case counts across Missouri counties. To do this, we used exponential random graph modeling.

ERGMs model relational data, and the outcome variable must be relational and binary, representing the presence or absence of a relationship. In our model, this was the presence or absence of a highly correlated COVID-19 case count trajectory between two given counties in Missouri. The model was built sequentially, starting with a null model. Next, all covariates except the SCI were entered into the model. Distance between every county pair was entered into the model as a relational term, meaning it represented a relationship between every county pair. Total county population and the percent of the county classified as rural were both entered into the model as object-level terms, meaning instead of the data representing a relationship between every county pair, these data were singular attributes of each county. After running this model, the SCI was entered into the last model and the Akaike information criterion (AIC) was used to compare overall model fit. Odds ratios (ORs) and 95% CIs are also reported.

Results

COVID-19 Case Count Trajectories

Missouri reported its first COVID-19 case on March 8, 2020, and at the time of analysis (September 30, 2020) the state had reported 129,733 cumulative cases with a 7-day average of 1127. Each county's average daily new case count data are available in [Multimedia Appendix 1](#).

Of the 6555 different county pairs ($[115 \text{ counties} * 114] / 2 = 6555$), the range of correlations was -0.09 to 0.90 with an average correlation of 0.36 . Of those, 1114 county pairs had COVID-19 case count trajectories correlated at 0.60 or greater. These 1114 county pairs then represented the relationship we predicted in the model.

Exponential Random Graph Model

The results of the sequential model building process are presented in [Table 1](#). In the final model, we assessed the likelihood that two counties in Missouri had highly correlated COVID-19 case count trajectories based on their level of social connectedness, controlling for the total population size of the counties, the percent of the counties that were rural, and the distance between the two counties. The model fit improved sequentially as evident by the decreasing AIC value as more covariates were entered into the model.

Table 1. Sequential building of an exponential random graph model of the relationship between the Social Connectedness Index (SCI) and Missouri counties' similar COVID-19 case counts from March to September 2020.

	Null model ^a		Model 1 ^b		Model 2 ^{c,d}		
	b (SE)	P value	b (SE)	P value	b (SE)	OR ^e (95% CI)	P value
Intercept	-2.40 (0.04)	<.001	-14.63 (0.12)	<.001	-14.85 (0.94)	0.00 (0.00-0.00)	<.001
Total county population (logged)	N/A ^f	N/A	0.44 (0.04)	<.001	0.51 (0.07)	1.66 (1.43-1.92)	<.001
Percent of county that is rural	N/A	N/A	0.01 (0.00)	.001	0.01 (0.00)	1.01 (1.00-1.01)	<.001
Distance in miles between county pairs	N/A	N/A	-0.62 (0.16)	<.001	-0.55 (0.15)	0.57 (0.43-0.77)	<.001
SCI ^g (logged)	N/A	N/A	N/A	N/A	0.16 (0.04)	1.17 (1.09-1.26)	<.001
Model fit							
Akaike information criterion	3251	N/A	2707	N/A	2691	N/A	N/A
Bayesian information criterion	3773	N/A	2741	N/A	2731	N/A	N/A
Log likelihood (<i>df</i>)	-1882.008 (1)	N/A	-1485.785 (4)	N/A	-1477.192 (5)	N/A	N/A

^aNull model with no covariates.

^bModel 1 included all covariates, except the SCI.

^cModel 2 included all covariates, including the SCI.

^dThe geometrically weighted edgewise shared partner term *gwesp* was included in models 1 and 2.

^eOR: odds ratio.

^fN/A: not applicable.

^gSCI: Social Connectedness Index.

All covariates were significantly associated with two counties having highly correlated COVID-19 case count trajectories. As the log of a county's total population increased, the odds of two counties having highly correlated COVID-19 case count trajectories increased by 66% (OR 1.66, 95% CI 1.43-1.92). (Log scales are commonly used when examining population growth; it also is helpful here for comparing changes in ratios or proportions.) As the percent of a county classified as rural increased, the odds of two counties having highly correlated COVID-19 case count trajectories increased by 1% (OR 1.01, 95% CI 1.00-1.01). As the distance (in miles) between two counties increased, the odds of two counties having highly correlated COVID-19 case count trajectories decreased by 43% (OR 0.57, 95% CI 0.43-0.77). For our main outcome, we found that as the log of the SCI increased between two counties, the odds of those two counties having highly correlated COVID-19 case count trajectories increases by 17% (OR 1.17, 95% CI 1.09-1.26), controlling for the counties' population size, rurality, and the distance between the two counties.

Discussion

Principal Findings

We found that as the likelihood of Facebook friendships between two counties increases, as measured with the SCI, the odds of those two counties having strong, positive correlations of their COVID-19 daily new case count trajectories also significantly increased. This relationship remained significant when

controlling for the distance between the two counties, their rurality, and their total population sizes.

These results build upon and align with prior, preliminary research using the SCI to understand COVID-19 spread. [8,11,12] These results also confirm the "signal" in the SCI "noise," meaning there is something uniquely captured in the SCI and Facebook friendships that cannot be explained by geography, distance, or population size.

The primary reasons for conducting this study were to assess if the relationship between the likelihood of Facebook friendships and COVID-19 spread could be explained by other factors. For example, it makes intuitive sense that two urban counties are more likely to have similar COVID-19 case count trajectories because, in general, urban areas had more cases earlier in the pandemic than rural areas [17]. It also makes sense that two urban counties would be more likely to share Facebook friendships than an urban and a rural county [18]. Likewise, it is reasonable to expect that two counties next to each other would be more likely to share Facebook friendships than two counties hundreds of miles apart [9]. Could the SCI signal as it relates to COVID-19 spread be explained by these other factors? Our results suggest there is something above and beyond these other factors that the SCI represents; however, it is not clear what exactly that is.

One possibility is that people tend to form friendships and social connections to those who share similar belief systems [12,19]. This could suggest that two counties with a greater likelihood

of sharing Facebook friendships means residents of those counties have a higher likelihood of sharing similar belief systems, in particular as they relate to COVID-19 and public health practices. For example, perhaps residents of two counties with a relatively high SCI value are equally likely to wear masks or not, restrict travel or not, etc. Residents sharing similar public health practices could explain why counties with relatively high SCI values are also more likely to have similar COVID-19 case count trajectories. Similar results have been found in earlier studies using the SCI [12].

Another possibility is that the SCI is picking up travel or movement data among county residents. People are more likely to form Facebook friendships with people they have offline connections with, and these offline connections may stem from a physical location such as a school, place of worship, or place of employment [20,21]. Therefore, a resident of one county may have a lot of Facebook friends in a neighboring county because that resident works at a large business in that neighboring county and travels there multiple days a week. That resident may also frequent restaurants and other businesses near their place of employment, increasing the opportunities to form friendships in this neighboring county. In rural areas, residents often travel long distances [22,23], so the SCI may indeed be capturing, in part, a person's likelihood of traveling to another county. This has relevance, of course, to COVID-19 spread.

In particular, the results of this study could be relevant for state and county public health departments in Missouri that are trying to implement COVID-19 prevention practices, such as setting event/business capacity limits or enacting mask requirements. Knowing that social connectedness, as measured through Facebook friendships, is associated with COVID-19 spread even after controlling for the distance between two counties might suggest that mitigation practices should extend beyond a regional approach and be implemented statewide.

Additional investigation is needed to more fully understand the SCI. Our study and others' prior work have demonstrated a signal, but now more research is needed to fully decipher that signal. We also encourage Facebook to continue to update and refine the SCI, so that researchers can understand more of what in the signal it is capturing and how it relates to COVID-19. However, while that work is underway, there may be utility in using the SCI in models of COVID-19 spread even without

knowing what it is capturing. In the case of a global pandemic, the need for timely data and models to inform mitigation efforts is critical. If including the SCI in these models can improve model fit and serve as a control for more fully understood variables, then it is worth including in the model.

Limitations

There are key caveats that must be acknowledged. First, more granular data are not included in the SCI, which would add greater clarity to the results. For example, we would have liked to have known the demographics of Facebook users in a given county and if the SCI was different for certain demographic subgroups in each county (eg, are older Facebook users in county 1 more likely to form friendships with older users in county 2). Second, the SCI was a cross-sectional data set created in March 2020, while our COVID-19 data were longitudinal from March to September 2020. It is unknown if, and by how much, the SCI changes over time and if this would impact our modeling. Third, we are network analysis and modeling experts; we are not epidemiologists or infectious disease experts. Therefore, we approached this study from a methodological perspective, not a public health perspective, and we acknowledge there are additional factors that should be studied before any policies or prevention practices are enacted based on these results.

Conclusions

This study further validated the signal raised by the SCI as it relates to COVID-19 spread. It is also the first study to use ERGM to model Facebook friendships as they relate to COVID-19 spread. We found that as the social connectedness increases between two counties, the odds of those two counties having highly correlated COVID-19 case count trajectories increases by 18%, controlling for the counties' population size, rurality, and the distance between the two counties. This suggests that the SCI is capturing a unique social connection phenomenon that is important in understanding disease transmission and is specifically relevant to COVID-19. Additional research is needed to better understand what the SCI is capturing practically and what it means for public health policies and prevention practices, but in the short term, researchers may consider adding it to other COVID-19 models to improve model fit.

Acknowledgments

The authors acknowledge the contribution of Dr Johannes Stroebel, who was involved in the creation of the SCI with Facebook and who provided clarifying guidance on the SCI for our use. Facebook and the Facebook Data for Good project did not review this manuscript before publication.

Authors' Contributions

BP was responsible for the study design. BP, KP, JML, and DAL were responsible for the data analysis. BP was responsible for drafting the manuscript, with KP, JML, and DAL revising.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Missouri Counties' COVID-19 Daily Case Counts.

[\[DOCX File , 22 KB-Multimedia Appendix 1\]](#)

References

1. Cinelli M, Quattrocioni W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. *Sci Rep* 2020 Oct 06;10(1):16598. [doi: [10.1038/s41598-020-73510-5](https://doi.org/10.1038/s41598-020-73510-5)] [Medline: [33024152](https://pubmed.ncbi.nlm.nih.gov/33024152/)]
2. Wong FHC, Liu T, Leung DKY, Zhang AY, Au WSH, Kwok WW, et al. Consuming information related to COVID-19 on social media among older adults and its association with anxiety, social trust in information, and COVID-safe behaviors: cross-sectional telephone survey. *J Med Internet Res* 2021 Feb 11;23(2):e26570 [FREE Full text] [doi: [10.2196/26570](https://doi.org/10.2196/26570)] [Medline: [33523825](https://pubmed.ncbi.nlm.nih.gov/33523825/)]
3. Tasnim S, Hossain MM, Mazumder H. Impact of rumors and misinformation on COVID-19 in social media. *J Prev Med Public Health* 2020 May;53(3):171-174. [doi: [10.3961/jpmph.20.094](https://doi.org/10.3961/jpmph.20.094)] [Medline: [32498140](https://pubmed.ncbi.nlm.nih.gov/32498140/)]
4. Karami A, Anderson M. Social media and COVID-19: characterizing anti-quarantine comments on Twitter. *Proc Assoc Inf Sci Technol* 2020;57(1):e349 [FREE Full text] [doi: [10.1002/prat.349](https://doi.org/10.1002/prat.349)] [Medline: [33173823](https://pubmed.ncbi.nlm.nih.gov/33173823/)]
5. Shen C, Chen A, Luo C, Zhang J, Feng B, Liao W. Using reports of symptoms and diagnoses on social media to predict COVID-19 case counts in mainland China: observational infoveillance study. *J Med Internet Res* 2020 May 28;22(5):e19421 [FREE Full text] [doi: [10.2196/19421](https://doi.org/10.2196/19421)] [Medline: [32452804](https://pubmed.ncbi.nlm.nih.gov/32452804/)]
6. Tayal P, Bharathi. S V. Reliability and trust perception of users on social media posts related to the ongoing COVID-19 pandemic. *J Hum Behav Soc Environment* 2021 Jan 04;31(1-4):325-339. [doi: [10.1080/10911359.2020.1825254](https://doi.org/10.1080/10911359.2020.1825254)]
7. Ahmed W, Vidal-Alaball J, Downing J, López Seguí F. COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. *J Med Internet Res* 2020 May 06;22(5):e19458 [FREE Full text] [doi: [10.2196/19458](https://doi.org/10.2196/19458)] [Medline: [32352383](https://pubmed.ncbi.nlm.nih.gov/32352383/)]
8. Kuchler T, Russel D, Stroebel J. JUE Insight: The geographic spread of COVID-19 correlates with the structure of social networks as measured by Facebook. *J Urban Economics* 2021 Jan:103314. [doi: [10.1016/j.jue.2020.103314](https://doi.org/10.1016/j.jue.2020.103314)]
9. Bailey M, Cao R, Kuchler T, Stroebel J, Wong A. Social connectedness: measurement, determinants, and effects. *J Econ Perspect* 2018;32(3):259-280. [Medline: [30362698](https://pubmed.ncbi.nlm.nih.gov/30362698/)]
10. Data for Good. URL: <https://dataforgood.fb.com/> [accessed 2021-05-18]
11. Charoenwong B, Kwan A, Pursiainen V. Social connections with COVID-19-affected areas increase compliance with mobility restrictions. *Sci Adv* 2020 Nov;6(47):eabc3054. [doi: [10.1126/sciadv.abc3054](https://doi.org/10.1126/sciadv.abc3054)] [Medline: [33097473](https://pubmed.ncbi.nlm.nih.gov/33097473/)]
12. Holtz D, Zhao M, Benzell SG, Cao CY, Rahimian MA, Yang J, et al. Interdependence and the cost of uncoordinated responses to COVID-19. *Proc Natl Acad Sci U S A* 2020 Aug 18;117(33):19837-19843 [FREE Full text] [doi: [10.1073/pnas.2009522117](https://doi.org/10.1073/pnas.2009522117)] [Medline: [32732433](https://pubmed.ncbi.nlm.nih.gov/32732433/)]
13. Valente TW. *Social Networks and Health: Models, Methods, and Applications*. New York, NY: Oxford University Press; 2010.
14. Explore census data. United States Census Bureau. URL: <https://data.census.gov/cedsci/> [accessed 2021-05-18]
15. TIGER/Line shapefiles. United States Census Bureau. 2020. URL: <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html> [accessed 2021-05-18]
16. Jones W, Ness A. *Statistics at Square One 10th Edition*. London: BMJ Books; 2003.
17. Morbidity and Mortality Weekly Report. COVID-19 stats: COVID-19 incidence,* by urban-rural classification - United States, January 22-October 31, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Nov 20;69(46):1753. [doi: [10.15585/mmwr.mm6946a6](https://doi.org/10.15585/mmwr.mm6946a6)] [Medline: [33211682](https://pubmed.ncbi.nlm.nih.gov/33211682/)]
18. Bailey M, Farrell P, Kuchler T, Stroebel J. Social connectedness in urban areas. *J Urban Economics* 2020 Jul;118:103264. [doi: [10.1016/j.jue.2020.103264](https://doi.org/10.1016/j.jue.2020.103264)]
19. Hristova D, Musolesi M, Mascolo C. Keep your friends close and your Facebook friends closer: a multiplex network approach to the analysis of offline and online social ties. 2014 Presented at: Eighth International AAAI Conference on Weblogs and Social Media; June 1-4, 2014; Ann Arbor, MI.
20. Ellison N, Steinfield C, Lampe C. The benefits of Facebook "Friends": social capital and college students' use of online social network sites. *J Computer-Mediated Commun* 2007;12(4):1143-1168. [doi: [10.1111/j.1083-6101.2007.00367.x](https://doi.org/10.1111/j.1083-6101.2007.00367.x)]
21. Subrahmanyam K, Reich SM, Waechter N, Espinoza G. Online and offline social networks: use of social networking sites by emerging adults. *J Appl Developmental Psychol* 2008 Nov;29(6):420-433. [doi: [10.1016/j.appdev.2008.07.003](https://doi.org/10.1016/j.appdev.2008.07.003)]
22. Arcury TA, Gesler WM, Preisser JS, Sherman J, Spencer J, Perin J. The effects of geography and spatial behavior on health care utilization among the residents of a rural region. *Health Serv Res* 2005 Feb;40(1):135-155 [FREE Full text] [doi: [10.1111/j.1475-6773.2005.00346.x](https://doi.org/10.1111/j.1475-6773.2005.00346.x)] [Medline: [15663706](https://pubmed.ncbi.nlm.nih.gov/15663706/)]
23. Rosenberg M, Corcoran SP, Kovner C, Brewer C. Commuting to work: RN travel time to employment in rural and urban areas. *Policy Polit Nurs Pract* 2011 Feb;12(1):46-54. [doi: [10.1177/1527154411411926](https://doi.org/10.1177/1527154411411926)] [Medline: [21733956](https://pubmed.ncbi.nlm.nih.gov/21733956/)]

Abbreviations

AIC: Akaike information criterion

ERGM: exponential random graph model

OR: odds ratio

SCI: Social Connectedness Index

Edited by T Sanchez; submitted 16.09.21; peer-reviewed by T Liu, A Chen; comments to author 01.10.21; revised version received 05.11.21; accepted 16.11.21; published 15.12.21

Please cite as:

Prusaczyk B, Pietka K, Landman JM, Luke DA

Utility of Facebook's Social Connectedness Index in Modeling COVID-19 Spread: Exponential Random Graph Modeling Study

JMIR Public Health Surveill 2021;7(12):e33617

URL: <https://publichealth.jmir.org/2021/12/e33617>

doi: [10.2196/33617](https://doi.org/10.2196/33617)

PMID: [34797775](https://pubmed.ncbi.nlm.nih.gov/34797775/)

©Beth Prusaczyk, Kathryn Pietka, Joshua M Landman, Douglas A Luke. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 15.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.