Aug 10th, 12:00 AM

Exploring Information Flow on Twitter: Social Network Analysis on Gender-Specific Medicine

Katharina Batzel
Universität Potsdam, kathi.batzel@gmail.com

Katharina Baum
Weizenbaum Institute, katharina.baum@uni-potsdam.de

Exploring Information Flow on Twitter: Social Network Analysis on Gender-Specific Medicine

Completed Research

Katharina Batzel
University of Potsdam
Kathi.Batzel@gmail.com

Katharina Baum
Weizenbaum Institute for the Networked Society / University of Potsdam
katharina.baum@uni-potsdam.de

Abstract

To date, sex and gender differences play only a minor role in medical research and practice, thereby putting individuals' health at risk. Gender-specific medicine, or the practice of taking these differences into account when conducting research and treating patients so far is being discussed primarily by experts. With people increasingly using social media such as Twitter for sharing and searching for health-related information online, Twitter can potentially educate about gender-specific medicine. However, little is known about the information circulation and the structure of interactions on the Twitter network discussing this topic. Results of a network analysis show that the network exhibits a community-structure, with information exchange being limited and concentrated in silos. This indicates that there is untapped potential for acquiring new information by users through interacting with individuals outside their community. Public health officials may benefit from this insight and tailor online campaigns to enhance awareness on gender-specific medicine.

Keywords

Gender-specific medicine, social network analysis, information flow, Twitter

Introduction

Both biologically determined factors (sex differences) and socio-cultural aspects (gender differences) that affect men and women differently are given little importance in health research and practice (Baggio et al. 2013). Historically, biomedical studies, clinical trials, and drug development have used male subjects, whether in studies with cells, mice, or humans (Clayton 2016). Under the assumption that human cells are identical, medical trials drew conclusions from their findings for both sexes. But medicine is neither sex- nor gender-neutral (Regitz-Zagrosek 2012). The Covid-19 pandemic presents the latest example of the importance of both sex and gender in medical research (Mauvais-Jarvis et al. 2020). Studies show that the virus is deadlier for men than for women, with an increased mortality rate of 0.9% in Chinese men and more severe cases in elderly European men (Gebhard et al. 2020). This difference is caused by sex-specific factors, such as hormone-driven immune response, as well as gender-specific factors such as lifestyle, stress, and socioeconomic conditions (Gebhard et al. 2020). Sex and gender differences occur in a variety of illnesses: cardiovascular disease, cancers, pulmonary disease, stroke, Alzheimer's, diabetes, or depression are all affected by sex or gender (Mauvais-Jarvis et al. 2020). The research stream concerned with the investigation of such disparities is termed gender-specific medicine (Legato 2003, 2009). Discussion of this phenomenon has been mainly confined to researchers and professionals in the field, meaning that there is a lack of inclusion and involvement by the wider public (Mauvais-Jarvis et al. 2020). One of the ways information about gender-specific medicine is disseminated to the wider public is through social networking sites (SNSs). SNSs enable users to access and disseminate information, exchange opinions, and experiences, to network, and form communities (Xu et al. 2015). These sites may therefore help educate the public by promoting health literacy (Haunschild et al. 2021). However, to date, no analysis has been published on the exchange of information on gender-specific medicine, though knowledge on the topic is essential to the provision of vital health care.

This study sets out to examine information circulation and the structure of interactions in the network discussing gender-specific medicine on Twitter. Insights from our research could benefit public institutions, such as ministries of health or medical unions, to gauge public knowledge and plan health literacy campaigns accordingly. Such campaigns could help increase awareness of the topic of gender-specific medicine, which could, for example, enable women to recognize differing symptoms of heart attacks. This study makes several theoretical and practical contributions. It adds to the literature on the dissemination of health-related content on SNSs (Singh et al. 2020, Roy et al. 2020). Further, it deepens our understanding of the discussions on gender-specific medicine on Twitter, which so far has been mainly addressed in the medical field (Mauvais-Jarvis et al. 2020). In addition to this, methodologically, results contribute to other studies that employ social network analysis to understand information dissemination through SNSs such as debates on immunization (Milani et al. 2020) or for campaign monitoring and the detection of key influencers in discussion around the World Breastfeeding Day (Moukarzel et al. 2020).

Based on a keyword search and by using social network analysis, we capture the network structure of the discourse on gender-specific medicine on Twitter. The analysis shows that the network is strongly decentralized and sparsely connected, with several moderately sized communities evolving around hubs. Due to its community structure, information exchange on gender-specific medicine is limited and concentrated in silos. We conclude that there is an untapped potential of acquiring new information for users in the network by interacting with individuals outside their community.

Theoretical Background

Twitter: A Platform with Unique Characteristics for Health Communication

The wide adoption of SNSs has fundamentally reshaped how we communicate and exchange information, including how we talk about health. Both individuals and healthcare providers use SNSs to exchange health-related information. For individuals, SNSs present a source of professional and patient knowledge (Himelboim and Han 2014), which may help them understand individual risk factors, grasp their diagnosis, decide on treatments, and assess their prognosis (Chen et al. 2018). Health-related organizations use SNSs to promote health literacy and engage with consumers (Park et al. 2013). Further, healthcare providers form virtual communities, sharing links to resources among one another and with students (Choo et al. 2015). Among social networks, Twitter in particular offers unique opportunities for the dissemination of health information since several affordances of the network promote information flow. For example, communication in the form of tweets as well as the use of links and hashtags enable efficient information intake by the user, as messages are easy to process and fast to read (Gleason 2013). Furthermore, the use of hashtags facilitates following, joining, and engaging in conversations around a specific topic, which may lead to virtual communities forming around a shared interest (Bruns and Burgess 2011; Xu et al. 2015). Interaction is further enabled through reply, retweet, quote, and mentions functions.

The particularities of Twitter described above enable communication on a variety of health-related topics that resulted in a large body of research examining these discussions. Especially in the past two years, online conversations on Covid-19 have attracted a substantial amount of research (Singh et al. 2020). These studies add to a broader epidemiologic literature concerned with information content and circulation on events such as the H1N1 outbreak in 2009 (Chew and Eysenbach 2010), or the Ebola epidemic between 2014 to 2016 in West Africa (Roy et al. 2020). Studies have further investigated online debates by pro-and anti-vaxxers (Himelboim et al. 2020), online discussions on breastfeeding (Moukarzel et al. 2020), the online discourse regarding tobacco use (Chu et al. 2019), or cancer (Wang et al. 2020). Further, given its large user base, Twitter presents a cost-effective way for public health organizations to reach and mobilize a large crowd. For example, in research conducted by Allen et al. (2020) on awareness of the HPV vaccine, the authors find that the awareness increased even for those individuals that did not possess a Twitter account. Moreover, as shown by research in other areas such as breastfeeding (#Breastfeed4Ghana), mental health (#MHAW, #WhyWeTweetMH) or cardiovascular disease (#CardioOncology), Twitter serves as a promising tool to promote health awareness (Berry et al., 2017; Conley et al., 2020; Harding et al., 2020;

Makita et al., 2021). To summarize, SNSs are increasingly being used for the exchange, dissemination, and promotion of health-related information, and research has followed these conversations. To date, however, no research has focused on the communication of gender-specific medicine on SNSs.

Social Network Analysis: The Study of Social Relations between Individuals

In the diffusion of health-related information, Twitter users do not act in isolation but are connected to others by their online communication. When it comes to studying how information disseminates through networks, different characteristics are import. Among them is the degree of centralization of the network structure. In highly centralized networks, only a few users contribute most of the content and therefore dominate the information in the network (Barabási 2009; 2016). Another network characteristic that is important for information spread is its distribution of connections, or degree. Typically, this exhibits a highly skewed pattern: A small number of nodes with many connections followed by a trailing tail of nodes with very few connections (Barabási 2016). Similarly, the level of density strongly affects the information flow of a network. In dense networks, individuals maintain close ties with others and form one or several strongly concentrated communities (Himelboim et al. 2017). This increases the efficiency of communication. Further, it is common to find individuals who do not communicate with peers (Wassermann and Faust 1994). These so-called isolates are disconnected and hence cannot receive information by social exchange (Haythornthwaite 1996).

Based on these network characteristics, Twitter networks will take one of six archetypes that describe how conversations spread (Smith et al. 2014). In their topology, Smith et al. (2014) differentiate between the six distinct patterns: (1) polarized crowds, (2) tight crowds, (3) fragmented brand clusters, (4) clustered communities, (5) broadcast and (6) support networks. Polarized crowds (1) are formed when the conversation happens in two or more groups that are densely connected within themselves but have few interactions with the opposing network and hence do not share information with them (Smith et al. 2014). In contrast, in unified crowds (2) the whole set of users is strongly connected. In brand clusters (3), a large group of users tweets about the same topic but does not interact with other users. Clustered community networks (4) arise from multiple small conversations in the network. Each sub-network has its influential users and information source. As multiple information sources are present, this network type indicates a multifaceted discussion of a topic where people allocate to different viewpoints (Himelboim et al. 2017). In a broadcast network (5), the connections are directed inwards towards a hub, with many users replicating the information shared by the central hub user. Lastly, with support networks (6), a single user will interact with a large number of people (Smith et al. 2014) and information spreads freely between users. The observations above show that identifying the network structure on discussions about gender-specific medicine on Twitter yields important insights into the information spread of the topic. Hence, we ask: *How is the network on gender-specific medicine on Twitter structured and how does information flow through the network?*

Methodology

Data Collection and Search Term Selection

We collected publicly available tweets containing 15 different search terms from Twitter from January to May 2021. We deliberately chose a longer sampling period compared to other literature from the health awareness field (e.g., Milani et al. 2020; Himelboim et al. 2020), which is often focused on events. Since the discussion on gender-specific medicine is an ongoing process of actors raising awareness (Legato 2003), we collected tweets in five sequential months to ensure the connectivity of the tweets and users. We collected only tweets written in English. To select the search terms, we adopted a search strategy similar to those used in literature reviews (e.g., Webster and Watson 2002), constituting of four steps: (a) Initial keyword search, (b) backward content search, (c) forward author search, and (d) forward publication search. Following this approach, we (a) started with the term "gender-specific medicine" in all its possible spellings. (b) We then scanned the tweets obtained from search term 1 for further clues of other popular terms and hashtags used in the field. This led to the identification of further seven search terms. (c) We then conducted an author-centric forward search by going through the Twitter profiles of 16 leading figures[1]

[1] The full list is available from the authors upon request.

and organizations in the field, which yielded seven additional search terms. (d) We inspected keywords of 60 research papers from 2021 covering the topic of gender-specific medicine authored or co-authored by the people identified in the previous step.

No.	Search Terms	Tweets	Users	Replies	Mentions	Retweets	Quotes
1	#HerHeartMatters	3,698	1,840	809	2,484	2,655	299
2	(#GenderBias OR "gender bias") * -#WomenInSTEM -leadership -promotion	3,673	4,322	1,118	1,290	2,449	127
3	"sex & gender" * -transphobia -trans -transgender -transsexual -dysphoria	2,114	2,804	975	1,317	1,217	79
4	#SABV	1,956	1,492	396	647	1,577	82
5	#SexDifferences OR #SexDifference	1,469	1,375	180	600	1,087	69
6	#SexMatters	673	680	132	166	529	42
7	("precision medicine" OR #PrecisionMedicine) (sex OR gender)	562	406	83	224	426	14
8	(#GenderDataGap OR "gender data gap") *	326	420	159	42	264	7
9	#SexAndGender *	279	318	50	176	183	24
10	(#GenderSpecificMedicine OR "gender specific medicine" OR #GenderMedicine OR #GenderedMedicine OR ("gender specific" medicine))	208	310	115	148	106	8
11	#GenderData *	114	103	2	26	77	7
12	#MedicalBias	71	145	33	34	8	6
13	#SexBias	57	72	22	17	43	1
14	(#GenderDifferences OR #GenderDifference)*	27	27	0	8	12	1
15	#HeartDiseaseInWomen	36	43	1	22	25	4
TOTAL		15,263	14,357	4,075	7,201	10,658	770
TOTAL (w/o duplicates)		15,061	12,603	4,014	6,950	10,526	769
* (rehabilitation OR vaccine OR diagnosis OR health OR medicine OR clinical OR therapy OR treatment OR disease OR illness OR symptoms OR prevention)							

Table 1. Search Terms and Descriptive Statistics of the Final Data Sample

None of those publications led to the identification of new hashtags, but previously identified terms could be detected. We regarded the re-occurrence of familiar content without the detection of new information in the new data to be an indicator of saturation (Webster and Watson 2002), leading to a final selection of 15 search strings. The final data sample consists of 15,601 tweets, and 12,603 users. The complete list of search terms, and the number of tweets, users and interactions can be found in Table 1. To make sure search results were concerned with the topic of gender-medicine, broad keywords that could potentially touch on a variety of adjacent topics (e.g., “gender bias”) were further refined by adding a search string with terms from the medical field (see Table 1).

Social Network Analysis

We used the Python package NetworkX to build the network (Hagberg et al. 2020). We included all the 12,603 users; self-loops were removed from the data. In our network, users serve as the nodes, and their interactions in the form of retweets, quotes, replies, and mentions are the edges. Since we regard different interaction forms as connections between users, the graph needs to visualize multiple ties between the same two actors. This characteristic is labeled multivariant and represented by a multigraph (Wassermann and Faust 1994). In the following, the graph describing this network will be abbreviated as multiG. In the visualization of the graph, NetworkX provides about 13 different layouts to choose from. Each of those layouts positions nodes and their edges differently based on the underlying algorithm. In this study, we use the spring layout for all graphs. This layout is based on the Fruchterman-Reingold force-directed algorithm which maps connected nodes closer to one another, than to disconnected ones (Fruchterman and Reingold, 1991). Two mechanisms are combined in any force-directed layout. Initially, nodes are pushed apart, then connected dots are pulled closer. Those types of layouts have the advantage of accommodating large networks and of clearly revealing community structure.

After constructing the graph using NetworkX, we followed the network classification approach by Himelboim et al. (2017) to detect the structure of discussions on gender-specific medicine. This classification approach is based on the six network archetypes described in the theory section (Smith et al. 2014). The approach consists of four steps that define the structure and topology of the network. In the following, we outline the steps involved: (1) In step 1, the centralization, measured as the sum of all nodes' degree centrality divided by the number of nodes, of the network is calculated. A centralization of 0 implies that all degrees are equal, meaning every user in the network maintains the same number of ties. In contrast, centralization is at 1 when all actors are connected to the same single node, as in a star-shaped network. Values between both extremes, 0 and 1, indicate variability in the range of centrality scores of the actors in the graph (Wassermann and Faust 1994). Values of 0.59 or larger indicate a high centralization, whereas networks with lower values are decentralized (Himelboim et al. 2017). (2) If the centralization of the network is lower than 0.59, density is measured in step 2. Density also varies between 0 and 1, where 0 indicates that no edges are present between any nodes in the graph and 1 represents a fully connected complete graph (Wassermann and Faust 1994). Networks with a density of 0.12 and higher are regarded as being densely connected (Himelboim et al. 2017). Networks with high density can be dominated by unified or divided clusters, which have contrary implications for the information flow. (3) If density is high, in step 3, network modularity is measured, which denotes the connectivity of the whole network. In this case, the value of 0.29 is used to differentiate between high and low modularity (Himelboim et al. 2017). (4) In the last step of the classification model, the share of isolates on the whole set of users is calculated. This is done to distinguish between sparse networks with a few connected communities (clustered), or networks with a large share of isolates and a few clusters (fragmented) (Himelboim et al. 2017). The share of isolates is calculated as the proportion of users without any interaction to other users. With the reference mean value being at 0.19, networks with a share of isolates of 20% are considered fragmented, while any value below 19% indicates a clustered structure of the network (Himelboim et al. 2017).

Results and Discussion

The Social Network of Gender-Specific Medicine

We created a graph comprising 12,603 nodes and 16,704 edges (replies: 2,240; mentions: 5,243; retweets: 9,221) to capture the network structure on gender-specific medicine conversations from January to May 2021. As they do not contribute to information circulation, 585 self-loops were removed. A total of 503 isolates were detected. In-degree and out-degree yield the same result, with an average of 1.6 links directed inwards and outwards between users. In summary, on average 3.22 interactions on the topic of gender-specific medicine take place between individuals over five months.

Network Structure and Information Flow

To answer our research question we applied Himelboim et al.'s (2017) classification approach presented above and sequentially calculated network centralization, density, and the isolates fraction. (1) Network centralization is very low (0.0002103). With values close to zero, the cut-off value of 0.59 for a highly centralized network is far from being met. Hence, our first observation is that the network on gender-specific medicine is strongly decentralized. This implies that most nodes in the network have an equal number of ties and that only a small fraction of users hold significantly more links than the rest. When further analyzing the distribution of relationships between users, we find that the weighted degree distribution of the graph is characterized by a high number of nodes with a degree of 1, and by a short right tail of nodes with a higher degree. Hence the network follows the power law. The observed degrees vary between 1 and 759, where most of the nodes (8,268 or 65.6%) have a degree of 1, another 1,626 users (12.9%) have a degree of 2. Most of the users thus interact with or are addressed by others only once, mutual interaction tends to occur rarely. (2) The next stage involved an examination of the graph density. As the network on gender-specific medicine is decentralized, it could either be structured as one unified accumulation of users or be divided into different camps (Himelboim et al. 2017). Results show that graph density is low (0.0001052), in fact, far below the threshold of 0.12. Hence, we observe only a small level of interconnectedness in comparison with other networks. To put this result into perspective, Himelboim (2020) calculated the density of the HPV vaccine debate on Twitter including approximately 39 thousand users with a graph density of 0.0003. Still, this value is 3-times larger than the density detected with around 0.0001 for a smaller sample size of 12,603 users. In the context of the vaccine debate, Milani et al. (2020) examined content posted by pro- and anti-vaxxers from June to October 2016. Here, density ranges from 0.0024 for a pro-vaccination network to 0.0011 for an anti-vaccination conversation in October. All values remain above the 0.0001 (0.01% of all possible connections) measured in the network on gender-specific medicine. Our second finding is therefore, that the network on gender-specific medicine is not only decentralized but also weakly connected. (3) Due to the low density of the network, step 3 was not performed, but instead, the fraction of isolates was calculated. (4) The fraction of isolates in the mutliG is at 3.99%. With the cut-off value for a high share of isolates being at 19%, this finding implies that most individuals and organizations interacted with at least one other actor in the network, even though the interaction is likely to have occurred only once. Given the low share of isolates, we conclude that the network on gender-specific medicine presents some form of group connectivity, where a few moderately sized communities form around hubs (Himelboim et al. 2017). See Figure 1 for a depiction of the steps in the analysis and resulting network structures.

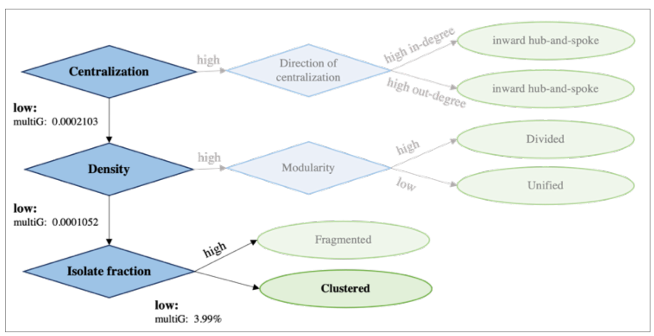

Figure 1. Network Classification Process, Illustration Based on Himelboim et al. (2017)

With respect to information flow, the detected network type of a clustered network indicates a diverse landscape of viewpoints between groups and a limited exchange of information between clusters. Specifically, the decentralized structure of conversations on gender-specific medicine indicates that users in the network do not rely on central actors for information. On the contrary, the exchange of information between users follows an egalitarian approach (Himelboim et al. 2017). Information originates from several different points of contact. Yet, users in the network are connected by the use of the same hashtags and thus share an overarching interest in gender-specific medicine. The boundaries and multitude of the groups create knowledge silos, revealing a landscape of different opinions and perspectives on the same topic. Further, the low density of the network points to slow and vulnerable information flow. Since the network is sparsely connected, users can only be reached through a few routes and strongly depend on users that connect them and provide access to the information network. This sparsity highlights missing coordinated activity by official sources or pioneering actors in that field (Himelboim et al. 2017).

Besides the entirety of the communication flow, we also analyzed differences in the frequency of communication forms. This observation is valuable since the interactions hold different implications for the information flow. The most popular way of communicating information on gender medicine is through retweets (9,221 edges). Thanks to Twitter's retweet and quote button, reposting content is extremely easy, and users in the network on gender medicine make ample use of it. In contrast to retweets, replies tend to be sent rarely (2,240 edges). While retweets are a form of replicating information, replies require active engagement with the content and thus present a higher activation barrier to overcome. The low share of replies implies that users in the network on gender medicine use Twitter mainly as a passive information source, without getting involved. Mentions connect users 5,243 times. We conclude that even though Twitter provides a multitude of interaction opportunities, users in the network on gender medicine prefer to remain passive, replicating existing information.

Interactions are Concentrated in a Giant Component

As the network analysis showed, discussion of gender-specific medicine on Twitter occurs within and among a few different communities. To further investigate the structure of those groups, we also analyzed the characteristics of those user crowds. This analysis was driven by the intention to gain additional insights into the information flow between actors, as communities hold unique structural properties. Compared to the overall network, communities are highly interconnected and denser, which enables easy and fast information exchange between its members (Barabási 2016). To grasp the communities, the network needed to be destructed into analyzable parts. The first stage of this process included the identification of disjunct sub-graphs, called components (Wassermann and Faust 1994). Calculating components in NetworkX led us to find that the network on gender-specific medicine is split into 1,364 components. The largest of these contains 8,503 nodes, which accounts for 67.5% of the total nodes. To place this size into perspective, the second and third largest components contain only 65 and 35 nodes respectively. A component that contains such a significantly large share of actors is called a giant component (Barabási 2016). Giant components can often be treated as a proxy for the whole graph (Barabási 2016). Indeed, the discussion on gender and sex bias in health is concentrated in the largest component. Outside of this structure, users present mixed thematic foci. Manually exploring the content of tweets from accounts in the respective group, we qualitatively detected the main conversation topics. Component 2 contained users located in India who are concerned with gender-biased laws to the disadvantage of men. The group used hashtags such as #JusticeForMen, #SpeakUpIndia, as well as #MensLivesMatter. In component 3, discussion revolves around the rigor and quality of scientific publications in ethics, philosophy, and psychology. The users are active academics, such as professors or else are affiliated with a university.

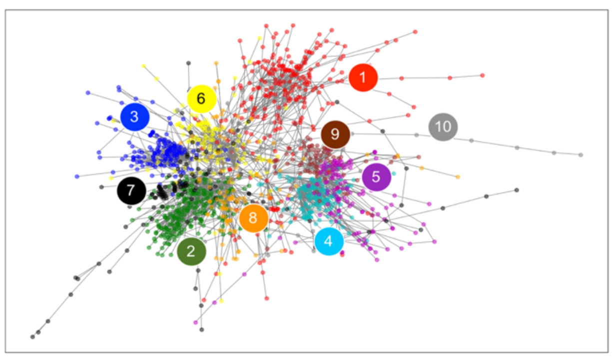

Figure 2. Top 10 Communities in the Giant Component

Communities in the Giant Component

To further examine the network structure of the giant component, we used community detection analysis. Community detection in social networks is a common problem, and several algorithms have been proposed (Newman 2004). One of the most widely implemented algorithms is the approach introduced by Clauset, Newman and Moore (2004). This algorithm finds communities by greedily optimizing modularity based on the pattern of connectivity between individuals. Communities are thus detected by grouping users who share connectivity patterns. To detect communities within the giant component, we applied the Clauset-Newman-Moore algorithm using NetworkX (NetworkX 2021). In total, 50 communities are located in the largest component, most being made up of up to 50 members. With 251 and 231 nodes respectively, two exceptionally large communities dominate the giant component. We extracted the top 10 largest communities from the giant component. The result is displayed in Figure 2. Within the communities, different domains of gender-specific medicine are emphasized. The content of conversations within the different communities was qualitatively assessed by examining prominent tweets. Community 1 (251 nodes) is composed of individuals critically reporting on the role of gender in varying realms of life. The second-largest community (231 nodes) is characterized by a few broadcasters and their larger audience. Content-wise, members of the community are concerned with the biological aspects of gender-specific medicine. This can be traced back to hashtags such as #SABV or #SexDifferences. Community 3 (157 nodes) centers around the NGO Women's Brain Project, which is an organization advocating for sex and gender differences between men and women in mental health and neuro medicine. Community 4 (129 nodes) focuses on cardiovascular diseases in women, with the top 5 users being cardiologists. To conclude, group 5 (129 nodes) includes mainly Canadian users, centering around the #WearRedCanada campaign. In respect of the analysis of the communities, we conclude that each group centers around a different thematic focus. The notion of homophily, therefore, does apply to the network on gender-specific medicine. Users tend to surround themselves with like-minded people. This further implies that knowledge is structured in silos, with little information variation within the communities.

Limitations, Implications, and Conclusion

This study was conducted to explore and understand the network structure and communication patterns in Twitter communication on gender-specific medicine. We find that limited information exchange takes place, where information circulation is restricted by the decentralization and sparsity of the network. Overall, the topic of gender-specific medicine attracts a large share of smaller communities (low centralization and low density), and few peripheral users (low isolate fraction). As a clustered network, it comprises medium-sized groups with different hubs that are surrounded by separate crowds. Information flow is characterized by the diversity of access points as well as by an overarching shared interest. The boundaries and multitude of the groups create knowledge silos, revealing a landscape of different opinions and perspectives. Overall, this analysis shows that the network is organized in community clusters and that the information flow is restricted.

This paper is not without limitations. Firstly, the data used for this paper is related to the hashtags and search terms that we used for data collection. Although we employed a thorough search process to detect all relevant key terms, it is possible that we missed less frequently used hashtags or that some of the search terms captured tweets on closely related topics. Further, choosing a variety of different search terms for the analysis might have resulted in the strong decentralization we observe. Xu et al. (2015), who followed a similar approach in their data collection by using many different hashtags, arrived at the conclusion that their network was decentralized as well. However, since the topic of gender-specific medicine covers a variety of sub-fields, the choice to include of a range of search terms seems appropriate. Over the course of time, these hashtags will change, and the list should therefore be adjusted for future research. Moreover, focusing on Twitter, the data does not reflect the general public. The conclusions drawn from our research therefore need to be treated with caution, bearing in mind that offline information flow on the same topic might be structured differently. Furthermore, by using English tweets, non-english, more localized communication was not grasped. Future research should investigate in more detail the specific discussion topics that are prominent in the network, for example with topic modeling. Furthermore, main sources of information and their quality could be analyzed.

This study holds important implications both for research as well as practitioners. Findings on community structure can be used to provide group-specific and tailored information to users. For example, community 3 of the giant component is concerned with neuroscience and the effect of sex and gender variable on the brain. Here, public health managers could seed information on the topic of cardiology to broaden information diversity for members of that community. From a user perspective, the results of the analysis enable existing Twitter users to re-position themselves to become more influential and reach a broader audience. Further, knowledge of information routes can also be used to strategically position and introduce new users. These should be located where there are structural holes and where no prior individual is located (Haythornthwaite 1996). Furthermore, the data could also be used to map influential users, which could then yield as multipliers of information as they are able to diffuse information much faster and more efficiently than other users in the network. In addition to this, results showed that some key terms lack widespread adaption. For example, the term #MedicalBias was only used in 71 English tweets in the five-month research period. To support coordinated interaction, online marketing agencies from national health institutes, for example, could launch a top-down distribution of a single hashtag through promotion and online marketing that bundles gender medicine content on Twitter.

In this paper, we investigated the circulation and dissemination of content on gender-specific medicine on Twitter by examining the network structure. As a dispersed network with a high community structure, information runs slowly through the network and largely remains within groups. Yet, its egalitarian structure reveals a diversity of input points, as users do not depend on single users for their information.

Acknowledgements

This work has been funded by the Federal Ministry of Education and Research of Germany (BMBF) under grant no. 16DII127 ("Deutsches Internet-Institut").

REFERENCES

Allen, J. D., Hollander, J., Gualtieri, L., Alarcon Falconi, T. M., Savir, S., and Agénor, M. (2020) Feasibility of a Twitter Campaign to Promote HPV Vaccine Uptake among Racially/Ethnically Diverse Young Adult Women Living in Public Housing, BMC Public Health, 20, 1, 830.

Baggio, G., Corsini, A., Floreani, A., Giannini, S., and Zagonel, V. 2013. "Gender Medicine: A Task for the Third Millennium," *Clinical Chemistry and Laboratory Medicine* (51:4), pp. 713–727.

Barabási, A.-L.2009. "Scale-Free Networks: A Decade and Beyond," *Science*, (325:5939), pp.412–413.

Barabási, A.-L. 2016. *Network Science*, Cambridge University Press.

Berry, N., Lobban, F., Belousov, M., Emsley, R., Nenadic, G., and Bucci, S. (2017) #WhyWeTweetMH: Understanding Why People Use Twitter to Discuss Mental Health Problems, Journal of Medical Internet Research, 19, 4, e6173.

Bruns, Axel, and Burgess, J. 2011. "The Use of Twitter Hashtags in the Formation of Ad Hoc Publics," in *Proceedings of the 6th European Consortium for Political Research General Conference, 2011*.

Conley, C. C., Goyal, N. G., and Brown, S.-A. (2020) "#CardioOncology: Twitter Chat as a Mechanism for Increasing Awareness of Heart Health for Cancer Patients," Cardio-Oncology, 6, 1, 19.

Chen, X., Hay, J. L., Waters, E. A., Kiviniemi, M. T., Biddle, C., Schofield, E., Li, Y., Kaphingst, K., and Orom, H. 2018. "Health Literacy and Use and Trust in Health Information," *Journal of Health Communication* (23:8), pp. 724–734.

Chew, C., and Eysenbach, G. 2010. "Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak," *PloS One* (5:11), p. e14118.

Choo, E. K., Ranney, M. L., Chan, T. M., Trueger, N. S., Walsh, A. E., Tegtmeyer, K., McNamara, S. O., Choi, R. Y., and Carroll, C. L. 2015. "Twitter as a Tool for Communication and Knowledge Exchange in Academic Medicine: A Guide for Skeptics and Novices," *Medical Teacher* (37:5), pp. 411–416.

Chu, K.-H., Allem, J.-P., Unger, J. B., Cruz, T. B., Akbarpour, M., and Kirkpatrick, M. G. 2019. "Strategies to Find Audience Segments on Twitter for E-Cigarette Education Campaigns," *Addictive Behaviors* (91), pp. 222–226.

Chung, J. E. 2017. "Retweeting in Health Promotion: Analysis of Tweets about Breast Cancer Awareness Month," *Computers in Human Behavior* (74), pp. 112–119.

Clauset, A., Newman, M. E. J., and Moore, C. 2004. "Finding Community Structure in Very Large Networks," *Physical Review E* (70:6), p. 066111.

Clayton, J. A. 2016. "Studying Both Sexes: A Guiding Principle for Biomedicine," *FASEB Journal* (30:2), pp. 519–524.

Freeman, L. 2004. *The Development of Social Network Analysis*, Empirical Press, Vancouver.

Fruchterman, T. M. J., and Reingold, E. M. (1991) Graph Drawing by Force-Directed Placement, Software: Practice and Experience, 21, 11, 1129–1164.

Gebhard, C., Regitz-Zagrosek, V., Neuhauser, H. K., Morgan, R., and Klein, S. L. 2020. "Impact of Sex and Gender on COVID-19 Outcomes in Europe," *Biology of Sex Differences* (11:29), p 1-13.

Gleason, B. 2013. "#Occupy Wall Street: Exploring Informal Learning About a Social Movement on Twitter," *American Behavioral Scientist* (57:7), pp. 966–982.

Hagberg, A., Schult, D., and Swart, P. 2020. *NetworkX*. (https://networkx.org/documentation/networkx-1.10/download.html, /, accessed February 18, 2022).

Haunschild, R., Bornmann, L., Potnis, D., and Tahamtan, I. 2021. "Investigating Diffusion of Scientific Knowledge on Twitter: A Study of Topic Networks of Opioid Publications," *ArXiv:2101.11483*

Harding, K., Aryeetey, R., Carroll, G., Lasisi, O., Pérez-Escamilla, R., and Young, M. (2020= Breastfeed4Ghana: Design and Evaluation of an Innovative Social Media Campaign, Maternal & Child Nutrition, 16, 2, e12909.

Haythornthwaite, C. 1996. "Social Network Analysis: An Approach and Technique for the Study of Information Exchange," *Library & Information Science Research* (18:4), pp. 323–342.

Himelboim, I., and Han, J. Y. 2014. "Cancer Talk on Twitter: Community Structure and Information Sources in Breast and Prostate Cancer Social Networks," *Journal of Health Communication* (19:2), pp. 210–225.

Himelboim, I., Smith, M. A., Rainie, L., Shneiderman, B., and Espina, C. 2017. "Classifying Twitter Topic-Networks Using Social Network Analysis," *Social Media + Society* (3:1), p. 1-13

Himelboim, I., Xiao, X., Lee, D. K. L., Wang, M. Y., and Borah, P. 2020. "A Social Networks Approach to Understanding Vaccine Conversations on Twitter: Network Clusters, Sentiment, and Certainty in HPV Social Networks," *Health Communication* (35:5), pp. 607–615.

Legato, M. J. 2003. "Beyond Women's Health: The New Discipline of Gender-Specific Medicine," *Medical Clinics of North America* (87:5), pp. 917–937.

Legato, M. J. (ed.). 2009. *Principles of Gender-Specific Medicine - Gender in the Genomic Era*, San Diego: Academic Press.

Mauvais-Jarvis, F., Merz, N. B., Barnes, P. J., Brinton, R. D., Carrero, J.-J., DeMeo, D. L., Vries, G. J. D., Epperson, C. N., Govindan, R., Klein, S. L., Lonardo, A., Maki, P. M., McCullough, L. D., Regitz-Zagrosek, V., Regensteiner, J. G., Rubin, J. B., Sandberg, K., and Suzuki, A. 2020. "Sex and Gender: Modifiers of Health, Disease, and Medicine," *The Lancet* (396:10250), pp. 565–582.

Makita, M., Mas-Bleda, A., Morris, S., and Thelwall, M. (2021) Mental Health Discourses on Twitter during Mental Health Awareness Week, Issues in Mental Health Nursing, 42, 5, 437–450.

Milani, E., Weitkamp, E., and Webb, P. 2020. "The Visual Vaccine Debate on Twitter: A Social Network Analysis," *Media and Communication* (8:2), pp. 364–375.

Moukarzel, S., Rehm, M., Fresno, M. del, and Daly, A. J. 2020. "Diffusing Science through Social Networks: The Case of Breastfeeding Communication on Twitter," *PLOS ONE* (15:8), p. e0237471.

NetworkX. 2021. "Networkx.Algorithms.Community.Modularity_max.Greedy_modularity_communities — NetworkX 2.6.2 Documentation," *NetworkX 2.6.2 Documentation*.

Newman, M. E. J. 2004. "Fast Algorithm for Detecting Community Structure in Networks," *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* (69:6), p. 066133.

Park, H., Rodgers, S., and Stemmle, J. 2013. "Analyzing Health Organizations' Use of Twitter for Promoting Health Literacy," *Journal of Health Communication* (18:4), pp. 410–425.

Regitz-Zagrosek, V. 2012. "Sex and Gender Differences in Health," *EMBO Reports* (13:7), pp. 596–603.

Roy, M., Moreau, N., Rousseau, C., Mercier, A., Wilson, A., and Atlani-Duault, L. 2020. "Ebola and Localized Blame on Social Media: Analysis of Twitter and Facebook Conversations During the 2014–2015 Ebola Epidemic," *Culture, Medicine, and Psychiatry* (44:1), pp. 56–79.

Singh, L., Bode, L., Budak, C., Kawintiranon, K., Padden, C., and Vraga, E. 2020. "Understanding High- and Low-Quality URL Sharing on COVID-19 Twitter Streams," *Journal of Computational Social Science* (3:2), pp. 343–366.

Smith, M. A., Rainie, L., Shneidermann, B., and Himelboim, I. 2014. "Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters," *Pew Research Center: Internet, Science & Tech*, (https://www.pewresearch.org/internet/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/, accessed August 15, 2021).

Wang, Y., McKee, M., Torbica, A., and Stuckler, D. 2019. "Systematic Literature Review on the Spread of Health-Related Misinformation on Social Media", *Social Science & Medicine*, 240, p. 112552.

Wang, X., Liang, G., Zhang, Y., Blanton, H., Bessinger, Z., and Jacobs, N. 2020. "Inconsistent Performance of Deep Learning Models on Mammogram Classification," *Journal of the American College of Radiology* (17:6), pp. 796–803.

Wassermann, S., and Faust, K. 1994. *Social Network Analysis - Methods and Applications*, (1st ed.), Cambridge University Press, Camebridge, England.

Webster, J., and Watson, R. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Quarterly* (26:2), pp. xiii–xxiii.

Xu, W. W., Chiu, I.-H., Chen, Y., and Mukherjee, T. 2015. "Twitter Hashtags for Health: Applying Network and Content Analyses to Understand the Health Knowledge Sharing in a Twitter-Based Community of Practice," *Quality & Quantity* (49:4), pp. 1361–1380.