



# The Tip of the Iceberg

## Medical Misinformation and the Internet

Jeanne Reppert

December 3, 2020

---

## 0.1 Project Abstract

This paper investigates content from online activities related to Joseph Mercola's website along with his social media activity on Facebook and Twitter. Mercola is one of many medical professionals who host websites that provide "alternative" health advice and frequently market health related products. During the Covid-19 pandemic, many websites such as Mercola's have used issues pertaining to Covid-19 to attract readers and followers. Often the information they share is not widely accepted by the larger medical community and is often viewed as medical misinformation. Prior to Covid-19, these same websites pressed an anti-vaccine agenda while also offering more innocuous advice about vitamins and broader health related practices. The full archive of articles found on Mercola's website have been scraped and a selection of them are studied to gain an understanding of popular content on the site. Additionally, all of Mercola's Facebook page content was investigated along with available Mercola tweets for 2020.

Three major observations were made regarding the data. First, Mercola's articles include a variety of more general health related material but also promote some very specific anti-vaccination content and content that embraces several common conspiracy theories related to Covid-19. Mercola's strategy appears to be to use his social media platforms to attract subscribers to his site so that he can share daily information through email and SMS messaging. Second, Facebook appears to have had some level of success in limiting the spread of misinformation by Mercola and its measures led to Mercola formally ceasing to use his group page in mid-2019 (although it still exists and directs users to his website.) Finally, Twitter appears to be less successful in stemming the propagation of misinformation by Mercola who utilizes the platform both to share a wide variety of Covid-19 misinformation and to direct followers to his website for more detailed misinformation and stories that feature common conspiracies.

This study serves as a primary descriptive tool in the investigation of medical professionals who have harnessed the internet and social media to spread misinformation regarding popular health issues and in 2020, the Covid-19 pandemic. Future studies should both focus on other popular figures who share responsibility for spreading anti-vaccination material and Covid-19 related conspiracy theories. Furthermore, additional studies are needed to analyze the activities of the receivers of the messaging promoted by these individuals.

## 0.2 Introduction

As the Covid-19 pandemic has unfolded, the popularity of websites reputed for spreading misinformation has grown significantly. Some attribute this growth to the initial lack of information about the virus and the adjustments that medical professionals made to public policies as new information was revealed. Distrust has grown as public policies encroached on citizen's perceived freedoms. Anti-vaxxers, who have strong organizational structures and an existing following, harnessed public distrust to amplify the uncertainty surrounding Covid-19 related policies. Many leaders propagate conspiracy theories and use fear tactics to gain support. Key public health officials and philanthropists such as Bill and Melinda Gates have been targeted as villains.

Identifying and understanding online misinformation is essential for public health governance. In addition to identifying and flagging problematic websites, public health officials must directly refute conspiracy theories and specific scientific misinformation spread online. By understanding rhetorical patterns and themes utilized, a set of standards might be developed to assess websites and rate their reliability and trustworthiness.

This study investigates the website sponsored by Joseph Mercola (Mercola.com) and its relationship to health misinformation to understand the interaction between public health issues and the polarization surrounding Covid-19 and vaccinations. Mercola.com is known as one of the largest purveyors of "alternative" health information. His website has shared information through the publication of articles since 2008. Mercola earns his living by directing his followers to an online store where vitamins and other health related products are sold. According to Stephen Barrett of Quackwatch.org (Kreidler, 2020), Mercola had a net worth in 2017 of more than \$100 million.

Mercola's website states that he has more than a million subscribers. A typical email subscriber will receive one daily newsletter with links to three articles. While two of those articles usually reference a short description about a supplement or health practice, the third article will often contain information that often would typically fall under the category of misinformation. Typically topics in this category would be common conspiracy theories, anti-vaccination information, and debunked ideas pertaining to the Covid-19 virus.

### 0.3 Methods

In order to investigate the tactics Mercola uses to gather followers and spread health information (and misinformation), the full corpus of archived articles published by Mercola from 2008 to 2020 were scraped from his website. Utilizing RSelenium and rvest, 9306 URL's were identified for potential scraping. Some of the URL's identified had broken links and were eliminated. Additionally, URL's that were not associated with health articles or with little text (either recipes, exercise illustrations or videos) were also omitted. Once these URL's were eliminated, the text, date, and title were scraped from a total of 8676 articles and organized into CSV files by year. Additionally, CrowdTangle was used to gather all available posts made on Mercola's group Facebook page and a history of tweets from Mercola were gathered. The data was investigated to note patterns and trends in the texts of the website articles and how those compared to the social media content.

### 0.4 Hypotheses

Websites that pass on medical misinformation have a tendency to follow a pattern of communication. Most websites have some innocuous articles that pertain to health issues including articles about supplements, healthy eating, exercise, and health lifestyles. These articles may be attractive to those who are mildly interested in health topic but not typically drawn to conspiracy theories that stretch the imagination more than the idea of trying a vitamin or other supplement. These types of articles may be a type of lure or "gateway drug" that draws the user in. Most websites press the viewer to share an email in order to view articles and once that email is captured, most websites will send a daily news digest that highlights additional articles. Thus over time, the user is exposed to a variety of topics including the more nefarious conspiracy theories that would be categorized as medical misinformation. One might expect then to observe in the text data, that during times when uncertainty is more prevalent due to a health crisis, these topics may skew more towards current events and during times when uncertainty in public health is less of a concern that topics will skew more towards articles about healthy lifestyles. One may also hypothesize that social media sites, for this type of misinformation, are more utilized to form connections to readers and to draw traffic to the website rather than to necessarily gain converts to the social media site itself. The real value in a tweet or a facebook post in this case is not necessarily to generate discussion or retweets but to gain clicks that will lead users to the website for the purpose of capturing their email and following up with additional emails and ultimately product advertisements.

### 0.5 Dataset descriptions

For this study, due to the large volume of text involved with 13 years of articles, four years of articles were chosen for study. Each year represents a time frame when the United States was tasked with fighting a new virus that emerged from worldwide sources. In 2009, the H1N1 flu virus was considered to be at pandemic levels in the United States. The Ebola virus reached the shores of the United States in 2012 albeit in very small numbers, and in 2016 the Zika virus raised concerns around health professionals. In 2020, the SarsCov2 virus sparked the most serious pandemic throughout the United States since the 1918 Spanish flu pandemic.

Since Mercola's site was first launched in 2008, his 2009 article archive is relatively small with a total of 231 articles. By 2012 his collection for the year had grown to 639 years. In 2016, 793 articles were published and made available in the archived material and in 2020 (up to the beginning of November) the number of articles available was already up to 730 and is certain to grow over the next few weeks. According to the

website, Mercola “fact checks” his archives which leads one to believe that he has a staff of ghost writers who are responsible for his website content.

The collection of Facebook archived posts are comprised of a csv file of 25,027 rows of data. Variables include number of page likes, number of comments, post likes, other post responses, shared links, videos, number of times videos are viewed, the post text and a score for overperformance. Most posts included range from mid-2009 to August 2019. His page has 1.78 million followers and 1.79 million likes. As of August 20, 2019, Mercola stopped posting on his Facebook page although posts by Mercola are still widely circulated on Facebook. He announced that he was leaving Facebook and includes a link to his website for more information. On his website, he encourages members to subscribe to his email, SMS messages and podcasts for daily information. Additionally, he encouraged his followers to find him on Parler. Mercola still maintains several Facebook groups in other languages. His Facebook page in Spanish has almost 1 million followers.

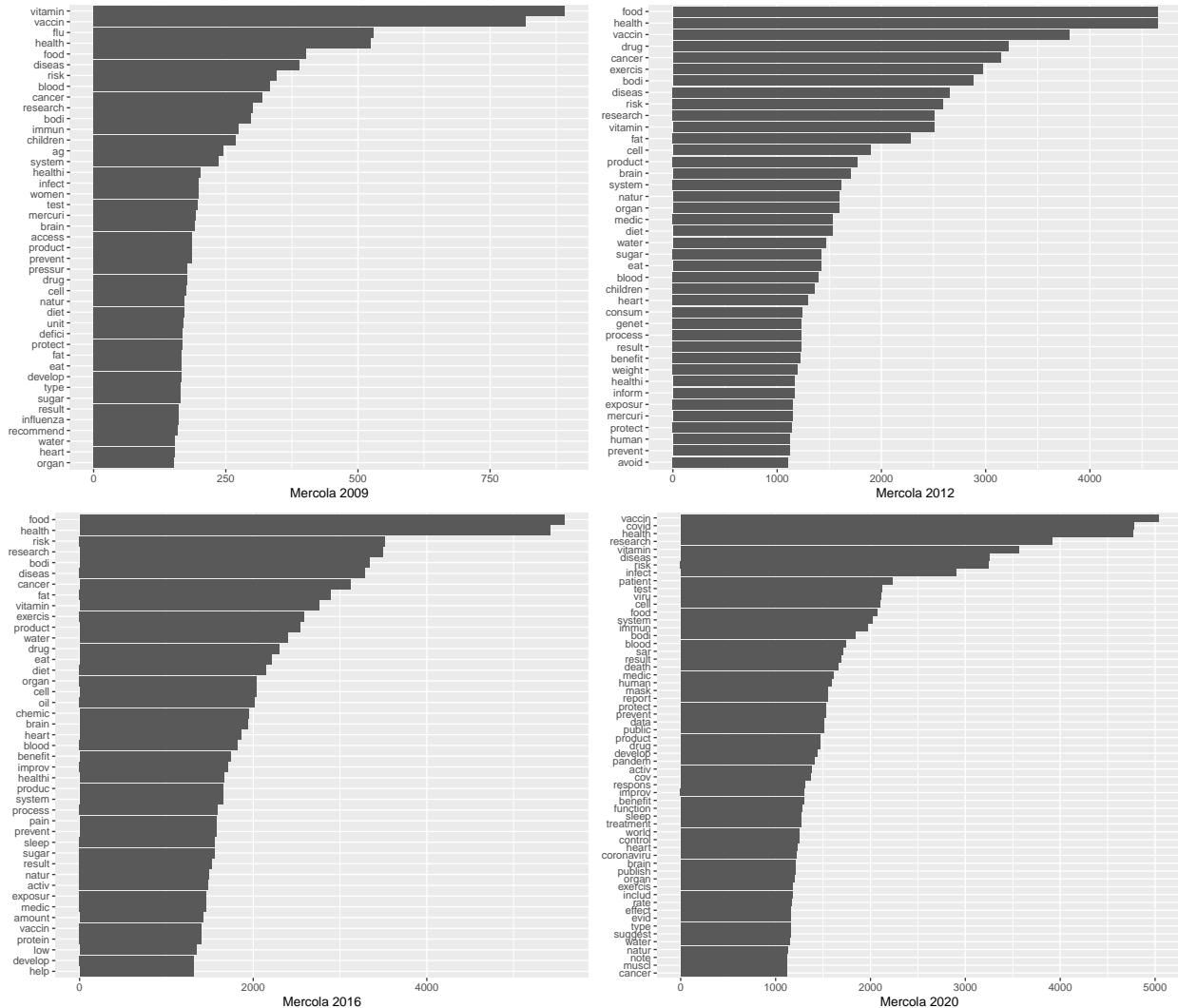
Twitter history for Mercola is more limited as Twitter limits the number of tweets that may be gathered to 3200. This smaller set for this reason, is comprised of tweets from July 4th, 2019 to November 13th, 2020. Variables include the tweet date, text, user description, total favorites, total followers, total friends, user location, user name, total statuses, tweet mentions, retweet mentions, tweeted urls’s, url retweets and total retweets. Mercola has a total of 290.1K followers are his Twitter page. Mostly, Mercola tweets out his web page article links on a daily basis.

## 0.6 Tools and Libraries Used

A variety of r packages have been used to gather and clean the data collected. RSelenium and rvest have been utilized to scrape data from the website. Massmine was used to gather tweets and CrowdTangle was used to gather data about Facebook posts. Tidyverse, tidytext, widyr, tm, dplyr, lubridate, readr and stringr were all utilized to process, clean and sort the data. Snowballc was used for stemming text and ggplot2, gridExtra, igraph, ggraph, wordcloud and plotly were all utilized to produce various visualizations. Sentiment analysis was performed utilized the syuzhet package and the package fs was used to load and manage files. For analysis of twitter co-occurrences the readtext and quanteda package were used.

## 0.7 Visualization of Mercola Website Articles: Top Words for Each Year

Given that Mercola had significantly fewer articles in 2009 and the articles in his first couple of years were less verbose, the data for 2009 is quite different in number in comparison to the other three years. This needs to be kept in mind when investigating the results. In all four years, the words vitamin, health and food all appear to figure heavily in the articles surveyed. Risk, disease and research also appear as some of the top words for all four years. In 2009, 2012, and 2020, the word vaccine appears as one of the top three words in each year with nearly 800 mentions (out of 231 articles), but more than 4500 mentions (out of 639 articles.) In 2016, when 793 articles were published, the number of mentions of vaccines drops to around 1000, but rises again in 2020, when greater than 5000 mentions of vaccines are made throughout 730 articles. In 2009 and 2012, closely associated with the occurrences of the word vaccine, the word mercury also occurs. In 2009 flu and its counterpart influenza is the third most mentioned word (the year when H1N1 peaked in the U.S.) However, neither Zika or Ebola are found in the top words in the years where this virus was an issue in the U.S. perhaps because not as many individuals were directly impacted by these viruses. In 2020, in particular, the words covid being the second most frequently mentioned word with more than 4500 mentions alone. Associated words such as sars, coronavirus, masks, pandemic, and virus also are top words. While not found in the list of top words in 2020, cancer is also a top mentioned word in 2009, 2012, and 2016.

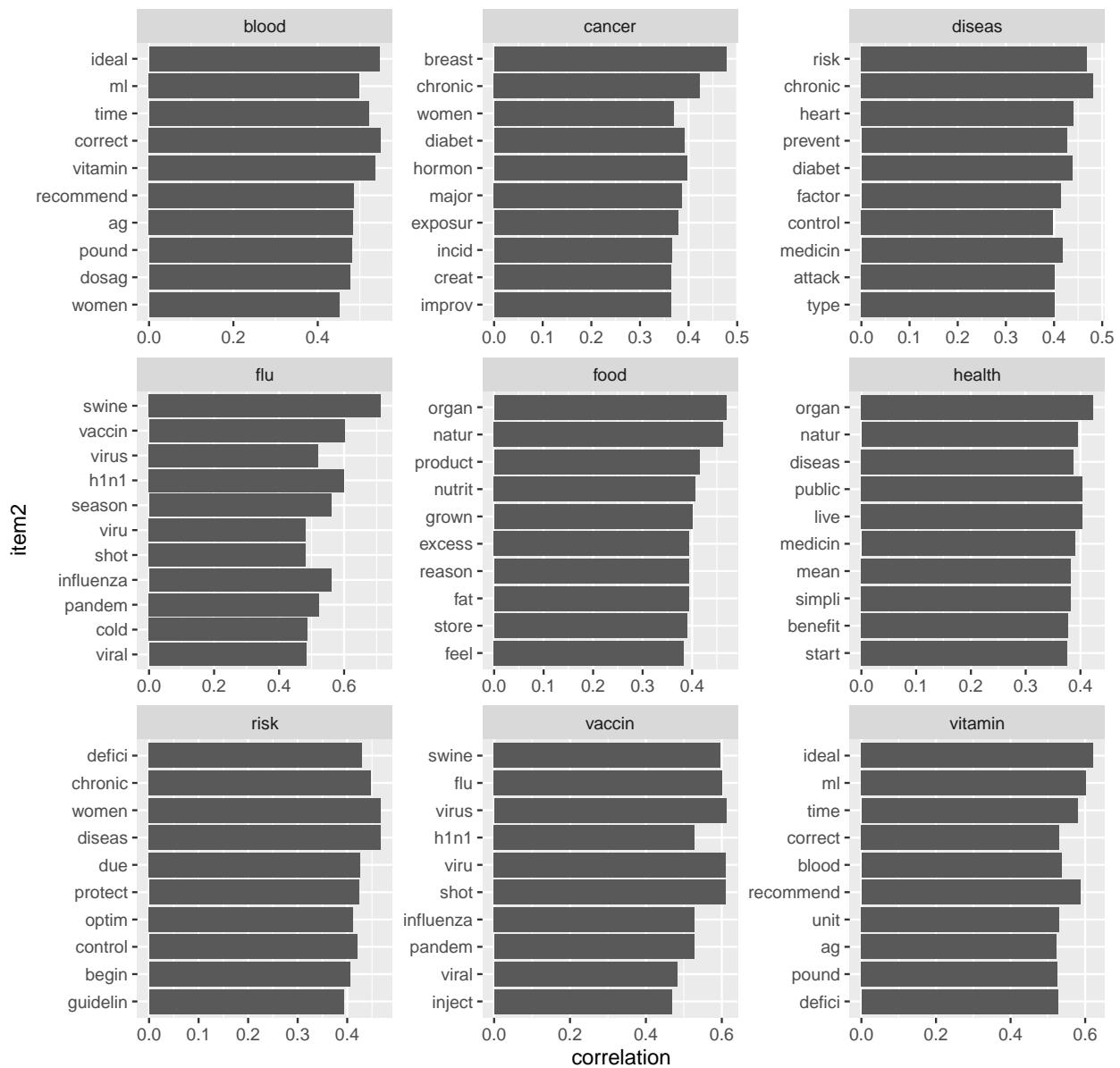


## 0.8 Visualization of Mercola Website Articles: Word Correlations with Top 9 Words

### 0.8.1 Top correlated words for Mercola Articles in 2009

Two primary topics appear to emerge when investigating word correlations for the Mercola 2009 dataset. The first topic appears to be centered around health related topics. For most of the top 9 words, correlated words are associated with major diseases, women's issues, vitamins, and food appear most frequently. Evidence of discussion regarding the H1N1 virus is also present and appears to center around the flu vaccine. Generally the correlated words do not appear to imply danger or any sort of ominous feeling.

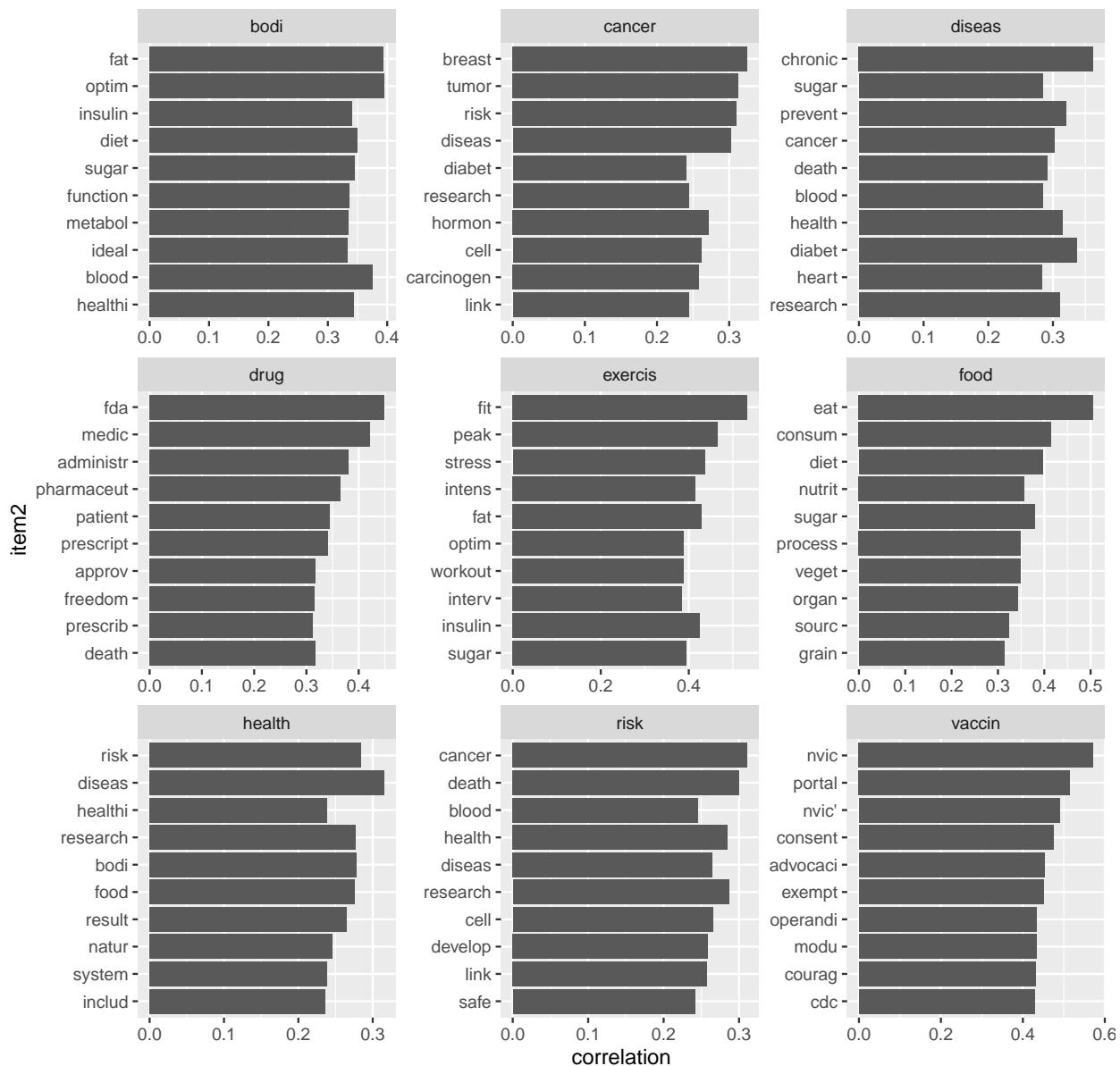
### Mercola 2009



### 0.8.2 Top Correlated Words for Mercola Articles in 2012

Top correlated words in this dataset appear to be more varied. Primary topics still appear to revolve around health related topics with some attention to diet, exercise and nutrition. However, in this dataset there appears to be a greater emphasis on health malfunction particularly as it pertains to cancer, diabetes, and heart issues. The words associated with cancer, drug, disease, risk and vaccine all seem to connotate a more somber discussion with associated words like death, safe, chronic, and carcinogen included in the correlated word list. Interestingly, the vaccine list of correlated words contains two references to nvic which is the vaccine study organization that is funded by Mercola. This group focuses on descriptions and arguments about the “dangers” of vaccines. Additionally, two other organizations are mentioned: the FDA and CDC.

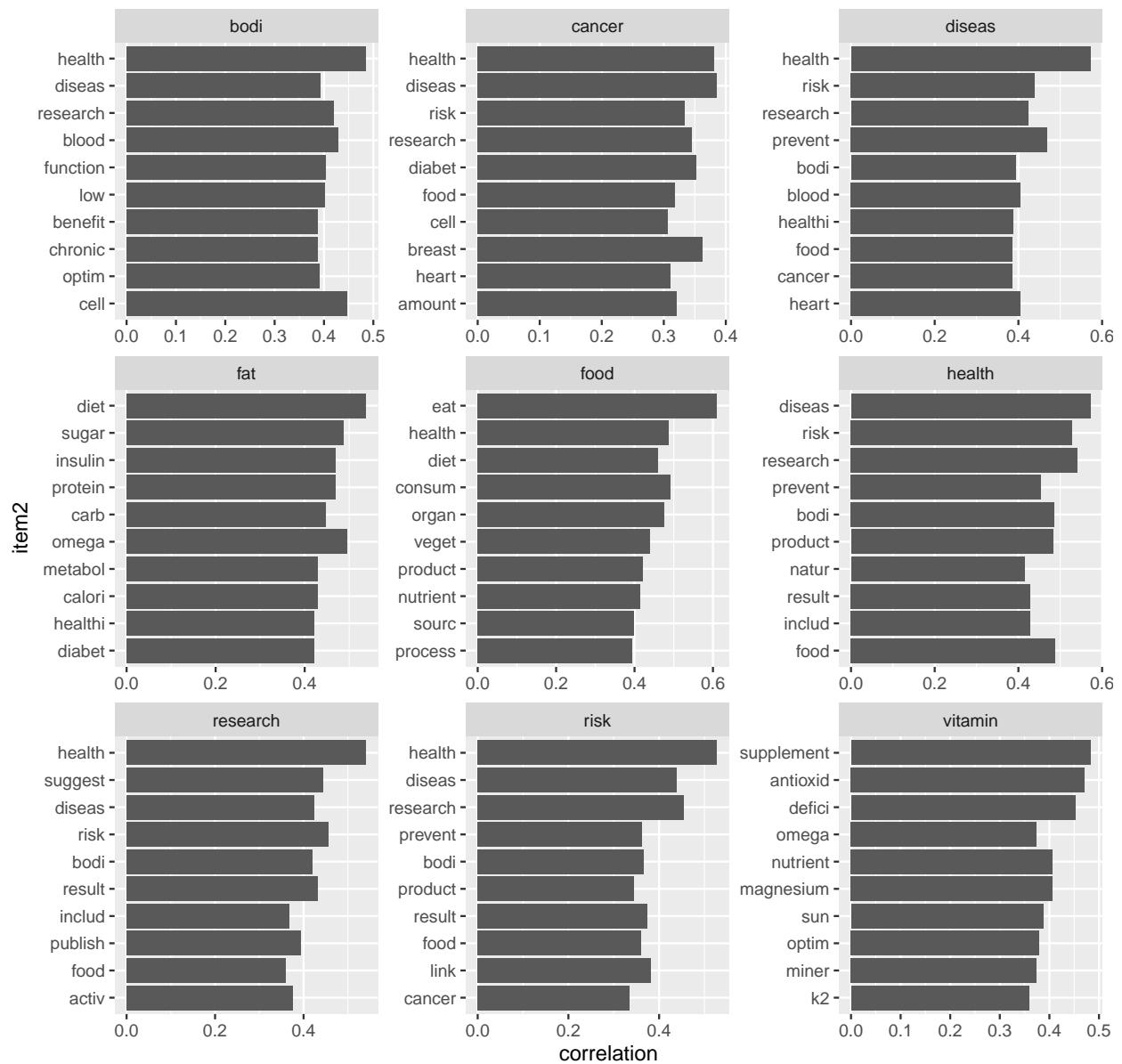
### Mercola 2012



### 0.8.3 Top Correlated Words for Mercola Articles in 2016

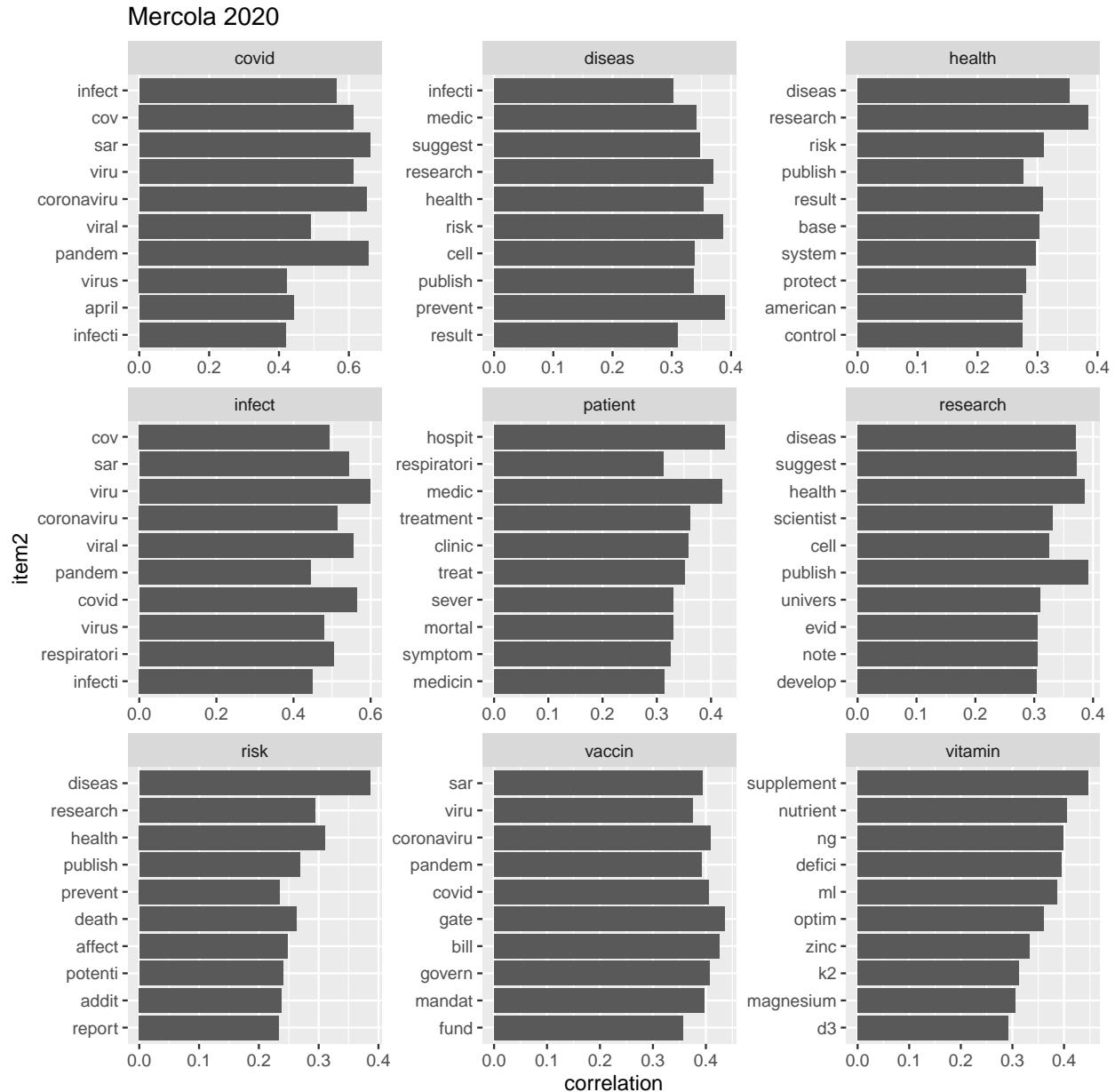
Top correlated words in the 2016 articles again appear to move towards more light hearted subjects. Emphasis is placed on diet, nutrition, natural products, and more specific mentions of vitamins. Even categories for cancer, risk and disease appear to have less of a variety of words and the topics still seem to revolve around food intake.

### Mercola 2016



#### 0.8.4 Top Correlated Words for Mercola Articles in 2020

Not surprisingly, 2020 has a unique list of top words in Mercola's articles and their correlated words connotate a far more serious discussion. The topics of the articles have almost fully departed from health, nutrition, and exercise topics and have switched to discussions that involve aspects of the pandemic. The correlated words for vaccine include both references to the pandemic but also to Bill Gates, the government and mandates. In the only mention of nutrition in the correlated words for vitamin, specific vitamins that have been popularly touted for immune building effects against covid are listed (vitamin D and Zinc.) More ominous words such as respiratory, severe, mortal, and infection and more commonly mentioned.



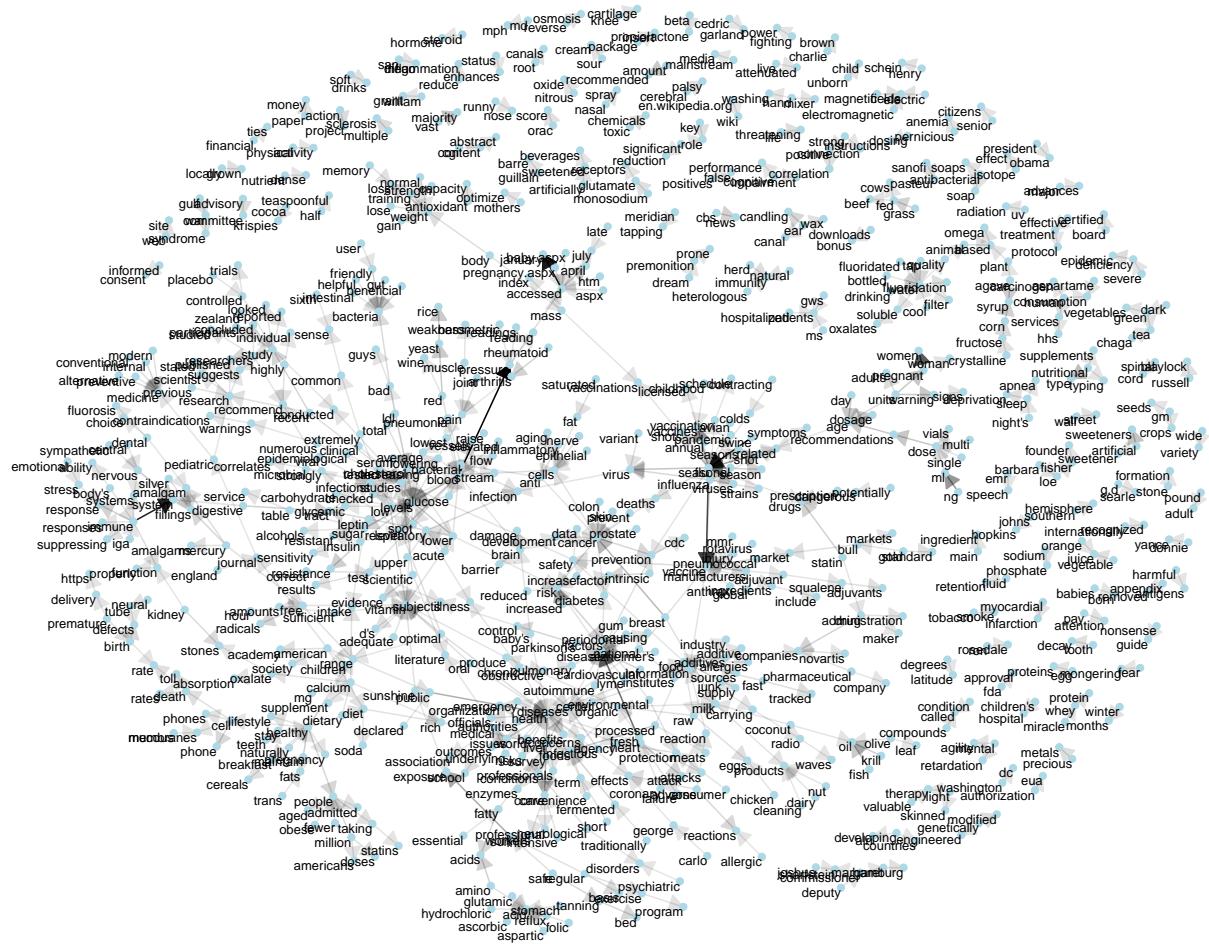
## 0.9 Visualization: Top Bigram Relationships for Mercola Website Articles

### 0.9.1 Bigram Relationships for 2009 Mercola Articles

Bigram relationships show the frequency that related words occur together in the scraped articles. Due to fewer articles in the 2009 collection, any bigrams that occurred greater than 3 times are included in the visualization. The bigram relationships for Mercola's 2009 articles closely mirror the observations made regarding the most frequent words and correlated words. Several closely related word groupings are noted. Each of the connected larger clusters appear to revolve around health issues with two of the larger ones addressing broader health topics and then two others referring to topics revolving around vaccinations, virus, and influenza. Closely associated in one of the vaccination groupings is the word "injury" indicating that these topics likely revolve around discussions of perceived vaccine dangers. Two of the less prominent

clusters discuss sleep and dental issues with references to amalgam fillings and mercury. The remaining smaller groupings tend to focus on nutrition, supplements, health, and exercise topics.

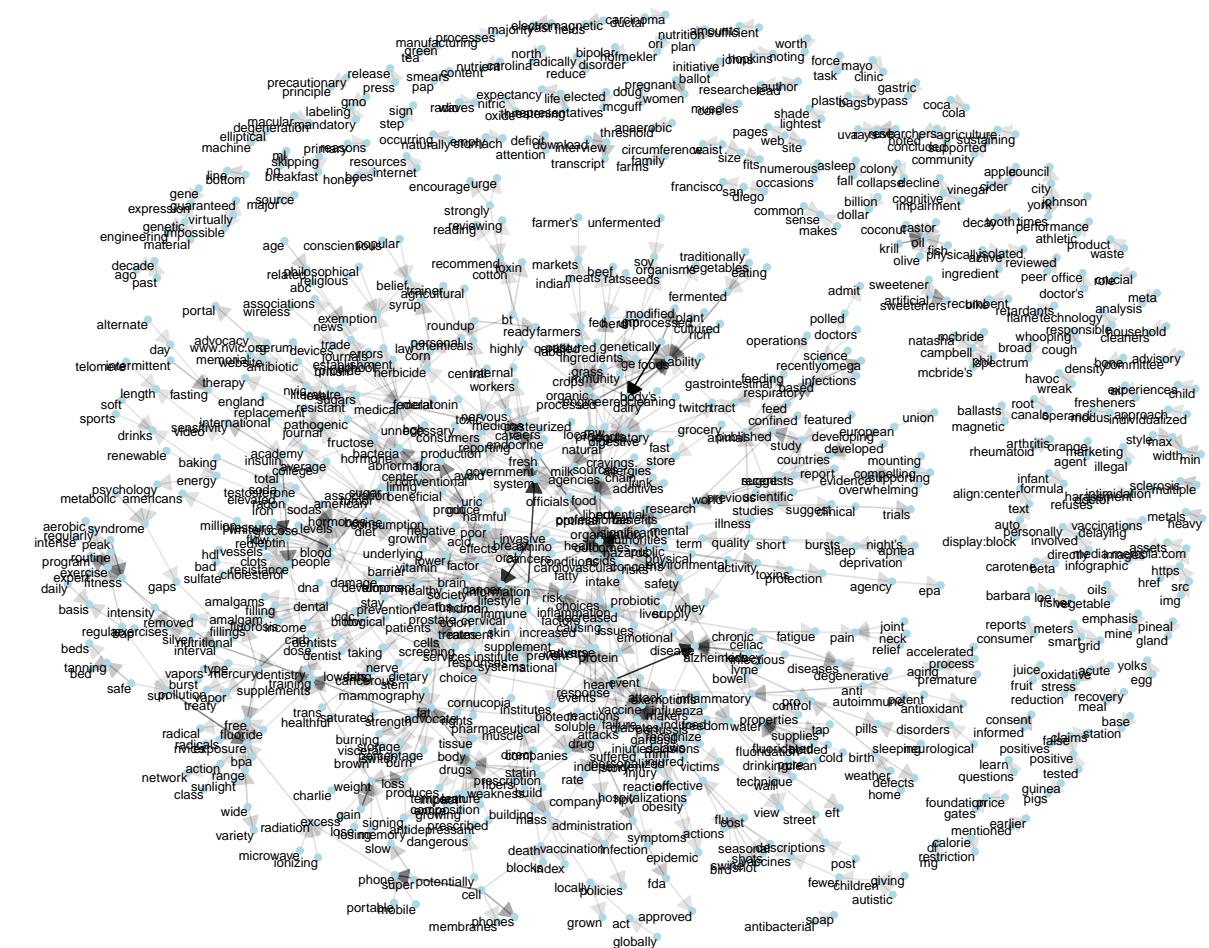
#### Mercola Bigrams 2009



#### 0.9.2 Bigram Relationships for 2012 Mercola Articles

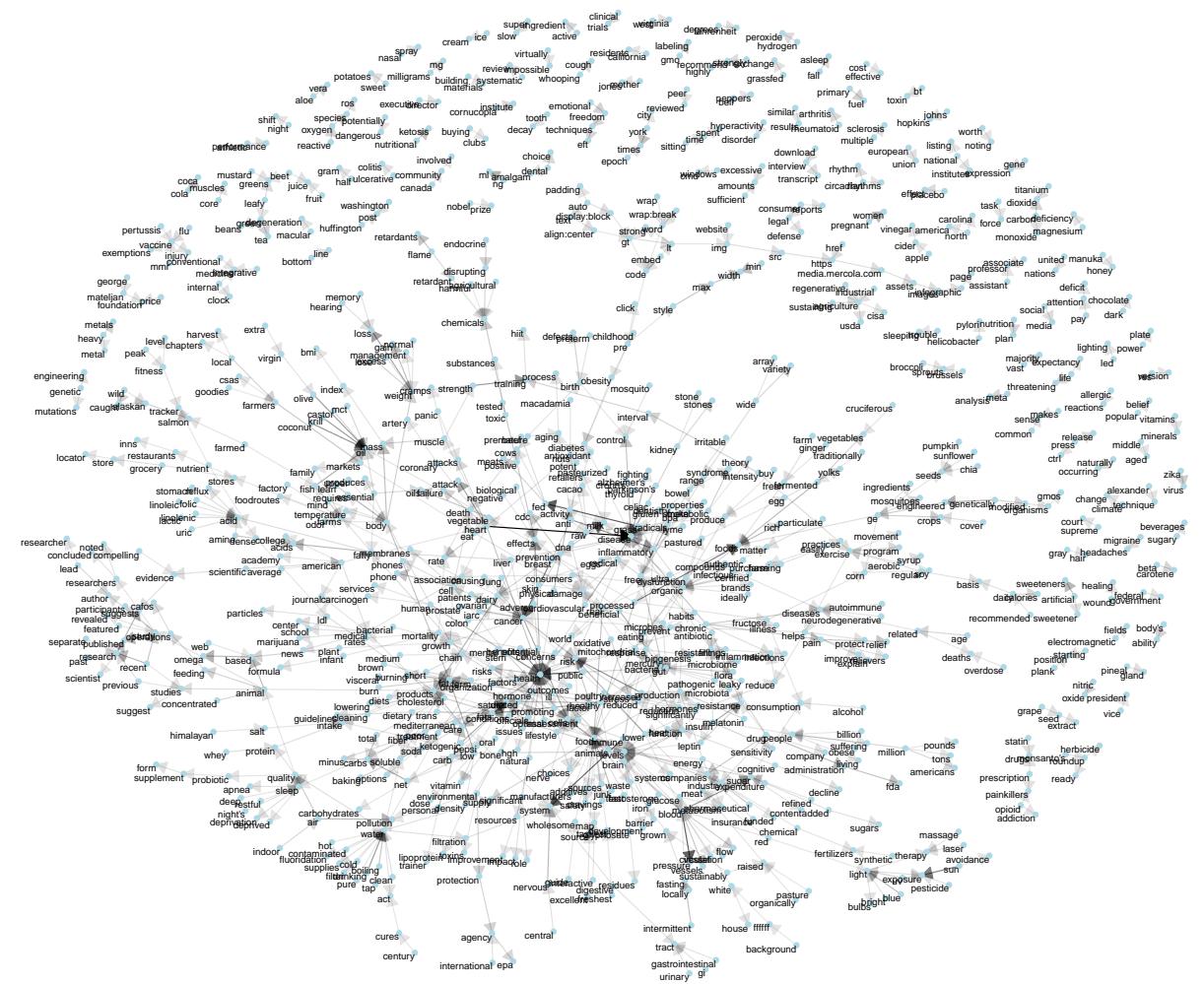
For the 2012 Mercola articles, the cutoff of bigram articles was increased to bigrams that occur more than 19 times. The central cluster for the 2012 bigram visualization appears to focus more broadly on the word “health” but with words like hazards, negative, attacks, failure, liberty, harmful, and cancer closely associated with the center of this cluster. Another cluster focuses on vaccines and include references to mmr, guardisal, epidemic,injuries, reactions, and influenza. Attacks, injury, victims, and reactions are all words that are also associated with this cluster. Other clusters appear to reference health problems with a focus on cancer and dental health including fluoride and mercury references. Another cluster of interest is centered around farm related topics and references words such as engineered, dairy, immunity, raw, allergies, pasteurized, and organic. Again there is many associated bigrams that reference topics such as vitamins, nutrition, and supplements. Of interest is a small cluster revolving around NVIC, the anit-vaxx organization that Mercola donates millions of dollars too and which references many of Mercola’s articles on their site (Satija & Sun, 2019.)

## Mercola Bigrams 2012



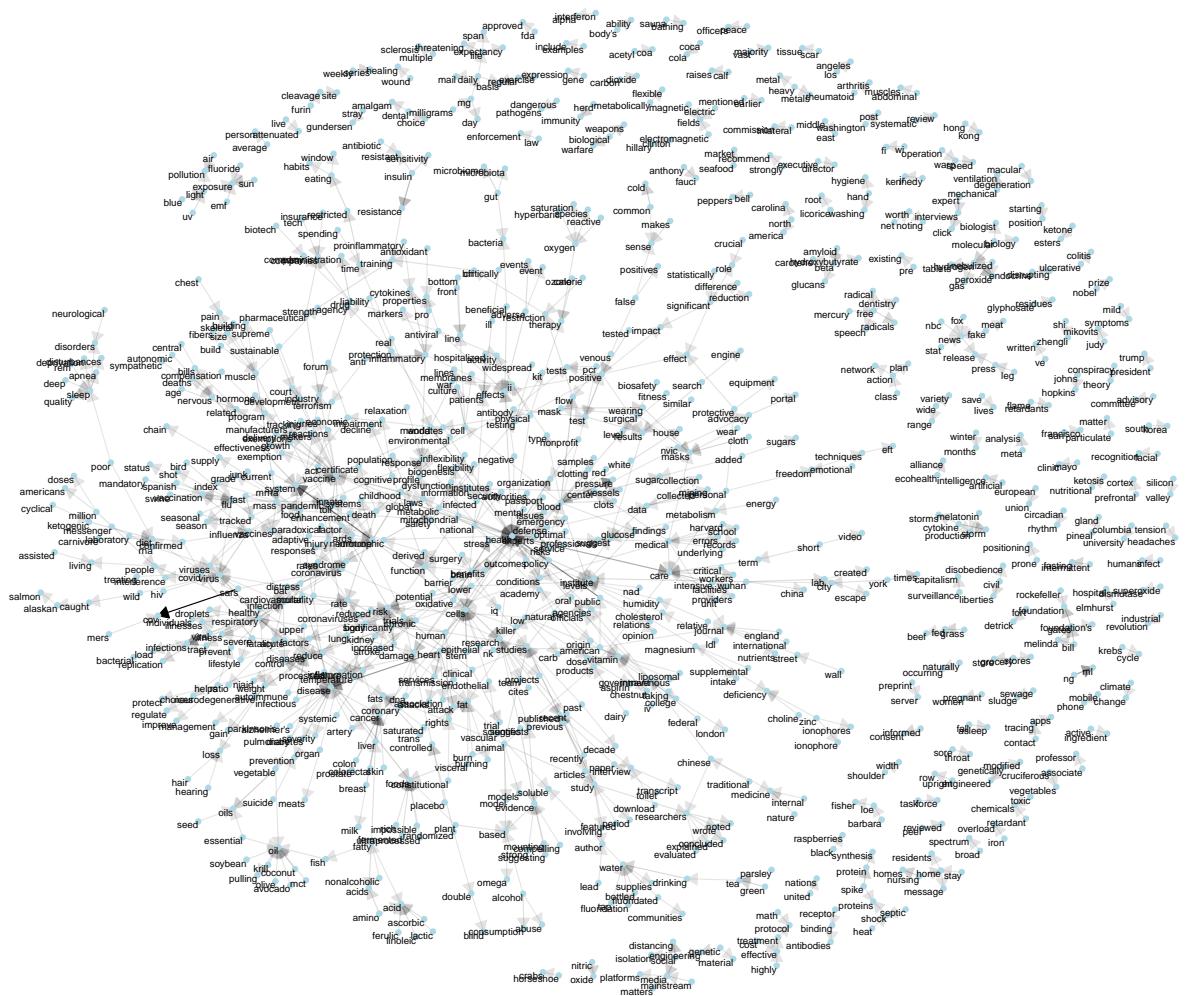
### 0.9.3 Bigram Relationships for 2016 Mercola Articles

The cutoff for bigram relationships in 2016 Mercola were set to relationships that occurred greater than 22 times. In 2016 the bigrams visualized in the primary clusters focus on broader and more general health topics. Few are associated with specific health problems and issues that were central in the 2012 visualizations like dental, sleep, fluoride, disease, and vaccine issues have moved to the periphery of the 2016 visualization. Bigrams related to dietary issues (carbs, fibers, ketogenic diets, hormones) and specific foods (fish, ginger, dairy, plants) appear more centrally within the visualization. The peripheral bigrams name a host of foods and supplements.



#### 0.9.4 Bigram Relationships for 2020 Mercola Articles

The cutoff used for bigram relationship for 2020 Mercola articles was bigram relationships that occurred greater than 18 times. Not surprisingly, this visualization illustrates the discussions around Covid-19 that understandably dominated the content of the 2020 archive of articles. The bigram relationships in the 2020 visualization appear to contain two primary densely populated clusters that are surrounded by a large variety of smaller clusters with a majority of those cluster that appear to focus on discussions surrounding the Covid-19 pandemic. Issues regarding pcr tests, antibodies, masks cytokine storms, vaccines, Wuhan, and transmissions all surround the center clusters. The peripheral bigrams also reflect covid related subjects. Mentions of Fauci and Bill Gates are noted along with mentions of tpics such as herd immunity, hand-washing, distancing, isolation, civil liberties, prine positioning, contact tracing, nursing homes, biological warfare, dangerous pathogens, operation warp speed, conspiracy theory, bilateral commission, Fox and NBC news, Trump, and Judy Mikovits (the main scientists mentioned in the plandemic film). Far fewer nutritional, exercise or health references are observed among the bigrams and many of the existing references revolve around supplements that Mercola recommends controversially to boost immunicy to Covid-19 including zinc, nitric oxide, and melatonin. The topical shift in this visualization of 2020 articles is quite evident.



## 0.10 Visualization: Word Cloud - Title Sentiment Analysis For Mercola Website Articles

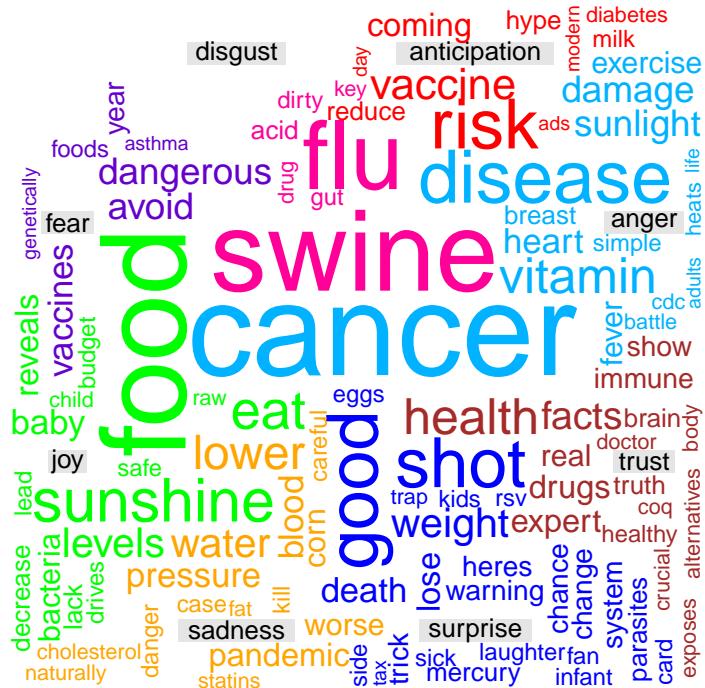
Another way to understand the content of the Mercola corpus is to investigate the titles of articles and their conveyed sentiment. This is of particular interest because these titles are what are displayed on social media posts and would be the attention grabbing material in daily emails and SMS texts. Like frequent words and bigrams, sentimental analysis of titles in the Mercola corpus also reflects shifts from one annual grouping to the next.

### 0.10.1 Title Sentiment Analysis Word Cloud for 2009 Mercola Website Articles

The 2009 Mercola titles predominantly reflect sentiments of anger and joy with some reference to subjects that reflect surprise. The emotions of anticipation, sadness, and trust contain fewer words. The sentiment of disgust and fear appear to contain the least amount of mentions. Food and its associated words connote joy while cancer is the most frequently occurring word connoting anger. The prominence of the discussion of the flu in 2009 along with discussions of vaccines (interestingly quite close to the word avoid) figure largely

in discussions that connote fear and disgust. Words that imply surprise and anticipation appear to be a mix of negative and positive sentiment.

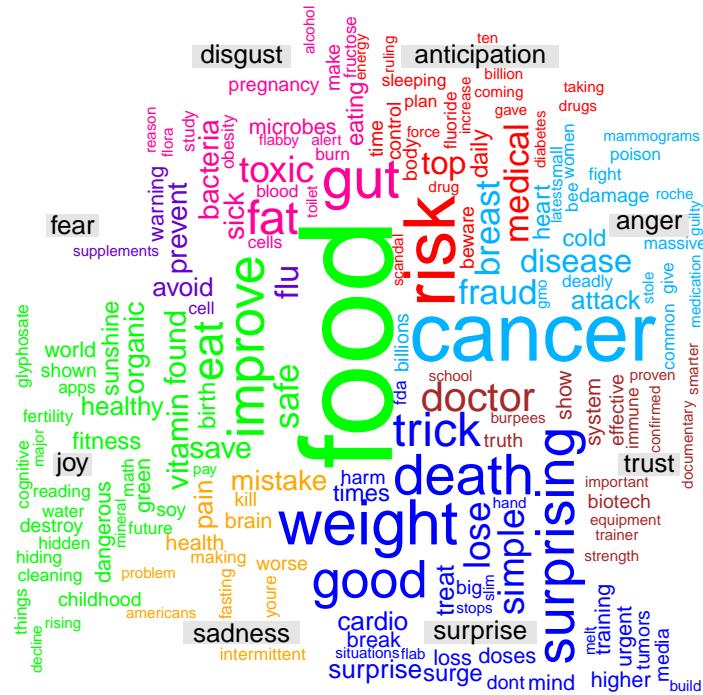
#### Mercola Website Articles – Title Sentiment for 2009



#### 0.10.2 Title Sentiment Analysis Word Cloud for 2012 Mercola Website Articles

The title sentiment for 2012 Mercola articles represent a small shift from 2009. Articles that represent disgust have increased. The word “flu” has shifted to the fear category. Food figures prominently in the joy category with more of a variety of associated words and cancer continues to be associated with anger. Attack and fraud are other prominent words in the anger category along with the category of breast cancer and mammograms. This may be indicative of the fact that Mercola has often touted the use of thermography as a substitute for mammograms and has been warned by the FDA to cease giving this medical guidance (Tsouderos, 2011). Words that connote trust, fear and sadness are relatively small categories. In this visualization vaccine shifts from a rather prominent position in the fear category in 2009 (the year of the H1N1 pandemic) to no longer being present in the chart. Once again the categories of anticipation and surprise appear to have a mix of words that contain negative and positive sentiments.

## Mercola Website Articles – Title Sentiment for 2012



### 0.10.3 Title Sentiment Analysis Word Cloud for 2016 Mercola Website Articles

The title sentiment visualization for Mercola articles in 2016 again represents a shift in topics. The sentiment categories of fear and trust are quite small with surprise being the largest category. Food and associated words figure prominently in the joy category while cancer and toxicity are most frequently mentioned in the disgust category. The word fraud has moved over to the anger category and the word vaccine has shifted into the anticipation category. Both of these categories appear to be a bit smaller than these categories in 2012.

## Mercola Website Articles – Title Sentiment for 2016



### 0.10.4 Title Sentiment Analysis Word Cloud for 2020 Mercola Articles

As in the other visualizations, the sentiment visualizations for 2020 Mercola articles shifts quite a bit. Food and its associated words still figure heavily in the joy category which is slightly smaller. Covid and death appear in the surprise category along with words like murder, pandemic, treat, death and research. In the disgust category infect, respiratory, immunity, viral and coronavirus are most frequently mentioned. While the fear category is fairly small the word prevent is most prominently followed by covid related words like shed, hazard, war, virus, avoid, and flu. Many words related to Covid-19 like patient, scientist, track, fraud, warp speed, Bill Gates, cytokine, danger, kill, and restrict all appear in the trust, anticipation, anger, and trust categories. Interestingly, the word vaccine appears in the trust category with few appearances than other words.

## Mercola Website Articles – Title Sentiment for 2020



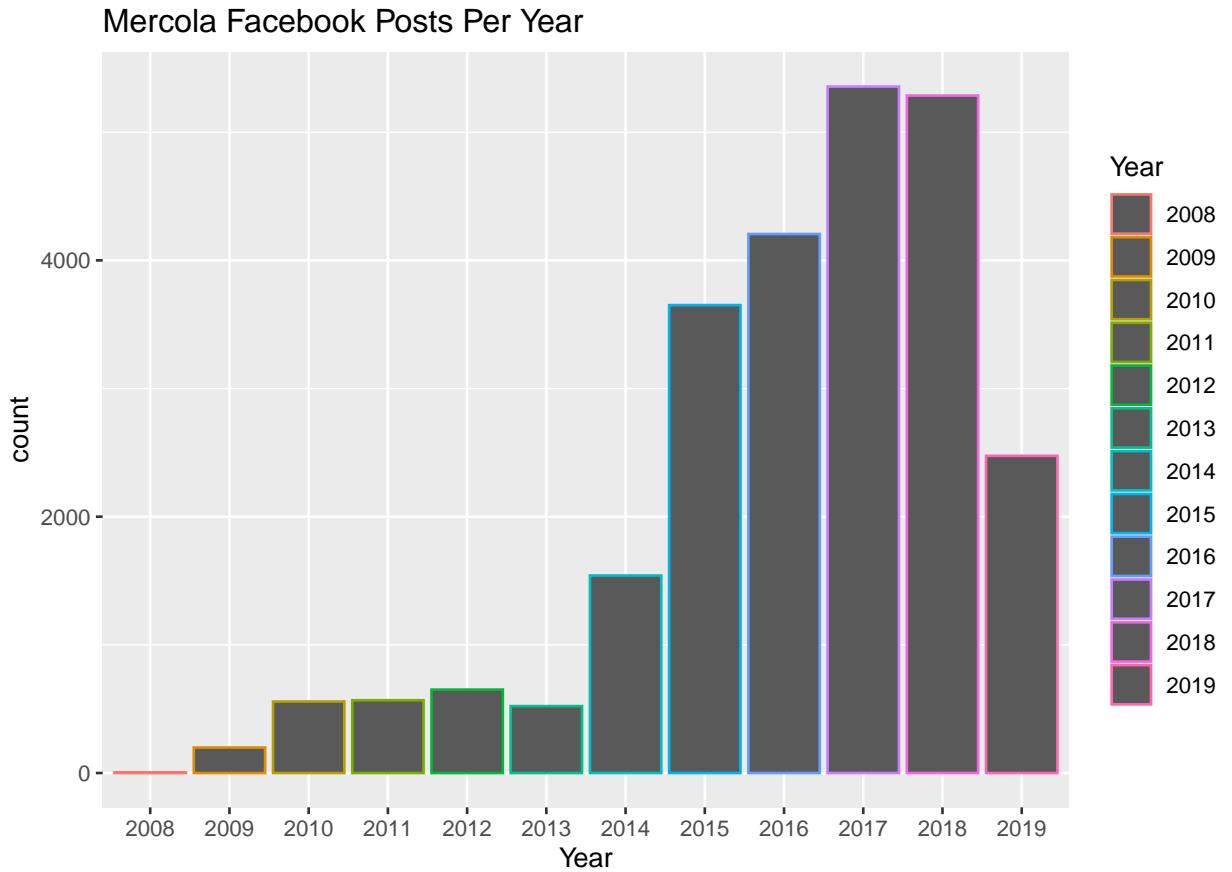
## 0.11 Social Media Material - Facebook

For Facebook data, CrowdTangle, a data service provided by Facebook was utilized to collect all post from Mercola's public website. This website contains posts that range from the year 2000 to 2020. Among the 25027 posts provided by CrowdTangle, only 12 posts exist for the years 2000-2008 and only 1 post is included for 2020. For this reason, those years are omitted from this study. In August, 2019 Mercola made the decision to cease activity on his Facebook group. During the last few weeks of activity, many posts discuss censorship and encourage users to connect with Mercola on his website and through his podcasts and SMS notifications. He also mentions that he may be found on Parler. As evidenced in the following visualizations of the data, while Mercola's activity peaked in 2017, the remaining years show a consistent and steady decline in comments, shares, and likes. It is quite possible that this decrease in activity is a function of attempts made by Facebook to limit the spread of misinformation after the 2016 elections particularly when posts refer to vaccine related topics.

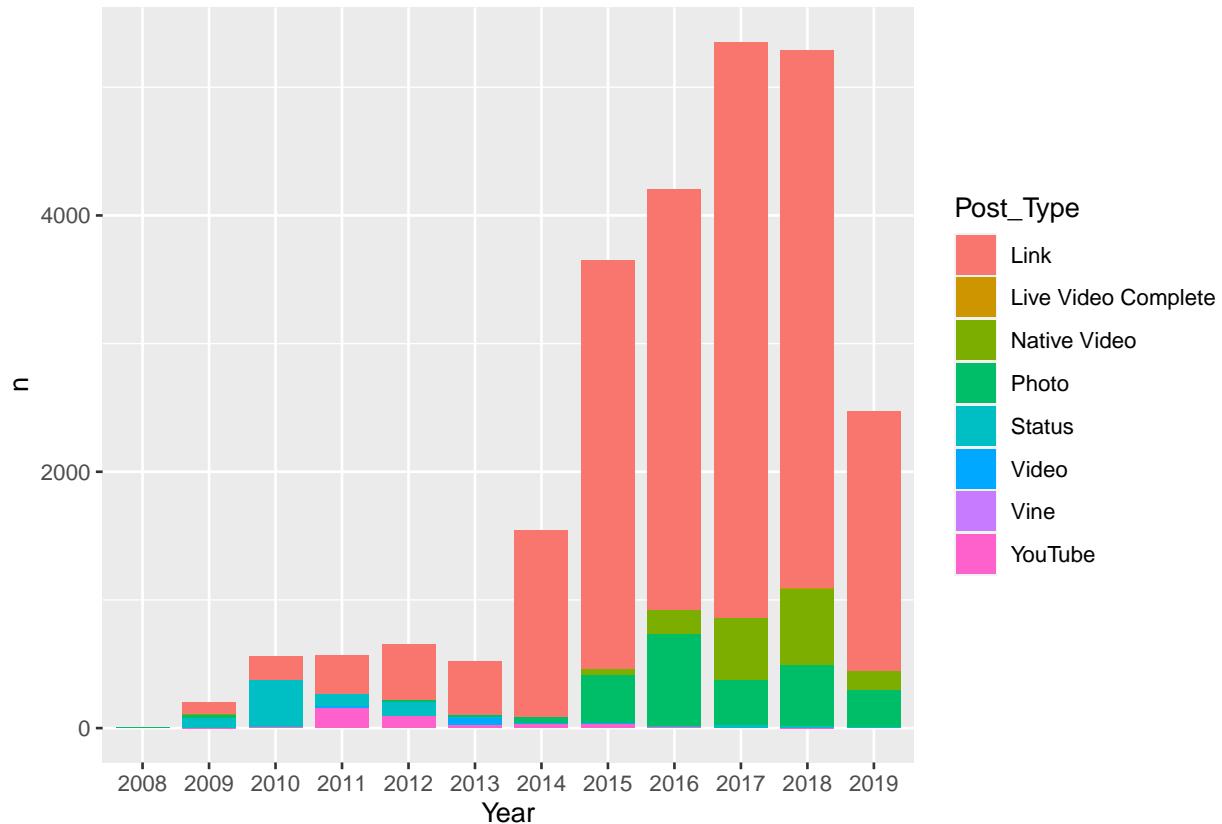
### 0.11.1 Overview of Mercola Facebook Post Statistics

The Mercola Facebook posts per year visualization shows a peak of posts in 2017 followed by a slight decrease in 2018. The significant decline in posts in 2019 is a function of a complete cessation of Facebook activity on the Mercola group page in 2019. The post type visualization shows the progression of simple status and link posts in 2009 and 2010 followed by more posts that included YouTube videos in 2011 and 2012. After 2011 posts that include links to the Mercola site increase significantly and peaking in 2016. In 2017 posts of links decreased slightly and then more significantly in 2019. The inclusion of photos in posts figure more

prominantly after 2015 and embedded videos (not from Youtube or Vine) begin to appear in posts after 2016. It does appear that Mercola continues to have some of a presence on Youtube with 371K subscribers to his main Youtube website.

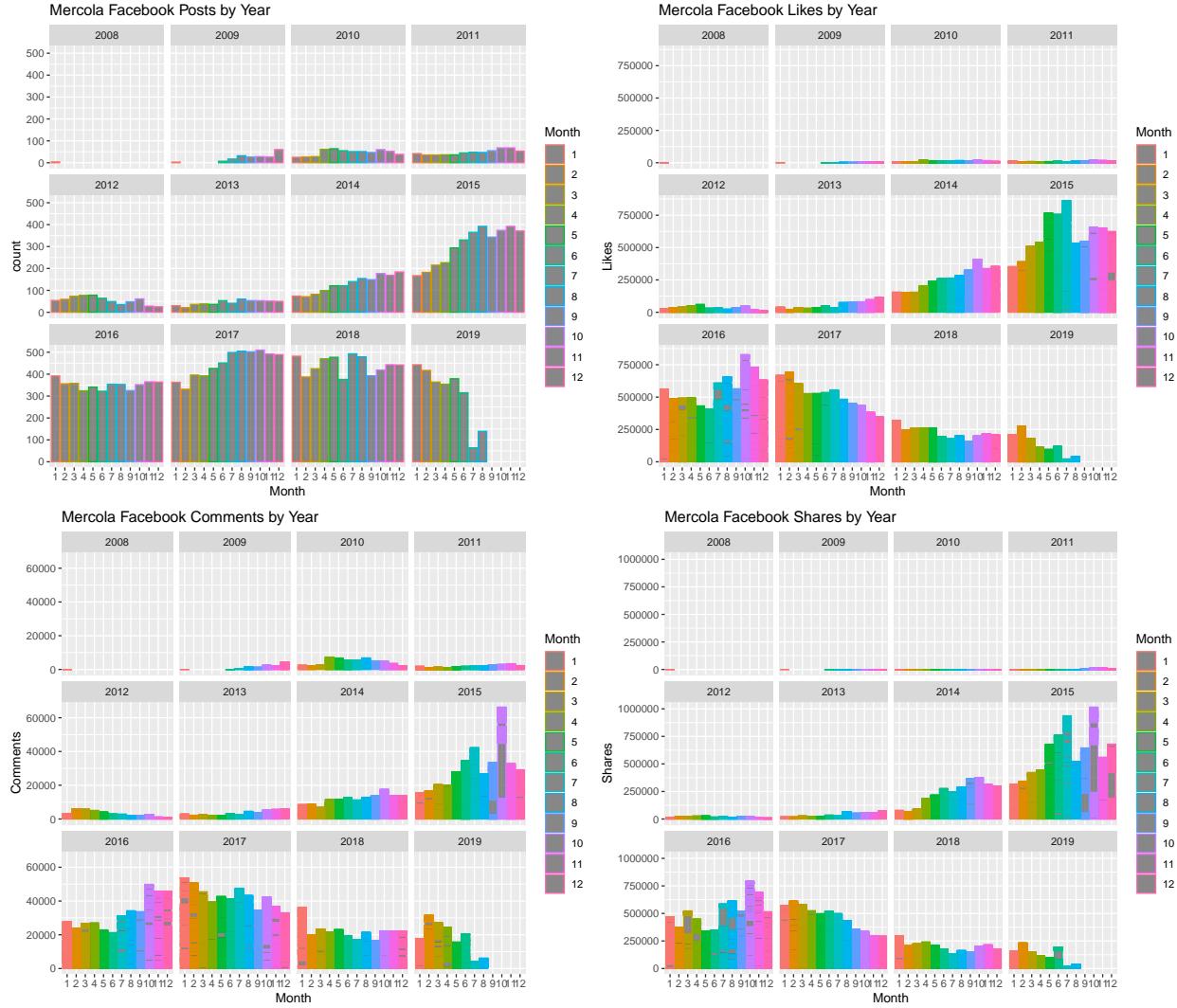


## Mercola Facebook Posts Sorted by Post Type and Year



## 0.12 Mercola Facebook Posts - Comments, Likes, and Shares

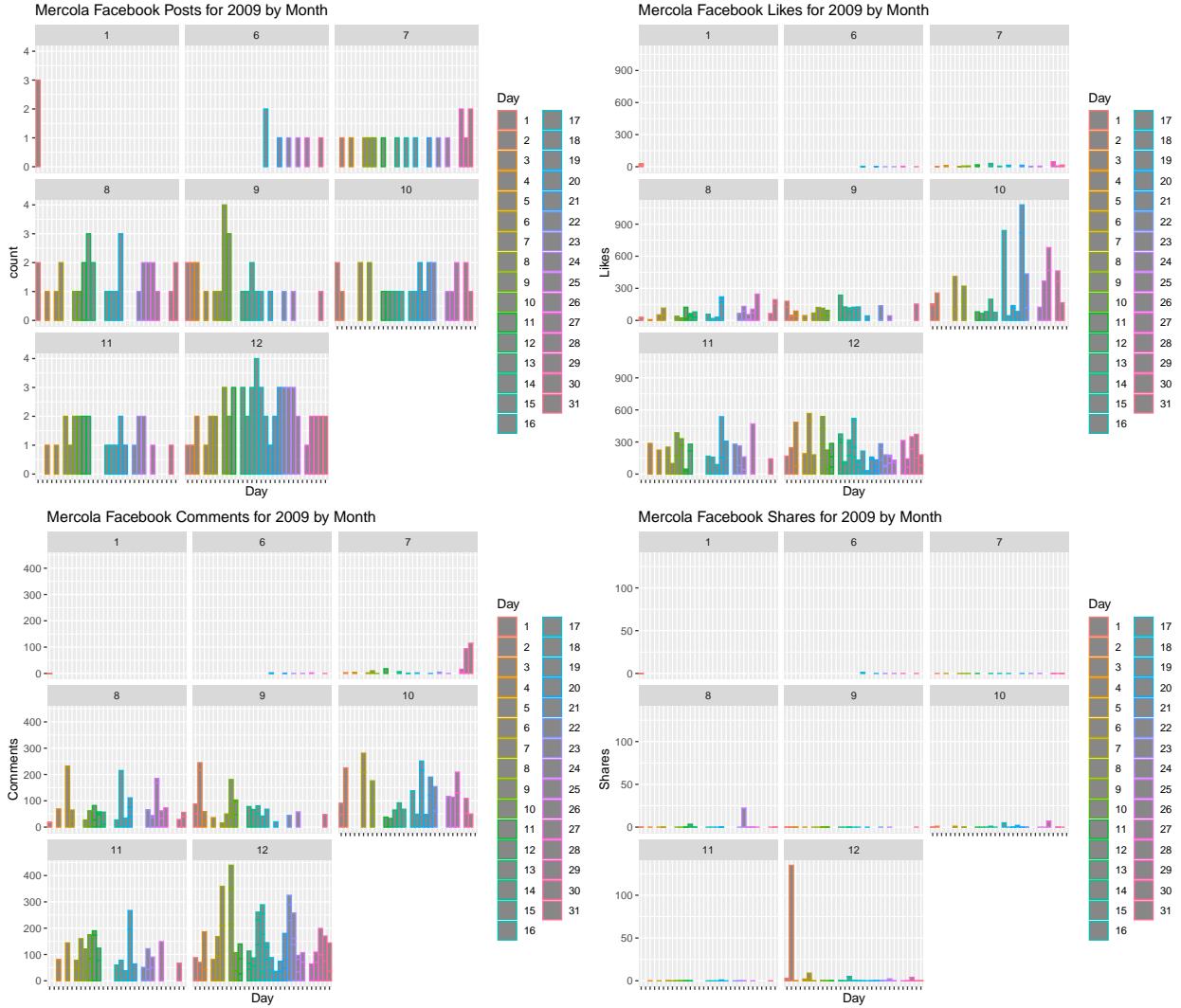
Posts made on the Mercola facebook group begin in earnest in mid-2009. They remain relatively static until 2014 when posts appear to grow significantly. The most consistent activity appears midway through 2017 and then decreases slowly until June of 2019. Final posts in July and August of 2019 are relatively smaller in number. Understandably, follower interactions with the Facebook group grow slowly and are not notable until 2013 and 2014. Comments, shares and Likes all appear to peak in 2015 and 2016. Afterwards, all three interaction types slowly decline until 2019. It appears that the decrease in interactions precedes the decrease in posts which may be indicative in mitigation efforts made by Facebook after 2016 to slow the circulation of medical misinformation (particularly as it pertains to vaccines). If this is indeed the case, then these visualizations show that shadowbanning practiced by Facebook is indeed effective.



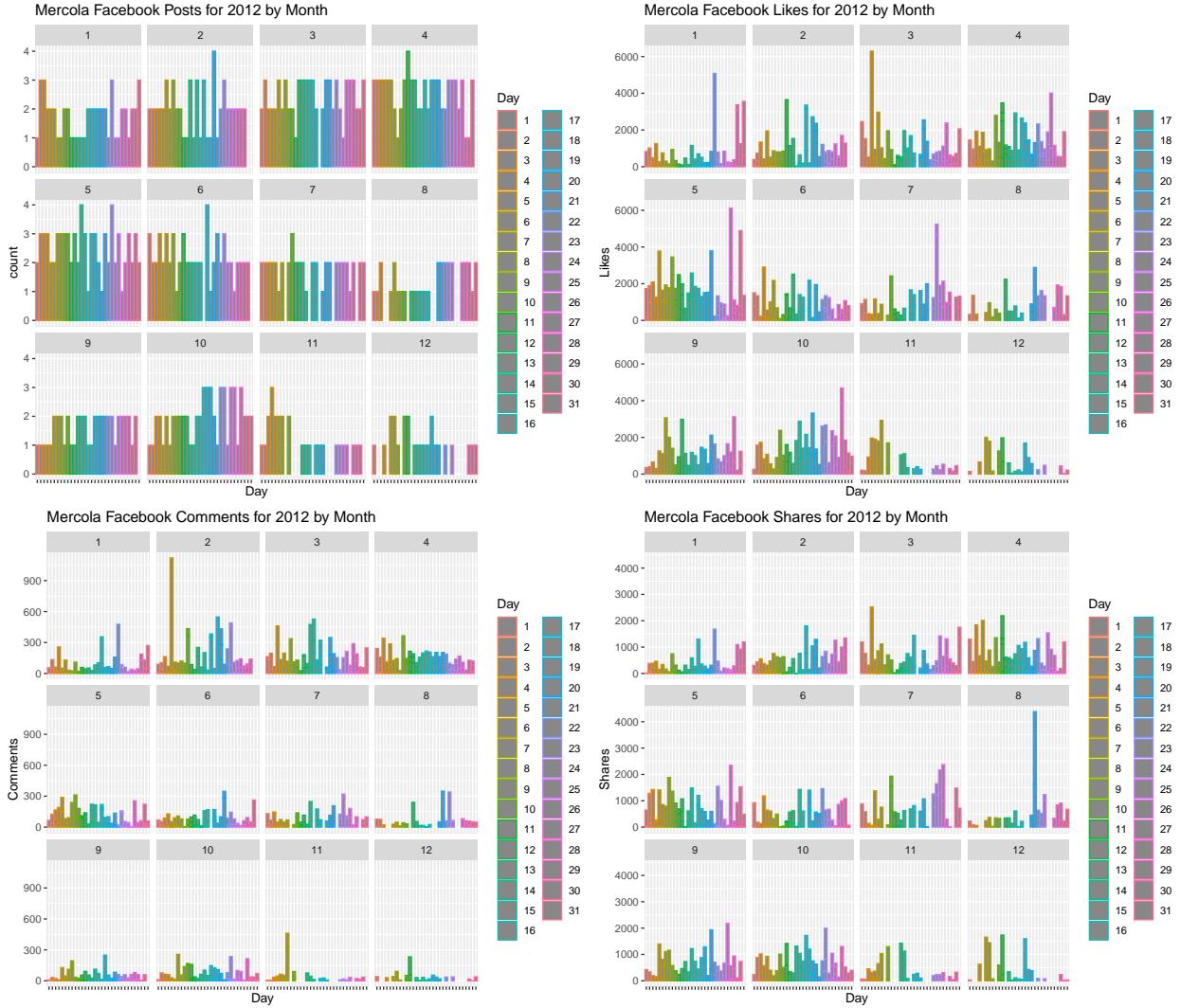
### 0.12.1 Investigating Facebook Posts, Likes, Shares, and Comments for 2009, 2012, 2016, and 2019

In order to compare Facebook data to the Mercola website articles, the Crowdtangle CSV file was filtered for 2009, 2012, 2016, and 2019. Since Facebook has only 1 post for 2020, posts by Mercola on Twitter will be used to compare with 2020 Mercola articles.

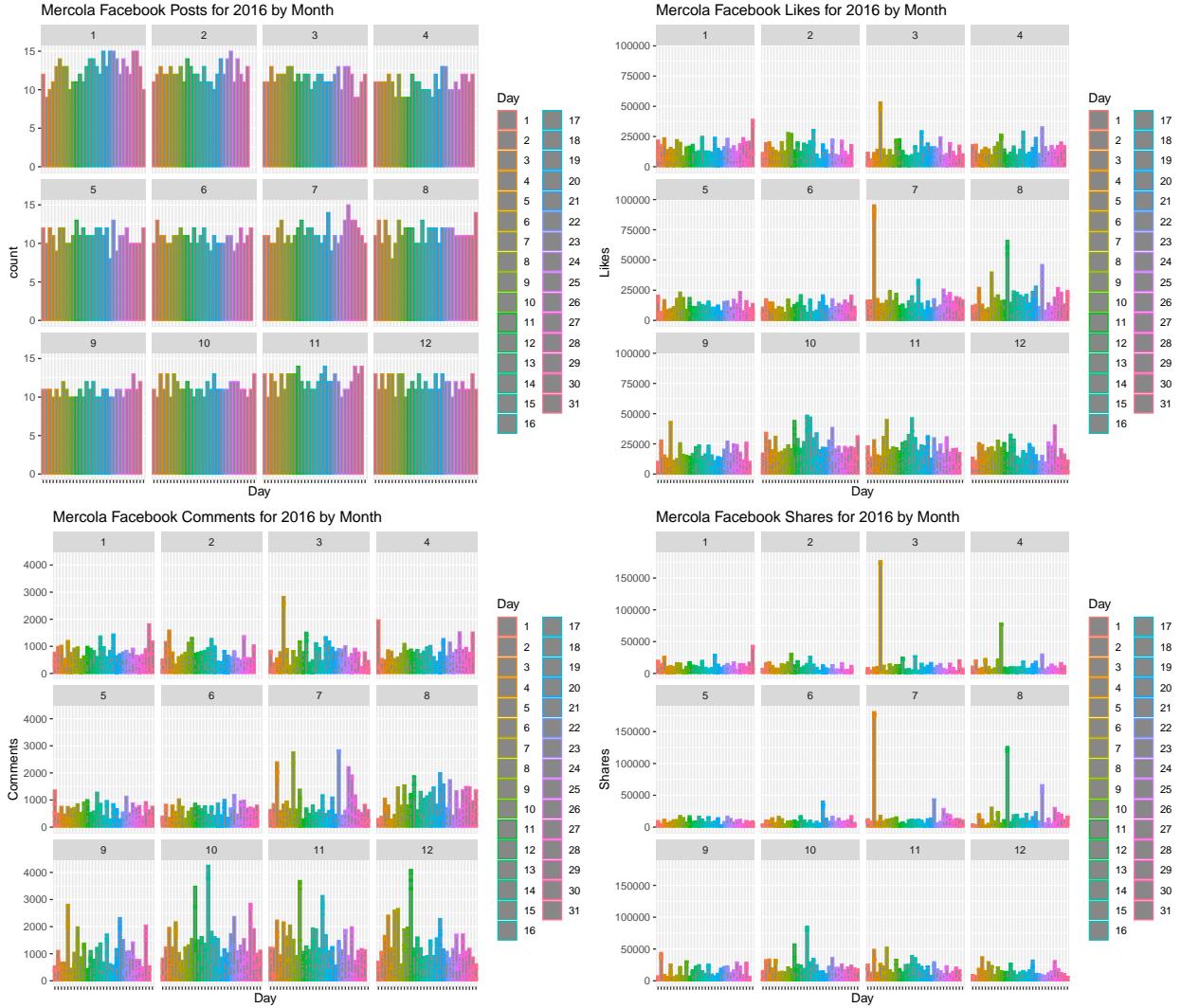
**0.12.1.1 Mercola Facebook Posts, Likes, Comments, and Shares for 2009** Posts in 2009 only began to appear in the second half of 2009. Likes and comments in 2009 are sparse with very few shares at all. October appears to be the month when pages were liked more frequently. Comments begin to appear in the second half of 2009.



**0.12.1.2 Mercola Facebook Posts, Likes, Comments, and Shares for 2012** Posts in 2012 remain steady throughout the first half of the year. During the second half of the year posts decrease a bit and become more infrequent towards the end of the year. The amount of likes for each month in 2012 varies greatly although understandably, the number of likes drops in November and December as posts drop. Comments follow a similar pattern as likes, however, they are still only a fraction of likes. The number of shares appears to be fairly steady throughout the year although there is one rather popular item in mid August that garners more than twice the average shares. Likewise shares in November and December drop off a bit.



**0.12.1.3 Mercola Facebook Posts, Likes, Comments, and Shares for 2016** Mercola's Facebook posts increasing fivefold from 2012 to 2016 and are evenly distributed throughout each month. Likes increase as well with one substantial peak in early July and a small handful of other peaks. Comment activity increases toward the end of the year while remain steady with two substantial peaks in early March and early July.

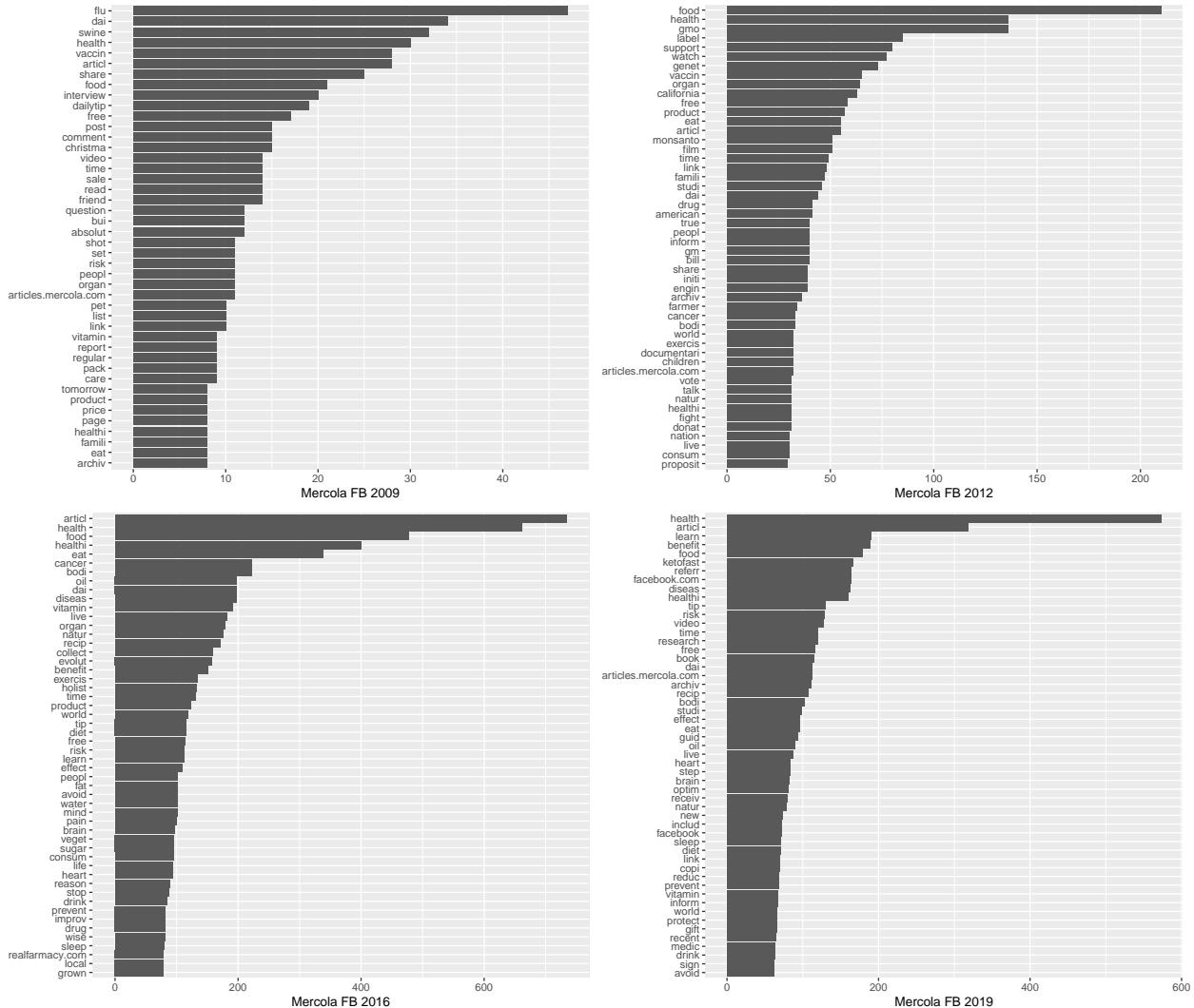


**0.12.1.4 Mercola Facebook Posts, Likes, Comments, and Shares for 2019** In 2019, during the first half of the year, post activity remains relatively steady with a small decrease between March and June. There is a gap of posts in July followed by another few weeks of posts. After the activity in August, there are no more posts for 2019. Overall, despite the same amount of posts in January and February as in 2016, likes drop quite a bit in the first two months and then fall at a greater rate the following four months. Comments decrease as well following a similar pattern. With the exception of a single peak in June, shares decrease significantly as well.



## 0.13 Visualization of Facebook Post Content: Top Words for Each Year

Top words in Facebook post content in 2009 reflect an emphasis on both the word “health” and the the swine flu pandemic which represent the most frequently used words in 2009. Posts containing the word “vaccine” show the word to have be the fifth most frequently shared word. In 2012 the two most frequently used words in posts are food and health. GMO and the word genetic (along with Monsanto - the company attacked by most who are against genetically modified foods). In 2012, vaccine is mentioned twice as frequently as in 2009. In 2016, health, food and article are the most frequently used words in posts. The remainder of the words tend to center around nutrition, food and other topics related to healthy living. Top words in the 2019 posts are similar to 2016 with health, article and learn being the top three words. Again most words in 2019 center around food, nutrition and articles pertaining to health living. In both 2016 and 2019 the word “vaccine” is not mentioned frequently. In all four years, the presence of the words dailytip, articles, video, sale, link, archive, film, and documentary alongs with the Mercola website address all point to the idea that most post are directing Facebook members to website content. The frequent mentions of Facebook in 2019 posts (which occur from January to mid-August) likely pertain to posts that inform group members that Mercola will no longer be posted on the Facebook website.

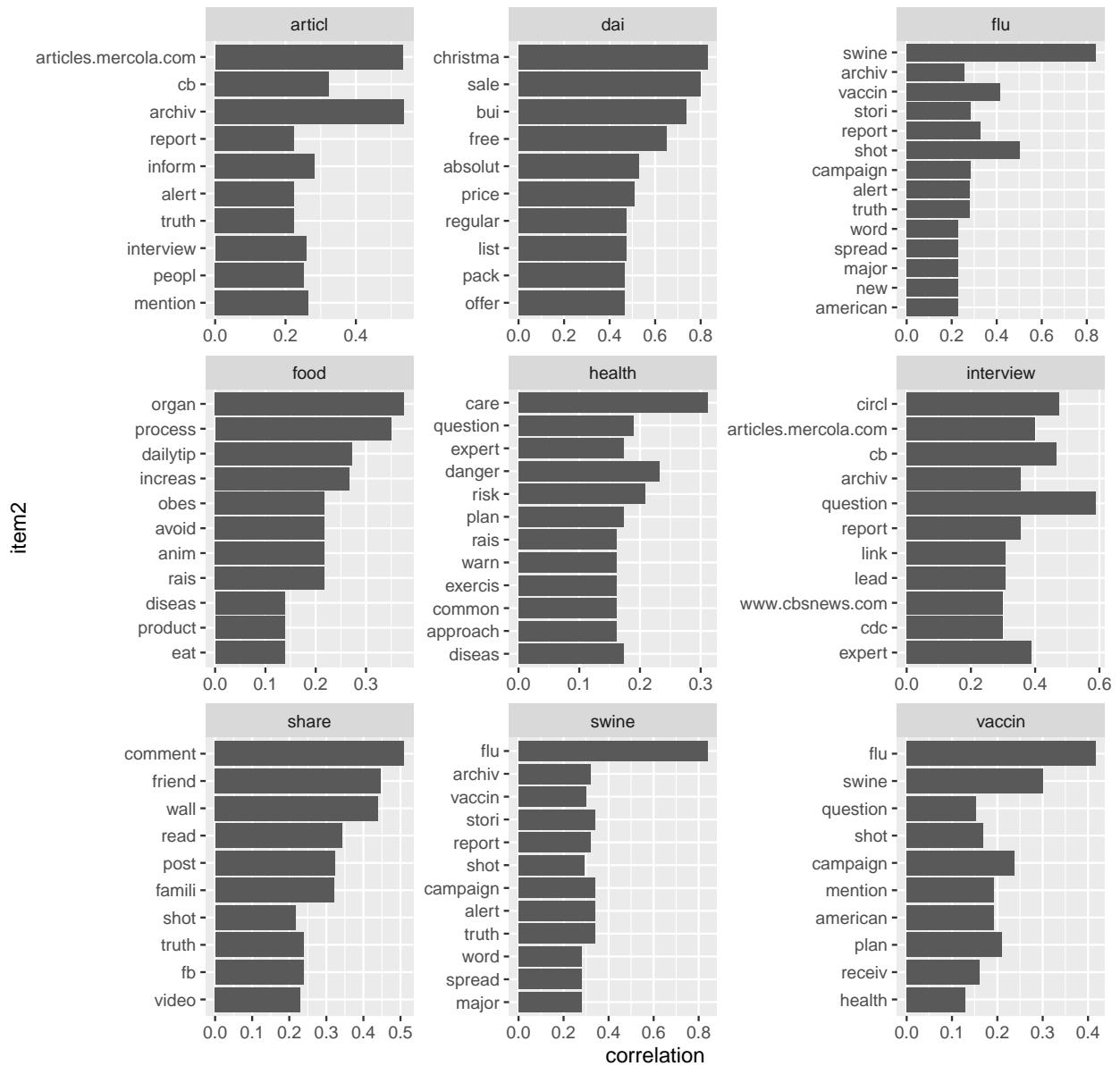


## 0.14 Visualization of Mercola Facebook Posts: Word Correlations with Top 9 Words

### 0.14.1 Top correlated words for Mercola Posts in 2009

In the top correlated words in Mercola Facebook posts in 2009, there are some interesting occurrences of words. In the context of the flu, the most frequently occurring words were swine and vaccine and its synonym "shot." The words truth, alert, and campaign seem to imply that there is an effort to create messaging around the idea of the flu either by Mercola or some other group or individual that he is referring to. There appears to be somewhat of an effort to promote articles with references to the website archives and some sales pitch related language within the "dai" category (for day). The reference to "circl" in the interview group relates to an inner circle group that was promoted early on for those who received special benefits when paying for a subscription on the website. There is a presence of discussion surrounding nutrition and health topics that appear to be cover a range of general subjects. The interview category includes a reference to CBS which likely reflects discussion regarding a CBS interview Mercola did in 2009 where he argues that the H1N1 virus is not dangerous (Swine Flu Cases: An Interview with Sharyl Attkisson of CBS News, 2009).

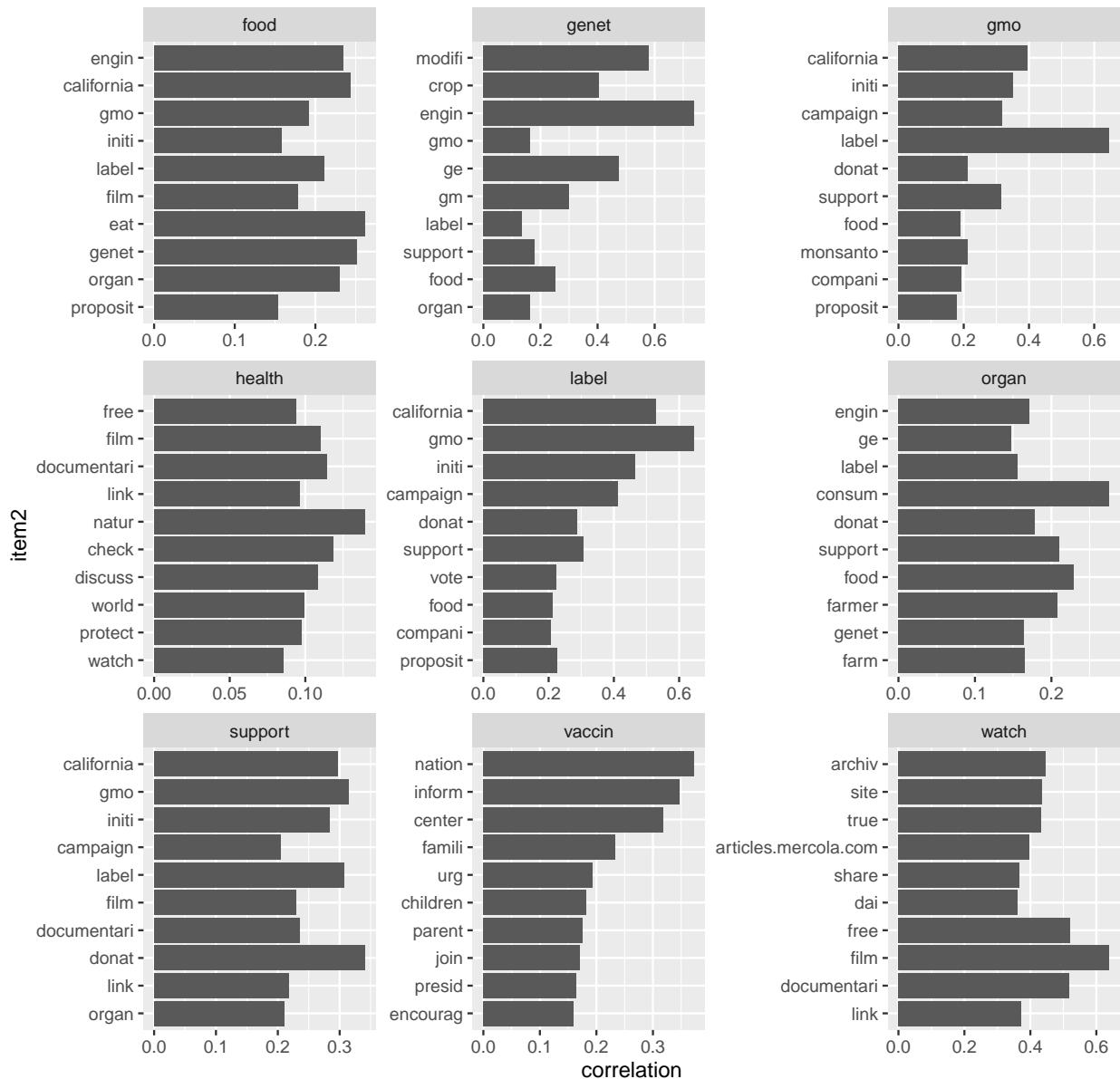
### Mercola Top Correlated Words for Facebook Posts – 2009



#### 0.14.2 Top correlated words for Mercola Posts in 2012

The discussion as indicated in Mercola Facebook posts in 2012 shifts to some degree. Genetically modified foods figure highly in discussions and appear as top words along with discussion regarding organic products (the category “organ”). Health discussions appear to revolve around free documentary films offered on the Mercola site. Vaccines comprise a smaller subset of the top words but still are prominent in discussions.

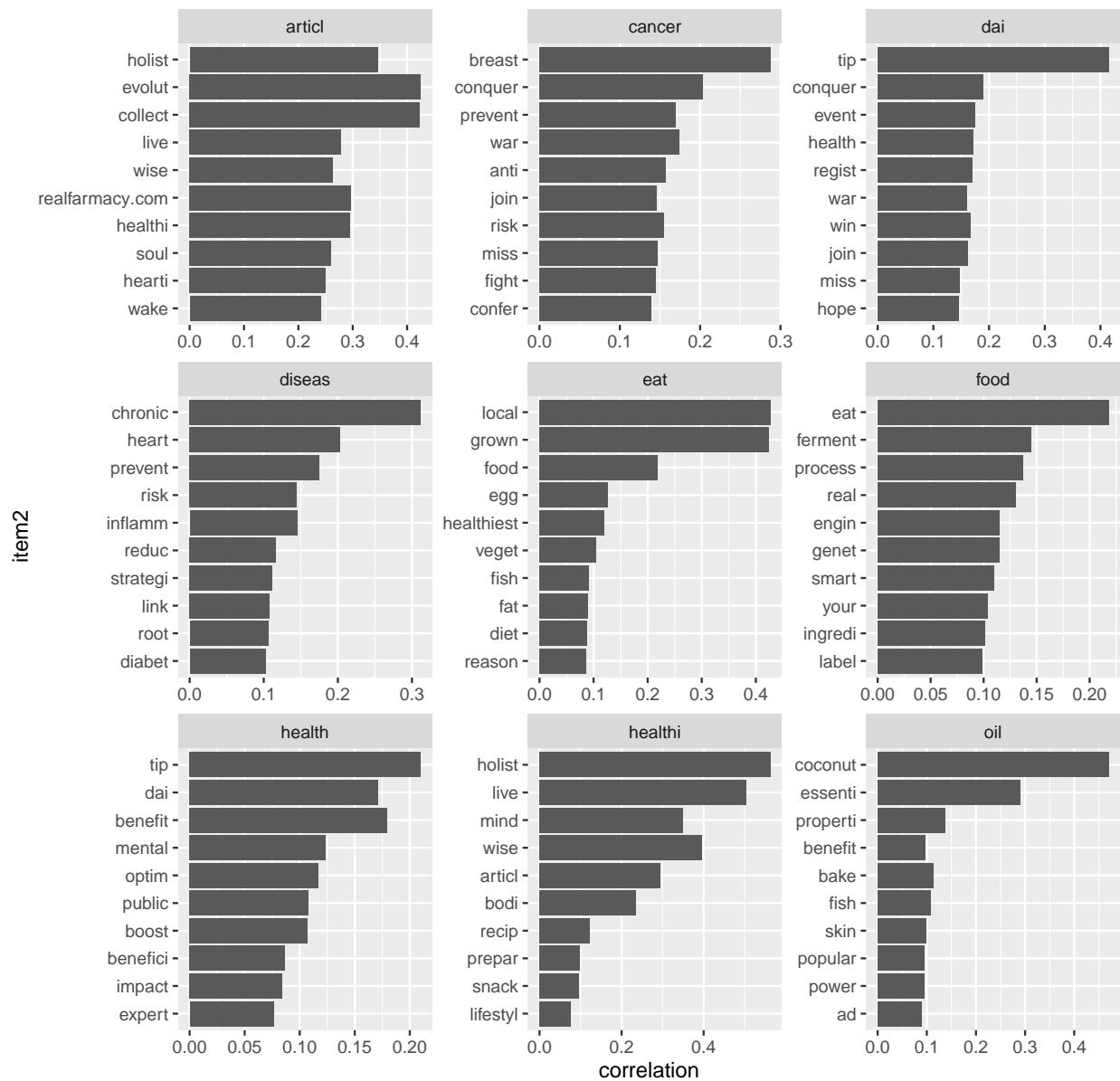
### Mercola Top Correlated Words for Facebook Posts – 2012



#### 0.14.3 Top correlated words for Mercola Posts in 2016

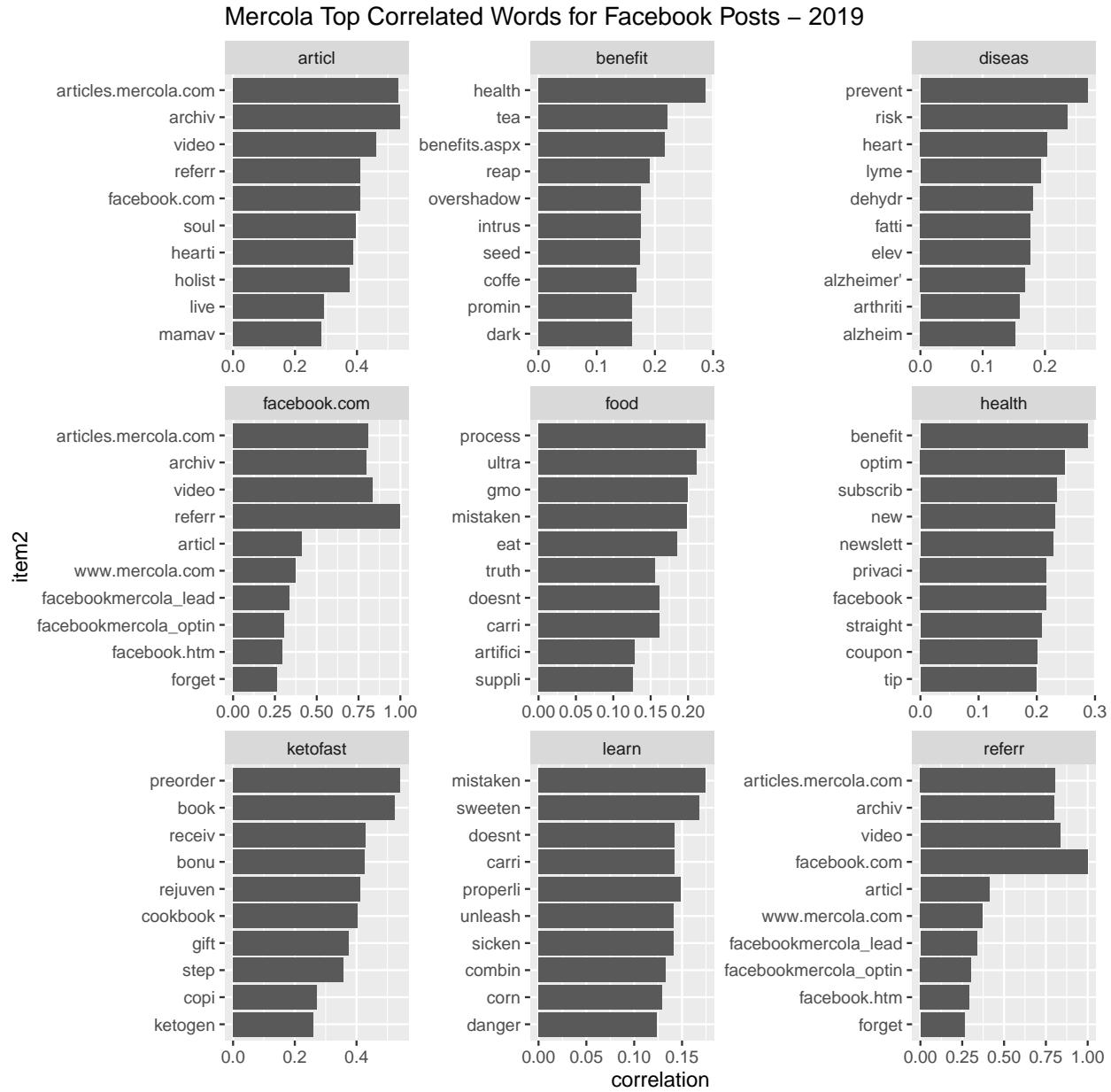
The top words and their correlated words in 2016 shift once again. In the article category, the website realfarmacy is referenced which is another popular website for health related misinformation. Cancer is discussed with correlated words like conquer, breast, war, fight, and risk having high occurrences. Food and nutrition appear to have a more prominent role in posts in 2016 with few references to GMO's as in 2012. Some general references to other diseases are made including diabetes and heart disease.

### Mercola Top Correlated Words for Facebook Posts – 2016



#### 0.14.4 Top correlated words for Mercola Posts in 2019

Facebook posts in 2019 change in tone for Mercola. The focus of health appears to correlate with his push to move his followers to signing up for a subscription instead of following him on Facebook. Privacy appears in the health category as this is his stated motive for leaving Facebook. Key diseases mentioned in the disease category are heart disease, lyme disease, alzheimer's and arthritis. Unlike other years, vaccines are not frequently mentioned. Mercola also appears to frequently promote his book on Ketofasting in posts. The presence of url's in the various groupings which do not appear in other years give the impression that Mercola is attempting to drive his audience away to his website.

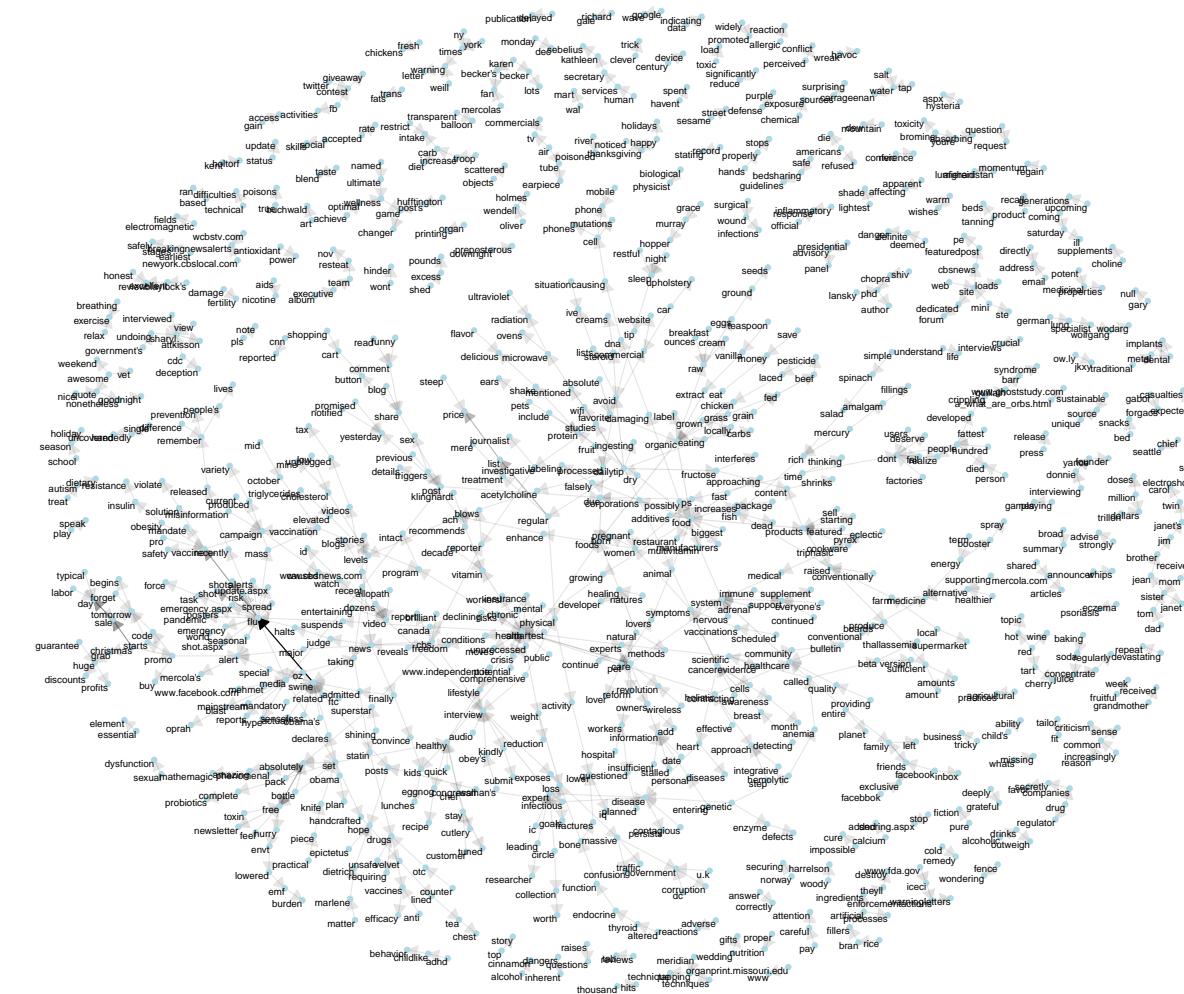


## 0.15 Visualization of Mercola Facebook Posts: Top Bigram Relationships

### 0.15.1 Bigram Relationships for 2009 Mercola Facebook Posts

Bigram relationships in 2009 Mercola Facebook Posts appear to have two primary centers with around five peripheral centers. One of the primary centers in the bigram diagram appears to contain words related to vaccines and the H1N1 pandemic. Interestingly, the words alert, safety, senseless, and mandate all appear around this grouping. The theme of the other grouping is a bit more difficult to discern but appears to include words relating to a mental health crisis. Other peripheral groupings appear to reference general topics related to nutrition, healthcare, and disease.

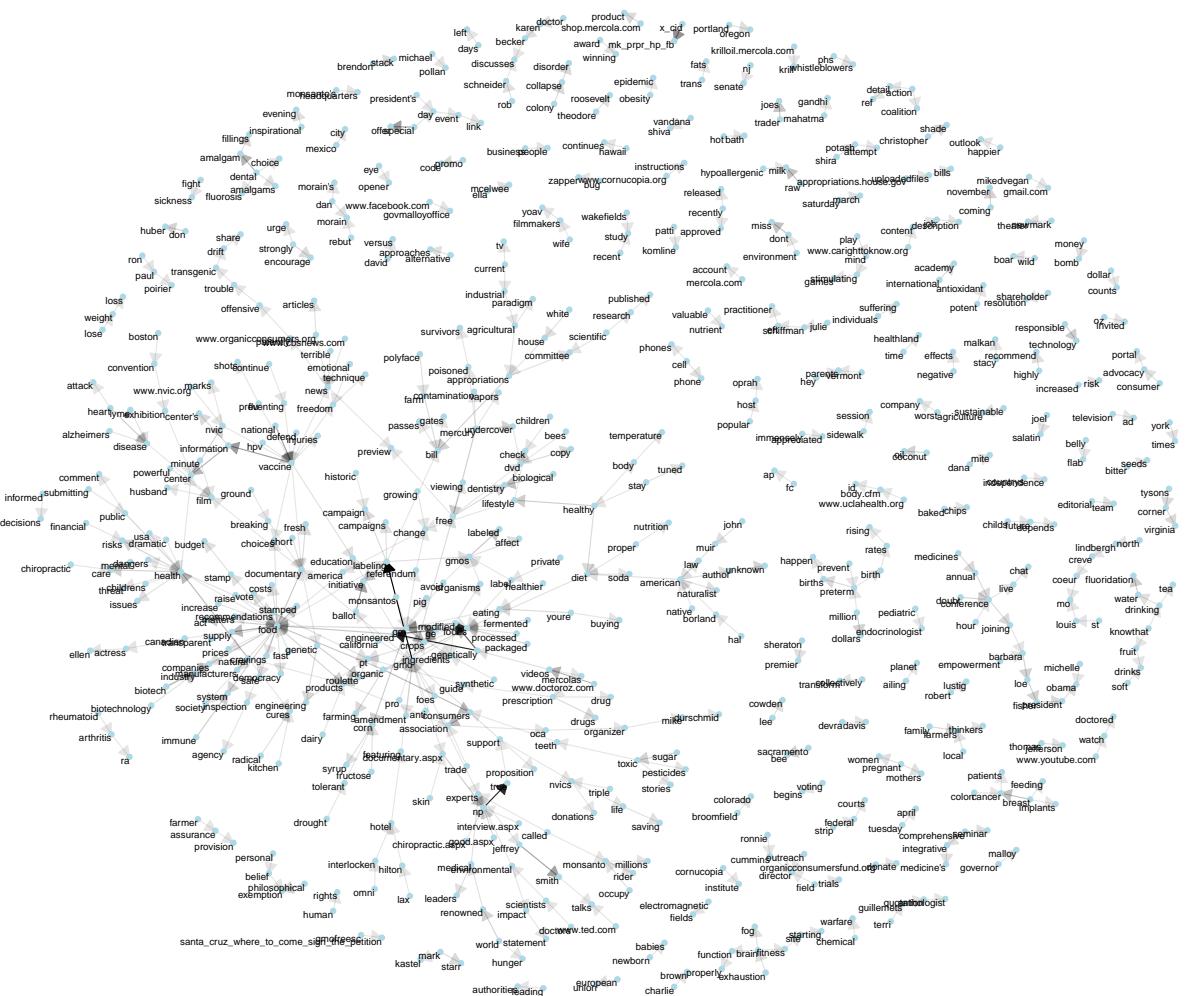
Mercola Facebook Bigrams 2009



### 0.15.2 Bigram Relationships for 2012 Mercola Facebook Posts

The center bigrams for 2012 Mercola Facebook posts tend to revolve around topics concerning genetically modified foods and organic crops. Other topics that are along the outside of the center involve topics such as vaccines, mercury, the flu, chiropractic medicine, and disease. There appears to be fewer references to health and nutrition in this visualization.

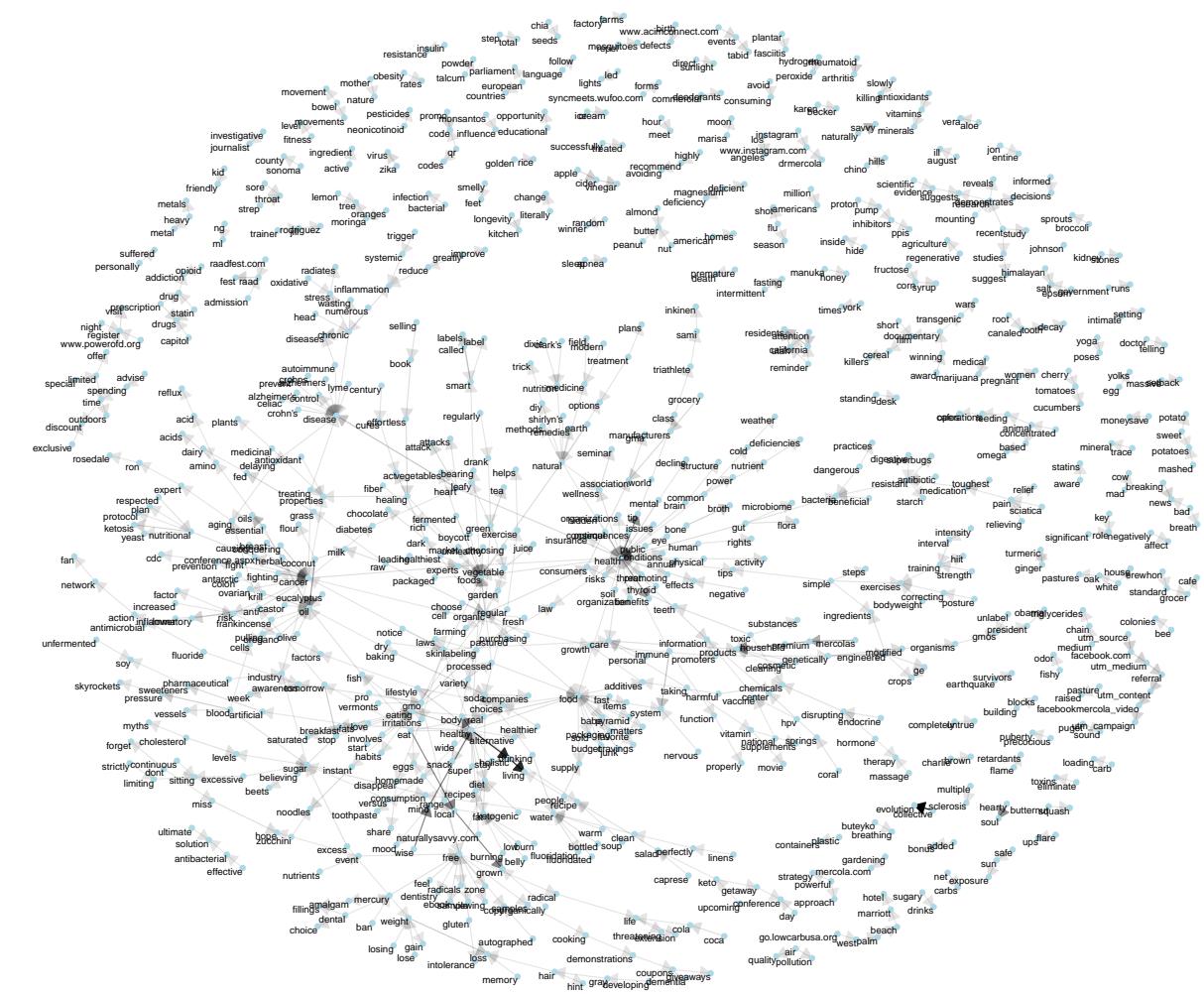
Mercola Facebook Bigrams 2012



### 0.15.3 Bigram Relationships for 2016 Mercola Facebook Posts

Three centers within the 2016 bigram relationship diagram appear to be centered around health (2 centers) and cancer. Another smaller center indicates bigrams that revolve around discussions of disease. Around these centers are other connected bigrams that refer to dental issues including fluoride, amalgam and mercury. Vaccine discussions (HPV) radiate from one health center and other discussions appear to exercise and nutrition.

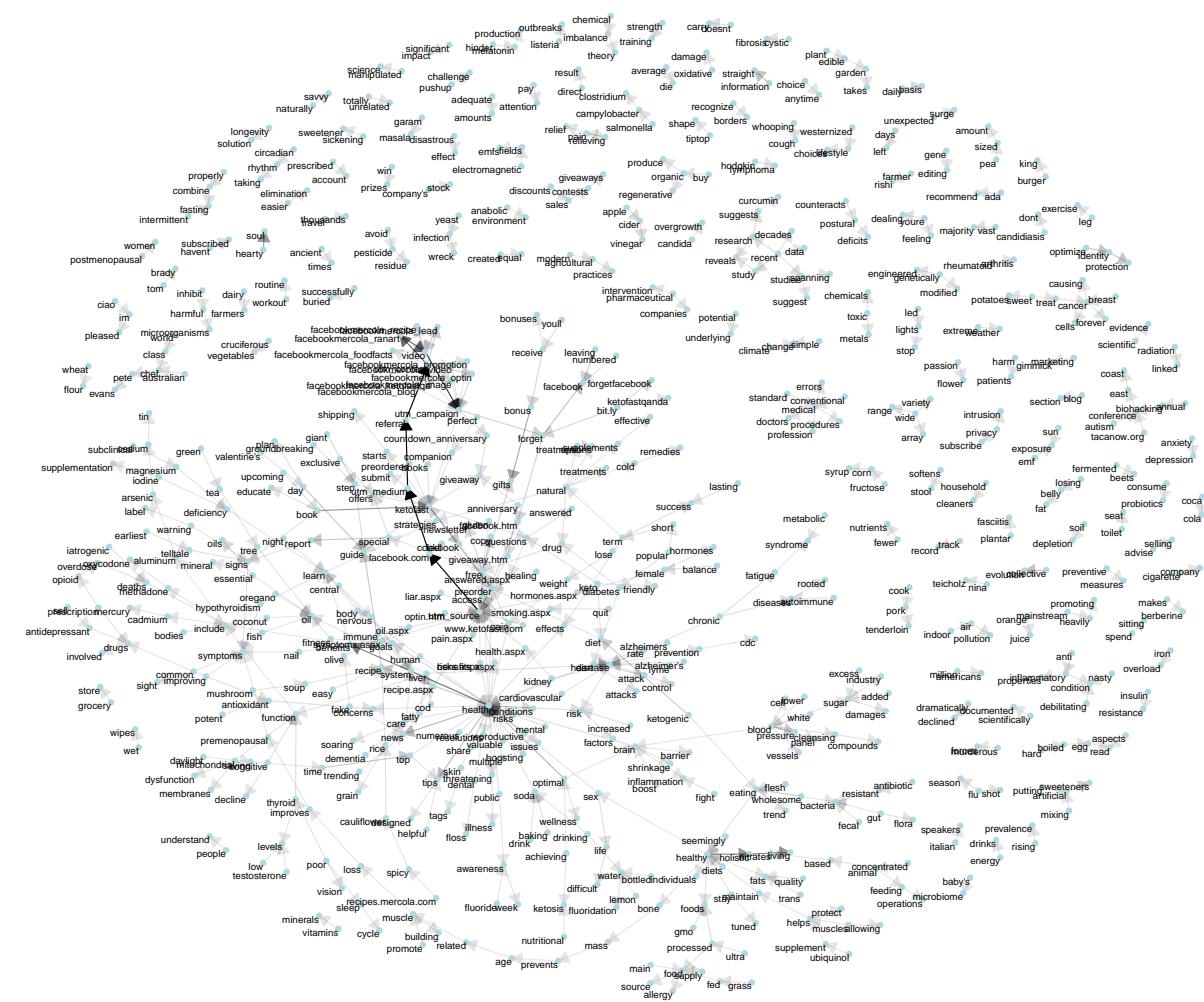
Mercola Facebook Bigrams for 2016



#### 0.15.4 Bigram Relationships for 2019 Mercola Facebook Posts

This final bigram centers almost exclusively on general health topics as related to nutrition with some references to health concerns. There is little reference to cancer, GMO's, dental issues or vaccine related issues. Of the two primary centers, one in particular indicates the final push in August to leave Facebook and subscribe to Mercola's website.

Mercola Facebook Bigrams for 2019

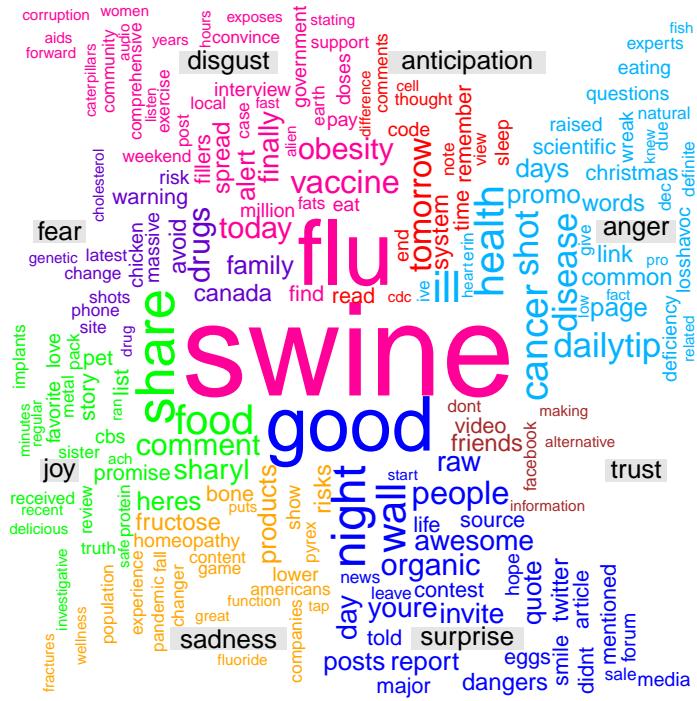


## 0.16 Visualization: Word Cloud - Sentiment Analysis For Mercola Facebook Posts

### 0.16.1 Sentiment Analysis Word Cloud for 2009 Mercola Facebook Posts

The word clouds included below represent sentiment analysis of message content of posts. For 2009 it is evident that many posts featured discuss the H1N1 pandemic and vaccines while pandemic appears in the sadness category. Cancer appears most frequently in the anger category along with health, ill, shot, dailytip, and disease. There are fewer words found in the categories of fear, trust, and anticipation. Words associated with disgust appear to be the most prominent and numerous collection followed by the categories of anger, joy, and surprise. As in the Mercola articles the word “food” figures prominently in the joy category. The next three most numerous words in this category are comment and share. Along with the word article, the words share and comment seem to indicate an attempt to direct viewers to interact with the material in a way that would move them or their friends to the website.

### Mercola Facebook Post Sentiment – 2009



#### 0.16.2 Sentiment Analysis Word Cloud for 2012 Mercola Facebook Posts

In 2012, the primary topics shift away from the H1N1 pandemic but food and health figure significantly in post sentiment during this year. Cancer, disease, obesity and dental topics figure heavily in the disgust category. Both vaccine and monsanto appear in the anger category indicating a negative view of these topics. Interestingly, the word “vote” appears in the surprise category which may indicate the presence of some political discussions. In the fear category, the word watch is the most prominent word. Other words in this category include documentary, government, dangerous, film and flu. During this year, many posts were made directing viewers to watch films about controversial topics (vaccines, chiropractic conferences, pharmaceuticals, etc.) that were posted on the website for free. They are now made available for sale on Mercola’s website.

### Mercola Facebook Post Sentiment – 2012



### 0.16.3 Sentiment Analysis Word Cloud for 2016 Mercola Facebook Posts

In 2016, the categories of surprise and joy focus upon topics related to food and nutrition with the primary words featured. Top words in the overall visualization are from these categories ("food", "good"). Cancer is also a frequent mention and is located in the anger category. In the sadness and disgust category words appear to be focused in the area of health issues, disease, and environmental threats although many of the words appear to be more general and not focused on a particular individual or institution. Overall the subject areas and words in this visualization are more general in nature and centered on health, nutrition, and wellness.

### Mercola Facebook Post Sentiment – 2016



#### 0.16.4 Sentiment Analysis Word Cloud for 2019 Mercola Facebook Posts

In the visualization for 2019, the word disease is the most prominent word and falls in the category of anger. This word is probably the most prominent among any of the sentiment visualizations. Also of interest is the emphasis in the anticipation category of discussions surrounding privacy and security. The word newsletter is also present in the anticipation category and along with the words latest, daily, news, and sign may signify the push mid-year to move away from Facebook to receiving daily email content. The groupings for trust and surprise also appear to contain words that also might be associated with this move as well. Under the category of disgust, most words are fairly distributed around issues of nutrition and chemicals while the anger category appears to focus on issues related to cancer and heart problems as well as some other forms of illnesses.

## Mercola FB Title Sentiment – 2019



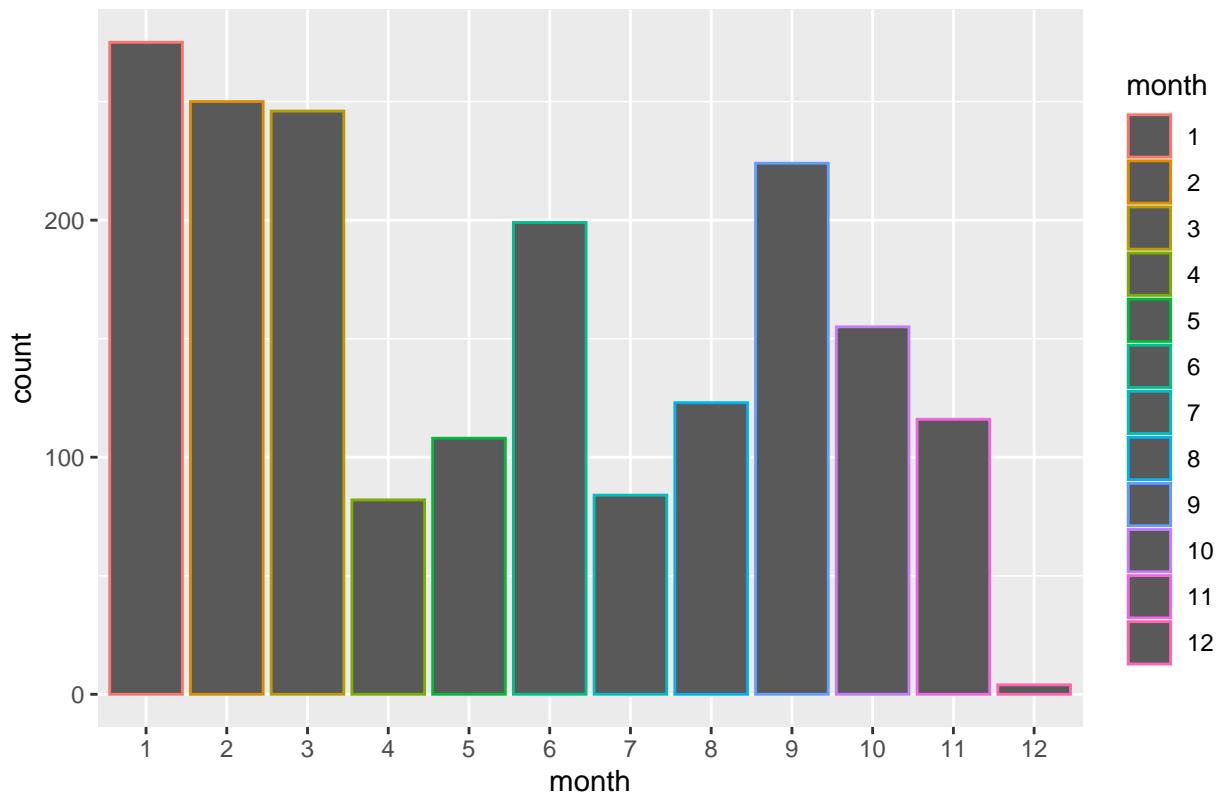
## 0.17 Social Media Material - Twitter Data

Given Mercola’s departure from Facebook in 2019, Mercola’s 2020 tweets were investigated to understand how his social media communication changed this year. Twitter permits a maximum of 3200 tweets to be gathered per user. In Mercola’s case the amount returned from a Massmine query was 3177 tweets which extended from the end of November, 2020 to July, 2019. These tweets were filtered so that only tweets from 2020 would be investigated. Once filtered, the 1866 tweets remained.

### 0.17.1 Mercola Tweets by Month for 2020

Mercola’s tweets peaked in the beginning of 2020 with roughly 250 tweets or more for the first three months of 2020. In April, his tweets dropped significantly and stayed low during May. In June, tweets peaked once again followed by a drop in tweets for July and August. September tweets neared the level of Mercola’s tweets in the first quarter but again dropped the following two months. Only one day of tweets were collected in December.

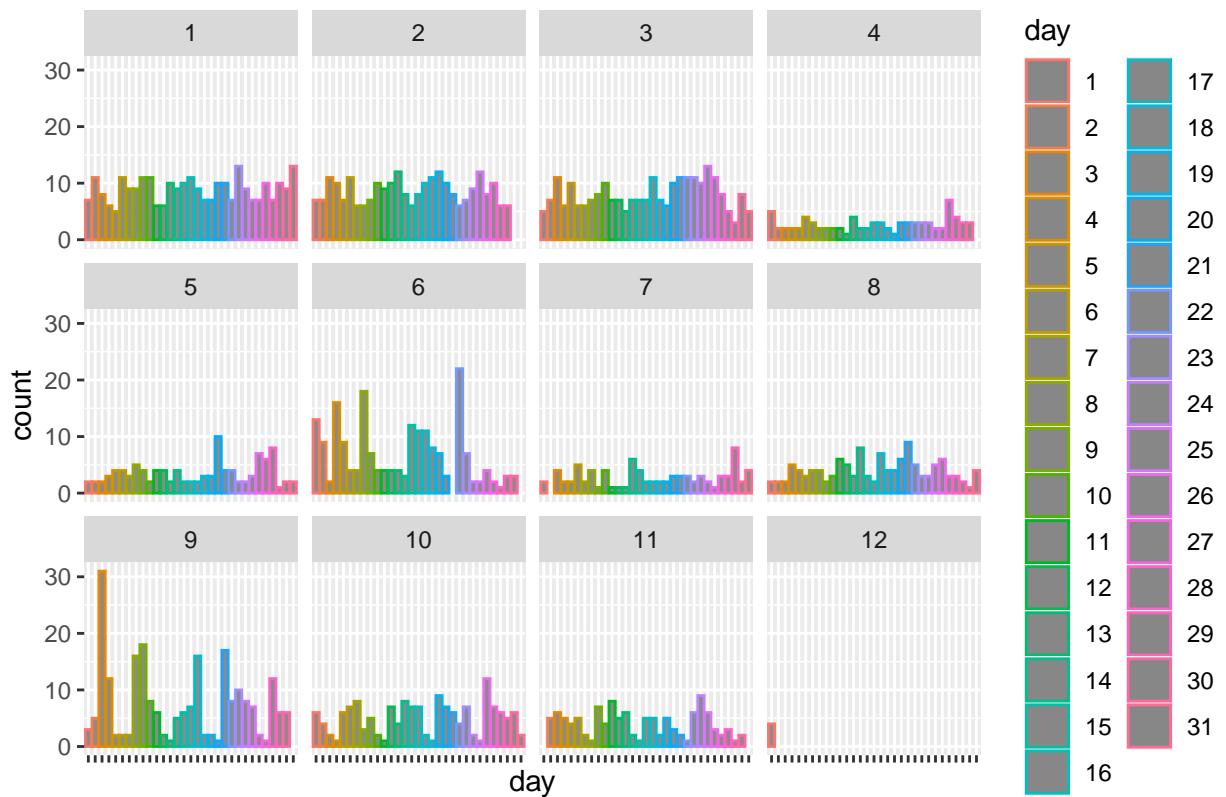
## Mercola Tweets Per Month in 2020



### 0.17.2 Mercola Tweets by Day and Month for 2020

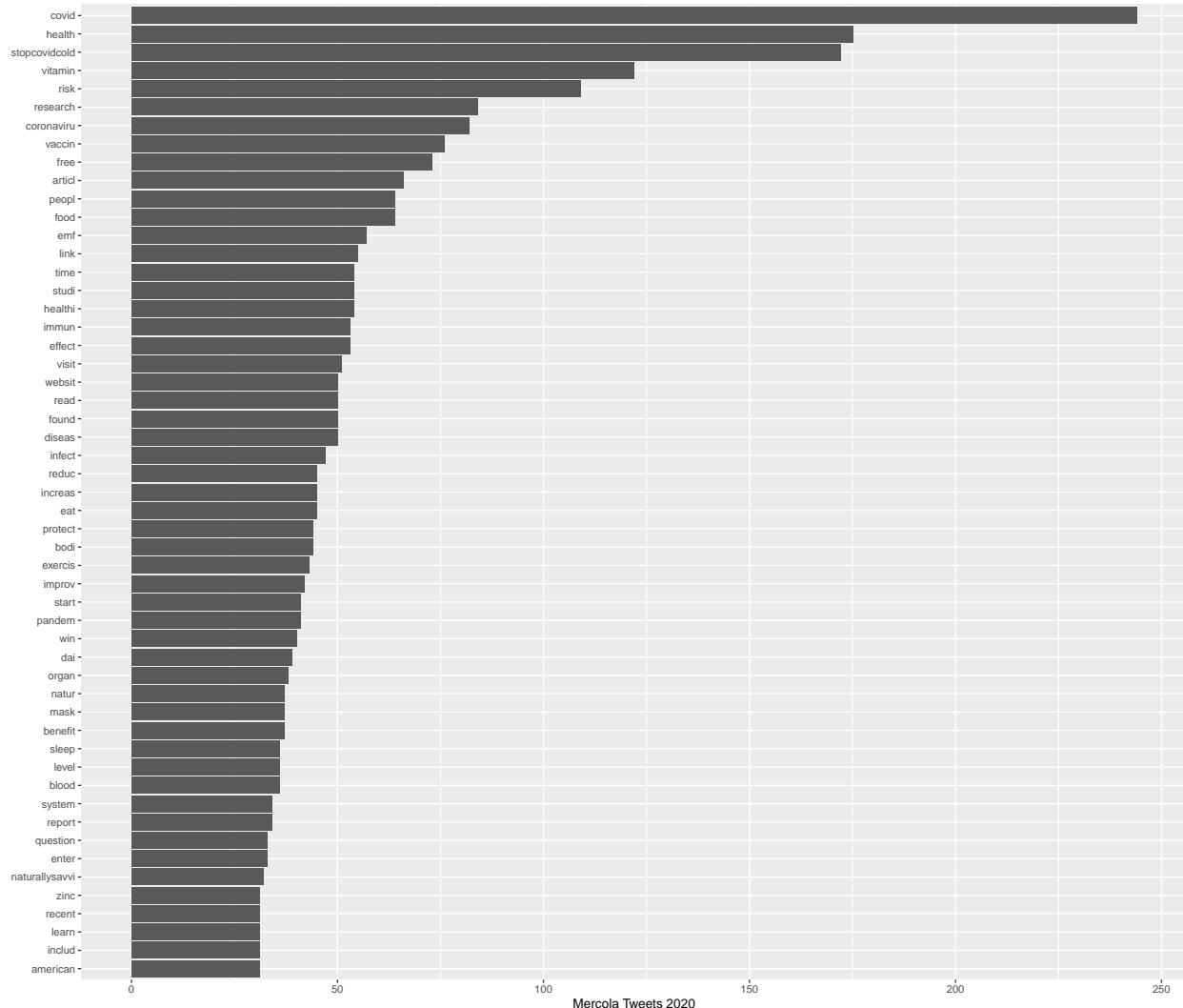
The amount of tweets for the first quarter of 2020 appear to be fairly steady. Once these tweets drop in April the average daily tweets appear to be consistent with an occasional peak in daily activity. June and September appear to have sporadic activity that may be worthy of further investigation.

## Mercola Tweets for 2020 by Month and Day



### 0.18 Visualization of Mercola Tweets: Top Words for 2020

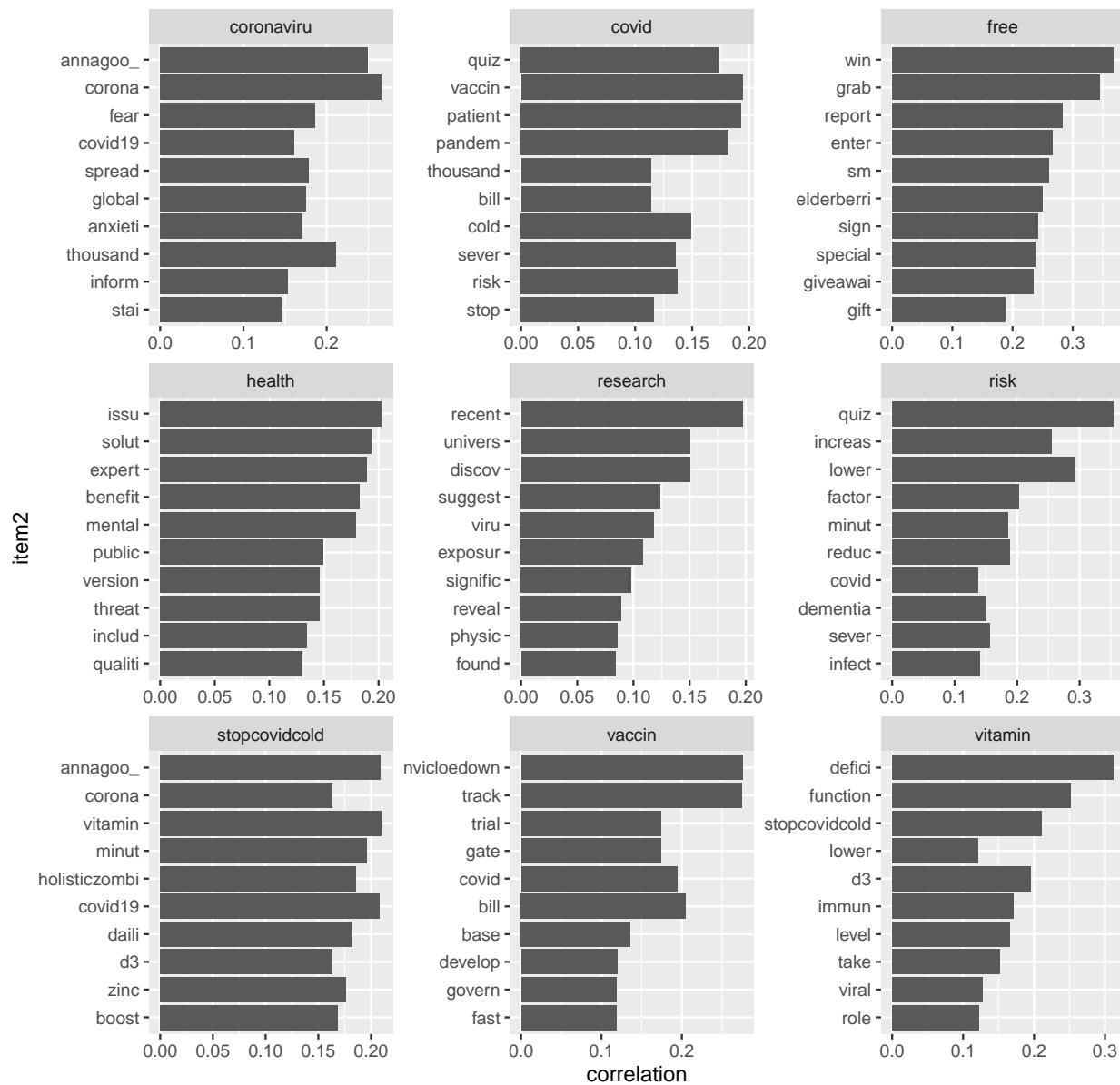
For the top words found in 2020 Mercola tweets, covid is the highest ranked with significantly more mentions than any other word. The phrase “stopcovidcold” is the third most frequent “word.” This phrase is a reference to another Mercola website by the same name. This website contains information that exclusively addresses purported ways to avoid contracting covid by taking supplements to “boost” the immune system or disputed ways to treat covid once it is contracted. The website is rather sparse but still features Mercola and four other colleagues. A handful of article links on the website point the reader back to Mercola’s website archive. Health is still a word that is frequently used along with vitamin and research. This is similar to findings in the articles and Facebook posts. Likewise, the word vaccine also appears relatively frequently. Several words (free, article, link, website, visit) imply that messages are pointing followers back to articles on Mercola’s website.



### 0.18.1 Top Correlated Words for Mercola Tweets in 2020

The top correlated word for coronavirus is “annagoo\_”. This is a handle for a twitter user that Mercola frequently retweets. She appears to follow him closely and retweet his material as well. She also frequently tweets the “stopcovidcold” hashtag along with an associated hashtag called “stopworldcontrol”. This website is associated with the former website but propagates the conspiracy that Covid-19 was a manufactured pandemic that was created to take over the world. The stopcovidcold category is associated with two vitamins that Mercola frequently touts for boosting immunity against Covid-19. The correlated words in the vaccine category are linked with frequent mentions of Bill Gates and building on the conspiracy theory that Gates is using Covid-19 to take over the world with vaccines. The user name “nvicloedown” is also most highly correlated with vaccines. Loe stands for Barb Loe who is the president of NVIC, an anti-vaccine organization to which Mercola is a major donor with a website that frequently refers readers back to Mercola’s website (Satija & Sun, 2019). The remainder of words either contain general references to broader health issues or appear to involve efforts to move followers to participate in a giveaway promotion of Mercola supplements. The url to this promotion leads to Mercola’s website and followers are entered into a contest when they share their personal information and sign up for SMS text messages.

## Mercola Top Correlated Words for Tweets – 2020

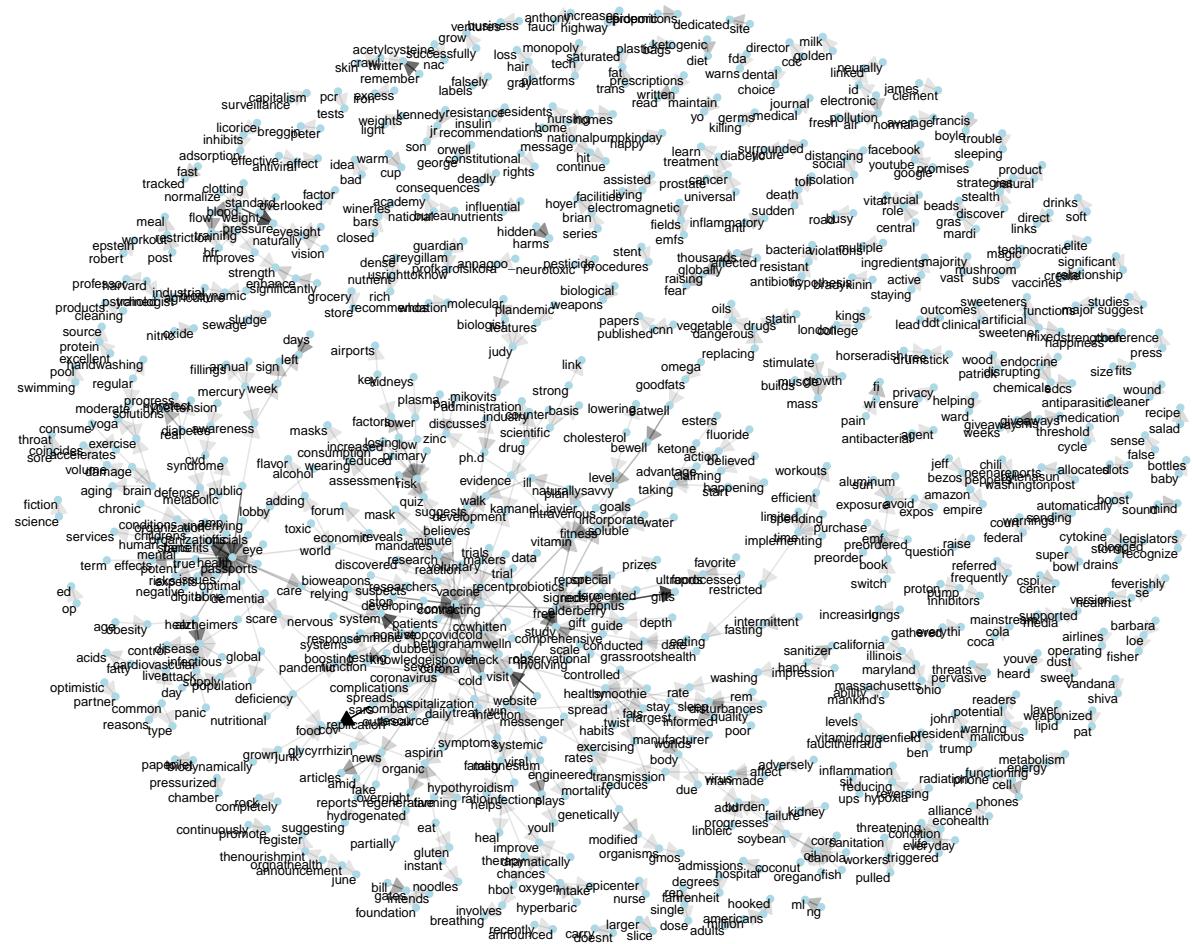


## 0.19 Visualization: Top Bigram Relationships and Sentiment Analysis for 2020 Mercola Tweets

### 0.19.1 Bigram Relationships for Mercola 2020 Tweets

The bigram visualization highlights most of the conspiracy theories and covid misinformation that is ultimately shared on Mercola's website. The center of the network is vaccines and stopcovidcold with two other areas that appear to focus on the giveaway offer and health issues. The focal point that is connected with the word health contains references to digital passports (the conspiracy theory that the vaccine will be used to control people's movements). Additionally, bigrams pertaining to Judy Mikovits and her film "Plandemic," mask mandates, bioweapons and man-made viruses are all noted in the outer connections.

## Mercola Tweet Bigrams for 2020



### 0.19.2 Sentiment Analysis for Mercola 2020 Tweets

Sentiments for 2020 Mercola tweets appear to be more negative over those that were looked at with the Mercola archive and Facebook posts. The fear and sadness categories are much larger than other visualizations and there are many more references to words in the disgust category. Pandemic, covid, risk and disease are all more frequently mentioned in the negative categories. The joy category is not quite as large as the other platform's joy category and the word "content" replaces "food" which was the most frequently occurring word in the joy category in prior visualizations (FB and articles). The word content, interestingly, is a reference to notices throughout Mercola's tweets that Twitter has hidden his content. His tweets encourage users to click through the link in order to be able to access the censored content which is typically a link to his website. The anticipation and trust category are also fairly small while the surprise category once again is indicative of the contests sponsored by Mercola in his tweets to capture SMS information from followers so that he can directly share his content with them through texts.

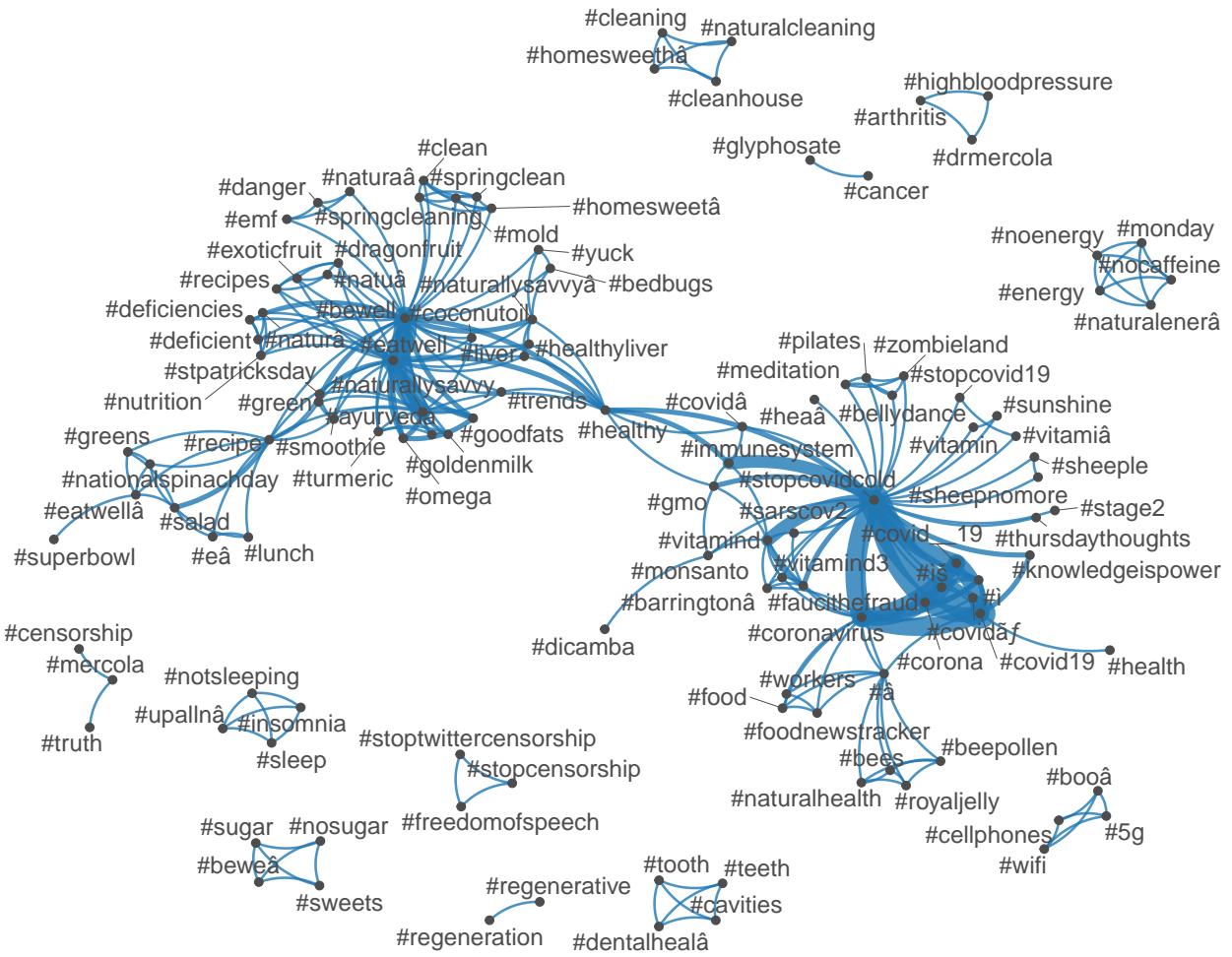
## Mercola Tweet Sentiment – 2020



## 0.20 Co-occurrence Graphs for 2020 Mercola Tweets

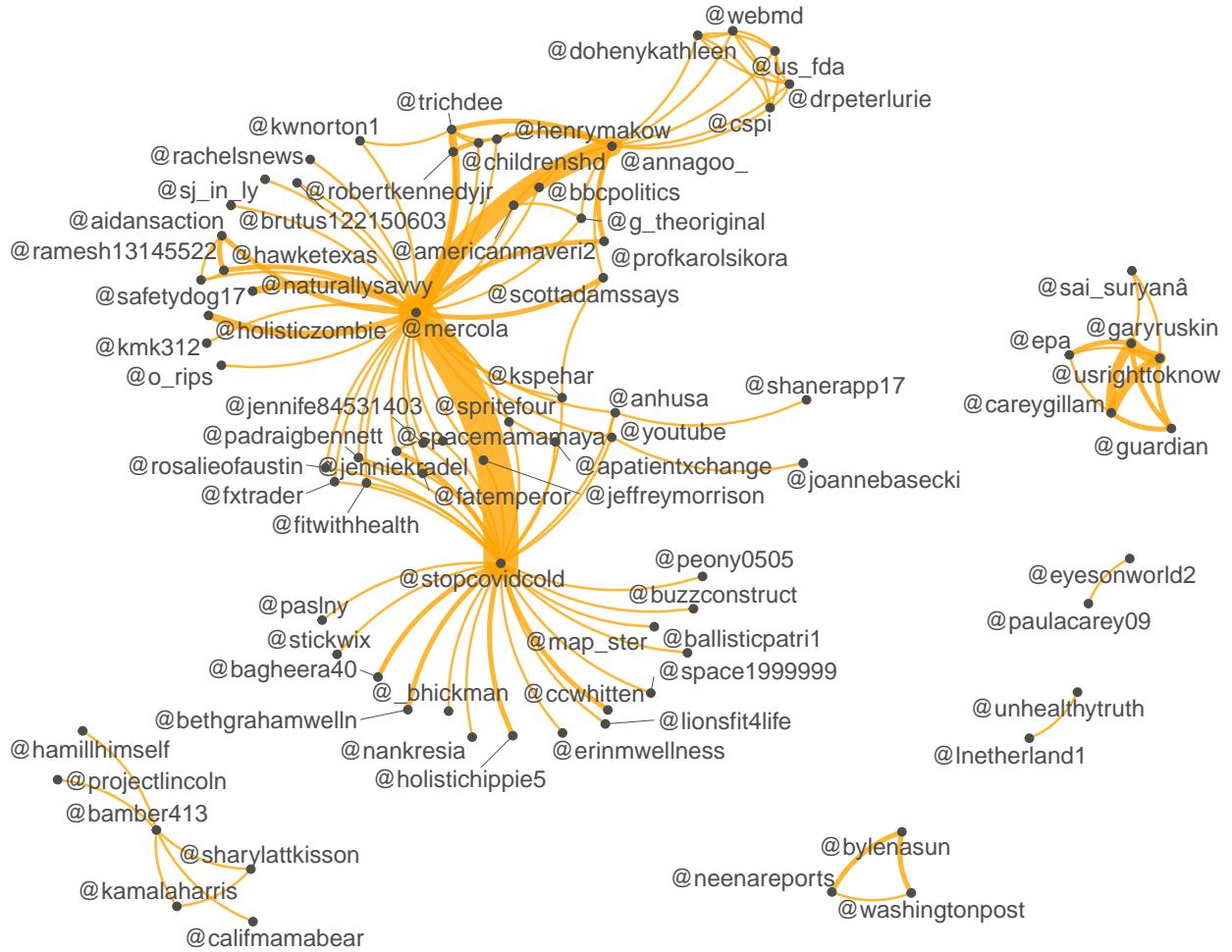
### 0.20.1 Hashtag Co-Occurrence Graphs for Mercola 2020 Tweets

The co-occurrence graphs for Mercola 2020 tweets also highlight the idea that messaging surrounding content is two-fold: one center of the hashtags refers to nutrition and health living topics while the other center refers to issues surrounding Covid-19 with some references to GMO issues.



### 0.20.2 Hashtag Co-Occurrence Graphs for Mercola 2020 Tweets

The co-occurrence graphs for Mercola 2020 tweets highlight individuals who Mercola interacts with frequently. Robert Kennedy Jr. (with website “childrenshd”) is known as a prominent anti-vaxxer whose website communicates similar medical misinformation. Frequently, Kennedy’s and Mercola’s website content refers to the other. “Annagoo\_” is frequently referenced with the stopcovidcold website/twitter handle.



## 0.21 Limitations/Next Steps

One of the limitations of the study was the vast amount of content by a single medical misinformer. While Mercola is interconnected with many of the major players in the anti-vaxx movement, gathering information from even the top propagators of medical misinformation will be an extensive effort. Additionally, many websites will focus on anti-vaxx and covid conspiracy issues while also branching off into other sub topics that merit following including GMO products, 5G and technocracy conspiracy theories. Some related websites are more political in content while others stick closer to health related issues. While this study does link Mercola with several medical misinformation topics it is still difficult to pinpoint the extent to which his websites influences public thought around these issues. More work will need to be done to connect the communicators and the message to the receivers in future studies. For next steps, the best way to share more descriptive information regarding internet misinformation would probably involve the construction of a database along with a website that would offer ways to explore data from the top participants in the anti-vaxx and covid conspiracy community. Studying trending hashtags related to vaccine and covid misinformation may also provide a way to pinpoint the interaction between followers and website subscribers and the message/messenger. Additionally, other platforms like Youtube, Instagram, and Parler would also need to be explored to fully understand patterns of communication within this community.

## **0.22 Conclusion**

The material explored in this study is clearly the tip of the iceberg. In fact, the depth and variety of communications found related to Mercola were considerably daunting. One mitigating factor, however, was the observation of how connected the major participants in this community are and how common threads are easily elicited from the material. It was also interesting to observe that Facebook's efforts to "shadowban" content is somewhat effective. Their efforts to crack down on individuals who share anti-vaxx material resulted in the pre-emptive avoidance of misinformation being spread by Mercola since he left the platform prior to covid in protest of Facebook's measures. While this outcome was unintended by Facebook, since these actions occurred prior to the pandemic, the outcome illustrates the importance of addressing misinformation whenever it appears, rather than waiting for a crisis to occur where misinformation becomes more prevalent and transmits more effectively. Twitter, on the other hand, appears to be less successful in discouraging medical misinformation on the internet. Its efforts to label content appear to be circumvented by Mercola and they still serve as a source for acquiring more subscribers. Nonetheless, Mercola's ultimate goal of attracting subscribers with whom he may directly communicate through a variety of venues indicates that social media is only one part of the problem of medical misinformation on the internet. Unfortunately, this creates an entirely different challenge in the effort to quell the flow of misinformation. Vigilance and knowledge of key players in various misinformation communities and a thorough knowledge of commonly held conspiracy theories at any given time are essential for pre-emptively and responsively sharing information in a way that contradicts and refutes harmful ideas that are shared in a variety of venues.

## **0.23 References**

- Kreidler, M. (2020, July 27). Dr. Joseph Mercola Ordered to Stop Illegal Claims | Quackwatch. <https://quackwatch.org/11ind/mercola/>
- Mercola.com—#1 Natural Health Website. Mercola.Com. Retrieved December 2, 2020, from <https://www.mercola.com/>
- Satija, N., & Sun, L. H. (2019, December 20). A major funder of the anti-vaccine movement has made millions selling natural health products. Washington Post. [https://www.washingtonpost.com/investigations/2019/10/15/fdc01078-c29c-11e9-b5e4-54aa56d5b7ce\\_story.html](https://www.washingtonpost.com/investigations/2019/10/15/fdc01078-c29c-11e9-b5e4-54aa56d5b7ce_story.html)
- Swine Flu Cases: An Interview with Sharyl Attkisson of CBS News. (2009, November 24). Mercola.com. <http://articles.mercola.com/sites/articles/archive/2009/11/24/superstar-cbs-reporter-blows-the-lid-off-the-swine-flu-media-hype-and-hysteria.aspx>
- Tsouderos, T. (2011, April 25). FDA warns doctor: Stop touting camera as disease screening tool. Chicago Tribune. <https://www.chicagotribune.com/lifestyles/ct-xpm-2011-04-25-ct-met-fda-warns-mercola-20110425-story.html>