Look who’s talking now: Cancer in primary care on Twitter. An observational study

Kristi M Milley1*, Sophie A Chima1, Kara-Lynne Cummings1, Jon D Emery1

1Primary Care Collaborative Cancer Clinical Trials Group, Centre for Cancer Research and Department of General Practice, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Victorian Comprehensive Cancer Centre, Melbourne, Australia

Abstract

Background: Twitter is a microblogging platform that helps share information. It is a dynamic tool that has been embraced by many user types including consumers and healthcare professionals (HCPs). Currently, there are no data on how cancer in primary care features on Twitter.

Aim: To explore the type of users and information shared about cancer in primary care on Twitter.

Design & setting: A descriptive exploratory study took place of publicly available Twitter data.

Method: Tweets were searched between July 2015 and June 2017 for ‘GP’, ‘general practice’, ‘primary care’, or ‘general practitioner’ in conjunction with ‘cancer’. A 20% random sample was coded for geographic location, user type, type of tweet, and theme. Tweet sentiment was analysed using R package sentimentr. Content that gained traction was compared by combining original tweets, retweets, favourites, and duration.

Results: There were a total of 3413 tweets from 1611 users in 44 countries. Consumers were the largest user group followed by health organisations, healthcare professionals, and the media. The most common theme across user types was diagnostic delay. Other themes that emerged included cancer screening, symptom awareness, and early diagnosis. Consumers published more negative tweets, particularly in relation to diagnostic delay. Health organisations focused on symptom awareness and screening. Over half of media tweets were stories that featured delayed diagnosis or screening.

Conclusion: A broad range of users engage with Twitter to share information about cancer in primary care. Content is different between user groups, but diagnostic delay and symptom awareness are common themes. Healthcare and professional organisations may need to consider approaches to counter negative messages about diagnostic delay.

How this fits in

Primary care plays an important role in the prevention, early detection, and diagnosis of cancer. There are also emerging opportunities to further integrate primary care in cancer follow-up and survivorship care. Twitter is a potential platform to communicate this message to different groups of stakeholders, including those at higher risk of cancer. It is an important social media tool that can be used for speed and brevity to communicate with a range of users. This is the first study to explore what content different users share about the role of primary care across the cancer continuum. It provides a foundation to better understand the public perception of primary care in this context.
This understanding will allow future research to leverage social media exposure to engage users with content designed to improve awareness of cancer and cancer symptoms, as well as the role of their GP in that journey from prevention through to survivorship.

**Introduction**

Social media are powerful communication platforms that are used by over 70% of the general public. At least half of adult internet users search for or share health-related information. Increasing knowledge and exchanging advice are two key reasons people engage with social media. There is great potential for education and communication with the public to disseminate key health messages. The most common social media platforms that feature in health are Facebook and Twitter. These applications provide a platform for creating and engaging different communities, improving information exchange and communication, and educating a wide range of users.

Twitter is a microblogging platform that helps capture and share bite-sized pieces of information. Users can share images and video with text up to 280 characters. Launched in 2006, by early 2019 Twitter had grown to nearly 321 million monthly users. The intersection of Twitter, oncology, education, and communication has been previously studied along with its impact and value for HCPs including oncologists or haematologists, radiologists, and surgeons. This has been reviewed in a variety of contexts including scientific conferences, networking and education, and improving information access.

In the context of public and patient engagement, there is evidence about the role of Twitter in cancer awareness, developing cancer social networks and reports of content analysis of cancer-related tweets. Twitter-based interventions have targeted cancer awareness, patient education, resource access and support, prevention, and screening. This evidence is almost exclusively from a secondary or tertiary care perspective, despite the important role primary care and GPs have in the prevention, detection, and early diagnosis of cancer.

As social media becomes increasingly popular it is important to understand how users promote and discuss specific topics. This understanding is an essential foundation to successfully developing targeted campaigns for cancer, which actively engage users and that translate into consumer actions. To date, there is no evidence on how the role of primary care and GPs in cancer are shared and discussed on Twitter. This exploratory study aimed to describe the different types of users and content shared about cancer in primary care, including what type of information generates more user engagement.

**Method**

**Search strategy and data collection**

Publicly available English language tweets between 30 June 2015 and 30 June 2017 were searched for the terms ‘GP’ or ‘general practitioner’ or ‘primary care’ and ‘cancer’. As collection and analysis of these publicly available tweets was an observational study and did not require interaction with any human subjects, it was exempt from institutional review. Tweets were identified, collected, and extracted through a data-services provider, Podargos (Calgary, Canada). A random number generator (Microsoft Excel 365) was used to randomly extract 20% of the tweets collected. Manual coding of the full sample set was not feasible, so a 20% sample size was selected guided by random extraction in existing studies.

**Data coding**

Tweets were assessed for relevance to cancer in primary care. Relevant tweets were then manually coded for the user type, tweet type, and geographic location. The user was determined by reading the biography and most recent tweets of the user account. To compare discussion of the most common cancer types, text was further searched for the terms ‘colorectal’, ‘bowel’, ‘colon’, ‘skin’, ‘melanoma’, ‘lung’, ‘prostate’, ‘prostate’ (sic), ‘ovar*’, ‘cerv*’, and ‘pancrea*’ (the asterisk acts as a wildcard character, to capture within the search results variants that begin with the specified letter string; for example, ‘cerv*’ would return cervix, cervical, etc). The dataset was coded by one researcher.
(KM) and the final complete dataset was reviewed by a second researcher (SC) to validate coder reliability. Any conflicts were resolved in consultation with a third team member (JE).

Data analysis

Descriptive analysis

Tweets were provided, coded, and analysed in Microsoft Excel. Descriptive statistics were used to summarise the user type, location, number of tweets per tweet type, and content theme. A quantitative content analysis was used to analyse tweets. A high-order review of tweets followed by an inductive coding approach was used to identify tweet themes. Given the brevity and succinct nature of tweets, each was coded for up to two themes.

Sentiment analysis

The overall sentiment of tweets was determined using RStudio (version 1.0.153) and the package sentimentr (version 2.6.1). A library of around 6800 English words, with predetermined sentiment scores, were used to determine whether the tweets were positive, negative, or neutral. The word ‘cancer’ was removed from the library before being analysed, as the sentiment library classified cancer as a negative word, which would skew analysis. The library also included incorrect spellings of words frequently misspelled on social media. The package scores a tweet between –6 to +6, with –6 being the most negative and +6 being the most positive score.

Content engagement

The top stories were determined by reviewing the tweet content and any external links to group together tweets about the same story. The total number of tweets for each story plus the total number of retweets and likes was used to rank the most popular stories. The duration of these stories was determined by calculating the number of active days on Twitter based on the date of the first and last tweet about each story.

Results

A total of 12,063 tweets were identified. Within the random sample \( (n = 2413) \), 11.8% \( (n = 284) \) were not related to cancer in primary care. The remaining 2129 tweets were produced by 1611 unique Twitter accounts (Table 1). Consumers were the largest user group, representing over a quarter of all users. In this context consumers were defined as patients with cancer, their carers, their family members, or members of the general public. Tweets were produced by users from 44 different countries, with the UK, US, Australia, and Canada producing 75% of the content (Table 2).

Table 1 Summary of the type of user and frequency of users and tweets

<table>
<thead>
<tr>
<th>User type</th>
<th>Users, n (%)</th>
<th>Tweets, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer</td>
<td>413 (25.6)</td>
<td>457 (21.5)</td>
</tr>
<tr>
<td>Health organisation</td>
<td>354 (22.0)</td>
<td>581 (27.3)</td>
</tr>
<tr>
<td>Healthcare professional</td>
<td>250 (15.5)</td>
<td>297 (14.0)</td>
</tr>
<tr>
<td>Media</td>
<td>229 (14.2)</td>
<td>350 (16.4)</td>
</tr>
<tr>
<td>Other</td>
<td>202 (12.5)</td>
<td>258 (12.1)</td>
</tr>
<tr>
<td>Unknown</td>
<td>163 (10.2)</td>
<td>186 (8.7)</td>
</tr>
<tr>
<td>Total</td>
<td>1611 (100)</td>
<td>2129 (100)</td>
</tr>
</tbody>
</table>

Table 2 Summary of tweet origins by geographic location

<table>
<thead>
<tr>
<th>Location</th>
<th>Tweets, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>1078</td>
</tr>
<tr>
<td>Unknown</td>
<td>370</td>
</tr>
<tr>
<td>US</td>
<td>308</td>
</tr>
<tr>
<td>Australia</td>
<td>139</td>
</tr>
<tr>
<td>Canada</td>
<td>74</td>
</tr>
<tr>
<td>Ireland</td>
<td>33</td>
</tr>
<tr>
<td>New Zealand</td>
<td>12</td>
</tr>
<tr>
<td>Spain</td>
<td>11</td>
</tr>
<tr>
<td>India, Nigeria, South Africa</td>
<td>10</td>
</tr>
<tr>
<td>Greece, Pakistan</td>
<td>8</td>
</tr>
<tr>
<td>Egypt</td>
<td>6</td>
</tr>
<tr>
<td>Belgium, France, Germany</td>
<td>4</td>
</tr>
<tr>
<td>Colombia, Indonesia, Japan</td>
<td>3</td>
</tr>
<tr>
<td>China, Italy, Nepal, Peru, Saudi Arabia</td>
<td>2</td>
</tr>
<tr>
<td>Cyprus, Czech Republic, Ethiopia, Israel, Jamaica, Kenya, Malaysia, Maldives, Malta, Mexico, Namibia, Norway, Oman, Philippines, Russia, Singapore, Sweden, Thailand, Uganda, Zimbabwe</td>
<td>1</td>
</tr>
</tbody>
</table>

Rows with more than one country indicate those countries produced the same number of tweets.
Overall, users’ tweets fitted into one of six categories. These were as follows: sharing news stories about cancer in primary care; promoting results of research studies; promoting a new resource, either for HCPs or consumers; promoting continuing education opportunities; promoting a funding opportunity; or sharing a personal statement.

Many of the tweets analysed mapped closely to a single aspect of the cancer care continuum. From this, themes that emerged included cancer symptom awareness, prevention, screening, diagnostic delay, survivorship, and palliative care. The most common theme across user types was about a delayed cancer diagnosis (19%). This was followed by symptom awareness (13%), cancer screening (9%), survivorship (5%), and early cancer diagnosis (4%). Other themes were about health system structure (4%) and patient barriers related to investigating possible cancer symptoms (2%) (Figure 1).

Some themes were more common to a specific user type. The most common theme of consumer tweets was a personal comment about their own experience of cancer or that of a friend or family member, in particular about perceived diagnostic delay (Figure 2). This was also a strong theme of media-user tweets (Figure 2). A second theme that featured strongly were stories around cancer screening, often highlighting the importance of primary care in increasing screening participation. HCPs also shared content about diagnostic delay, as well as more content about screening and survivorship (Figure 2). The strongest theme in health organisation tweets was cancer symptom awareness (Figure 2).

A third of tweets contained text about a specific tumour (n = 739/2129). Colorectal cancer (33%) was the most common tumour type followed by breast (18%); skin (17%), including both melanoma and non-melanoma skin cancers; lung (11%); and prostate (9%).

Overall, all user groups, except HCPs, produced tweets that were classified as mildly negative with their average score being close to neutral (0). Consumers’ tweets had the most negative scores (Table 3).

The top three stories shared by users that produced the greatest engagement were as follows: (1) ‘GP criticises colleagues who failed to diagnose her cancer’; (2) ‘GPs being paid not to refer cancer patients’; and (3) ‘Patients suspecting cancer put GP listening skills ahead of shorter wait times’. These three stories featured on Twitter for between 6 and 160 days (Figure 3).

Discussion

Summary
This is the first study to explore the content of tweets about cancer in primary care. Taking into consideration the complexity of a search strategy that could identify all relevant tweets, the number of tweets identified over the 2-year search period highlights that cancer in primary care is a small, niche...
Themes of tweets shared by users suggests a negative view around the role of primary care in cancer, particularly during cancer diagnosis. Content shared by both consumers and the media often appeared to perpetuate attitudes and expressions of blame towards GPs. This negative attitude was

<table>
<thead>
<tr>
<th>User type</th>
<th>-6</th>
<th>-5</th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Average score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>18</td>
<td>44</td>
<td>120</td>
<td>192</td>
<td>57</td>
<td>17</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.42</td>
</tr>
<tr>
<td>HCP</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>9</td>
<td>50</td>
<td>131</td>
<td>71</td>
<td>18</td>
<td>9</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0.18</td>
</tr>
<tr>
<td>Health organisation</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>16</td>
<td>44</td>
<td>117</td>
<td>278</td>
<td>89</td>
<td>24</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-0.16</td>
</tr>
<tr>
<td>Media</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>26</td>
<td>71</td>
<td>176</td>
<td>59</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.21</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>19</td>
<td>46</td>
<td>130</td>
<td>42</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.16</td>
</tr>
<tr>
<td>Unknown</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>11</td>
<td>35</td>
<td>112</td>
<td>31</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.21</td>
</tr>
</tbody>
</table>

HCP = healthcare professional.

**Figure 2** Emergent themes in cancer in primary care tweets by user type: consumers, media, healthcare professionals, and health organisations
also clearly evidenced by the tweets that gained the most engagement. These stories highlighted that even though trending on Twitter is often short lived, it does have the capacity to continue to propagate negative content for months.

The negative sentiment expressed by consumers and media agencies warrants further consideration, particularly in the context of shaping the public’s view of the importance of primary care across the cancer continuum. Negative bias presented by the media towards primary care in general has previously been demonstrated. A review of newspaper articles about neurological illnesses found that over 20% of articles contained medical errors, exaggerations, or stigmatising language. Primary care nurses have also reported their experiences of the influence of health stories reported in the media, which affected patient perceptions of care and the decisions they made about their care. In this respect, social media may both perpetuate the problem and be a solution to mediate public perception of cancer in primary care.

The content shared by HCPs suggest that this group use Twitter in a limited capacity. The themes in their tweets suggest they use it as a platform to share educational opportunities and resources with other HCPs. This is in line with other medical disciplines where the focus is on the potential of social media for HCPs in the context of educational and professional opportunities. The content of HCPs tweets also suggests little communication between GPs and consumers. Instead, the consumer relationship appeared to be nurtured by healthcare organisations, such as cancer-specific charities. Healthcare organisations also appeared to be the only group targeting multiple user groups, where their communications were targeted at HCPs, consumers, and the media.

Quantifying the impact of this conversation on Twitter is difficult, but it is reasonable to consider that a persistent negative perception could have an impact on the views and confidence of the general public in both their personal general practice and primary care more broadly with regards to cancer management, especially early diagnosis. The limited content identified about survivorship highlights there may be an opportunity to increase awareness of the role of primary care and GPs in this space.
Strengths and limitations
This is the first study to explore the content and themes of tweets shared on Twitter about cancer in primary care. There are limited studies that have undertaken content analysis of tweets and the methodology of these studies is varied. Capturing all involved stakeholders and tweets is difficult given the multitude of potential search terms and combinations to capture relevant tweets.

This study deliberately used the search terms ‘general practitioner’ and ‘GP’, which is a common term in Australia and the UK, but may be less commonly used in other countries to describe primary care physicians. The use of the broader ‘primary care’ term could potentially relate to medical professional groups other than GPs or family physicians in the US. There is evidence in other oncology fields that the development of a specific taxonomy and/or hashtags can help focus discussion and identify relevant content. Promoting primary care-specific hashtags may be a way to better centralise and identify this conversation.

The study presented was performed retrospectively and used available data for a 2-year period up to mid-2017. Consequently, it may not represent the rapidly growing content on Twitter. Additionally, given English language was an inclusion criterion, the analysis of the countries where tweets were posted will be skewed towards English-speaking countries. There are limited data on the validity of using sentiment analysis in this context. Lastly, the analysis did not attempt to identify tweets from ‘bot’ accounts. Up to 15% of Twitter accounts may be automated ‘bot’ accounts. The study was unable to estimate the impact of these accounts on propagating tweets or specific tweet themes. The reasoning was that this study was about what content was shared and engaged with by users. As such, there was no impact on the analysis if the content was produced by an automated account or a live person.

Comparison with existing literature
This is the first study to explore the content of conversations on Twitter about cancer in primary care. A 2017 systematic review of Twitter as a tool for health research found that only 5% of included studies were focused on cancer. None of the studies identified explicitly involved primary care. The most common fields were public health, infectious disease, behavioural medicine, and psychiatry.

Implications for research and practice
This study raises a question about if there is a need for primary care organisations to counter negative opinions and content around the role of primary care in cancer presented on Twitter. There is limited data on the impact of information shared by users on Twitter. In the age of influencers, understanding the spread and impact of content may help inform the development of targeted messages or awareness campaigns. It may also help identify primary care champions, users that are successfully supporting the role of primary care in cancer. In other disciplines, tweet chats have been used to engage in discussions with patients, although this has been limited to patients with breast and lung cancer.

Cancer in primary care is a niche conversation on Twitter but one in which the media and consumers, and their cancer experiences feature prominently. Predominately, these experiences were negative. There was a distinct absence of substantial positive promotion of primary care. Themes that emerged centred around perceived diagnostic delay, screening, and symptom awareness. As the role of primary care in cancer survivorship continues to grow there is an opportunity to increase awareness through Twitter and other social media platforms. Future research should explore the impact of negative sentiment on public perceptions of primary care and how organisations could influence sentiment within this platform to promote the importance of high-quality primary care across the cancer continuum.

Funding
Not applicable.

Ethical approval
As collection and analysis of these publicly available tweets was an observational study and did not require interaction with any human subjects it was exempt from institutional review.

Provenance
Freely submitted; externally peer reviewed.

Acknowledgements
This work was supported by the Primary Care Collaborative Cancer Clinical Trials Group (PC4).

References


