

Breast cancer prevention and treatment misinformation on Twitter: An analysis of two languages

DIGITAL HEALTH
Volume 9: 1–11
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076231205742
journals.sagepub.com/home/dhj



Izzati Yussof¹ , Nur Fa'izah Ab Muin² , Masnizah Mohd³,
Ernieda Hatah¹, Nor Asyikin Mohd Tahir¹ and Noraida Mohamed Shah¹ 

Abstract

Objective: To determine the prevalence and types of misinformation on Twitter related to breast cancer prevention and treatment; and compare the differences between the misinformation in English and Malay tweets.

Methods: A total of 6221 tweets related to breast cancer posted between 2018 and 2022 were collected. An oncologist and two pharmacists coded the tweets to differentiate between true information and misinformation, and to analyse the misinformation content. Binary logistic regression was conducted to identify determinants of misinformation.

Results: There were 780 tweets related to breast cancer prevention and treatment, and 456 (58.5%) contain misinformation, with significantly more misinformation in Malay compared to English tweets (OR = 6.18, 95% CI: 3.45–11.07, $p < 0.001$). Other determinants of misinformation were tweets posted by product sellers and posted before the COVID-19 pandemic. Less misinformation was associated with tweets utilising official/peer-reviewed sources of information compared to tweets without external sources and those that utilised less reliable information sources. The top three most common content of misinformation were food and lifestyle, alternative medicine and supplements, comprising exaggerated claims of anti-cancer properties of traditional and natural-based products.

Conclusion: Misinformation on breast cancer prevention and treatment is prevalent on social media, with significantly more misinformation in Malay compared to English tweets. Our results highlighted that patients need to be educated on digital health literacy, with emphasis on utilising reliable sources of information and being cautious of any promotional materials that may contain misleading information. More studies need to be conducted in other languages to address the disparity in misinformation.

Keywords

Health literacy, social media, infodemiology, digital divide, breast neoplasms

Submission date: 8 May 2023; Acceptance date: 18 September 2023

Introduction

Over the last few decades, the internet has become an important source of health information for the public. Physicians may be regarded as the most trusted source of information, but many patients turn to the internet instead as it is convenient and readily accessible.^{1,2} Internet users increasingly use social media such as Twitter, Facebook, and TikTok to seek and share health information. These platforms have gained participation from all social groups including health professionals and organizations that use

¹Centre for Quality Management of Medicines, Faculty of Pharmacy, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

²Oncology and Radiotherapy Department, Hospital Canselor Tuanku Muhriz, Cheras, Malaysia

³Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

Corresponding author:

Noraida Mohamed Shah, Centre for Quality Management of Medicines, Faculty of Pharmacy, Universiti Kebangsaan Malaysia, Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia.
Email: noraida_mshah@ukm.edu.my



these platforms to disseminate health-related knowledge.³ Although social media platforms provide immense opportunities for people to engage with each other in beneficial ways, they also allow misinformation to flourish as studies suggested that false information may spread more easily than true information.⁴ The fundamental role of health misinformation on social media has been highlighted by the COVID-19 pandemic when the World Health Organization (WHO) declared an ‘infodemic’ due to the overload of information following the pandemic in 2020.⁵

Concerns regarding the impact of online misinformation on public health spurred various research on this topic. A systematic review found that health misinformation was most common on Twitter and that most studies analysed posts only in one language.³ Online misinformation occurs on multiple health-related topics, from vaccine, to communicable and non-communicable diseases.³ Breast cancer is of particular concern, as it is a growing health problem and is currently the most commonly diagnosed cancer worldwide among females.⁶ Previous researchers have studied misinformation about breast cancer on various platforms including Pinterest, Facebook, Twitter, and Weibo.^{7–9} Posts containing misinformation ranged from 30 to 51%, comprising different topics from risk factors to treatment and prevention of breast cancer. Misinformation related to prevention and treatment poses a substantial concern among healthcare providers as it can potentially cause treatment delay and non-adherence.^{10,11} Online health-information seeking behaviour was associated with non-adherence to endocrine therapy for breast cancer in a prior study,¹² but it was also associated with better adherence in other disease conditions,¹³ suggesting that other factors may have contributed to the differences in outcome, such as prevalence of misinformation.

Despite the large number of studies on misinformation in social media, many of these focused only on one language, especially English.^{3,14–17} The problem of health misinformation is not limited to English-language content, and it can be even more challenging in languages other than English. Malaysia is a multiracial country that uses Malay as the national language. The use of English and Malay are both prevalent in the country, but the level of English proficiency varies. For people with limited English proficiency, social media may become a major source of information as various content from fellow native speakers is readily available. Unlike traditional media, content is created by individuals who may have limited knowledge of medicine with no factual verification or accountability, thus presenting a higher risk of misinformation.

Therefore, this study aims to determine the prevalence and types of misinformation on Twitter related to breast cancer prevention and treatment, and compare the differences between the misinformation in English and Malay tweets. Several different terminologies exist when discussing the spread of false information on the internet. Misinformation is regarded

as an inadvertent spread of false information, while disinformation involves knowingly sharing or creating false information to cause harm.¹⁸ Following previous works, we use the term ‘misinformation’ as an umbrella term to describe health-related false information, giving the benefit of the doubt to the users involved, as the intent is not always clear.¹⁰ To our knowledge, this is the first study investigating misinformation on social media in Malay language regarding breast cancer prevention and treatment. This study is important in order to understand the content of misinformation and identify the determinants that can predict misinformation.

Methods

Study design and data collection

This quali-quantitative study examined data from Twitter to compare the extent of misinformation on breast cancer treatment and prevention between English and Malay tweets and explore their characteristics. Permission to use Twitter public Application Programming Interface (API) endpoints was obtained prior to data collection,¹⁹ using web scraper scripted in Python to collect related tweets that are publicly available. The following information was collected: the date of post, text content, URL to the original tweet, and URL to external sites. Ethical approval to conduct the study was obtained from the Research Ethics Committee, Universiti Kebangsaan Malaysia (JEP-2022-240). The need for informed consent was waived by the ethics committee because this study used only publicly available data published voluntarily by Twitter users. The Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist was used as a guide for reporting the qualitative aspect of the study (Appendix 1).²⁰

Inclusion criteria are tweets containing information regarding prevention and treatment for breast cancer, and posted between January 2018 until March 2022 in Malay or English language. We excluded tweets that contain more than 50% of words in languages other than English or Malay, and those that are inaccessible due to broken links. We aimed for at least 3000 tweets for each language,⁹ using the keywords ‘breast cancer’ for English tweets and ‘kanser payudara’ or ‘kanser payu dara’ for Malay tweets. A preliminary search showed considerably more English than Malay tweets, with over 30,000 English tweets in one month. Therefore, all Malay tweets from the data collection period were included for screening, and tweets in English were randomly selected from different periods based on computer-generated random sequences using Python. Tweets from October were purposely collected each year to explore the characteristics of tweets shared during the Breast Cancer Awareness Month (BCAM) in October.

Categorisation and content analysis

The tweets were first screened to remove duplicates and identify tweets not written in English or Malay languages.

Table 1. Tweet categorisation: definition and examples.

Category	Definition	Example
Tweet categories		
Prevention and treatment	Tweets about prevention or treatment of breast cancer	<i>FDA Approves Adjuvant Olaparib in BRCA-Mutated HER2-Negative High-Risk Early Breast Cancer.</i>
Other medical information	Tweets about other breast cancer information related to epidemiology, screening, risk factors and symptoms of breast cancer	<i>Exposure to common chemical during pregnancy may reduce protection against breast cancer: Research suggests propylparaben is an endocrine disruptor.</i>
Medically irrelevant	Tweets not containing any medical-related information such as personal stories about breast cancer, solidarity, and fundraising	<i>Rocking pink in honour and memory of everyone that has won or lost the fight of Breast Cancer.</i>
Type of misinformation		
Inaccurate	Tweets containing inaccurate, misleading, or exaggerated benefits or dangers compared to actual scientific evidence	<i>Turmeric has phenomenal anti-cancer properties and has been known to help to inhibit breast cancer.</i>
Fabricated	Tweets with no evidence of the claim in the scientific literature	<i>Cure breast cancer through yoga meditation. No diet. No medicine.</i>

Note: Tweets in Malay were translated to English. Some tweet examples in the manuscript were rephrased to protect the identity of users and hide the name of products promoted. Spelling corrections were made for better clarity in tweets with short forms and spelling errors.

Duplicates were removed by identifying duplicate tweet URLs using the Excel ‘Remove duplicate’ function. The remaining were then assessed for medical relevance, and those with broken links were removed. Tweets that contain medical-related information were assessed to identify content related to the treatment and prevention of breast cancer. These categorizations were conducted by a single researcher (IY).

Tweets on prevention and treatment of breast cancer were further evaluated and categorised based on accuracy (true information/misinformation). Two female coders—a pharmacist (IY) and an oncologist (NFAM)—independently evaluated the tweets on prevention and treatment for the accuracy of the information, types, and content of misinformation. Any discrepancies were resolved by a third person (a female pharmacy lecturer, NMS). For tweets containing links to external sources (e.g.: webpages, articles, images, videos, podcasts and social media posts in other platforms), accuracy of information in these were also verified. Information accuracy was judged based on experience and clinical knowledge whenever possible, or cross-referenced with reliable sources including the clinical practice guidelines on breast cancer and peer-reviewed journal articles. The inter-rater reliability was calculated using Cohen’s Kappa, which showed a good agreement level at 0.874 and 0.887 for English and Malay tweets, respectively.

Tweets were also categorised based on the source of information, presence of attempts to sell services or products, and the timing—whether they were posted during

BCAM in October, or whether tweets were posted before or after the COVID-19 pandemic (tweets before March 2020 were considered pre-pandemic). Using content analysis, tweets that contain misinformation were further categorised based on the type and content of misinformation. Categorisation was done based on previous breast cancer misinformation studies,^{7,8} and adapted to suit our data. Table 1 summarises the definition and examples of tweet from each content category.

Statistical analysis

Data were categorised in Microsoft Excel 2019 and quantitatively analysed using IBM SPSS Statistics version 26 (IBM Corporation, Armonk, NY). Continuous data were presented with means and standard deviations, and categorical data with frequencies and percentages. Pearson Chi-square test was used to identify the association between misinformation and tweet language, the timing of posts, sellers-associated tweets, and the source of information. Binary logistic regression was used to estimate the odd ratios and 95% confidence intervals of these variables.

Results

Identification and screening of tweets

A total of 3167 Malay tweets were collected from the keywords ‘kanser payudara’ ($n = 2759$) and ‘kanser payu dara’

($n = 408$) between the years 2018 and 2022. For the English language, 344,736 tweets were randomly collected from the same period, of which 3054 were randomly selected. A total of 6221 tweets authored by 4245 unique users were collected and analysed. More than half of the tweets were not medically related ($n = 3261$, 52.4%). Among those that were medically related, information on epidemiology, screening, risk factors and symptoms ($n = 1646$, 61.2%) was more prevalent than information on prevention and treatment ($n = 780$, 29.0%). A flow diagram summarising the identification and screening of tweets is shown in Figure 1.

Categorisation and sources of misinformation

A summary of tweet categories is shown in Table 2. From the 780 tweets analysed for content accuracy, 456 (58.5%) misinformation was identified. More than half of the tweets contain links to external sources of information ($n = 512$, 65.6%). External links found in the tweets included

websites managed by credible health-related organisations and those peer-reviewed by medical professionals (e.g.: National Institutes of Health, Mayo Clinic, WebMD), blogs, private company websites (e.g.: AstraZeneca, Pfizer, Farmashealth), news portals (e.g.: Yahoo News, Reuters, Daily Mail, Bernama, Berita Harian), journal articles, social media sites (e.g.: Facebook, Instagram), and videos or podcasts from external websites (e.g.: YouTube, Vimeo). For analysis, sources of information were further grouped into official/peer-reviewed sources, miscellaneous sources, and no external source. Tweets categorised as official/peer-reviewed were posted by Twitter accounts affiliated with health organisations (e.g.: Ministry of Health, hospitals, cancer-related organisations) or those containing links to websites affiliated with health organisations or contain information that are peer-reviewed (e.g.: news portals, journal articles, health portals). Extensive lists of official websites and peer-reviewed health-related sources are available in Appendices 2 and 3. Tweets by personal accounts without external sources were grouped as

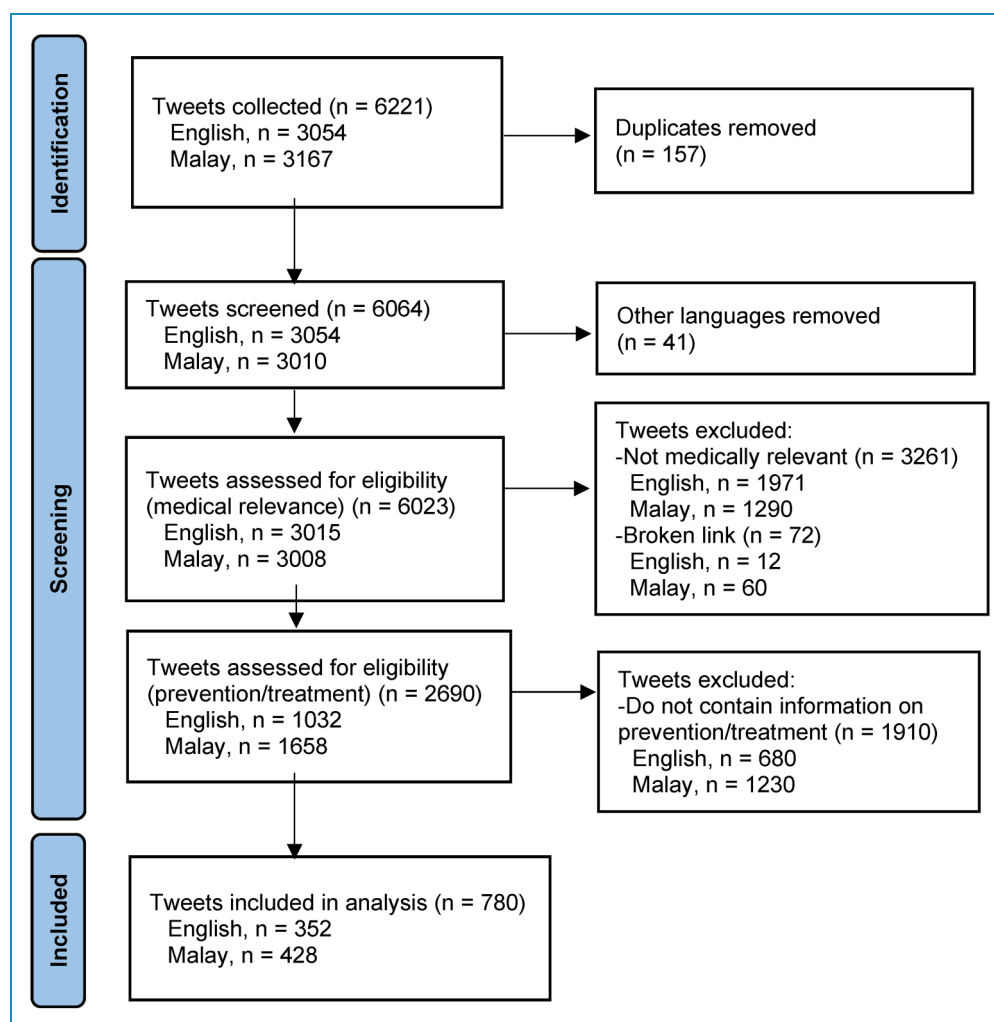


Figure 1. Flow diagram summarising identification and screening of tweets.

Table 2. Summary of information sources used in tweets.

Information sources	English		Malay		Total	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
No external source	96	27.3%	172	40.2%	268	34.4%
Official/peer-reviewed sources	200	56.8%	54	12.6%	254	32.5%
Official organisation/peer-reviewed websites	89	25.3%	15	3.5%	104	13.3%
News	53	15.0%	38	8.9%	91	11.6%
Research/journal articles	58	16.5%	1	0.2%	59	7.6%
Miscellaneous sources	56	15.9%	202	47.2%	258	33.1%
Blogs/unofficial websites	32	9.1%	147	34.3%	179	23.0%
Company websites	15	4.3%	12	2.8%	27	3.5%
Facebook/Instagram	0	0.0%	37	8.7%	37	4.7%
YouTube videos/podcasts	9	2.5%	6	1.4%	15	1.9%
Grand total	352	100.0%	428	100.0%	780	100.0%

‘No external source’. The others were categorised as miscellaneous sources (e.g.: blogs, company websites, social media sites, websites that are not peer-reviewed). This is shown in Table 2, which also compares the sources of information used in English and Malay tweets.

Approximately one-quarter of the tweets ($n = 199$, 25.5%) were posted by sellers attempting to promote their products or services, mostly consisting of Malay tweets ($n = 175$, 87.9%). As shown in Table 3, misinformation was more prevalent in tweets from the following categories: Malay language, posted by product sellers, posted before COVID-19, posted in months other than BCAM, tweets without any external sources, and containing information from miscellaneous sources. A chi-square analysis revealed a significant association between all these categories and the presence of misinformation.

A binary logistic regression was performed to ascertain the effects of language, product seller, information source, and timing of tweet in relation to COVID-19 and BCAM on the likelihood that the tweet contains misinformation. The logistic regression model was statistically significant, $\chi^2(6) = 643.4$, $p < 0.001$. The model explained 75.6% (Nagelkerke R^2) of the variance in misinformation and correctly classified 90.0% of cases. Sensitivity was 96.1%, specificity was 81.5%, positive predictive value was 88.0% and negative predictive value was 93.6%. Four of the five predictor variables analysed were statistically significant, as shown in Table 4.

Content of misinformation

More than half of the tweets on breast cancer prevention and treatment contained misinformation. Among these, 233 (51.1%) contained fabricated information, and 223 (48.9%) were considered inaccurate. In the English tweets, 68 (69.4%) out of 98 misinformation were categorised as fabricated and 30 (30.6%) were inaccurate information. For Malay tweets, 165 (46.1%) out of 358 were fabricated, and 193 (53.9%) were inaccurate. The content of misinformation was further categorised into themes. The top three most common themes for misinformation were food & lifestyle (i.e. tweets that advocate specific food or lifestyle for breast cancer), alternative medicine (i.e. tweets that advocate the use of herbs or health products as a treatment or cure for breast cancer), and supplements (i.e. tweets that advocate the use of herbs or health products to improve breast cancer symptoms/outcome). Most of the misinformation related to supplements and alternative medicine was made by product sellers ($n = 153/196$, 78.1%). Figure 2 summarises the distribution of misinformation content by language.

The majority of the misinformation regarding food and lifestyle were categorised as inaccurate ($n = 175/240$, 72.9%), consisting of exaggerated claims of anti-cancer properties of food such as soy, carrot, turmeric, *bambangan*, mango, fig, and olive; and various substances such as lycopene, resveratrol, and curcumin.

Table 3. Categories of tweets and presence of misinformation.

Categories	n	Misinformation		p-value
		Yes (n = 456)	No (n = 324)	
Language				
Malay	428	358 (83.6%)	70 (16.4%)	<0.001
English	352	98 (27.8%)	254 (72.2%)	
Tweets posted by product sellers				
No	581	267 (46.0%)	314 (54.0%)	<0.001
Yes	199	189 (95.0%)	10 (5.0%)	
Tweets posted before or after COVID-19				
Before COVID-19	455	319 (70.1%)	136 (29.9%)	<0.001
After COVID-19	325	137 (42.2%)	188 (57.8%)	
Tweets posted during BCAM				
No	637	398 (62.5%)	239 (37.5%)	<0.001
Yes	143	58 (40.6%)	85 (59.4%)	
Information sources				
No external sources	268	231 (86.2%)	37 (13.8%)	<0.001
Miscellaneous sources	258	215 (83.3%)	43 (16.7%)	
Official/peer-reviewed sources	254	10 (3.9%)	244 (96.1%)	

Note: Percentages were calculated based on the content of misinformation for each subcategory.

Research showed that drinking a glass of soy milk can prevent breast cancer by 30%.

The results showed that high carrot intake was associated with a 21% decreased risk of breast cancer.

There were also exaggerated claims on the benefits of yoga, meditation, and veganism for breast cancer prevention or treatment.

Just go raw vegan. I know four people who have cured stage 4 breast cancer by going vegan.

We observed many repetitive claims that hard-boiled eggs could cut the risk of getting breast cancer tweeted to various users by the same account between 2020 and 2021:

You should eat 2-3 hard-boiled eggs every day for breakfast because they cut your risk of breast cancer by 44% and heart disease by 26%.

In contrast, misinformation on alternative therapy was mainly fabricated information (n = 134/158, 84.8%), comprising claims of breast cancer cure for various food, supplements, traditional, and plant-based products, such as Sabah snake grass, cannabidiol oil, aromatic ginger leaf, keladi tikus (*Typhonium flagelliforme*), butterfly wing plant, and Cascara sagrada.

Sabah snake grass. Cure for cancer and other chronic illnesses.

Leukaemia | Prostate | Brain cancer | Breast cancer | and many more. Like this page < link to Facebook > .

Table 4. Binary logistic regression model for the determinants of misinformation.

Variable	Coef (B)	SE	Odds ratio (Exp (B))	95% CI		p-value
				Lower	Upper	
Language						
English			1			
Malay	1.821	0.297	6.178	3.449	11.065	<0.001 ^a
Tweet posted by product sellers						
No			1			
Yes	1.395	0.409	4.034	1.808	9.002	0.001 ^a
Source of information						
No external source			1			
Official/peer-reviewed sources	-4.981	0.435	0.007	0.003	0.016	<0.001 ^a
Miscellaneous sources	-1.007	0.325	0.365	0.193	0.691	0.002 ^a
Timing in relation to COVID-19						
After COVID-19			1			
Before COVID-19	0.947	0.318	2.578	1.382	4.812	0.003 ^a
Timing in relation to BCAM						
Non-October			1			
October	0.517	0.360	1.677	0.828	3.396	0.151

SE: standard error; CI: confidence interval.

^aStatistically significant.

Cannabidiol oil can cure almost all kinds of cancer and can be also helpful by many other diseases.

Almost half of the information on alternative therapy was found in promotions for amygdalin or laetrile, also commonly described as ‘Vitamin B17’, which was claimed to be an effective alternative to chemotherapy:

Free from breast cancer with <product name>. ‘I had to undergo chemotherapy 6 times but I refused and chose to take <product name>.’

For supplements, the misinformation content was evenly categorised as fabricated ($n = 20/38$, 52.6%) and inaccurate ($n = 18$, 47.4%). The tweets in these categories did not make any claim for cure, but claims were mostly made

for their positive effects on immune system and various body functions, as well as their ability to prevent cancer. These include promotions for omega-3 and docosahexaenoic acid (DHA) supplements, various herbal tea, glutathione supplement, soy-based products, and ‘super juices’:

DHA functions: Prevents asthma. Increases sexual function. Prevents skin problems. Prevents breast cancer. Prevents prostate cancer. Protects from migraine and numbness. Prevents coldness on feet and hands.

Jane says our non-genetically modified, soy-free <product name> has helped to restore her energy after going through breast cancer treatment, giving her an overall feeling of well-being and helping to clear her brain fog. (*name has been changed to protect identity)*

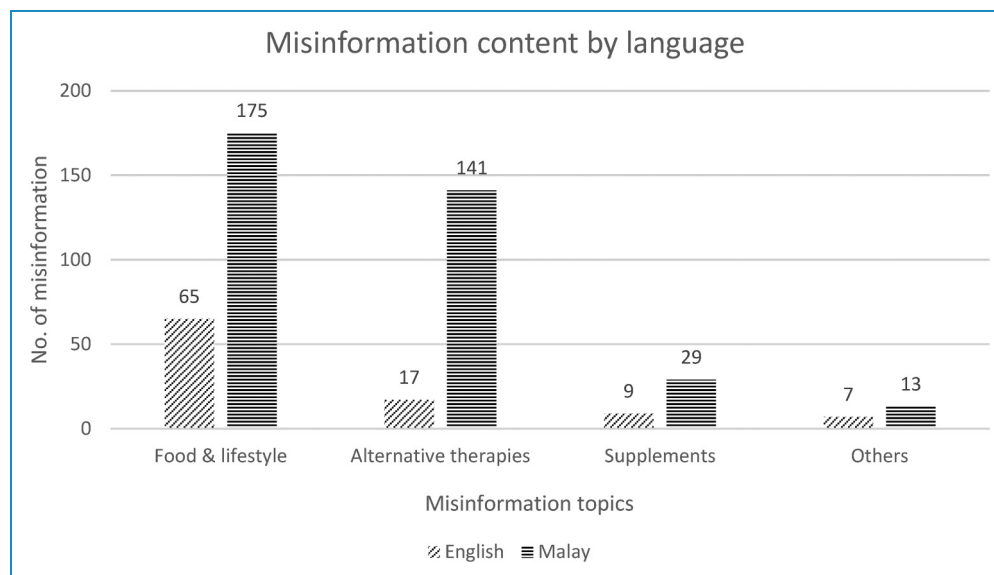


Figure 2. Summary of misinformation content by language.

Misinformation under the ‘Others’ category includes misinformation about hormone therapy, faith, surgery, gene testing, topical application, cupping, and conspiracy theory:

Only 5% of women benefit from extended anti-estrogen therapy after the first 5 years. 100% of women deserve to know what is right for them. Do you know what’s right for her? Order <product name>.

<cream product> contains piperine (known to prevent cancer and has anti-breast cancer properties), vitamin C, vitamin A, flavonoid, carotenoid, antioxidant (fights free radicals and protect the body from diseases).

We had the cure for #breastcancer. But once it’s under control, it’s no longer a money-maker. The suits in ivory towers who don’t deal with patients changed the paradigm by lying to you.

Discussion

Our study evaluated the distribution and content of misinformation on breast cancer prevention and treatment in English and Malay tweets. To our knowledge, this is the first study that compared misinformation regarding breast cancer in two languages, and we found substantially more misinformation in Malay tweets compared to English. This could be attributed to the sources used for the tweets, as Malay tweets tend to use information from less reliable sources such as blogs, non-peer-reviewed websites, and other social media sites such as Facebook and Instagram. Misinformation was often created by individuals

with no official institutional affiliation; therefore, it is important to check the credibility of information source.^{3,10} Another important aspect to consider is the more extensive research on algorithms to detect misinformation in English compared to Malay language.^{14–17} Social media sites such as Twitter attempted to combat misinformation by having specific features such as labelling content that may contain problematic information and giving prompts when users engage with a potentially misleading post.^{21,22} However, these features will be obsolete if the algorithm to detect misinformation in the Malay language is deficient.

The most common types of misinformation were on food, lifestyle approaches, and supplements that can purportedly prevent or improve breast cancer outcome. These types of misinformation were also found in previous studies on breast cancer misinformation.^{7–9} Many patients tend to look beyond conventional treatment for cancer prevention and treatment, with much preference for natural-based and non-pharmacological options.²³ It was alarming to note that approximately one-third of the misinformation found in the present study were advocating the use of these unproven options as alternatives to evidence-based conventional treatment, and majority of these were present in the Malay tweets. In past studies, the use of complementary treatment was associated with refusal of conventional treatment, and breast cancer patients who opted for alternative treatment were found to have a five-fold increase in the risk of death.^{24,25} Delay in seeking health treatment and diagnosis is a major concern in Malaysia, where the prevalence of delay was higher than in other developed and developing countries.²⁶ This problem was associated with several factors, including the use of alternative therapy and negative attitude towards the treatment.²⁶ Within the

Malaysian context, belief in traditional healers is highest among the Malay ethnic group, which contributes to the delay in getting treatment, causing a poor outcome and a low survival.²⁷ The widespread misinformation on breast cancer treatment alternatives on social media may be a contributing factor that needs to be intervened to tackle this issue. The use of concurrent complementary therapy with conventional cancer treatment could potentially jeopardise the effectiveness of the treatment.^{28–30} It could also affect kidney and liver functions, thus further complicating the course of treatment.³¹

Users trying to promote their products or services spread much of the misinformation related to using supplements and health products. Our results showed that product sellers were four times more likely to spread misinformation compared to general users. Similar findings were also noted in past studies, where misinformation was disseminated by private companies promoting their products.^{3,32–34} There were instances where products registered as supplements were promoted as a prevention or alternative ‘cure’ for cancer. This is especially common in the Malay tweets, despite it being illegal in Malaysia to advertise any product for cancer treatment to the general public.³⁵ Promotions on the Internet may have contributed to the increased popularity of these products among cancer patients. The use of complementary and alternative approach among cancer patients has increased from 25% in the 1970s to over 49% after 2000, according to data collected in Europe, Australia and North America.²³ The prevalence of CAM use among Malaysian cancer patients was reportedly higher at 61.2%.³⁶ A recent meta-analysis found that the prevalence of herbal medicine usage in cancer patients was higher among African and Asian countries, and prevalence was also higher in patients from low- and middle-income countries.³⁷ A local study found that many Malaysian women perceived herbal medicines as being safer and more effective than modern medicines.³⁸ This is often how health products were promoted on social media, and they are often accompanied by testimonials claimed to be from other patients facing the same health condition. The effectiveness of these online promotions may be attributed to the preferences of certain groups of Asian patients to obtain information through person-to-person communication, as opposed to Caucasian patients preferring objective, scientific information obtained from research institutions.³⁹ Healthcare providers in a country with multicultural population need to be aware of the diversity in patient beliefs and preferences, so that healthcare education can be tailored accordingly.

The WHO declaration on the ‘infodemic’ following COVID-19 has raised awareness and concerns on the effect of misinformation on public health, spurring various research on this topic. However, our results show that misinformation on breast cancer prevention and treatment was more common before the COVID-19 pandemic. Similarly, a study on COVID-19 misinformation found

that posts about COVID-19 included less misinformation than other health-related posts prior to the pandemic.⁴⁰ Although this contradicted our initial assumption that COVID-19 catalysed the propagation of misinformation, the declaration by WHO on the dangers of ‘infodemic’ may have created awareness that resulted in better detection and prevention of online misinformation. A previous study on breast cancer misinformation found an increase in news stories classified as ‘verified’ in BCAM of October compared to other months.⁷ In our data, chi-square analysis revealed a significant association, but it was not significant in the logistic regression analysis.

Research and practice implication

Our study showed that misinformation on breast cancer prevention and treatment was common on social media, and this problem is especially prominent in Malay tweets. It can potentially distort patients’ belief about their treatment, thus worsening the problem with poor treatment adherence and poor outcome. Most of the misinformation observed may have stemmed from lack of knowledge in evidence-based medicine and unfamiliarity with best practices when sharing online information. Individuals with low digital health literacy may be more prone to judge information credibility based on superficial qualities such as image quality, position in search engine results and celebrity endorsements.^{2,41} The importance of using official and peer-reviewed sources for health-related information needs to be advocated. Health literacy and digital literacy education should be improved, especially for older patients who may be more susceptible to misinformation.^{42,43} Furthermore, there is a need to increase accessibility of reliable information sources in Malay, as our observation revealed that there are less sources available in Malay compared to English. There is also a need for stricter monitoring of online advertisements for healthcare products. Social media is a popular medium for product promotion and the volume of content may make it difficult for manual monitoring. This provides opportunities for future research, particularly in the application of machine learning to detect patterns and identify posts that may contain unethical product promotion and flag them for review by human moderators.

Study limitations

Data for Malay tweets were relatively scarce compared to English, as Twitter is not the most commonly used social media platform by Malaysians.⁴⁴ However, Twitter was chosen as it is relatively more difficult to access data for certain social media due to concerns over data privacy that led governments to tighten regulation on data access and storage, and issues involving data breaches such as the Cambridge Analytica scandal.^{3,45} Despite this, we obtained rich data from the Malay tweets, which also

contained links to more commonly used social media sites such as Facebook, Instagram, and YouTube. Twitter is usually used by people from the younger generation, and may have limited representation of online information shared by older population. As breast cancer is more prevalent in older population, misinformation circulated in this subgroup may have been missed. Despite these limitations, we believe that this study contributes meaningful findings as it highlights prevalence of misinformation in a relatively less-studied language. In addition, data were collected over several years, providing opportunities to analyse misinformation on continuously evolving topics of discussion.

Conclusion

Misinformation on breast cancer prevention and treatment is prevalent on social media. This study showed that there was twice as much misinformation in Malay compared to English tweets, thus highlighting the need to address misinformation in other languages. Malay tweets utilised less information from official and peer-reviewed sources compared to English tweets, which may have contributed to the extensive degree of misinformation. Content of misinformation predominantly comprised of information on supplements, food and lifestyle, and alternative therapies. Tweets posted by product sellers were found to be a significant determinant of misinformation. These findings suggest that patients need to be educated on digital health literacy, with emphasis on utilising the right information sources, critical appraisal of information and being wary of any promotional materials that may contain misleading information.

Acknowledgements: We would like to thank Mr Nishakaran Pushpa Rajah for his assistance in the data collection process.

Contributorship: IY, NMS, EH and NAMT researched the literature and conceived the study, and were involved in gaining ethics approval. IY, NMS and MM were involved in protocol development and data collection. IY, NMS and NFAM were involved in data analysis. IY wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Ethical approval: Permission to conduct this study was obtained from the Universiti Kebangsaan Malaysia (UKM) Research Ethics Committee (REC number: JEP-2022-240).

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Universiti Kebangsaan Malaysia Research Grant (grant number GUP-2020-004).

Guarantor: NMS

ORCID iDs: Izzati Yussof  <https://orcid.org/0000-0003-0801-4792>

Nur Fa'izah Ab Muin  <https://orcid.org/0000-0001-9755-8032>

Noraida Mohamed Shah  <https://orcid.org/0000-0002-2509-4213>

Supplemental material: Supplemental material for this article is available online.

References

- Hesse BW, Nelson DE, Krepes GL, et al. Trust and sources of health information: the impact of the internet and its implications for health care providers: findings from the first health information national trends survey. *Arch Intern Med* 2005; 165: 2618–2624. J Archives of Internal Medicine.
- Petric G, Cugmas M, Petric R, et al. The quality of informational social support in online health communities: a content analysis of cancer-related discussions. *Digit Health* 2023; 9: 20552076231155681. 2023/02/25.
- Suarez-Lledo V and Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res* 2021; 23: e17187. 2021/01/21.
- Vosoughi S, Roy D and Aral S. The spread of true and false news online. *Science* 2018; 359: 1146–1151. 2018/03/29.
- Organisation WH. Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation, <https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation> (2020).
- Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; 71: 209–249. 20210204.
- Biancovilli P, Makszin L and Csongor A. Breast cancer on social media: a quali-quantitative study on the credibility and content type of the most shared news stories. *BMC Womens Health* 2021; 21: 202. 2021/05/17.
- Wilner T and Holton A. Breast cancer prevention and treatment: misinformation on Pinterest, 2018. *Am J Public Health* 2020; 110: S300–S304. 2020/10/02.
- Chen L, Wang X and Peng TQ. Nature and diffusion of gynecologic cancer-related misinformation on social media: analysis of tweets. *J Med Internet Res* 2018; 20: e11515. 2018/10/18.
- Wang Y, McKee M, Torbica A, et al. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med* 2019; 240: 112552. 2019/09/29.
- Nelson AJ, Puri R and Nissen SE. Statins in a distorted mirror of media. *Curr Atheroscler Rep* 2020; 22: 37. 2020/06/20.
- Quinn EM, Fleming C and O'Sullivan MJ. Endocrine therapy adherence: a cross-sectional study of factors affecting adherence and discontinuation of therapy. *Ir J Med Sci* 2016; 185: 383–392. 2015/05/14.
- Lim HM, Dunn AG, Lim JR, et al. Association between online health information-seeking and medication adherence:

- a systematic review and meta-analysis. *Digit Health* 2022; 8: 20552076221097784. 2022/05/20.
14. Di Sotto S and Viviani M. Health misinformation detection in the social web: an overview and a data science approach. *Int J Environ Res Public Health* 2022; 19(4):2173. 2022/02/26.
 15. Alameri SA and Mohd M. Comparison of Fake News Detection using Machine Learning and Deep Learning Techniques. In: *2021 3rd International Cyber Resilience Conference (CRC)* 29-31 Jan. 2021 2021, pp.1–6.
 16. Alonso MA, Vilares D, Gómez-Rodríguez C, et al. Sentiment analysis for fake news detection. *Electronics* 2021; 10(11): 1348.
 17. Ramakrishnan K and Balakrishnan V. Health Misinformation in the Covid-19 Era - Detecting Misinformation on Bi-lingual Corpora using Lexical Features. In: *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)* 16-18 Nov. 2022 2022, pp.1-5.
 18. Wardle C. Fake news. It's complicated., <https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79> (2017, accessed November, 21 2021).
 19. Twitter. Developer Platform, <https://developer.twitter.com/en/docs> (2022, accessed 9 June 2022 2022).
 20. Tong A, Sainsbury P and Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007; 19: 349–357.
 21. Ozturk P, Li H and Sakamoto Y. Combating Rumor Spread on Social Media: The Effectiveness of Refutation and Warning. In: *2015 48th Hawaii International Conference on System Sciences* 5-8 Jan. 2015 2015, pp.2406–2414.
 22. Vraga E, Tully M and Bode L. Assessing the relative merits of news literacy and corrections in responding to misinformation on Twitter. *New Media & Society* 2022; 24: 2354–2371.
 23. Horneber M, Bueschel G, Dennert G, et al. How many cancer patients use complementary and alternative medicine: a systematic review and metaanalysis. *Integr Cancer Ther* 2012; 11: 187–203. 2011/10/25.
 24. Johnson SB, Park HS, Gross CP, et al. Use of alternative medicine for cancer and its impact on survival. *J Natl Cancer Inst* 2018; 110: 121–124.
 25. Johnson SB, Park HS, Gross CP, et al. Complementary medicine, refusal of conventional cancer therapy, and survival among patients with curable cancers. *JAMA Oncol* 2018; 4: 1375–1381. 2018/07/22.
 26. Norsa'adah B, Rampal KG, Rahmah MA, et al. Diagnosis delay of breast cancer and its associated factors in Malaysian women. *BMC Cancer* 2011; 11: 141. 2011/04/19.
 27. Abdullah N and Mohamed N. Influence of cultural practices on breast cancer risks, stage at presentation and outcome in a multi-ethnic developing country. *Oncol Lett* 2021; 22: 806. 2021/10/12.
 28. Borrelli F and Izzo AA. Herb-drug interactions with St John's wort (*Hypericum perforatum*): an update on clinical observations. *AAPS J* 2009; 11: 710–727. 2009/10/28.
 29. Jung AY, Cai X, Thoene K, et al. Antioxidant supplementation and breast cancer prognosis in postmenopausal women undergoing chemotherapy and radiation therapy. *Am J Clin Nutr* 2019; 109: 69–78. 2019/01/23.
 30. Ambrosone CB, Zirpoli GR, Hutson AD, et al. Dietary supplement use during chemotherapy and survival outcomes of patients with breast cancer enrolled in a cooperative group clinical trial (SWOG S0221). *J Clin Oncol* 2020; 38: 804–814. 2019/12/20.
 31. Ernst E. Risks of herbal medicinal products. *Pharmacoepidemiol Drug Saf* 2004; 13: 767–771. 2004/09/24.
 32. Allem JP, Ferrara E, Uppu SP, et al. E-Cigarette surveillance with social Media data: social bots, emerging topics, and trends. *JMIR Public Health Surveill* 2017; 3: e98. 2017/12/22.
 33. Merianos AL, Gittens OE and Mahabee-Gittens EM. Depiction of health effects of electronic cigarettes on YouTube. *J Subst Use* 2016; 21: 614–619. 2017/02/22.
 34. O'Connor C, O'Grady C and Murphy M. Spotting fake news: a qualitative review of misinformation and conspiracy theories in acne vulgaris. *Clin Exp Dermatol* 2022; 47: 1707–1711. 2022/04/19.
 35. Medicines (Advertisement & Sale) Act 1956. Malaysia.
 36. Hamed Abdalla MEA, Ali AM and Loong L. The use of complementary and alternative medicine (CAM) among cancer patients at a tertiary hospital in Malaysia. *Complement Ther Med* 2020; 50: 102343. 2020/05/24.
 37. Asiimwe JB, Nagendrappa PB, Atukunda EC, et al. Prevalence of the use of herbal medicines among patients with cancer: a systematic review and meta-analysis. *Evidence-Based Complementary Altern Med* 2021; 2021: 9963038.
 38. Tengku Mohamad TAS, Islahudin F, Jasamai M, et al. Preference, perception and predictors of herbal medicine use among malay women in Malaysia. *Patient Prefer Adherence* 2019; 13: 1829–1837.
 39. Kakai H, Maskarinec G, Shumay DM, et al. Ethnic differences in choices of health information by cancer patients using complementary and alternative medicine: an exploratory study with correspondence analysis. *Soc Sci Med* 2003; 56: 851–862. 2003/02/01.
 40. Broniatowski DA, Kerchner D, Farooq F, et al. Twitter and Facebook posts about COVID-19 are less likely to spread misinformation compared to other health topics. *PLoS One* 2022; 17: e0261768. 2022/01/13.
 41. Keselman A, Arnott Smith C, Murcko AC, et al. Evaluating the quality of health information in a changing digital ecosystem. *J Med Internet Res* 2019; 21: e11129. 2019/02/09.
 42. Brashier NM and Schacter DL. Aging in an era of fake news. *Curr Dir Psychol Sci* 2020; 29: 316–323. 2020/09/25.
 43. Choukou MA, Sanchez-Ramirez DC, Pol M, et al. COVID-19 infodemic and digital health literacy in vulnerable populations: a scoping review. *Digit Health* 2022; 8: 20552076221076927. 2022/03/01.
 44. Salim H, Shariff Ghazali S, Lee PY, et al. Health literacy levels and its determinants among people with asthma in Malaysian primary healthcare settings: a cross-sectional study. *BMC Public Health* 2021; 21: 1186.
 45. Fast I, Sorensen K, Brand H, et al. Social media for public health: an exploratory policy analysis. *Eur J Public Health* 2015; 25: 162–166. 2014/06/20.