

# Learning from social media: utilizing advanced data extraction techniques to understand barriers to breast cancer treatment

Rachel A. Freedman<sup>1</sup> · Kasisomayajula Viswanath<sup>2,3</sup> · Ines Vaz-Luis<sup>1</sup> · Nancy L. Keating<sup>4,5</sup>

Received: 6 May 2016 / Accepted: 17 June 2016 / Published online: 23 June 2016  
© Springer Science+Business Media New York 2016

**Abstract** Past examinations of breast cancer treatment barriers have typically included registry, claims-based, and smaller survey studies. We examined treatment barriers using a novel, comprehensive, social media analysis of online, candid discussions about breast cancer. Using an innovative toolset to search postings on social networks, message boards, patient communities, and topical sites, we performed a large-scale qualitative analysis. We examined the sentiments and barriers expressed about breast cancer treatments by Internet users during 1 year (2/1/14–1/31/15). We categorized posts based on thematic patterns and examined trends in discussions by race/ethnicity (white/black/Hispanic) when this information was available. We identified 1,024,041 unique posts related to breast cancer treatment. Overall, 57 % of posts expressed negative sentiments. Using machine learning software, we assigned treatment barriers for 387,238 posts (38 %). Barriers included emotional (23 % of posts), preferences and spiritual/religious beliefs (21 %), physical (18 %), resource (15 %),

healthcare perceptions (9 %), treatment processes/duration (7 %), and relationships (7 %). Black and Hispanic (vs. white) users more frequently reported barriers related to healthcare perceptions, beliefs, and pre-diagnosis/diagnosis organizational challenges and fewer emotional barriers. Using a novel analysis of diverse social media users, we observed numerous breast cancer treatment barriers that differed by race/ethnicity. Social media is a powerful tool, allowing use of real-world data for qualitative research, capitalizing on the rich discussions occurring spontaneously online. Future research should focus on how to further employ and learn from this type of social intelligence research across all medical disciplines.

**Keywords** Breast cancer · Underserved populations · Race and ethnicity · Adherence · Qualitative research

## Introduction

Racial disparities in outcomes for women with breast cancer are persistent and well documented [1, 2]. Evidence suggests that differences in treatment receipt and adherence for black and white patients significantly contribute to disparities in survival [3–8]. Numerous studies have cited complex reasons for the lower observed rates of treatment and adherence for black women, including socioeconomic status (SES), insurance, provider factors, system failures, as well as beliefs and mistrust in providers [3–7, 9–12]. To date, most studies examining barriers to care for diverse populations have been conducted within registry- or claims-based cohorts. Additional smaller studies using surveys, focus groups, and medical records are often limited to a single geographic area or institution and may not necessarily generalize across diverse populations. Furthermore, most surveys have structured formats and are subject to recall bias.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10549-016-3872-2) contains supplementary material, which is available to authorized users.

✉ Rachel A. Freedman  
rafreedman@partners.org

- <sup>1</sup> Department of Medical Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215, USA
- <sup>2</sup> Center for Community-Based Research, Dana-Farber Cancer Institute, Boston, MA, USA
- <sup>3</sup> Department of Society and Behavioral Sciences, Harvard School of Public Health, Boston, MA, USA
- <sup>4</sup> Department of Health Care Policy, Harvard Medical School, Boston, MA, USA
- <sup>5</sup> Division of General Internal Medicine, Brigham and Women's Hospital, Boston, MA, USA

Recently, social media has been recognized as a potential source of important data from patients who may be under-represented in studies using conventional research methodologies, emerging as a rich yet largely untapped resource for understanding what patients are candidly saying about their experiences and treatments [13–15]. Currently, over 85 % of Americans use the Internet regularly, with nearly half of them using at least one social networking site (e.g., Facebook, Twitter, etc.), and the number of individuals using social media is increasing rapidly [16, 17]. Approximately 50 % of Internet users are over age 35 [18], and the online community is growing in diversity [19]. In 2011, 80 % of white, 71 % of black, and 68 % of Hispanic American adults regularly used the Internet—double the rates in 2000 [17]—and recent reports suggest that minorities and lower SES groups with digital access use social media as much as other groups [19, 20]. Specifically, the use of social media with user-generated content such as Facebook, Twitter, and other outlets (e.g., chat rooms, blogs) has also increased among diverse populations, providing a tremendous opportunity to study frank and open conversations among people with similar concerns [19, 21, 22].

In this study, we utilized machine learning, a subfield of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence. Machine learning explores the study and construction of algorithms that can learn from and make predictions on data. Although social intelligence research is a relatively new scientific methodology and exploratory in nature, this type of content was utilized for a report published by the Institute of Medicine to explore the concerns of Gulf War veterans and provided powerful perspective on veterans' experiences, journeys, and concerns [15]. Here, we explored content shared by Internet users with breast cancer and performed a novel social media analysis to examine barriers to treatment. Specifically, we conducted a large-scale, qualitative content analysis of online postings about breast cancer using an advanced software platform developed by ConsumerSphere Social Intelligence. We scoured all available Web sites to gain insights into barriers to care that may be more difficult to collect using traditional qualitative and quantitative methodologies. We aimed to identify key issues and themes that patients with breast cancer were sharing online, focusing on barriers to treatment.

## Methods

### Search tools, data extraction, and data collection

ConsumerSphere uses an advanced software platform to mine and structure unstructured, qualitative data for insight and intelligence. The software 'listens' to conversations

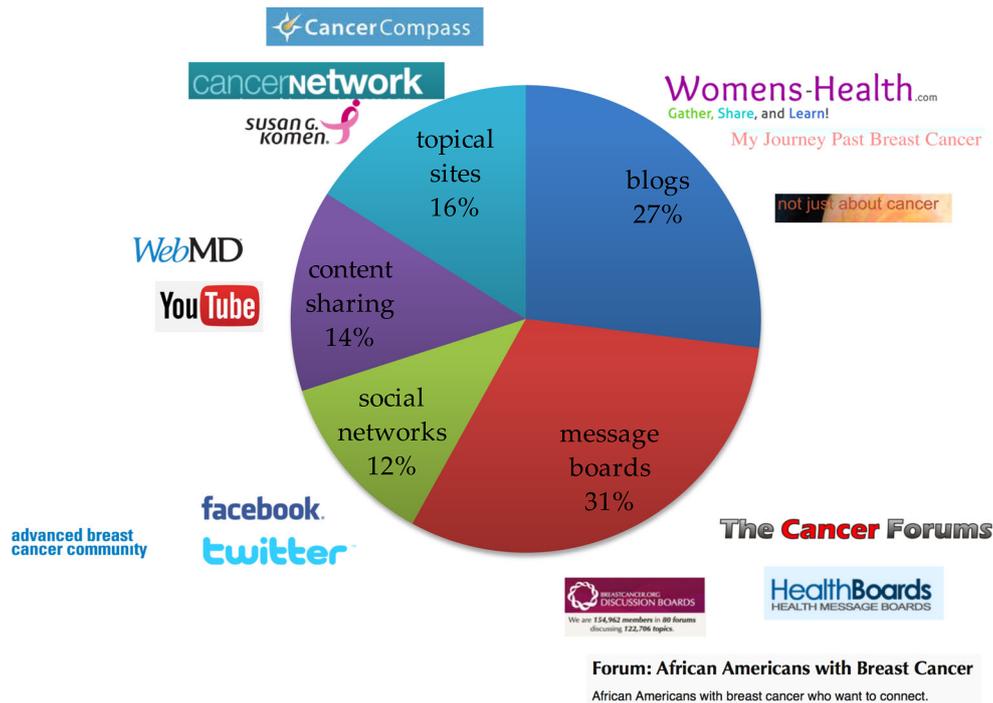
online wherever they are occurring, and examines who is talking, where users are talking, and what they are talking about. We provided ConsumerSphere a list of standard search terms relevant to breast cancer, including treatments (radiation, surgery, hormonal therapy, chemotherapy) and previously demonstrated barriers to treatment completion [3, 4, 7, 12, 23–28] (Supplemental Table). Because we wanted to capture postings about treatments for non-metastatic breast cancer, we included terms about local therapy and chemotherapy, including specific agents administered in the neo/adjuvant setting. Advanced search techniques were applied using Web spiders, crawlers, and site scraping. ConsumerSphere extracted topical data, tagged data with the origin and user, and created a large, unstructured 'big' dataset. Data collection occurred across 2021 sites, over a complete range of social discussion channels, including sites directed toward minority women (Fig. 1) and was directed by (but not limited to) our pre-defined keywords, phrases, topics, and questions.

After completion of the comprehensive data collection, natural language processing, text analytics, and social data mining were employed to examine previously described and undescribed patterns in data. These analyses were human-assisted and included repeated training, testing, and reviewing of the program output by ConsumerSphere. In this thematic analysis, we tagged and sorted data, determined key motivations of topics being discussed, and assigned an underlying treatment barrier when possible.

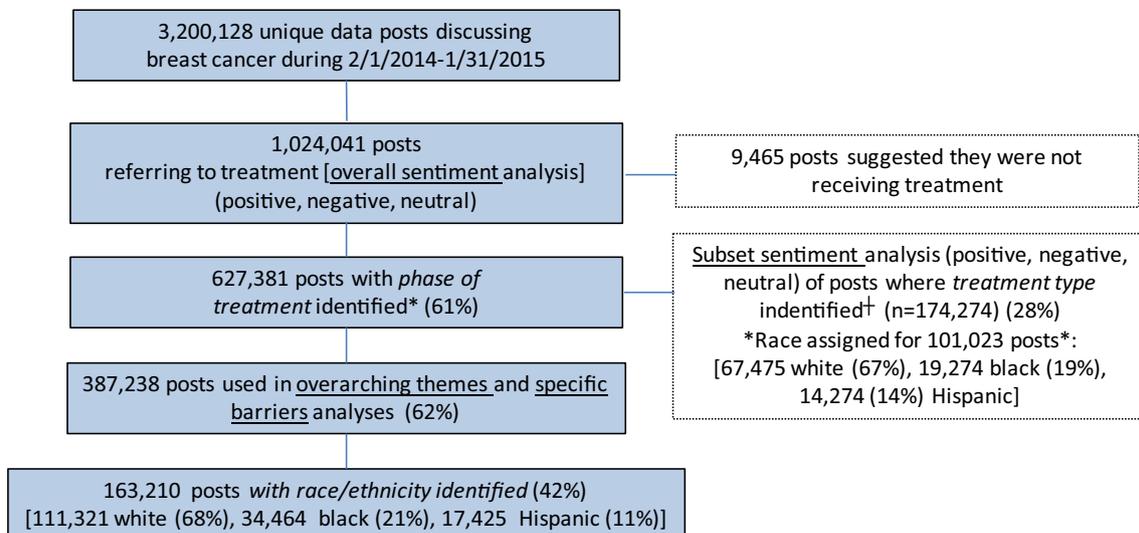
### Sites and users

We examined postings from a 365-day period, ending on January 31, 2015, on message boards, blogs, topical sites, content sharing sites, and social networks (Fig. 1). We identified 3,200,128 unique posts that discussed breast cancer, and we limited our analyses to the 1,024,041 (32 %) about treatment (Fig. 2). Internet users could have more than one posting included, but *only* if it was a part of a unique post (a single user with multiple posts within a conversation was counted once, but users posting multiple unique comments across discussions/sites were counted for each comment). A single comment appearing repeatedly through sharing/linking was counted once.

Figure 2 displays the schema of included posts. When possible (627,381/1,024,041 posts; 61 %), we identified a phase of treatment (pre-diagnosis, diagnosis, assessment, decision to treat, treatment) by tagging posts based on cues for a user's current situation through topical keywords and relevant self-reported experiences. Among the 627,381 posts, we assigned overarching themes *and* treatment barriers for 387,238 (62 % of 627,381). Because most disparities have been primarily described for black and Hispanic women, we were specifically interested in



**Fig. 1** Types and breakdown of Internet sites included in the analysis ( $n = 2021$  sites)



**Fig. 2** Schema of included posts for each analytic component\*. \*Treatment ‘phase’ = pre-diagnosis, diagnosis, assessment, decision to treat, and treatment. †Treatment ‘type’ = overall treatment, surgery, radiation, chemotherapy, hormonal therapy, ‘drugs’ in general

examination of posts for these subgroups of users when possible. Of 387,238 posts with specific themes or barriers assigned, 163,210 had an identified race/ethnicity of the user (white, black, or Hispanic) based on information in the individual’s profile, self-report in posts, or if it was apparent because a post came from a site targeted to Hispanic or black populations. When a user’s race/ethnicity could not be identified, the post was included in ‘overall’ results only.

### Content analyses

We first summarized the general attitudes/sentiments (negative/positive/neutral) for all 1,024,041 posts referring to treatment *overall* and the 174,274 posts assigned to a particular treatment (surgery, radiation, chemotherapy, hormonal therapy, ‘drugs’) (subset sentiment analysis) (Fig. 2). Second, we categorized overarching themes discussed by *treatment phase* (pre-diagnosis to treatment) when evident

( $n = 387,238$ ) to examine how barriers might differ over the treatment trajectory. We identified four themes: (a) organizational (previous experiences with healthcare, systems issues), (b) sociocultural (beliefs, family, spiritual, cultural), (c) psychological (emotional), and (d) situational (relating to job, insurance, income, access, other responsibilities). To explore barriers further, we then assigned more specific treatment barriers for the same 387,238 posts *across all treatment phases combined*, including physical, resource, and healthcare perceptions, using the keywords/themes in Supplemental Table. The overarching theme and specific barrier analyses were not mutually exclusive and were overlapping. Finally, we separately examined the subset of 9465 posts (of 1,024,041) that suggested users refused treatment, in case barriers for this group were distinct from those among women undergoing treatment. For all analyses, we examined findings overall and by race/ethnicity when possible (Fig. 2).

Analyses are descriptive in nature and exploratory and no formal statistical tests could be performed. Because we had no identifiable participant information, the study was considered exempt by the Office for Human Research Studies at Dana-Farber Cancer Institute.

## Results

### Overall and subset sentiment analyses

Among the 1,024,041 posts referring to breast cancer treatment in general, 54 % asked questions ('*Anyone else have path results that were unexpected?*'), 33 % shared content ('*Like you, I am on pins and needles...I am prepared for a cancer diagnosis.*'), and 13 % answered questions ('*Everyone is different, but my port was painful for just a day or two...*'). Overall, 57 % of posts skewed negatively, 31 % were neutral, and 12 % were positive. In a subset sentiment analysis of the 627,381 posts with phase of treatment (pre-diagnosis, diagnosis, etc.) identified, 172,274 posts described a specific treatment (surgery, radiation, etc.). Posts discussed chemotherapy (35 %), surgery (33 %), radiation (15 %), 'drugs' (10 %), and hormonal therapy (7 %) (Fig. 3). Approximately 50 % of all treatment-specific posts skewed negatively; these were most common among posts about surgery or 'drugs'. When race/ethnicity was identified ( $n = 101,023$  posts), black users more frequently posted negative comments (66 % of posts) than Hispanic and white users (55 % of posts for each) (Fig. 3).

### Overarching themes by treatment phase

The four overarching themes expressed (organizational, sociocultural, psychological, situational as defined above)

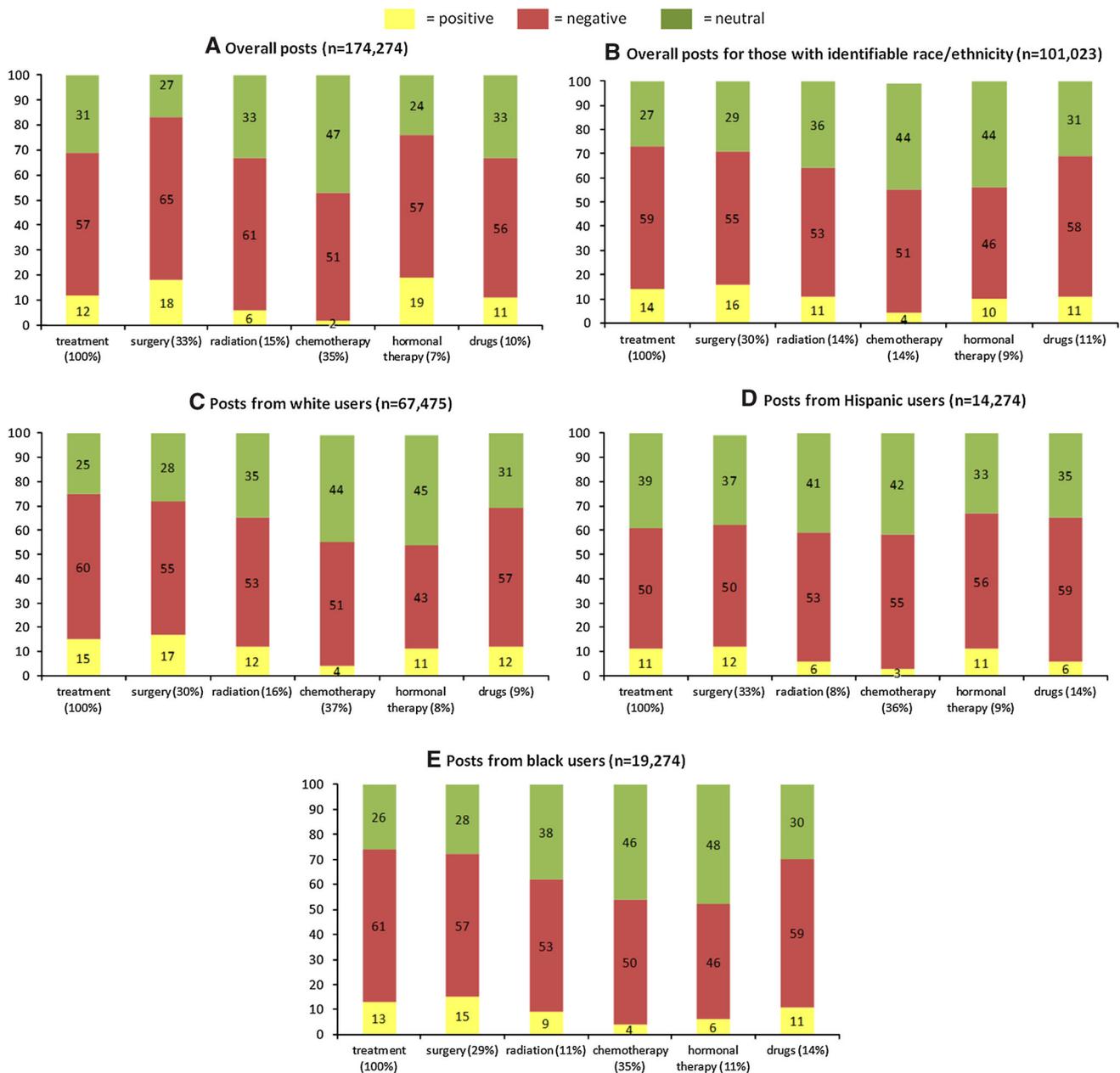
by treatment phase (pre-diagnosis to treatment) for the 387,238 posts are shown in Fig. 4. Organizational barriers generally increased from pre-diagnosis (6 % of posts) to diagnosis (13 %) and remained high during assessment (28 %), decisions to treat (21 %), and treatment (29 %). Sociocultural barriers decreased over the treatment trajectory (24 % of posts in the pre-diagnosis phase to 18–20 % of posts about treatments) as did psychological barriers (43 % to 19–25 %). Situational barriers remained relatively constant over the treatment trajectory and were reported in a quarter of posts. Among patients with race/ethnicity identified, psychological, situational, and organizational barriers were most frequent for white users, sociocultural and situational barriers were most frequent for Hispanic users, and situational and organizational barriers were most frequent for black users. Notably, Hispanic and black users experienced more organizational barriers than white users around the time of diagnosis.

### Specific barriers to treatment across all treatment phases

A more detailed examination of specific barriers (in contrast to the overarching themes above) among the 387,238 unique posts is shown in Fig. 5. Barriers expressed included emotional (including anxiety, fears, denial, depression; 23 % of posts), personal beliefs (including misinformation, healthcare preferences, spiritual/religious/cultural; 21 % of posts), physical concerns (including limitations, body changes, side effects; 18 % of posts), resource barriers (including costs, logistics, insurance; 15 % of posts), healthcare perceptions (including trust, communication, negative experiences, accessibility of services; 9 % of posts), issues with treatment processes/duration (including complexity, regimen; 7 % of posts), and relationships (including children, friends, intimacy; 7 % of posts).

Numerical differences were noted by race/ethnicity for the 163,210 posts from users with identifiable race/ethnicity, with posts by black and Hispanic users reporting more barriers related to beliefs (24–25 % vs. 21 % of posts by whites) and fewer posts related to emotional (14–17 % vs. 31 %) and relationship (6–9 % vs. 14 %) barriers. Hispanic and black users also reported more barriers related to healthcare perceptions (13 % posts vs. 5 % by white users). Resource barriers were reported with similar frequency for all groups (22–24 % of posts from white/black/Hispanic users), and overall, 49, 46, and 43 % of all posts from black, Hispanic, and white users were related to either resources or beliefs. Table 1 shows representative quotes from each specific barrier.

For *emotional* barriers, most conversations reported fears, anxiety, denial, and depression. Fear was the most common emotional sentiment expressed (35 % of posts)



**Fig. 3** Percent of general sentiments (positive/negative/neutral) by treatment type [overall (3A,  $n = 174,274$  posts) and by race/ethnicity among the users for whom treatment type could be identified\* (3B–3E [ $n = 101,023$  posts]). \*The percent under each treatment bar

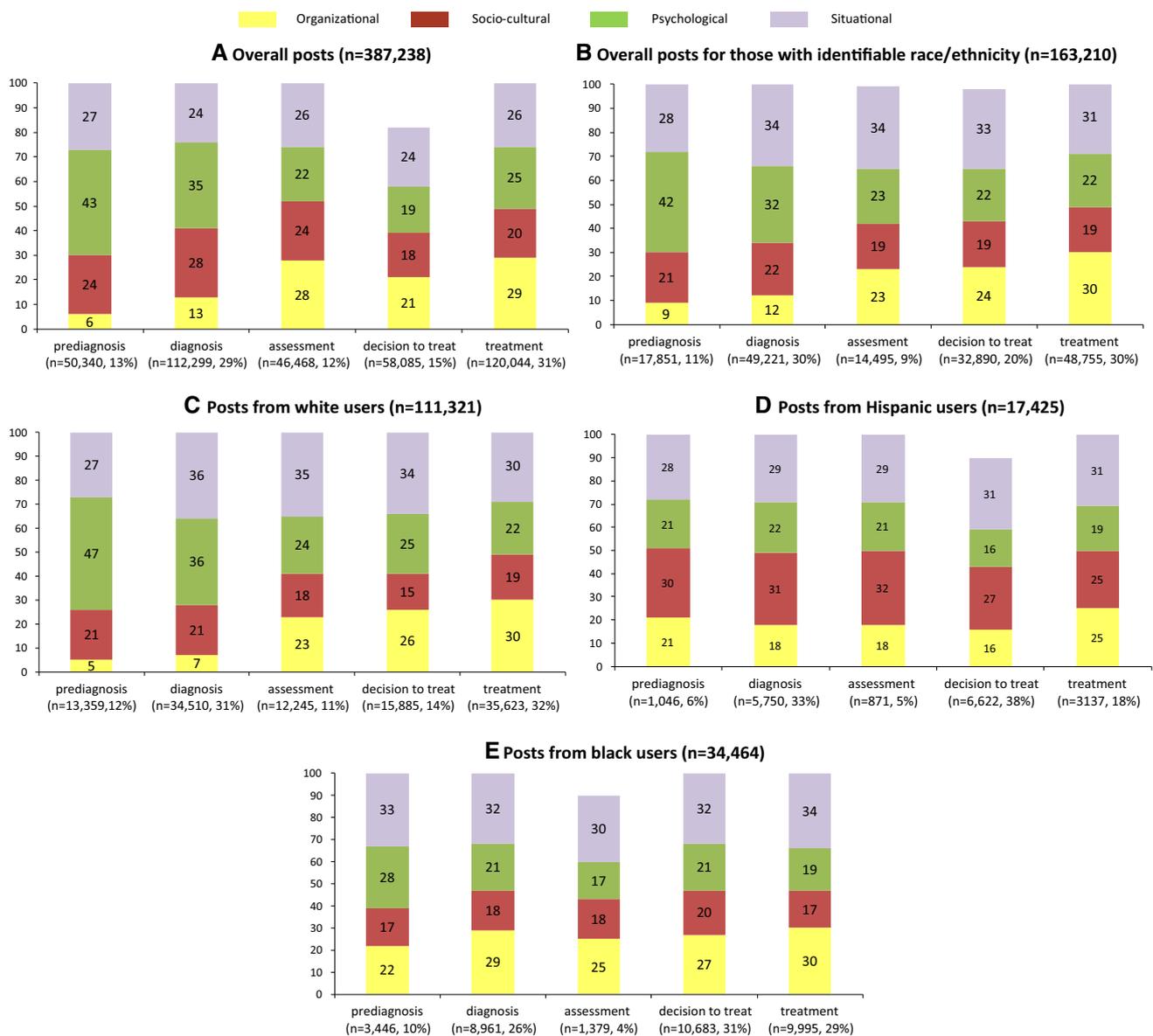
and this was the most common emotion expressed by Hispanic users (37 vs. 27 % of black user posts and 33 % of posts by whites). Denial was the most prominent emotion described by black users (32 % of posts), compared with 26 % Hispanic users' posts and 10 % of white users' posts. Posts related to anxiety were most common from white users (31 % posts vs. 20 and 25 % of posts by Hispanic and black patients, respectively).

With regard to *beliefs*, the most common sentiments were spiritual/religious (41 %), although other prominent

represents the percent of posts about treatment that were related to each component of therapy. The *yellow*, *red*, and *green bars* represent the proportion of sentiments that were positive, negative, and neutral, respectively

themes included misinformation (30 %) and preferences/perceptions (29 %). Hispanic and white users were more likely than blacks to report spiritual or religious sentiments about treatment (41 and 38 % vs. 31 % of posts from black users) and black users were more likely to express issues with perceptions/preferences (43 % of posts by black users vs. 36 and 33 % posts from Hispanic and white users, respectively).

The most common *physical* concerns expressed were side effects (40 %), followed by physical limitations



**Fig. 4** Profile of barriers<sup>§</sup> by treatment phase and race/ethnicity for 387,238 posts (a) and for those whose race/ethnicity could be identified ( $n = 163,210$ , 4B–4E)\*. <sup>§</sup>Organizational = related to experiences with healthcare and systems issues, sociocultural = related to family/spiritual/cultural influences, psychological = emotional concerns, situational = related to income/insurance/

(31 %) and body changes (29 %). Side effect comments were less frequent in posts by black (30 %) and Hispanic users (29 %) compared with 43 % posts by white users. Body image was most apparent in posts by Hispanics (36 %), while physical limitations were most frequently reported in posts by black users (38 vs. 33 % and 35 % from white and Hispanic users).

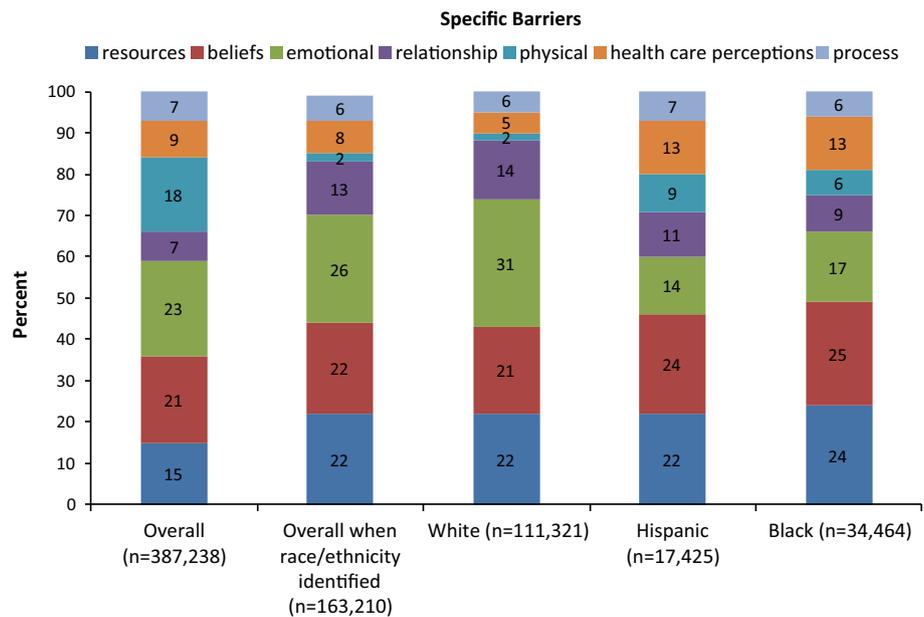
*Resource* concerns included posts about insurance (49 %), costs (33 %), and logistics of treatment (18 %). Black and Hispanic user posts more often included cost barriers (45 % of posts from black and 41 % of posts from

accessibility/employment/responsibilities. \*The percent under *each bar* represents the percent of posts about that phase. The *yellow bars* represent organizational, the *red bars* represent sociocultural, the *green bars* represent psychological, and the *purple bars* represent situational

Hispanic users vs. 32 % of posts from whites) and logistics barriers (28 and 25 % vs. 16 % of posts from white users).

Dominant concerns raised within posts about *health-care perception* barriers included poor communication (36 %), trust (22 %), accessibility of services (21 %), and negative experiences (21 %). Concerns about trust were more commonly mentioned by black and Hispanic users (31 and 29 % of posts) vs. white users (26 %); accessibility concerns were raised in 29 % of posts by black and Hispanic users compared with 20 % of posts by white users.

**Fig. 5** Specific barriers to treatment (%) ( $n = 387,238$ ) and by race/ethnicity when identified ( $n = 163,210$ )



Among posts related to *relationship* barriers, the most dominant issues included problems with intimacy (35 %), friends (34 %), and children (31 %), with intimacy issues being most commonly reported in posts by black and Hispanic users (39 and 37 % vs. 29 % of posts by whites), while issues with friends and children were more frequently reported in posts by whites (71 % of posts) than minority users (61–65 % of posts).

*Duration and process* barriers were categorized as issues with the regimens prescribed (41 %), duration of treatment (23 %), aftereffects of treatment (19 %), and complexity of care (17 %). Complexity of care was more prominently reported in posts by black and Hispanic users (29 and 23 % vs. 19 % of posts by white users).

### Users refusing treatment

In 9465 posts, users suggested that they refused recommended treatments for their breast cancer. Dominant themes in these conversations included fear of side effects (*...enduring treatment would be worse than death for me.*), denial (*I decided this is MY life and I WILL have it MY WAY!*), holistic beliefs (*I'm a 54 year old who refuses treatment because the protocols of treatment are barbaric to me. I've studied a nutritionally holistic approach and I firmly believe I must refuse the traditional approach.*), preferences (*I will not put myself or my family through the seemingly endlessness of treatment, the visits to hospitals, and the entire set of difficult logistics just to prolong my life.*), fatalism (*I think chemo is the biggest scam in the world and millions of people are being told to take it when in fact it will do nothing.*), and faith (*All I need is prayers*

*and strength from my family, friends, and my church family*).

### Discussion

Using a novel, comprehensive analysis of over 1 million posts about breast cancer treatments from a diverse population of social media users, we observed frequently discussed barriers over the course of treatment, with discussions often skewing negatively. The barriers expressed by users differed somewhat for black and Hispanic users compared with white users and were more often related to preferences, perceptions and cultural/religious/spiritual beliefs, costs of therapy, and logistical barriers.

Racial disparities in breast cancer are well documented and the reasons for treatment differences and outcomes are complex [3–5, 7, 23, 29–38]. Not surprisingly, our findings have suggested similar themes in barriers to care as raised by others in the literature, including access [7, 10–12], system failures [23], mistrust of providers [4], and psychological issues [39–41]. However, some of our findings should be highlighted. Misperceptions, healthcare preferences, and spiritual/cultural/religious beliefs comprised nearly one quarter of the barriers to treatment reported in our study, with physical barriers such as side effects comprising <10 % of posts by black and Hispanic women and only 18 % overall, although we were unable to fully assess whether treatments were not completed or not initiated at all. Further, modifiable factors such as resource barriers were frequently reported by users, and among

**Table 1** Representative quotes from each barrier type identified\*

Specific barrier	Representative quotes
Emotional ( $n = 89,064$ )	
Anxiety (29 %)	I get my results tomorrow too. I'm getting more anxious by the minute
Fears (35 %)	I am so afraid that I could cry
Denial (13 %)	I have no time for potential troubles, having enough actual ones
Depression (23 %)	I've never been this depressed in my life. I really don't know where to begin
Beliefs ( $n = 81,319$ )	
Misperceptions (29 %)	Breast cancer does not run in my family so I am not at risk
Healthcare preferences (30 %)	We don't go to the doctor until we're very sick
Spiritual/religious/cultural (41 %)	Whatever our religious beliefs, if you believe in a great Creator, this is the part of the plan and my plan for health Breast cancer is a white woman's disease
Physical ( $n = 69,702$ )	
Side effects (40 %)	The digestive effects of this chemo are horrid
Limitations (31 %)	Breast cancer treatment just destroys the body? It feels like it will is simply not possible to ever regain any fitness once you've lost it?
Body change (29 %)	My hair is becoming so thin and going away I won't feel like a whole woman anymore
Resources ( $n = 58,085$ )	
Insurance (49 %)	I have no insurance and have just been diagnosed. What are my options?
Costs (33 %)	I'm not sure I can afford to have breast cancer. With the hormones and all the meds, I'm gonna go broke
Logistics (18 %)	I only have access to public transportation and it's so hard to coordinate all my appointments I don't always have someone to watch their kids when they go in for an appointment
Healthcare perceptions ( $n = 34,851$ )	
Poor communication (36 %)	My oncologist doesn't tell me anything. I get most of the important info and instruction from his RN
Negative experiences (21 %)	They were rude the whole time, made me feel ignorant and I am not sure I want to proceed with it!! I had my doctor switched to a woman doctor so I wouldn't be as embarrassed or feeling violated
Trust (22 %)	I don't 'trust' doctors anymore, but am confident my current care team is competent. My last one wasn't. If you have doubts go elsewhere, preferably a cancer center
Accessibility of services (21 %)	The doctor's office is so far away and inconvenient, it's almost impossible for me to get there
Relationships ( $n = 27,106$ )	
Intimacy (35 %)	...it saddens me that there are men out there who cannot support their wives through this terrible ordeal
Friends (34 %)	A life changing diagnosis of breast cancer has shown me who my friends are and who I can count on in my life
Children (31 %)	(This has) caused feelings of worry and guilt about my role as a mother
Duration/process ( $n = 26,987$ )	
Regime (41 %)	I had 33 rads, 28 reg and 5 boosts. I took a 10 break for vacation about mid-way through but still was tired
Duration of treatment (23 %)	I'm not sure if I can handle this anxiety for results for five more years
Aftereffects (19 %)	I just can't take the time off...I'm afraid I'm going to lose my job. It's my life and I need it
Complexity (17 %)	My cancer treatment didn't come easily...it involved of a complex series of treatments and was a very difficult process

\*  $n$ 's represent the numbers of unique posts expressing these sentiments

users reporting not receiving treatment for their breast cancer, preferences/perceptions and religious/cultural/spiritual beliefs, worry about side effects, denial, and fatalism were the most commonly reported barriers. Further,

organizational barriers were more frequently discussed by minority (vs. white) users during pre-diagnosis and diagnosis, suggesting obstacles with system factors. All of these findings suggest that tolerability of treatments is not a

predominant issue limiting treatment receipt and that addressing beliefs and logistical barriers has potential to impact receipt of care. These mutable factors should be surmountable with the right patient education, support, and services for patients (e.g., navigators, coordinators, patient assistance funds).

Using this type of ‘social intelligence’ for research is a new, iterative research discipline that mines the vast repository of unstructured big data for insight into patients’ concerns and experiences. It does not rely on pre-defined content or rules-based programming and is instead driven by pattern recognition and adaptability to thematic content. In contrast, conventional research methods are more structured and work to formalize relationships between variables, providing robust tests for statistical significance. Further, traditional research methods typically rely on model assumptions and have the risk of suggesting the wrong study conclusions if the underlying assumptions are wrong.

Finding ways to optimally capitalize on the immense power of online candid patient interactions and conversations provides a new method of conducting qualitative and eventually quantitative research across many medical disciplines. Here, we captured spontaneous, real-time conversations in a non-intrusive way and in a natural setting. Further, we included over 50,000 posts from users who self-identified as black or Hispanic, representing a substantial proportion of the posts analyzed in our analysis, and demonstrating the ability to reach diverse users with this type of research. Social media represents an essentially untapped resource of big data with vast potential, particularly in patients who are difficult to reach using traditional methods. In particular, this type of research may prove optimal when examining the late and longer-term impact of our treatments among cancer survivors, when active clinical follow-up becomes more limited.

Despite the novel and exciting nature of this research, we recognize several challenges with this type of analysis. First, we lacked demographic information on many users, and although race/ethnicity was self-reported, it is possible that we misclassified some users. Second, although we focused search terms on treatments administered for curative intent, we could not distinguish metastatic from non-metastatic users. However, understanding barriers to treatment regardless of cancer stage is important. Third, the views expressed online may be skewed negatively because struggling patients may be more likely to engage in a community for support, although these may be the patients who are most important to reach with this type of study. In addition, the feelings expressed by a social media user may be ‘different’ from those expressed in other situations (e.g., at an office visit with a provider), but may not necessarily be more representative of the truth for that patient. Fourth,

it is possible that some users posted within multiple different conversation threads, and although multiple posts within a single discussion/conversation by a user were only included once, some users may have posted on various discussions or sites. Fifth, we had no information on specific treatments recommended or treatment adherence.

In conclusion, we harnessed real-world data using this novel modality for qualitative research, capitalizing on the rich conversations occurring online for patients with breast cancer. We learned about barriers to care for a large and diverse population of users and will use these data to inform an upcoming survey to further explore the issues identified. We observed a smaller than expected proportion of users reporting physical barriers to treatment as a limiting factor, while modifiable factors such as resource, organizational, trust, and beliefs about treatment were more predominant. Future research should further focus on how to further employ and learn from this type of social intelligence research.

**Acknowledgments** We thank Patrick Furey at ConsumerSphere for his thoughtful assistance with this analysis.

**Funders** The funders for this study are American Cancer Society (RAF), Susan G. Komen (RAF), K24CA181510 (NLK) from the National Cancer Institute, Fundacao para a Ciencia e Tecnologia [HMSP-ICS/0004/2011, Career Development Award, (IVL)].

#### Compliance with ethical standards

**Conflicts of interest** Rachel A Freedman has no relevant conflicts; research funding (to institution only) was received from Eisai Co., Ltd, and Genentech, Puma Biotechnology. Kasisomayajula Viswanath, Ines Vaz-Luis, and Nancy L. Keating have no conflicts to disclose.

#### References

- Hunt BR, Whitman S, Hurlbert MS (2014) Increasing Black: White disparities in breast cancer mortality in the 50 largest cities in the United States. *Cancer Epidemiol* 38(2):118–123
- Zeng CZ, Wen W, Morgans AK, Pao W, Shu XO, Zheng W (2015) Disparities by Race, age, and sex in the improvement of survival for major cancers: results From the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) Program in the United States, 1990 to 2010. *JAMA Oncol* 1(1):88–96
- Bickell NA, Wang JJ, Oluwole S, Schrag D, Godfrey H, Hiotis K, Mendez J, Guth AA (2006) Missed opportunities: racial disparities in adjuvant breast cancer treatment. *J Clin Oncol* 24(9):1357–1362
- Bickell NA, Weidmann J, Fei K, Lin JJ, Leventhal H (2009) Underuse of breast cancer adjuvant treatment: patient knowledge, beliefs, and medical mistrust. *J Clin Oncol* 27(31):5160–5167
- Hershman D, McBride R, Jacobson JS, Lamerato L, Roberts K, Grann VR, Neugut AI (2005) Racial disparities in treatment and survival among women with early-stage breast cancer. *J Clin Oncol* 23(27):6639–6646

6. Hershman DL, Shao T, Kushi LH, Buono D, Tsai WY, Fehrenbacher L, Kwan M, Gomez SL, Neugut AI (2011) Early discontinuation and non-adherence to adjuvant hormonal therapy are associated with increased mortality in women with breast cancer. *Breast Cancer Res Treat* 126(2):529–537
7. Freedman RA, Virgo KS, He Y, Pavluck AL, Winer EP, Ward EM, Keating NL (2011) The association of race/ethnicity, insurance status, and socioeconomic factors with breast cancer care. *Cancer* 117(1):180–189
8. Freedman RA, He Y, Winer EP, Keating NL (2013) Racial/Ethnic differences in receipt of timely adjuvant therapy for older women with breast cancer: are delays influenced by the hospitals where patients obtain surgical care? *Health Serv Res* 48(5):1669–1683
9. Keating NL, Kouri E, He Y, Weeks JC, Winer EP (2009) Racial differences in definitive breast cancer therapy in older women: are they explained by the hospitals where patients undergo surgery? *Med Care* 47(7):765–773
10. Roetzheim RG, Gonzalez EC, Ferrante JM, Pal N, Van Durme DJ, Krischer JP (2000) Effects of health insurance and race on breast carcinoma treatments and outcomes. *Cancer* 89(11):2202–2213
11. Ward E, Halpern M, Schrag N, Cokkinides V, DeSantis C, Bandi P, Siegel R, Stewart A, Jemal A (2008) Association of insurance with cancer care utilization and outcomes. *CA Cancer J Clin* 58(1):9–31
12. Wu XC, Lund MJ, Kimmick GG, Richardson LC, Sabatino SA, Chen VW, Fleming ST, Morris CR, Huang B, Trentham-Dietz A et al (2012) Influence of race, insurance, socioeconomic status, and hospital type on receipt of guideline-concordant adjuvant systemic therapy for locoregional breast cancers. *J Clin Oncol* 30(2):142–150
13. Casanas ICC, Wanat M (2015) Using social media in supportive and palliative care research. *BMJ Support Palliat Care* 5(2):138–145
14. Nwosu AC, Debattista M, Rooney C, Mason S (2015) Social media and palliative medicine: a retrospective 2-year analysis of global Twitter data to evaluate the use of technology to communicate about issues at the end of life. *BMJ Support Palliat Care* 5(2):207–212
15. Institute of Medicine of the National Academies. Report Brief: Gulf War and Health: Treatment for Chronic Multisymptom Illness. <http://www.iom.edu/Reports/2013/Gulf-War-and-Health-Treatment-for-Chronic-Multisymptom-Illness.aspx>. Accessed 9 Jan 2013
16. PewResearchCenter: Internet, Science, & Tech. <http://www.pewinternet.org/2015/01/09/social-media-update-2014/>. Accessed 4 June 2015
17. PewResearchCenter: Internet, Science % Tech. Digital differences. <http://www.pewinternet.org/2012/04/13/digital-differences/>. Accessed 4 June 2014
18. PewReserachCenter: Internet, Science, & Tech. Who are social networking site users? <http://www.pewinternet.org/2011/06/16/part-2-who-are-social-networking-site-users/>. Accessed 4 June 2015
19. Viswanath K, McCloud R, Minsky S, Puleo E, Kontos E, Bigman-Galimore C, Rudd R, Emmons KM (2013) Internet use, browsing, and the urban poor: implications for cancer control. *J Natl Cancer Inst Monogr* 47:199–205
20. Kontos EZ, Emmons KM, Puleo E, Viswanath K (2010) Communication inequalities and public health implications of adult social networking site use in the United States. *J Health Commun* 15(Suppl 3):216–235
21. Pew. Global digital communication: texting, social networking popular worldwide 2011. <http://www.pewglobal.org/2011/12/20/global-digital-communication-texting-social-net-working-popular-worldwide/>. Accessed 12 March 2016
22. Pew Research Center. African Americans and technology use: a demographic portrait. Posted January 4, 2014. <http://www.pewinternet.org/files/2014/01/African-Americans-and-Technology-Use.pdf>. Accessed 11 March 2016
23. Bickell NA, LePar F, Wang JJ, Leventhal H (2007) Lost opportunities: physicians' reasons and disparities in breast cancer treatment. *J Clin Oncol* 25(18):2516–2521
24. Freedman RA, He Y, Winer EP, Keating NL (2013) Racial/Ethnic differences in receipt of timely adjuvant therapy for older women with breast cancer: are delays influenced by the hospitals where patients obtain surgical care? *Health Serv Res* 48:1669–1683
25. Hack TF, Degner LF, Parker PA, Team SC (2005) The communication goals and needs of cancer patients: a review. *Psychooncology* 14(10):831–845 (**discussion 846–837**)
26. Sheppard VB, Adams IF, Lamdan R, Taylor KL (2010) The role of patient-provider communication for black women making decisions about breast cancer treatment. *Psychooncology* 20:1309–1316
27. Chen JY, Diamant AL, Thind A, Maly RC (2008) Determinants of breast cancer knowledge among newly diagnosed, low-income, medically underserved women with breast cancer. *Cancer* 112(5):1153–1161
28. Sedjo RL, Devine S (2011) Predictors of non-adherence to aromatase inhibitors among commercially insured women with breast cancer. *Breast Cancer Res Treat* 125(1):191–200
29. Freedman RA, He Y, Winer EP, Keating NL (2009) Trends in racial and age disparities in definitive local therapy of early-stage breast cancer. *J Clin Oncol* 27(5):713–719
30. Freedman RA, Hughes ME, Ottesen RA, Weeks JC, He Y, Wong YN, Theriault R, Keating NL (2013) Use of adjuvant trastuzumab in women with human epidermal growth factor receptor 2 (HER2)-positive breast cancer by race/ethnicity and education within the National Comprehensive Cancer Network. *Cancer* 119(4):839–846
31. Hershman D, Weinberg M, Rosner Z, Alexis K, Tiersten A, Grann VR, Troxel A, Neugut AI (2003) Ethnic neutropenia and treatment delay in African American women undergoing chemotherapy for early-stage breast cancer. *J Natl Cancer Inst* 95(20):1545–1548
32. Gross CP, Filardo G, Mayne ST, Krumholz HM (2005) The impact of socioeconomic status and race on trial participation for older women with breast cancer. *Cancer* 103(3):483–491
33. Gross CP, Smith BD, Wolf E, Andersen M (2008) Racial disparities in cancer therapy: did the gap narrow between 1992 and 2002? *Cancer* 112(4):900–908
34. Lund MJ, Brawley OP, Ward KC, Young JL, Gabram SS, Eley JW (2007) Parity and disparity in first course treatment of invasive breast cancer. *Breast Cancer Res Treat* 109:545–557
35. Griggs JJ, Culakova E, Sorbero ME, Poniewierski MS, Wolff DA, Crawford J, Dale DC, Lyman GH (2007) Social and racial differences in selection of breast cancer adjuvant chemotherapy regimens. *J Clin Oncol* 25(18):2522–2527
36. Griggs JJ, Sorbero ME, Stark AT, Heininger SE, Dick AW (2003) Racial disparity in the dose and dose intensity of breast cancer adjuvant chemotherapy. *Breast Cancer Res Treat* 81(1):21–31
37. Silber JH, Rosenbaum PR, Fox KR (2013) Differences in breast cancer survival by race—reply. *JAMA* 310(22):2457
38. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, Karaca G, Troester MA, Tse CK, Edmiston S et al (2006) Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* 295(21):2492–2502

39. Deshields T, Tibbs T, Fan MY, Taylor M (2006) Differences in patterns of depression after treatment for breast cancer. *Psychooncology* 15(5):398–406
40. Mausbach BT, Schwab RB, Irwin SA (2015) Depression as a predictor of adherence to adjuvant endocrine therapy (AET) in women with breast cancer: a systematic review and meta-analysis. *Breast Cancer Res Treat* 152(2):239–246
41. Ell K, Sanchez K, Vourlekis B, Lee PJ, Dwight-Johnson M, Lagomasino I, Muderspach L, Russell C (2005) Depression, correlates of depression, and receipt of depression care among low-income women with breast or gynecologic cancer. *J Clin Oncol* 23(13):3052–3060