



Understanding Breast Cancer Survivors' Information-Seeking Behaviors and Overall Experiences: A Comparison of Themes Derived from Social Media Posts and Focus Groups

Journal:	<i>European Health Psychology Society</i>
Manuscript ID	GPSH-2019-0524.R2
Manuscript Type:	Psychology and Health
Keywords:	Breast Cancer, social support, focus groups, Information-seeking

SCHOLARONE™
Manuscripts

1
2
3 ***Authors:***
4

5 *Corresponding Author

6 Jennifer Currin-McCulloch, PhD, LSW
7 Colorado State University School of Social Work
8 jen.currin-mcculloch@colostate.edu
9

10
11 Amelia Stanton, PhD
12 Harvard Medical School Department of Psychiatry
13

14 Ryan Boyd, PhD
15 Lancaster University Psychology Department
16

17
18 Margaret Neaves, LMSW
19 Satellite Healthcare Department of Social Work
20

21 Barbara Jones, PhD, MSW, FNAP
22 University of Texas at Austin Steve Hicks School of Social Work
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Title:** Understanding Breast Cancer Survivors' Online Information-Seeking Behaviors and Overall
4 Experiences: A Comparison of Themes Derived from Social Media Posts and Focus Groups
5

6
7
8 **Abstract**

9
10 **Objective:** Using two different analysis techniques, this study explored differences and similarities
11 in information-seeking discourse and overall breast cancer experiences between posters to a Reddit
12 board and breast cancer survivor focus groups.
13
14

15
16
17 **Design:** This study incorporates two qualitative methods for determining themes in breast cancer
18 survivors' information-seeking behaviors and overall cancer experiences. First, posts from a breast
19 cancer-specific Reddit community were extracted and analyzed using the meaning extraction
20 method (MEM) to determine core themes. Then, investigators performed a thematic analysis of
21 two focus groups of breast cancer survivors (N = 18). Finally, themes derived from each analysis
22 method were compared.
23
24
25
26
27
28
29

30
31 **Main Outcome Measures:** Outcome measures include themes extracted from Reddit posts and
32 themes generated from breast cancer survivor focus groups.
33
34

35
36 **Results:** Findings between qualitative methodologies represent similar yet nuanced themes in
37 survivors' discourse. The MEM resulted in seven themes: diagnosis, treatment process, social
38 support, existentialism, risk, information-seeking, and surgery. Focus groups revealed the same
39 initial four MEM themes plus the following: disclosure, coping, and fears.
40
41
42
43

44
45 **Conclusions:** The MEM is a cost-effective research mechanism for informing common themes of
46 experiences of cancer patients and survivors and may offer initial data to guide psychosocial
47 oncology research design and recruitment.
48
49
50

51
52 **Key Words:** breast cancer; meaning extraction methods; social support; focus groups; information-
53 seeking
54
55
56
57
58
59
60

1
2
3 **Word Count:** 4982
4
5
6
7

8 **Introduction**

9

10 When diagnosed with breast cancer, women receive an onslaught of information about
11 medical treatments, potential changes to their body, and shifts in their familial, social, and
12 occupational roles (Boyages et al., 2016; Brandão, Schulz, & Matos, 2017; Male, Fergus, &
13 Cullen, 2016). They encounter numerous healthcare providers and well-meaning loved ones who
14 share varying opinions on what types of treatments they should or should not receive (Berger et
15 al., 2017; Laidsaar-Powell et al., 2017). Medical advances increase burdens of decision-making: a
16 multitude of decisions involving surgery, reconstruction, radiation, or promising clinical research
17 must be made. If a woman has a family history of cancer, genetic counseling and testing may also
18 enter into her decision-making algorithm (Wevers et al., 2017). Women often weigh benefits
19 between future health and cosmetic results with impacts of family and careers (Fasse et al., 2017;
20 Swanberg, Nichols, Ko, Tracy, & Vanderpool, 2017; Wallner et al., 2017). Sorting through the
21 possibilities and the potential future impacts of their treatment-related decisions often produces
22 tremendous stress and anxiety (Drageset, Lindstrom, & Underlid, 2009). The complex decision-
23 making process can become overwhelming and increase a woman's proclivity to negative coping
24 (Drageset et al., 2009; Williams & Jeanetta, 2015).
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

44 Breast cancer survivors frequently turn to online forums for decision-making and social
45 support. The number of online breast cancer forums has increased exponentially since 2006, with
46 significantly more posts on breast cancer forums appearing each month (Quinn et al., 2013). Online
47 forums and support groups provide emotional benefits to cancer survivors, including information
48 sharing, receiving support, and decreased isolation through social interaction (Kim et al., 2012).
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Groups also promote hope and nurture a sense of altruism (Shaw, McTavish, Hawkins, Gustafson,
4 & Pingree, 2000), and they may offer patients some control over their decisions and the ability to
5 transition from a passive to active role in their care (Quinn et al., 2013). Though the abundance of
6 information about breast cancer in these forums may help women choose among viable treatment
7 options and formulate appropriate questions for their providers, easy access to facts, figures, and
8 personal narratives may also lead to information burden.
9

10
11 In recent years, social scientists have used natural language data, collected through
12 traditional experimental studies as well as from online support groups and social media sites, to
13 study complex psychological phenomena, such as personality (Yarkoni, 2010), core values (Boyd
14 et al., 2015), and sexual self-schemas (Stanton, Boyd, Pulverman, & Meston, 2015). Natural
15 language data offers a number of unique advantages to researchers (see Boyd, 2017 for a thorough
16 review of these advantages). The volume of text that is publicly available on these sites is
17 unprecedented (Tausczik & Pennebaker, 2010), and researchers can easily extract the user-
18 generated data at little to no cost. Other benefits of language data include reliability over time,
19 internal consistency, inherent ecological validity, low cost, and considerable variety between
20 people (Boyd & Pennebaker, 2017). Though there have been concerns about self-selection biases
21 and data quality in online settings (Peek, Holmes, Sun, & 2014; Strasser et al., 2012), research
22 suggests that social media sites can be used effectively as data sources, specifically for data on
23 health-related outcomes (Alshaikh, Ramzan, Rawaf, Majeed, 2014). Text extracted from social
24 media sites may be used as a precursor or even an alternative to laboratory measurement of certain
25 complex phenomena that are difficult to measure (Stanton, Meston, & Boyd, 2017).
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 One family of methods commonly used to analyze natural language data is broadly known
52 as topic modeling. The Meaning Extraction Method (MEM; Chung & Pennebaker, 2008), a
53
54
55
56
57
58
59

1
2
3 specific type of topic modeling, identifies clusters of words that tend to co-occur across a corpus
4
5 of text in an automated fashion. These word clusters represent latent components or “themes” that
6
7 are present within the corpus. Using the MEM, themes in a given body of text can be quantified
8
9 and analyzed over time and across groups (e.g., by demographic characteristics, between clinical
10
11 populations and non-clinical populations, comparing individuals randomized to a treatment versus
12
13 control group, etc.). One strength of the MEM rests in its ability to combine the wide sampling
14
15 power that is typically associated with quantitative data with the detail and nuance that is
16
17 characteristic of qualitative data (Stanton et al., 2015).
18
19

20
21
22 Though content analyses of online interactions related to breast cancer have been
23
24 conducted, the MEM has not yet been used to isolate themes of women’s experiences with breast
25
26 cancer. In a previous content analysis that used qualitative methods to assess breast cancer
27
28 survivors’ coping mechanisms and physical adaptations, topics in online posts revolved around
29
30 information seeking, with a focus on diagnosis and pathology as well as on suggestions for
31
32 symptom management and healthy living (Rubenstein, 2015). Other posts were intended to
33
34 provide encouragement and empathy (Rubenstein, 2015). Although the Rubenstein paper provides
35
36 critical information about the ways in which social support and health information are exchanged
37
38 online, the researchers’ postings and interactions with participants within the online forum may
39
40 have influenced participants’ posts and study findings. Language analysis techniques like the
41
42 MEM enable researchers to examine themes across massive corpora of natural language data,
43
44 which is challenging to accomplish using traditional qualitative methodology. Given the quantity
45
46 of data that is available online, the MEM may be helpful in extracting information that
47
48 complements or expands upon the findings of previous qualitative studies.
49
50
51
52
53
54
55
56
57
58
59

1
2
3 The current study sought to build upon past work that has demonstrated significant overlap
4 in the language content/themes found in controlled laboratory environments and social media (e.g.,
5 Stanton, Meston, & Boyd, 2017). Specifically, we sought to compare the themes of breast cancer
6 survivorship that emerged from social media data with those that emerged via content analysis of
7 focus group data. Interviews and focus groups are the most common methods of collecting data in
8 qualitative healthcare research (Gill, Stewart, Treasure, & Chadwick, 2008). To our knowledge,
9 themes extracted from natural language data via the MEM have yet to be compared to themes
10 gleaned from focus groups, which are useful in generating a nuanced understanding of
11 participants' experiences and beliefs about a certain topic. The current study offers insight on the
12 association between themes of breast cancer survivorship that are expressed on social media and
13 the self-reported experiences that breast cancer survivors share with researchers and/or medical
14 providers.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

30 **Methods**

31 **Social Media Data**

32
33 Archival data was collected from the Reddit website. An entertainment, social networking,
34 and news website, Reddit is similar to an online bulletin board and other internet forums; registered
35 users can submit various forms of content (e.g., hyperlinks, personal narratives) to topic-specific
36 discussion boards, which are voted up or down by other users to determine the positions of the
37 posts on each of the site's pages. We extracted data from the *r/breastcancer* subforum, which has
38 over 3,000 members and is described as a "community of support and information for those
39 affected by this disease through their personal struggle or that of a loved one." Each Reddit
40 subforum is structured such that users can either leave 1) an original post (OP) to which other users
41 can reply, or 2) a comment on those posts written by others. We included both OPs and comments
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 in response to the OPs (excluding hyperlinks) to parallel the structure of a focus group, in which
4
5 participants actively engage with each other in a group context.
6

7
8 The data from the *r/breastcancer* board was extracted using a custom script that pulled
9
10 archival data from Reddit's application programming interface (API). This Reddit API is publicly
11
12 accessible and allows researchers to acquire language data directly from the site without using the
13
14 typical web interface. Unfortunately, Reddit does not collect thorough demographic data on the
15
16 site's users, so we cannot describe the characteristics of the sample. The data posted through March
17
18 2016 were extracted chronologically. Ultimately, the sample included 1051 OPs and comments.
19
20

21
22 *Social media data analysis.* We used the MEM to extract the core themes of breast cancer
23
24 survivorship. The details of the MEM's analytic technique have been described thoroughly
25
26 elsewhere (Blackburn, Yilmaz, & Boyd, 2018; Boyd, 2017; Chung & Pennebaker, 2008;
27
28 Pulverman, Boyd, Stanton, & Meston, 2017), but we offer a brief description here. First, all 1051
29
30 OPs and comments were analyzed to determine word prominence. The words in the corpus were
31
32 "lemmatized," or converted to their basic inflections (e.g., "eats", "ate", and "eating" are converted
33
34 to "eat"), using the Meaning Extraction Helper (Boyd, 2014). Then, common closed-class (i.e.,
35
36 "function") words and uncommon open class (i.e., "content") words were removed. Content words
37
38 were considered uncommon if they appeared in less than 7.5% of all observations; this percentage
39
40 falls within the recommended range proposed by Chung and Pennebaker (2008) when they
41
42 developed the MEM. All remaining content words that are used by a set minimum of participants
43
44 (5%) were then given a binary score (0 = absent, 1 = present) for each observation. Finally, these
45
46 scores were submitted to a principle components analysis (PCA), the results of which revealed
47
48 clusters of words (which we refer to as "themes") that commonly co-occurred across observations.
49
50
51
52

53 54 **Focus groups**

55
56
57
58
59
60

1
2
3 To compare themes generated by computerized data extraction methods with those
4 described by breast cancer survivors through interpersonal group dialogues, we held two focus
5 groups. We partnered with a local breast cancer support organization to recruit breast cancer
6 survivors for these groups. Table 1 displays sociodemographic and cancer treatment variables from
7 participants in the focus groups. To arrange the initial focus group, the organization emailed
8 women in their database inviting them to participate in the study during the time normally allotted
9 for a survivor support group. To set up the second focus group, the support organization invited
10 their staff of breast cancer navigators to participate, thus enabling us to gather a broader perspective
11 of survivors' experiences through thousands of client interactions. The staff of the support
12 organization were all breast cancer survivors and professionally trained breast cancer navigators.
13 This navigator focus group occurred at an offsite facility to foster a sense of safety for sharing
14 experiences.

15
16
17 To ensure adherence to the research aims and questions, two of the authors (AS and JC)
18 facilitated both group sessions. Prior to the start of the focus groups, each participant completed a
19 one-page demographic survey that asked about their age, race, marital status, education, cancer
20 diagnosis and treatment modalities, and online presence. The first focus group included eleven
21 women and the second group included seven women. Each focus group lasted 90 minutes and
22 drew from an interview guide that included four main topics: (1) brief description of their
23 experiences at diagnosis and during breast cancer treatment; (2) engagement and/or disengagement
24 of internal and external resources during diagnosis, treatment, and recovery; (3) experiences in
25 accessing the internet for breast cancer information or support; (4) discussion of themes that
26 emerged from computerized data extraction methods to discover whether themes matched
27 individual experiences. After analyzing the Reddit data with the MEM and the focus group data

1
2
3 with thematic analysis, the coding team compared the themes that were derived from the two
4 methods. The objective of the comparison was not to find whether one method produced superior
5 themes; but rather, to perform a narrative synthesis of the data emerging from each method,
6 highlighting the ways in which themes converged or diverged. Comparison dialogues centered
7 around the emergence of the themes within each data set, as well as the words within each theme
8 (within the Reddit data) that had the highest loadings, noting the context of the dialogues across
9 the two samples.

10
11
12
13
14
15
16
17
18
19 The interview guide included the same topics for each group; however, the facilitators
20 encouraged the participants in the navigator group to respond from a broader perspective of trends
21 they encounter in caring for breast cancer survivors; however, some of the navigators briefly
22 shared personal experiences.

23
24
25
26
27
28
29 *Focus group data analysis.* Following each focus group, one investigator (MN) watched
30 the video-recorded session and transcribed the focus group data into a word document. The data
31 was then transferred into an Excel document for coding. Two of the investigators (AS and JC)
32 performed data analysis utilizing thematic analysis methodology as described by Guest,
33 MacQueen and Namey (2012). Analysis began with reading through each transcription to gather
34 an overall impression of the data. Data was segmented based on interview questions to assist with
35 the identification of descriptive themes related to diagnosis, treatment, social support, and coping
36 processes. Two investigators (AS and JC) coded the first focus group individually then met to
37 discuss coding methods and variance in codes. An analysis of variance between coders rendered
38 (83.4%) agreement in codes from the first focus group and (84.6%) agreement in codes for the
39 second focus group. Both coders discussed differences in codes until they achieved consensus.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

1
2
3 Following discussion and consensus on codes from the first focus group, the investigators
4 created a codebook. Data analysis for the second focus group followed that of the first focus group.
5
6 They entered data from the second focus group into the existing codebook, noting new codes that
7
8 arose within the data. Next, one of the investigators reviewed the master codebook, highlighting
9
10 coding nuances and commonalities. Both investigators then reviewed codes to discover key themes
11
12 within the data. While coding the data from the second interview, the researchers discovered that
13
14 participants' experiences closely matched those of the women in the first focus group. Specifically,
15
16 no new themes or meaning codes arose within the second focus group to further elaborate the first
17
18 focus group's core seven themes. Therefore, the researchers determined that they had achieved
19
20 thematic and meaning saturation and ended participant recruitment. Lastly, they compared final
21
22 themes from the focus group data with themes from the computer extraction method.
23
24
25
26
27

28 *Qualitative Comparison of Themes:* After analyzing the Reddit data with the MEM and the
29
30 focus group data with thematic analysis, the coding team compared the themes that were derived
31
32 from the two methods. The objective of the comparison was not to find whether one method
33
34 produced superior themes; but rather, to perform a narrative synthesis of the data emerging from
35
36 each method, highlighting the ways in which themes converged or diverged. Comparison
37
38 dialogues centered around the emergence of the themes within each data set, as well the as words
39
40 within each theme (within the Reddit data) that had the highest loadings, noting the context of
41
42 the dialogues across the two samples.
43
44
45
46

47 *Ethics:* The investigators received approval from their academic institutional review board
48
49 prior to initiation of participant recruitment. Each participant reviewed the research protocol,
50
51 signed the informed consent, and provided permission for video-taping. Participants received a
52
53 \$20 gift card to offset travel expenses. The data that support the findings of this study are available
54
55
56
57
58
59
60

1
2
3 on request from the corresponding author. The data are not publicly available due to privacy or
4
5 ethical restrictions.
6

7 8 **Results** 9

10 The presentation of results begins by portraying thematic commonalities between both
11
12 qualitative methods and then shifts to the unique presentations of themes within each analytic
13
14 method. Ten distinct themes emerged from the data sets: (1) screening/diagnosis, (2) treatment
15
16 process, (3) social support, (4) existentialism, (5) disclosure, (6) coping, (7) fears, (8) risk, (9)
17
18 information-seeking, and (10) surgery. Notably, four themes emerged in both samples and centered
19
20 on physical and psychosocial experiences during screening, diagnosis, and treatment, as well as
21
22 the ways in which women found social support and existential meaning throughout their illness.
23
24
25

26 Screening and diagnosis dialogues centered around discussions of cancer testing and
27
28 emotionally disturbing diagnosis dialogues with healthcare providers. Women also revealed
29
30 insurance and financial barriers, as well as challenges in finding medical providers to perform their
31
32 mammograms and biopsies. The women encountered troubling times in deciding which treatment
33
34 option(s) would best fit their personal and familial needs. Several women also expressed worries
35
36 at the time of diagnosis about the potentially negative consequences their illness may have on their
37
38 families, in particular, their children, regardless of the child's age. Treatment-related dialogues
39
40 focused on physical and psychosocial side-effects and the consequences on self and loved ones
41
42 such as depression, anxiety, fatigue, chemo brain, body image changes, nausea, and hair loss.
43
44
45
46

47 Social support appeared similarly for women from Reddit and the focus groups with many
48
49 turning to family, friends, faith communities, and other breast cancer survivors. Intimate and social
50
51 relationships experienced some strain due to physical and emotional changes, as well as
52
53 communication barriers. Social support appeared differently for survivors based on the
54
55
56
57
58
59
60

1
2
3 communities in which they identified. For example, women in the Latina/Latinx communities had
4
5 extensive social support systems, while older women struggled to find transportation and
6
7 caregivers to provide physical care following surgeries and chemotherapy or radiation treatments.
8
9
10 Discussions emerged similarly around role adaptations at work and other social settings.
11
12 Resoundingly, social support arose as instrumental in fostering survivors' motivation, emotional
13
14 well-being, and management of physical needs.
15
16

17 Although cancer presented numerous challenges for posters to the Reddit forum and
18
19 women in the focus groups, many participants found existential benefits from their diagnosis. They
20
21 described new-found personal strength and clarity in the people and activities in their life that
22
23 brought joy. The “silver linings” of cancer appeared as acknowledgment of gratitude and a
24
25 reprioritizing of relationships and wellness activities that might reduce stress and risks for cancer
26
27 metastasis. Survivors also explored the deeper, existential sides of themselves including their
28
29 spirituality and higher purpose, as well as creative outlets for self-expression through writing and
30
31 art.
32
33
34

35 **Social Media Results**

36
37 The MEM analysis resulted in a model with seven distinct components. Seven theme labels
38
39 were chosen by the experimenters to describe the co-occurring words that were associated with
40
41 each of the seven distinct themes: (1) diagnosis, (2) social support, (3) risk, (4) existentialism, (5)
42
43 treatment process, (6) information-seeking, and (7) surgery (see Table 2 for MEM themes and
44
45 loading values). Each observation was given a score for each theme by summing the total
46
47 percentage of theme-relevant words in each observation. Table 3 displays quoted text from
48
49 observations with high scores for each of the seven themes extracted from all Reddit posts, as well
50
51 as corresponding text from focus group participants.
52
53
54
55
56
57
58
59
60

1
2
3 The themes that were selected for inclusion had Eigenvalues above 4.25 and significantly
4 contributed to the total variance of the model. Words that were retained within each of the seven
5 themes had loadings of .25 or higher. Overall, the seven identified themes accounted for 16.13%
6 of the total variance. This percentage may seem small in the context of traditional factor analysis,
7 but it falls well within the range of variances that have been reported in other studies that applied
8 principle components analyses to natural language data (Chung & Pennebaker, 2008; Stanton et
9 al., 2015; Wolf, Chung, & Kordy, 2010).

19 Of the seven themes that emerged from the Reddit data, three (risk, information seeking,
20 and surgery) were unique to the online context. The first unique theme, risk, captured descriptions
21 that discussed genetic and other health risks associated with increased likelihood of a breast cancer
22 diagnosis. Posts that had high scores on the risk theme highlighted the link between the BRCA
23 gene, inherited risk, and family histories of cancer. Other posts discussed the importance of genetic
24 counseling and the ways in which counseling can improve coping for those who carry high genetic
25 risk. Posts that scored highly on information seeking, the second unique theme extracted from the
26 Reddit data, addressed strategies for finding accurate information about breast cancer and
27 associated treatments, sometimes providing specific resources for readers. In addition, some posts
28 with high scores on the information seeking theme offered support on each stage of the diagnosis
29 and treatment process, noting specific resources that apply to certain phases but not others. Finally,
30 posts that scored highly on the last unique theme to emerge from the Reddit data, surgery,
31 described the various surgical treatment options and offered personal anecdotes about these
32 options, sometimes delving into the emotional tolls of each procedure. Other posts with high scores
33 on the surgery theme described not only the initial surgeries (e.g., lumpectomies, mastectomies)
34 but also the subsequent surgeries (e.g., reconstruction, surgeries to remove defective implants) and
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59

1
2
3 the ways in which these surgeries were associated with hope for recovery and pain, both physical
4 and emotional.
5
6

7 **Focus Group Results**

8
9
10 Seven main themes arose from the focus group data; the first four listed were the same as
11 the MEM-derived themes and the final three represent themes that were unique to the focus groups.
12
13 The themes from the groups included (1) screening/diagnosis, (2) treatment, (3) social support, (4)
14 existentialism, (5) disclosure, (6) coping, and (7) fears. Table 4 displays the focus group themes
15 and codes within these themes.
16
17
18
19
20

21 The first unique theme from the focus groups, disclosure, related to women's comfort in
22 sharing her breast cancer diagnosis with family, friends, work colleagues, and current or future
23 intimate partners. The women mostly based their disclosure decisions on potential impacts of
24 others. Several of the women withheld their diagnosis from loved ones for extended periods of
25 time to protect them from the emotional burden of their diagnosis. Specifically, those with young
26 children or ailing parents strategized the best ways to disclose their diagnosis and sometimes hid
27 the information by disguising their hair loss and body changes or by avoiding visits with parents.
28 These women shared that social media offered the opportunity to disclose their diagnosis in a safe
29 setting and also the chance to gather insight about ways to share their news with loved ones. The
30 single women in the group shared an additional concern about when and how to tell people that
31 they date about their breast cancer. The process of disclosure depleted women's emotional energy
32 and appeared as one of the most challenging aspects of their cancer diagnosis.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 The women expressed many fears related to changes in their body, independence, and
50 relationships with others. Almost unanimously, the women talked about their fears of their cancer
51 returning or dying from their illness. Due to the extensive surgeries and side effects from treatment,
52
53
54
55
56
57
58
59
60

1
2
3 many of the women feared body image changes, not feeling themselves or embarrassing others
4 with their looks, such as their young children. They worried about becoming dependent and
5 requiring help from others. Many feared that illness progression would leave them debilitated like
6 other women they witnessed in treatment. Younger women expressed worries about dating as well
7 as infertility from their chemotherapy treatments. Lastly, the women shared concerns that changes
8 to their physical abilities would leave them less valuable to their children, partners, or friends.
9

10
11
12 Finding ways to cope with the distress of waiting for a cancer diagnosis, managing
13 treatment-related side effects, and balancing the uncertainties of survivorship transitions required
14 the women to draw on their consistent support mechanisms, as well as develop new coping
15 strategies. Most of the women turned to their existing support mechanisms to draw strength. They
16 also benefited from becoming immersed in their roles within family or careers. Caregiving roles,
17 especially as parents, enabled women to refocus their thoughts onto other's needs. They sought
18 ways to maintain a sense of normalcy and attempted to hold onto their daily routines. The ability
19 to remove negative forces in their lives such as relationships and distressing cancer information
20 afforded them a clearer space to process their thoughts. A small group of the women engaged a
21 hypervigilant stance and sought as much medical information as they could about their disease and
22 treatment options. In telling stories of their coping processes, the women most commonly chose
23 to explore positive aspects of their illness and participated in benefit-finding to decipher small and
24 large existential outcomes from their illness experience.
25
26

27 **Discussion**

28
29 Our findings reveal many similarities and some notable differences between both breast
30 cancer support formats as well as between both thematic analysis techniques. The seven themes
31 arose from the MEM (diagnosis, social support, risk, existentialism, treatment process,
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

1
2
3 information-seeking, and surgery) and seven themes were derived from focus groups
4
5 (screening/diagnosis, treatment, support systems, existential concerns, disclosure, coping and
6
7 fears). Women in the focus groups shared experiences with diagnosis, treatment and social support
8
9 that were similar to the posters on the Reddit forum; however, the in-person dialogue presented a
10
11 more nuanced and vulnerable description of the participant's coping processes, fears of
12
13 dependency and dying, and experiences surrounding disclosure of their diagnosis. Marshall and
14
15 colleagues (2016) conducted a comparison of online text extracted from a social media outlet with
16
17 items endorsed on breast cancer symptom checklist. Utilizing the K-medoid clustering method,
18
19 the study analyzed over 50,000 messages generated by users of the breast cancer forum on
20
21 MedHelp.org. Overall, the online data significantly overlapped with the checklist data. Findings
22
23 from this study mirror that of Marshall and colleagues (2016) and demonstrate the effectiveness
24
25 of topic modeling in this domain.
26
27
28
29

30
31 Focus groups have historically served as the gold standard for psychosocial cancer research
32
33 in developing, testing, and confirming soundness of research questions; however, focus groups
34
35 require extensive resources and introduce methodological limitations, including experimenter bias.
36
37 For example, focus groups require extensive time in planning, recruitment, and transcription.
38
39 Validity of findings become dependent upon participant recruitment and the influence of
40
41 normative discourse or group power dynamics (Smithson, 2000). Language analysis techniques
42
43 like the MEM enable researchers to examine themes across massive corpuses of natural language
44
45 data, which becomes challenging using traditional qualitative methodology. Data collected from
46
47 focus groups and individual interviews may be affected by experimenter bias, in that the questions
48
49 and associated probes that are used in these methodologies are generated by the researcher or the
50
51 moderator. Indeed, by facilitating interaction among members, offering transitional summaries,
52
53
54
55
56
57
58
59

1
2
3 and noting non-verbal responses, the moderator plays a large role in the focus group (McLafferty,
4 2004); language data extracted from online forums like Reddit are not shaped by a moderator or
5
6 by the experimenter. The MEM can also be used to track discourse over time, which may highlight
7
8 treatment-related changes, problematic thinking styles, and the emergence of adaptive coping
9
10 strategies (Pulverman et al., 2016).
11
12
13

14 Findings from this study highlight the benefits of the MEM in advancing psychosocial
15 cancer research. However, several limitations should be acknowledged. The focus group sample
16
17 had limited racial/ethnic and educational diversity. Women in the focus groups may have met
18
19 previously in their support group and have shared a familiar and trusting environment in which to
20
21 share more vulnerable experiences. As well, focus group participants had familiarity with the
22
23 researchers' backgrounds in counseling fields and breast cancer and this may have increased their
24
25 likelihood to share sensitive topics. Focus group attendees were affiliated with a breast cancer
26
27 support organization; thus, their responses may more likely reflect experiences of well-informed
28
29 and socially connected survivors. Reddit, as an online social media platform, may cater to younger
30
31 or more technologically savvy and educated individuals. Given our focus group samples' young
32
33 age and high educational attainment, our findings between methodologies may be comparable.
34
35 Importantly, the posts extracted from Reddit may have been written by a range of individuals—
36
37 those who were recently diagnosed, are currently undergoing treatment, are long-term survivors,
38
39 as well as family members, friends, and colleagues of individuals across the diagnosis, treatment,
40
41 and survivorship phases.
42
43
44
45
46
47
48

49 This kind of diversity speaks to another potential confound that may also be considered a
50
51 strength of this mixed-methods approach. Given the different populations assessed by the two
52
53 thematic analysis methods, different themes are likely to emerge from each method. Of the 10
54
55
56
57
58
59
60

1
2
3 resultant themes from this study, four themes co-occurred in both samples, with slight variations
4 potentially due to the context of dialogues among participants in each setting. Although there was
5
6 variability in the themes that emerged from each method, this variability does not discredit the
7
8 validity of either approach; rather, it highlights relevant nuances in experiences that may be driven
9
10 by contextual factors, including familiarity among members, willingness to express vulnerability,
11
12 and cultural norms in online versus in-person settings. For example, fears, one of the themes that
13
14 emerged from the focus groups but was not found in the Reddit data, may have been easier to
15
16 discuss in person among fellow survivors than to share over social media. Similarly, a recent
17
18 linguistic analysis of posts on breastcancer.org revealed that content shied away from challenging
19
20 topics, such as the fear of dying, and instead focused on protecting social relationships and personal
21
22 well-being (Malloch & Taylor, 2019). The purpose of this comparison is therefore not to
23
24 demonstrate the superiority of one method over the other, as we must expect some variation given
25
26 the characteristics of each technique; rather, it is to build a rich and nuanced understanding of the
27
28 breast cancer journey, acknowledging that both methods are ultimately useful in identifying the
29
30 specific needs of this population and tracking improvements as well as persistent concerns.
31
32
33
34
35
36

37
38 Passive online sampling through language analysis techniques could offer oncology
39
40 researchers additional mechanisms for exploring relationships that warrant more attention,
41
42 relationships that could be challenging to assess in a laboratory setting, or relationships that may
43
44 not be as strong as initially hypothesized. Researchers may consider engaging in more active
45
46 sampling using this methodology; that is, posing an open-ended question to an online group and
47
48 then analyzing written responses with the MEM. This approach has already been applied to written
49
50 responses on sensitive topics, such as childhood sexual abuse and sexuality in general, to
51
52 characterize experiences and to describe complex sensations (Handy, Stanton, & Meston, 2019;
53
54
55
56
57
58
59

1
2
3 Kilimnik, Boyd, Stanton, & Meston, 2018). Given the physical and psychosocial barriers that
4 cancer poses for research and data collection, the MEM opens opportunities to locate large samples
5 of participants that may historically be hard to find. The MEM, and social media data in general,
6 may also enable the inclusion of voices of cancer survivors who have physical or social limitations
7 that hinder in-person research participation.
8
9
10
11
12
13

14 **Acknowledgements**

15
16
17 The first author was supported by a Doctoral Training Grant in Oncology Social Work
18 Renewal, DSWR-18-087-03, from the American Cancer Society and a PEO Scholar Award. The
19 second author was supported by an NIMH T32 training grant (5T32MH116140-02). The authors
20 would express their gratitude for the breast cancer survivors for sharing their experiences.
21
22
23
24
25

26 **Disclosure statement**

27
28 The authors report no financial and/or business interests.
29
30

31 **Data availability**

32 Supporting data for this study can be requested through the corresponding author.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- Alshaikh, F., Ramzan, F., Rawaf, S., & Majeed, A. (2014). Social network sites as a mode to collect health data: A systematic review. *Journal of Medical Internet Research*, *16*(7), e171.
- Berger, A. M., Buzalko, R. J., Kupzyk, K. A., Gardner, B. J., Djalilova, D. M., & Otte, J. L. (2017). Preferences and actual chemotherapy decision-making in the greater plains collaborative breast cancer study. *Acta Oncologica*, *56*(12), 1690-1697.
- Blackburn, K. G., Yilmaz, G., & Boyd, R. L. (2018). Food for thought: Exploring how people think and talk about food online. *Appetite*, *123*.
<https://doi.org/10.1016/j.appet.2018.01.022>
- Boyages, J., Kalfa, S., Xu, Y., Koelmeyer, L., Mackie, H., Viveros, H., ... & Gollan, P. (2016). Worse and worse off: The impact of lymphedema on work and career after breast cancer. *Springerplus*, *5*(1), 657.
- Boyd, R. L. (2014). MEH: Meaning Extraction Helper (Version 1.2.42) [Software].
- Boyd, R. L. (2017). Psychological Text Analysis in the Digital Humanities. In S. Hai-Jew (Ed.), *Data Analytics in Digital Humanities* (pp. 161–189). https://doi.org/10.1007/978-3-319-54499-1_7
- Boyd, R. L., & Pennebaker, J. W. (2017). Language-based personality: A new approach to personality in a digital world. *Current Opinion in Behavioral Sciences*, *18*, 63–68.
<https://doi.org/10.1016/j.cobeha.2017.07.017>
- Boyd, R. L., Wilson, S. R., Pennebaker, J. W., Kosinski, M., Stillwell, D. J., & Mihalcea, R. (2015, April). Values in words: Using language to evaluate and understand personal values. In *International Conference of Web and Social Media* (pp. 31-40).

- 1
2
3 Brandão, T., Schulz, M. S., & Matos, P. M. (2017). Psychological adjustment after breast cancer:
4
5 A systematic review of longitudinal studies. *Psycho-oncology*, 26(7), 917-926.
6
7
8 Browall, M., Kenne Sarenmalm, E., Persson, L.-O., Wendstrom, Y., & Gaston-Johansson, F.
9
10 (2015). Patient-reported stressful events and coping strategies in post-menopausal women
11
12 with breast cancer. *European Journal of Cancer*, 25, 324-333.
13
14
15 Chung, C. K., & Pennebaker, J. W. (2008). Revealing dimensions of thinking in open-ended self-
16
17 descriptions: An automated meaning extraction method for natural language. *Journal of*
18
19 *Research in Personality*, 42(1), 96-132.
20
21
22 Drageset, S., Lindstrom, T. C., & Underlid, K. (2009). Coping with breast cancer: Between
23
24 diagnosis and surgery. *Journal of Advanced Nursing*, 66(1), 149-158.
25
26
27 Fasse, L., Flahault, C., Vioulac, C., Lamore, K., Van Wersch, A., Quintard, B., & Untas, A. (2017).
28
29 The decision-making process for breast reconstruction after cancer surgery:
30
31 Representations of heterosexual couples in long-standing relationships. *British Journal of*
32
33 *Health Psychology*, 22(2), 254-269.
34
35
36 Fiszer, C., Dolbeault, S., Sultan, S., & Brédart, A. (2014). Prevalence, intensity, and predictors of
37
38 the supportive care needs of women diagnosed with breast cancer: A systematic review.
39
40 *Psycho-Oncology*, 23(4), 361-374.
41
42
43 Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative
44
45 research: interviews and focus groups. *British Dental Journal*, 204(6), 291-295.
46
47
48 Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied thematic analysis*. Los Angeles:
49
50 Sage Publications.
51
52
53
54
55
56
57
58
59

- 1
2
3 Hamama-Raz, Y., Pat-Horenczyk, R., Roziner, I., Perry, S., & Stemmer, S. M. (2019). Can
4
5 posttraumatic growth after breast cancer promote positive coping?—A cross-lagged study.
6
7
8 *Psycho-oncology*, 28(4), 767-774.
9
- 10 Handy, A. B., Stanton, A. M., & Meston, C. M. (2019). What does sexual arousal mean to you?
11
12 Women with and without sexual arousal concerns describe their experiences. *The Journal*
13
14 *of Sex Research*, 56(3), 345-355.
15
16
- 17 Kilimnik, C. D., Boyd, R. L., Stanton, A. M., & Meston, C. M. (2018). Identification of
18
19 nonconsensual sexual experiences and the sexual self-schemas of women: Implications for
20
21 sexual functioning. *Archives of Sexual Behavior*, 47(6), 1633-1647.
22
23
- 24 Kim, E., Han, J. Y., Shaw, B., Shah, D., McTavish, F. & Gustufson, D. (2012). The process and
25
26 effect of supportive message expression and reception in online breast cancer support
27
28 groups. *Psycho-oncology*, 21, 531-540.
29
30
- 31 Laidsaar-Powell, R., Butow, P., Charles, C., Gafni, A., Entwistle, V., Epstein, R., & Juraskova, I.
32
33 (2017). The TRIO Framework: Conceptual insights into family caregiver involvement and
34
35 influence throughout cancer treatment decision-making. *Patient Education and*
36
37 *Counseling*, 100(11), 2035-2046.
38
39
- 40 Maheu, C., Hébert, M., Louli, J., Yao, T. R., Lambert, S., Cooke, A., ... & Kyriacou, J. (2019).
41
42 Revision of the fear of cancer recurrence cognitive and emotional model by Lee-Jones et
43
44 al with women with breast cancer. *Cancer Reports*, e1172.
45
46
- 47 Male, D. A., Fergus, K. D., & Cullen, K. (2016). Sexual identity after breast cancer: sexuality,
48
49 body image, and relationship repercussions. *Current Opinion in Supportive and Palliative*
50
51 *Care*, 10(1), 66-74.
52
53
54
55
56
57
58
59

- 1
2
3 Malloch, Y. Z., & Taylor, L. D. (2019). Emotional self-disclosure in online breast cancer support
4 groups: Examining theme, reciprocity, and linguistic style matching. *Health*
5 *Communication, 34*(7), 764-773.
6
7
8
9
10 Marshall, S. A., Yang, C. C., Ping, Q., Zhao, M., Avis, N. E., & Ip, E. H. (2016). Symptom clusters
11 in women with breast cancer: An analysis of data from social media and a research study.
12 *Quality of Life Research, 25*, 547-557.
13
14
15
16
17 McLafferty, I. (2004). Focus group interviews as a data collecting strategy. *Journal of Advanced*
18 *Nursing, 48*(2), 187-194.
19
20
21
22 Peek, N., Holmes, J. H., & Sun, J. (2014). Technical challenges for big data in biomedicine and
23 health: data sources, infrastructure, and analytics. *Yearbook of Medical Informatics,*
24 *23*(01), 42-47.
25
26
27
28
29 Pulverman, C. S., Boyd, R. L., Stanton, A. M., & Meston, C. M. (2017). Changes in the sexual
30 self-schema of women with a history of childhood sexual abuse following expressive
31 writing treatment. *Psychological Trauma: Theory, Research, Practice, and Policy, 9*(2),
32 181–188. <https://doi.org/10.1037/tra0000163>
33
34
35
36
37
38 Quinn, E. M., Corrigan, M. A., McHugh, S. M., Murphy, D., Mullane, A. D., & Redmond, H. P.
39 (2013). Who's talking about breast cancer? Analysis of daily breast cancer posts on the
40 internet. *The Breast, 22*, 24-27.
41
42
43
44
45 Rubenstein, E. L. (2015). "They are always there for me": The convergence of social support and
46 information in an online breast cancer community. *Journal of the Association for*
47 *Information Science and Technology, 66*(7), 1418-1430.
48
49
50
51
52
53
54
55
56
57
58
59

- 1
2
3 Shaw, B., McTavish, F., Hawkins, R. P., Gustafson, D., & Pingree, S. (2000). Experiences of
4 women with breast cancer: Exchanging social support over the CHESS computer network.
5
6 *Journal of Health Communication, 5*, 135-149.
7
8
9
10 Smithson, J. (2000). Using and analysing focus groups: limitations and possibilities. *International*
11
12 *Journal of Social Research Methodology, 3*(2), 103-119.
13
14 Stanton, A. M., Boyd, R. L., Pulverman, C. S., & Meston, C. M. (2015). Determining women's
15 sexual self-schemas through advanced computerized text analysis. *Child Abuse &*
16
17 *Neglect, 46*, 78-88.
18
19
20 Stanton, A. M., Meston, C. M., & Boyd, R. L. (2017). Sexual Self-Schemas in the Real World:
21
22 Investigating the Ecological Validity of Language-Based Markers of Childhood Sexual
23
24 Abuse. *Cyberpsychology, Behavior, and Social Networking, 20*(6).
25
26 <https://doi.org/10.1089/cyber.2016.0657>
27
28
29
30 Strasser, S. M., Smith, M., Pendrick-Denney, D., Boos-Beddington, S., Chen, K., & McCarty, F.
31
32 (2012). Feasibility study of social media to reduce intimate partner violence among gay
33
34 men in metro Atlanta, Georgia. *Western Journal of Emergency Medicine, 13*(3), 298-304.
35
36
37 Swanberg, J. E., Nichols, H. M., Ko, J., Tracy, J. K., & Vanderpool, R. C. (2017). Managing cancer
38
39 and employment: Decisions and strategies used by breast cancer survivors employed in
40
41 low-wage jobs. *Journal of Psychosocial Oncology, 35*(2), 180-201.
42
43
44 Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and
45
46 computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1),
47
48 24-54.
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Tighe, M., Molassiotis, A., Morris, J., & Richardson, J. (2011). Coping, meaning and symptom
4 experience: A narrative approach to the overwhelming impacts of breast cancer in the first
5 year following diagnosis. *European Journal of Oncology Nursing*, *15*, 226-232.
6
7
8
9
10 Wallner, L. P., Li, Y., McLeod, M. C., Hamilton, A. S., Ward, K. C., Veenstra, C. M., ... & Hawley,
11 S. T. (2017). Decision-support networks of women newly diagnosed with breast
12 cancer. *Cancer*, *123*(20), 3895-3903.
13
14
15
16
17 Wevers, M. R., Aaronson, N. K., Bleiker, E. M., Hahn, D. E., Brouwer, T., van Dalen, T., ... &
18 Vrouenraets, B. C. (2017). Rapid genetic counseling and testing in newly diagnosed breast
19 cancer: patients' and health professionals' attitudes, experiences, and evaluation of effects
20 on treatment decision making. *Journal of Surgical Oncology*, *116*(8), 1029-1039.
21
22
23
24
25
26 Williams, F., & Jeanetta, S. C. (2015). Lived experiences of breast cancer survivors after diagnosis,
27 treatment and beyond: Qualitative study. *Health Expectations*, *19*, 631-642.
28
29
30
31 Wolf, M., Chung, C. K., & Kordy, H. (2010). Inpatient treatment to online aftercare: E-mailing
32 themes as a function of therapeutic outcomes. *Psychotherapy Research: Journal of the*
33 *Society for Psychotherapy Research*, *20*(1), 71-85.
34
35
36
37
38 Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word
39 use among bloggers. *Journal of Research in Personality*, *44*(3), 363-373.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1

Focus Group Participant Demographics and Cancer Treatment Characteristics (n=18)

<u>Category</u>	<u>N(%)</u>	<u>M(SD)</u>	<u>Range</u>
Current age		53 (9.05)	39-72
Age at diagnosis		48 (9.53)	36-71
Time since diagnosis (in years)		4.5 (4.99)	1-17
Stage of cancer			
0	3 (16.7)		
I	3 (16.7)		
II	8 (44.4)		
III	3 (16.7)		
IV	0 (0.0)		
Other	1 (5.6)		
Surgery			
No	1 (5.6)		
Yes	17 (94.4)		
Radiation			
No	7 (38.9)		
Yes	11(61.1)		
Chemotherapy			
No	6 (33.3)		
Yes	12 (66.7)		
Hormones			
No	9 (50.0)		
Yes	9 (50.0)		
Genetic Testing			
No	5 (27.8)		
Yes	13 (72.2)		
Education			
Some high school or less	1 (5.6)		
High school graduate/GED	0 (0.0)		
Some college	4 (22.2)		
4 years of college	11(61.1)		
Advanced degree	2 (11.1)		
Relationship status			
Single, not dating	3 (16.7)		
In a committed relationship	2 (11.1)		
Married	12 (66.7)		
Separated/divorced	1 (5.6)		
Widowed	0 (0.0)		
Ethnicity			
Hispanic/Latina	5 (27.8)		
Not Hispanic/Latina	13 (72.2)		

1		
2		
3	Race	
4	American Indian/Alaska native	0 (0.0)
5	African American/Black	2 (11.1)
6	White	10 (55.6)
7	Asian American	1 (5.6)
8	Pacific Islander	0 (0.0)
9	Other	4 (22.2)
10		
11	Internet usage	
12	I have never used the internet	0 (0.0)
13	I use the internet a couple of times a	1 (5.6)
14	month	
15	I use the internet on a weekly basis	17 (94.4)
16	I use the internet daily	0 (0.0)
17		
18	Which internet sites do you use?	
19	Cancer information sites	
20	No	4 (22.2)
21	Yes	14 (77.8)
22		
23	Breast cancer support sites	
24	No	6 (33.3)
25	Yes	12 (66.7)
26		
27	Breast cancer blogs	
28	No	12 (66.7)
29	Yes	6 (33.3)
30		
31	Medical center websites	
32	No	5 (27.8)
33	Yes	13 (72.2)
34		
35	Other sites	
36	No	13 (72.2)
37	Yes	5 (27.8)
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		

Table 2

Breast Cancer Themes, including Most Commonly Co-Occurring Words and Loading Values, from the Reddit Sample-Extraction

<u>Theme 1: Diagnosis</u>		<u>Theme 2: Social Support</u>		<u>Theme 3: Risk</u>		<u>Theme 4: Existentialism</u>		<u>Theme 5: Treatment Process</u>		<u>Theme 6: Information Seeking</u>		<u>Theme 7: Surgery</u>	
Code	Loading Value	Code	Loading Value	Code	Loading Value	Code	Loading Value	Code	Loading Value	Code	Loading Value	Code	Loading Value
Lump	0.67	Husband	0.50	BRCA	0.76	Live	0.41	Lymph	0.73	Diagnosis	0.39	Double	0.55
Ultrasound	0.62	Recovery	0.39	Test	0.71	Die	0.41	Node	0.72	Question	0.34	Mastectomy	0.53
Mammogram	0.53	Friend	0.38	Genetic	0.66	Issue	0.46	Radiation	0.51	Answer	0.33	Negative	0.52
Biopsy	0.51	Expect	0.35	Mutation	0.65	Life	0.39	Chemo	0.43	Story	0.33	Stage	0.49
Tissue	0.50	Time	0.33	Gene	0.63	Cancer	0.36	Spread	0.42	Information	0.30	Reconstruction	0.36
Told	0.45	Give	0.32	Risk	0.53	Effect	0.37	Lumpectomy	0.36	Group	0.29	Surgery	0.32
Said	0.42	Work	0.32	History	0.50	People	0.31	Treatment	0.33	Read	0.26	Week	0.31
Size	0.40	Stress	0.31	Insurance	0.44	Side	0.31	Mom	0.30	Learn	0.26	PM	0.31
Breast	0.38	Kid	0.30	Positive	0.43	Year	0.29					Wife	0.30
Doctor	0.36	Care	0.30	Family	0.41	Change	0.29					Month	0.29
Worry	0.34	Sound	0.29	High	0.35	Woman	0.28					Diagnose	0.29
Small	0.34	Hard	0.29	Mother	0.33	Important	0.28						
Tumor	0.34	Stay	0.29	Cover	0.32	Treat	0.27						
Left	0.33	Deal	0.28	Age	0.28	Mean	0.26						
Right	0.33	Share	0.28	Chance	0.27								
Large	0.33	Hospital	0.28										

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Table 3

Quotes Representing Each Theme

<u>Theme</u>	<u>Meaning Extraction Method</u>	<u>Focus Groups</u>
Diagnosis	Tuesday, I go to see my doctor. He gives me a clinical breast exam. He feels the lump but he also told me something completely unexpected. He started to feel on the outer side of my left breast and asked me if I had noticed how thick and firm the tissue was on that side, going into my underarm. I simply thought I would go in there and he would tell me not to worry.	You get the diagnosis you have breast cancer, but actually that is only one part of the story... and there's this really insane rollercoaster that happens, where you get one call about one test result and another call about another and you don't know half the time what it means when they tell you... There's just like that terrorizing few weeks in the beginning that are just universally just so hard on everybody who gets a diagnosis.
Social support	It's awesome that your husband is supportive, and support him right back. It's not easy for anyone - but supporting each other will help.	... [I]t's the immediate family, the Spanish community is very close, and you never see them come alone to a meeting, they always have somebody that they want to bring, even a cousin a neighbor. The Hispanic community help each other, and you know, when they don't have financials they actually, the family and friends and neighbors will get together to pitch in and everybody give a dollar.
Risk	I was just diagnosed this year and have learned a lot in a short amount of time... While you can inherit the risk for breast cancer from your mother, the most telling would be if she's had genetic testing done for BRCA 1/2 and a few other genes that carry a high risk. If she's positive for those genes, you might want to look into being tested to see if you've inherited them.	—
Existentialism	At least personally, I feel like the main benefit of living a long and healthy life is that I would have extra time to spend with my partner.	And I've seen the spiritual go both ways. It can be an existential crisis for a lot of people. I've seen people both embrace the spiritual and the religious life and others move away from it completely.
Treatment process	A cancer diagnosis is always scary, but doctors have a lot more tools for fighting it... They are waiting on the biopsy because the info they get from that will determine what treatment plan they suggest.	[T]here are just a lot of issues that happen ongoing and it does not end when your active treatment ends and I think there's also this misperception, I think, you know, that doctors are not good at explaining to patients that are going through treatment. You know, they're trying to keep them positive to get them through treatment,

but then when treatment ends, they just think they're cancer free and walk away. And it's all done. It's over. And it's not over.

Information-seeking Radiation and chemotherapy are the main treatments, these have also changed over the years, so don't let people's horror stories get to you. Even if someone went through cancer 5 years ago, things have changed. Good luck to your mom in her fight. If she needs it, there is probably a support group in your area. For now, here is an internet hug for you and her. If you want to read some info these are good sites: [[breastcancer.org](http://www.breastcancer.org/)](<http://www.breastcancer.org/>) [CDC site](http://www.cdc.gov/cancer/breast/basic_info/) [[cancer.net](http://www.cancer.net/cancer-types/breast-cancer)](<http://www.cancer.net/cancer-types/breast-cancer>)”

Surgery She opted for a nipple/skin sparing double mastectomy and her cosmetic results look fantastic (there are some scary mastectomy pics out there). Most of those are older procedures, look for a surgeon who knows how to do his procedure (if your cancer locale permits it). My wife's breasts have some scars but by far the best looking boobs I've seen post mastectomy.

Disclosure ——— Of the young women I think two things, their fear is that they don't want to scare their children, they don't want to look different for their children, they want it to be as normal as possible. And then fear that, if they're single, they are not going to be able to date, right? What are they gonna, how are they gonna tell them, you know, their breasts look different? And when do they tell them.

Coping ——— I was diagnosed October 2nd of '15 and I took half days off to have chemo. I took three weeks to have surgery and then I just had to keep going, you know? The energizer bunny. But I did have a huge support group at my [job]. And I have a huge family so they're a support group too.

Fears ——— But I'm scared for recurrence. I'm more scared. I was not scared at chemo. I was not scared with radiation. I'm scared what will happen in future. How we can prevent a recurrence.

Table 4

Themes Derived from Breast Cancer Focus Groups

<u>Screening/Diagnosis</u>	<u>Treatment</u>	<u>Social Support</u>	<u>Disclosure</u>	<u>Existentialism</u>	<u>Coping</u>	<u>Fears</u>
Delays	Surgery	Friends	Not comfortable sharing	Prioritizing	Normalcy	Asking for help
Mammogram	Chemotherapy	Family	Family	Gratitude	Reframing	Dying
Young Age	Radiation	Significant other	Kids	Wellness	Accepting help	Recurrence
Stage	Reconstruction	Church	Dating	Self-discovery	Strength from others	Not being needed
Tests/labs	Motivated by fear of death	Neighbors	Engaging with social media	Personal strength	Roles	Body changes
Decision-making	Physical and emotional healing	Community members	Timing	Self-efficacy	Work	Not being self
Fear	Side effects	Internalized stigma	Consider family's needs	Higher purpose	Caregiving	Dating
Avoidance of support	Chemo brain	Support groups		Role model	Removing negative energy	Infertility
Withdrawal	Complications			Comforting others	Hypervigilance	Embarrassing others
Medical distrust					Humor	Comparing to others
Shame					Pain meds	
Feeling a burden					Seeking information	
Protecting others						