CHAPTER 3

Spontaneously published illness stories on a website for young women with breast cancer: do writers and themes reflect the wider population?

Regina Overberg[1]
Andries de Man[1]
Ron Wolterbeek[2]
Wilma Otten[3]
Bertie Zwetsloot-Schonk[1]

[1] Clinical Informatics Group, Leiden University Medical Center, Leiden
[2] Department of Medical Statistics, Leiden University Medical Center, Leiden
[3] Department of Prevention and Care, TNO Quality of Life, Leiden

Informatics for Health and Social Care 2013; 38(1): 54-66

Abstract

This study examined writer characteristics and themes written about in a set of 167 spontaneously published stories on a Dutch website for young women with breast cancer. The stories were coded for six disease characteristics and sixteen themes. Coding results were compared to the characteristics of young women with breast cancer in a hospital cancer register and to the frequency of problems among young breast cancer patients participating in quantitative studies. We found that writer characteristics were diverse. Yet, logistic regression showed that women were more likely to be a writer if they were diagnosed at a younger age (OR 0.82; 95% CI (0.78, 0.85)), underwent a mastectomy (OR 4.63; 95% CI (2.59, 8.26)), or were in the first treatment period (OR 2.83; 95% CI (1.44, 5.58)). All sixteen themes were present in the stories, but some themes were addressed less often than their frequency among participants of quantitative studies suggested. The findings indicate that a set of spontaneously published stories might not completely reflect the characteristics and themes of the wider population of young women with breast cancer. Websites with spontaneously published stories should inform readers about this.

Keywords: Internet, breast cancer, illness stories, writer characteristics, story themes

1 Introduction

Young breast cancer patients can encounter problems with respect to a wide range of issues, such as: body image and sexuality [1]; concerns about pregnancy [2], fertility issues [3,4] and heredity [5,6]; and, communication problems with physicians [7] or with family and friends [8]. For some patients it is helpful to write about their experiences with their illness [9,10]. The process of constructing one's illness story can help with making sense, asserting control, transforming identity and making decisions [11]. Patients can put their stories on the Internet, where others -such as family, friends or other patients- can read them.

For readers of the stories it might be most helpful when a set of stories on a website gives an acceptable impression of the various topics that are relevant for coping with the disease: the set of stories should be a good starting point for a fellow patient, relative or friend to get an impression of what it means to suffer from the disease. Thus, the stories should be about topics that are generally considered important for coping with the disease.

Some of the websites with closed individual stories are provided by research groups, who have carefully selected the set of stories that is presented on the website. The research group who composed the Health Talk Online website (formerly DIPEx) [12], for example, used a method called maximum variation sampling to ensure that a wide range of perspectives and experiences is represented [13]. There is a clear consensus among an international team of patients, practitioners, policy makers, and researchers that such a range of perspectives and experiences should be included [14].

Besides research groups, there are active groups of patients who provide websites with closed individual stories. On many of these websites patients can spontaneously publish their own story. In this way a set of stories develops which is not based on any formal sampling methods. A number of concerns might rise at this point. Butow et al. [15] mention that it might be that only articulate patients or patients who have strong motivations to tell their story spontaneously publish their stories. Moreover, in an earlier study we found that patients wanted to read about themes they experienced problems with [16]. If this also applies to writing, than patients might only spontaneously write about themes they experienced problems with. As a consequence, it might be that a set of spontaneously published stories is written by a specific group of patients and that only specific themes are addressed in the stories.

The present study examines the set of spontaneously published stories on the website of the *Amazones* foundation. This is a Dutch website for and by young women with breast cancer [17]. Our study entailed a content analysis of all 167 spontaneously published stories on this website and encompassed the following questions: 1) What are the characteristics of the patients who spontaneously publish their story on this website?, and 2) Which themes do they address in their stories? In order to answer these questions, we used a coding scheme with six disease characteristics and sixteen themes. Furthermore, we examined whether the writer characteristics reflect the characteristics of the wider population of young women with breast cancer. We used the cancer register of the Leiden University Medical Centre (LUMC) to get an indication of the characteristics of the wider population of young women with breast cancer. Finally,

for five clearly defined themes we examined whether the extent to which these themes were written about in the stories reflect the prevalence of these themes among the wider population of young women. We used quantitative studies that examined prevalences of these five themes among young women with breast cancer to get an indication of the extent to which these themes are prevalent in the wider population.

2 Methods

2.1 Main empirical material

In January 2007 we downloaded all 167 stories from the *Amazones* website to analyse them. This was the total number of stories posted since women started to submit anonymously (by using nicknames) their stories for publication on the website in 2004. There were no restrictions for the women as to length, structure and content of their stories. Submitted stories were put on the website by the board of the foundation. The board did not exclude any stories from publication. In this article, the women who posted their story are referred to as the writers.

The stories were all in the Dutch language and written in the first person. Most stories followed a chronological order: discovery of breast cancer, treatments and the restitution process. All stories were closed individual stories: each of the stories was published in one time as a whole, i.e. no additions or changes were made after publishing a story online and the stories were not part of an interactive forum. One third of the writers stated their region of origin: more than 90% of them lived in the Netherlands.

2.2 Coding writer characteristics

In order to code the writer profiles a coding scheme with six characteristics was developed. The six characteristics were based on how other websites disclose patient stories [12] and on studies about the aspects patients appreciate in websites with stories [16,18]. The characteristics in the coding scheme were: 1) age at diagnosis, 2) time since diagnosis, 3) partner status, 4) children status, 5) treatment received, and 6) phase in the course of disease.

The continuous variables 'age at diagnosis' and 'time since diagnosis' were coded with a value (years and months, respectively) or with 'unknown'. 'Time since diagnosis' was defined as the number of months between being diagnosed and writing one's story. The other writer characteristics were all categorical variables. 'Partner status' and 'children status' were coded with 'yes', 'no' or 'unknown'. 'Treatment received' was divided into: a) lumpectomy, b) mastectomy, c) axillary lymph node removal, d) radiation therapy, e) chemotherapy, f) immunotherapy, g) hormonal therapy, and h) breast reconstruction. When writers mentioned some of these treatments, these were coded with 'yes' and the ones that were not mentioned, with 'no'. When none of the treatments were mentioned, all eight were coded with 'unknown'. For 'phase in the course of disease' the categories and code instructions were: 1) in first treatment period (treatments a-f), 2) free of cancer (undetectable cancer or treatments g-h), 3) breast cancer for second time (recurrence of the same breast tumour or a new primary breast tumour), 4) metastasized cancer

(cancer in other parts of the body than the breasts), 5) passed away (obituary at the end of a story written by the website moderators), and 6) unknown.

One coder coded all the stories for writer characteristics, and a second coder coded 15% (n=25). Besides the information in the story itself, the information in the so-called *Amazones* profiles on the website was used: 97 writers (58%) completed such a profile. The inter-rater reliability for each writer characteristic was calculated using the Intraclass Correlation Coefficient (ICC). The ICC's ranged from 0.607 to 1.000 (Table 1). Two ICCs for 'age at diagnosis' were calculated. These indicate the inter-rater reliability for i) assigning a value for age versus assigning 'unknown', and ii) the values that were assigned for age, after excluding the stories with 'age unknown' (n=6). For 'months since diagnosis' one ICC was calculated, since no 'unknown' was assigned for the 25 stories that were coded by two coders.

2.3 Coding story themes

In order to code the themes in the stories a coding scheme with sixteen themes was developed (Table 1). The sixteen themes were divided in four domains: 1) Diagnosis, 2) Treatment, 3) Health care system, 4) Living with it. For the development of the coding scheme, we studied which themes were disclosed in breast cancer stories on two different websites [12,19]. Both were developed by research groups and aim to represent a wide range of perspectives and experiences [13,20]. Two independent coders unfamiliar with the *Amazones* website coded all the stories for themes. The coders were trained using an identical sub-sample (n=13) of the stories under study. A story could receive multiple codes. A code was assigned when women wrote about a theme, irrespective of how they experienced the theme. The inter-rater reliability for each theme was calculated using Cohen's kappa (Table 1).

2.4 LUMC cancer register

To examine how the characteristics of the writers relate to those of the wider population of young women with breast cancer, the characteristics of the writers were compared to those of young women with breast cancer registered at the Leiden University Medical Centre (LUMC). The LUMC is one of eight university hospitals in the Netherlands with a capacity of 882 beds. All cancer patients who consult the LUMC are entered into the LUMC cancer patients register. This register contains information per tumour as to its location, the date of its discovery, the ways in which it was treated, whether it recurred locally/regionally and/or whether it metastasized, and if so when, the age of the patient at diagnosis, and whether or not the patient has died. The register does not include information on partner and children status, and any decisions as to breast reconstruction.

In order to make a useful comparison with the *Amazones* writers, we selected from the register women with breast cancer, diagnosed before the age of 50, who contacted the LUMC between 1997 and 2007. This age range was selected because the *Amazones* present themselves as women younger than 50 [17]. The time range 1997-2007 was selected to ensure that all phases in the course of disease were represented. The selection consisted of n=1160 women with a mean age at diagnosis of 42.4 years (SD=5.6) and a mean time since diagnosis of 56.7 months (SD=36.1). The treatment that most of the

women had undergone was radiation therapy (82%; n=952), followed by axillary lymph node removal (73%; n=849), chemotherapy (72%; n=838), mastectomy (48%; n=560), hormonal therapy (43%; n=494) and immunotherapy (6%; n=73). Almost one third of the women was free of cancer (31%; n=359) and one in five had passed away (20%; n=228). The remaining women were in first treatment period (7%; n=76), had cancer for a second time (6%; n=73), had metastasized cancer (6%; n=70) or their phase in the course of disease was unknown (31%; n=354).

2.5 Quantitative studies

In June 2008 PubMed, Embase and PsychInfo were searched for quantitative studies about problems experienced by young women with breast cancer in order to examine how the frequency with which themes were written about spontaneously in the *Amazones* stories relates to the extent to which themes are prevalent in the wider population of young women with breast cancer. In the three literature databases we searched for the themes 'contact with medical staff', 'family and friends', 'body image and sexuality', 'pregnancy issues' and 'concerns about heredity' by combining the search terms *breast cancer* and *young women* successively with *health care system, family (support), social network, social contacts, social behaviour, child(ren), body image, sexuality, pregnancy, heredity, genetic counseling*. All studies that reported prevalences were used for the comparison. Also, for a more useful comparison the theme 'pregnancy issues' was divided into the subthemes 'wanting to become pregnant after treatments' and 'pregnant at diagnosis'.

Table 1 Coding schemes for writer profile and story themes and inter-rater reliability.

Coding scheme writer profile 15% (n=25) of stories were coded by a second coder		Intraclass Correlation Coefficient (ICC)
Age at diagnosis (in years)	Assigning a value for age versus assigning 'unknown'	0.759
	The values that were assigned for age, after excluding the stories with 'age unknown' (n=6)	0.849
Time since diagnosis (in months)	The values that were assigned for time since diagnosis (for none of the 25 stories 'time since diagnosis unknown' was assigned)	0.971
Partner status, n (%)		0.800
Children status, n (%)		1.000

Table 1 (Continued)

Coding scheme writer profile 15% (n=25) of stories were coded by a second coder		Intraclass Correlation Coefficient (ICC)
Treatment received, n (%)	Lumpectomy	0.823
	Mastectomy	0.809
	Axillary lymph node removal	0.914
	Radiation therapy	0.839
	Chemotherapy	1.000
	Immunotherapy	0.657
	Hormonal therapy	0.844
	Breast reconstruction	0.607
Phase in the course of disease, n (%)		0.642
Coding scheme story themes All stories (n=167) were coded by a second coder; the training sub-sample of stories (n=13) was left out when calculating Cohen's kappa's		Cohen's kappa
Diagnosis	Feelings about diagnosis	0.70
Treatment	Decision-making	0.69
	Coping with treatment	0.43
	Side effects	0.79
Health care system	Delay in/errors at diagnosis	0.43
	Waiting for test results	0.68
	Contact with medical staff	0.63
	How treatment was performed	0.39
	Second opinion	0.72
Living with it	Work and insurance	0.57
	Family and friends	0.67
	Body image and sexuality	0.65
	Pregnancy issues	0.88
	Coping with breast cancer	0.57
	Practical advices	0.49
	Concerns about heredity	0.65

2.6 Statistical analysis

The analyses included descriptive statistics of writer characteristics and of the frequency of themes in the stories. Furthermore, a multivariate logistic regression analysis that included all disease characteristics simultaneously was performed to compare the writers with the women in the LUMC cancer register. Group (writers vs. LUMC register women) was included in the model as the dependent variable; disease characteristics were included as independent variables. The two dichotomous variables lumpectomy (yes/no) and mastectomy (yes/no) were combined to form a new dichotomous variable (mastectomy vs lumpectomy only). For the women in the cancer register the time since diagnosis was defined as the period between the date of diagnosis and the date on which the women were selected from the register. For 'phase in the course of disease' a category 'unknown' was included in the model, since for almost one third of the register women this variable was unknown. Confounding was assessed by evaluating changes in the relevant regression coefficients (beta's). Goodness-of-fit was determined by the Hosmer and Lemeshow Test. In the final model non-significant predictors ($p>0.05$) were removed.

2.7 Ethical considerations

The board of the *Amazones* foundation gave us permission to conduct the study. In order to maximize confidentiality, we report our results on an aggregated level without quoting the writers [21]. The LUMC Medical Ethical Committee concluded that our study involved no *medical* intervention and that we could proceed.

3 Results

3.1 Characteristics of the writers

Writer characteristics are shown in Table 2. The mean age at diagnosis was 36 years (SD=6.2). The mean time since diagnosis was 27 months (SD=37.4). The treatments most often undergone were: chemotherapy (77%), mastectomy (68%) and radiation therapy (61%). Few writers received immunotherapy (5%). Most writers were 'free of cancer' (46%) or 'in first treatment period' (34%). In addition, most writers had a partner (76%) and children (74%), and almost one-third of the writers who had received a mastectomy decided to undergo breast reconstruction (31%).

Table 2 Characteristics of the women who published their story on the *Amazones* website; i.e. the *Amazones* writers (n=167).

Characteristics		Amazones writers (n=167)
Age at diagnosis (in years)[a]	Mean (SD)	36 (6.2)
	Minimum	22
	Maximum	52
Time since diagnosis (in months)[b]	Mean (SD)	27 (37.4)
	Minimum	<1
	Maximum	252
Partner, n (%)	Yes	127 (76)
	No	18 (11)
	Unknown	22 (13)
Children, n (%)	Yes	123 (74)
	No	23 (14)
	Unknown	21 (13)
Treatment received, n (%)[c]	Lumpectomy only	43 (26)
	Mastectomy	113 (68)
	Axillary lymph node removal	91 (55)
	Radiation therapy	102 (61)
	Chemotherapy	129 (77)
	Immunotherapy	8 (5)
	Hormonal therapy	67 (40)
	Breast reconstruction	35 (31[d])
Phase in the course of disease, n (%)	In first treatment period	56 (34)
	Free of cancer	76 (46)
	Breast cancer for second time	14 (8)
	Metastasized cancer	11 (7)
	Passed away	5 (3)
	Unknown	5 (3)

[a] age at diagnosis unknown n=33
[b] months since diagnosis unknown n=8
[c] all eight treatments unknown n=6, type of surgery unknown n=1
[d] percentage based on the writers who had a mastectomy (n=113)

Table 3 Final logistic regression model for being a writer on the *Amazones* website.[a]

Disease characteristics		OR[b]	95% CI[c]
Age at diagnosis (years)		0.82	(0.78-0.85)
Time since diagnosis (months)		0.98	(0.97-0.99)
Treatment underwent[d]	Lumpectomy only	1.00	
	Mastectomy	4.63	(2.59-8.26)
	No axillary lymph node removal	1.00	
	Axillary lymph node removal	0.45	(0.26-0.79)
	No hormonal therapy	1.00	
	Hormonal therapy	1.65	(1.00-2.71)
	No immunotherapy	1.00	
	Immunotherapy	0.16	(0.06-0.47)
Phase in the course of disease	Free of cancer	1.00	
	In first treatment period	2.83	(1.44-5.58)
	Cancer for second time	0.64	(0.28-1.48)
	Metastasized cancer	0.85	(0.36-2.01)
	Passed away	0.04	(0.01-0.15)
	Unknown	0.00	(0.00-)

[a] The model includes n=1243 complete cases (n=127 writers and n=1116 cancer register women)
[b] OR=odds ratio; OR's are adjusted for all covariates in the model
[c] CI=confidence interval
[d] Chemotherapy and radiation therapy were not included in the model due to non-significance

3.2 Comparison with the LUMC cancer register women

The results of the logistic regression analysis are shown in Table 3. There was no confounding and goodness-of-fit was satisfactory (Hosmer and Lemeshow Test, p=0.995). The clearest results were that women were more likely to be a writer if they were diagnosed at a younger age (OR=0.82; 95% CI (0.78, 0.85)), underwent a mastectomy (OR=4.63; 95% CI (2.59, 8.26)) and/or were in the first treatment period (OR=2.83; 95% CI (1.44, 5.58)). Radiation therapy and chemotherapy were not significant predictors, nor were three categories of 'phase in the course of disease' (cancer for a second time, metastasized cancer, and unknown).

3.3 Frequency of themes in the stories

All sixteen themes were present in the stories (Figure 1). Many stories contained themes in the domains 'Diagnosis' and 'Treatment'. In 53-61% of the stories women wrote about their feelings on the diagnosis. Treatment decision-making appeared in 47-54% of the stories and side effects were disclosed in 40-45% of the stories. Also, two themes in the domain 'Living with it' were disclosed in many stories: family and friends appeared in 40-47% of the stories, and coping with breast cancer in 41-61% of the stories. The themes least often disclosed were practical advices (written about in 1-4% of the stories) and work and insurance (written about in 4-6% of the stories).

3.4 Comparison with the quantitative studies

The comparison of the prevalences of the themes in the set of stories with the prevalences in the quantitative studies is shown in Table 4. The prevalence of three of the themes was almost equal between the stories and the quantitative studies: 'family and friends' (40-47% vs 13-53%), 'pregnant at diagnosis' (7% vs 4-6%), and 'concerns about heredity' (10-11% vs 7-9%). In the stories the prevalence of the theme 'contact with medical staff' (26-27% vs 37%) was somewhat lower than the prevalence in the quantitative study. The difference in prevalence was substantial for the themes 'body image and sexuality' (20-25% vs 50%) and 'wanting to become pregnant after treatments' (9% vs 54-57%). None of the themes had a greater prevalence in the set of stories than in the quantitative studies.

4 Discussion

We examined the writer characteristics and themes written about in the spontaneously published *Amazones* stories. We found that the stories were written by a diverse group of young breast cancer patients: the group of writers was characterized by a diversity of ages, number of months since diagnosis, treatments and phases in the course of disease. The group of writers slightly deviated from the wider population of young women with breast cancer: in comparison to the women in the LUMC register who were diagnosed with breast cancer before the age of 50 relatively many writers were of a younger age, underwent a mastectomy and were in the first treatment period. With respect to the themes we found that the set of stories contained a broad range of themes: all sixteen themes under study were present in the set of stories. Yet, some themes were present in only a few stories. Furthermore, for two of the five themes we examined closer it was found that these themes did not appear in the set of stories in the extent that could be expected on the basis of quantitative studies among young women with breast cancer.

Table 4 Prevalence of themes in *Amazones* stories (n=167) and in quantitative studies on young women with breast cancer.

Domain	Theme	Stories %	Other studies %	Measurement	Design	Participants	Study
Health care system	Contact with medical staff	26-27	37	Reported to have had communication problems with their physicians	In-person interviews, consisting of closed as well as open-ended questions; most questions were adopted from validated questionnaires	185 women diagnosed before the age of 50 who were cancer-free 5 yrs later (median age at the beginning of the study (5 yrs before) was 45 years)	Bloom et al., 2004 [7]
Living with it	Family and friends	40-47	22-53	Experienced talking about death/ cancer/ wills and finances/ fears/ future/ feelings with partner as problematic	Structured, postal one-time survey including subscales about relationship with partner and children (adopted from a validated questionnaire)	204 women diagnosed at age 50 or younger who were at least 3 months post-diagnosis (mean current age 43.5 yrs, SD=6.2)	Walsh et al., 2005 [8]
			13-20	Experienced talking about cancer with children/ helping children cope/ caring for children as problematic			
	Body image and sexuality	20-25	50	Experienced ≥2 body image problems some of the time or at least one problem much of the time	Structured, in-person interviews; most questions were adopted from validated questionnaires	549 women aged 22-50 yrs in a stable (un)married relationship within 7 months of diagnosis (no mean age was reported)	Fobair et al., 2006 [1]

Table 4 (Continued)

Domain	Theme	Stories %	Other studies %	Measurement	Design	Participants	Study
Living with it	Pregnancy issues – Wanting to become pregnant after treatments	9	54	Had the question whether becoming pregnant after treatment was possible	Mailed, self-report questionnaire; participants could indicate for a list of fertility- and menopause related questions whether they had had that question or not	228 women diagnosed at age 40 or younger who were 6 to 60 months after diagnosis (mean age at diagnosis 35.5 yrs, SD=4.0)	Thewes et al., 2005 [4]
			57	Recalled substantial concern at diagnosis about becoming infertile with treatment	Structured, e-mailed, one-time survey about fertility issues; questions were as far as possible adopted from validated questionnaires	657 women diagnosed at age 40 or younger who were premenopausal (mean current age 35.8 yrs)	Partridge et al., 2004 [3]
	Pregnancy issues – Pregnant at diagnosis	7	4	Excluded from analysis of the study (see above) due to being pregnant at diagnosis	See above	860 informed consent givers	Partridge et al., 2004 [3]
			6	Was pregnant when breast cancer was diagnosed	Documentation of the clinical presentation and course of breast cancer using data from patient records	407 women aged 20-29 years at diagnosis (median age 28 yrs)	Guinee et al., 1994 [2]
	Concerns about heredity	10-11	7	Was found to carry BRCA1 or BRCA2 deleterious mutations	Screening for BRCA1 and BRCA2 mutations	136 women diagnosed before age 46 years in Spain (no mean age was reported)	De Sanjose et al., 2003 [5]
			9	Was associated with a germline mutation in BRCA1 or BRCA2	Molecular genetic study in which germline mutations in BRCA1 and BRCA2 were investigated	234 women diagnosed under the age of 41 years in southern Sweden (no mean age was reported)	Loman et al., 2001 [6]

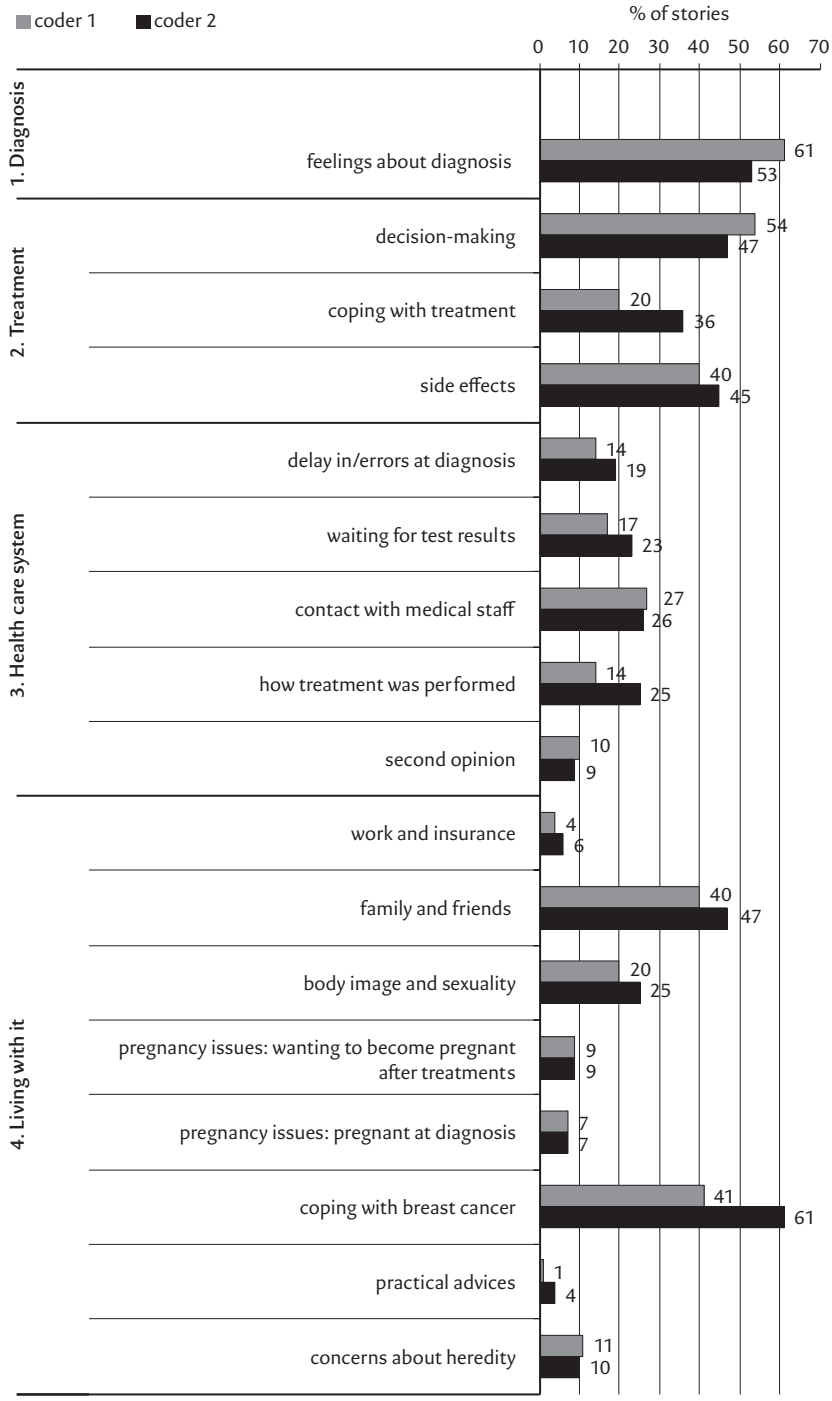

Figure 1 Frequency of themes in the Amazones stories (n=167).

An explanation for our finding that relatively many writers were in the first treatment period might be that the period right after hearing the diagnosis is the most difficult for a patient. In that period many things happen that disturb the normal daily routine. The patient has to find a way to cope with her altered health situation and with the uncertainty this involves. As a result the need to write about it might be relatively high right after hearing the diagnosis.

Patients might predominantly write about the themes that are keeping them busy at the time of writing their story. Since relatively many writers were in the first treatment period, our finding that the themes most frequently written about were the themes directly related to survival (feelings about diagnosis, decision-making about treatment, side effects of treatment) and the effects of illness on daily life (family and friends, coping with breast cancer) are in line with what could be expected. It also may explain our finding that in only few stories themes such as 'work and insurance' and 'practical advices' were addressed. Such themes may become more important after treatments.

The themes 'body image and sexuality' and 'contact with medical staff' did not appear in the set of stories in the extent that could be expected on the basis of quantitative studies among young women with breast cancer. Several explanations can be given for this finding. First of all, whether a patient writes about a theme might be dependent on the patient's demographic and disease characteristics and on the patient's phase in the course of disease. As mentioned above, a patient who is in the first treatment period will be most concerned with surviving and less with -for example- body image and sexuality. Secondly, it may be that certain themes do not come to a patient's mind at the time of writing her story, because she has already coped with these themes or these themes are less important to her. For some patients the theme 'contact with medical staff' might fall in this category. In this regard the act of writing a personal story is a different type of self-disclosure than filling out a structured questionnaire or answering questions in a structured interview: in a questionnaire or interview a patient is actively asked whether or not he/she experiences problems with a certain theme. A third explanation might be that a certain theme does come to a patient's mind at the time of writing her story, but that she decides not to write about it because in her opinion the theme is too personal which makes her not feel comfortable to publish her experiences with this theme on a freely accessible website. This explanation might hold for the theme 'body image and sexuality'. However, when publishing a story on a website the writer can stay anonymous (for example by using a nickname). So, it can also be argued that the Internet is ideally suited to write about personal themes [22].

The question arises whether patients write about those themes other patients want to read about. An earlier study of us showed that breast cancer patients who were asked about what themes they want to read in stories of other patients ranked the themes 'coping with emotions' and 'impact on daily life' high [16]. The present study showed that the themes 'feelings about diagnosis', 'coping with breast cancer' and 'family and friends' are present in many stories. Thus, with regard to these themes, there is quite a match between the themes patients write about and the themes patients want to read about. Yet, in our earlier study [16] patients mentioned they also wanted to read about 'practical information' and 'work rehabilitation'. The present study showed that these two themes are present in only a few stories.

When a theme is present in only a few stories this can be problematic in the case a website with a set of stories does not have a search facility on story topics: the reader has to scroll through all stories to find the few stories in which the theme is present. Yet, when a website with a set of stories does have a search facility on story topics it may be less problematic that a theme is present in only a few stories: the few stories in which the theme is present will be found relatively quick and easy. However, there are two other possible problems in the case a theme is only present in a few stories. First of all, the range of perspectives and experiences in respect of that theme might be limited. Second, a reader might get the idea that a theme that is written about in only a few stories will also be rare in real life. The latter need not to be true.

Organizations/patients who offer websites with stories should be motivated to describe how the set of stories on their website comes into being and how likely it is that a set of stories contains a broad range of themes and a broad range of perspectives and experiences in respect of these themes [14]. Such a description should certainly be applied to websites with spontaneously published stories, because -as the present study showed- on such websites the range of writer characteristics, perspectives and experiences may not fully reflect the wider population. Yet, when a patient has specific questions to other patients he/she can also consider to pose the question on a forum or to search on a forum for the answer instead of searching for stories. However, Van Uden-Kraan et al. [23] found in their study about health-related Internet use by patients with somatic diseases that only a small proportion of patients had ever read along with an online patient support group or had send postings to such a group.

5 Limitations

The two coders did not completely agree with respect to the writer profiles and the themes. The relatively low inter-rater reliability for 'immunotherapy' might be caused by the fact that few writers underwent immunotherapy. The two coders disagreed about 'immunotherapy' in only one of the 25 stories they both coded for writer characteristics. In addition, the relatively low inter-rater reliability for 'phase in the course of disease' might be caused by the six categories in which phase could be coded: this is quite a large number of categories. The two coders disagreed for 'phase in the course of disease' only in two of the 25 stories. For four of the themes there was some disagreement between the two coders (Cohen's Kappa <0.50). The coding instructions for these themes might not have been completely clear.

In order to examine how the writer characteristics relate to those of the wider population of young women with breast cancer, we compared the writer characteristics to the characteristics of young women with breast cancer in the LUMC cancer register. The LUMC cancer register was the best source we could think of to get insight in the wider population, but it should be noted that a university hospital register might contain a relatively large amount of more complicated cases. In order to make the comparison as useful as possible we selected a group of breast cancer patients from the register based on age, sex and year.

In order to examine how the frequency themes were written about in the stories relates to the frequency of these themes in the wider population, we compared the frequency of themes in the stories with the frequency of problems among young breast cancer patients participating in quantitative studies that use physiological outcome measures (pregnancy, BRCA-genes) or structured validated questionnaires. The background characteristics of participants in some of the quantitative studies differed slightly from those of the writers, which may have biased the comparison. In addition, the quantitative studies focused on difficulties participants had in respect of a particular theme, whilst we focused on whether the writers wrote about a theme at all.

Finally, it should be taken into account that websites are constantly changing. Patients will keep publishing stories on the *Amazones* website, which makes the composition of the set of stories dynamic.

6 Conclusion

This study shows that the set of spontaneously published stories on the *Amazones* website is written by a diverse group of writers and consists of a broad range of themes. Readers of the *Amazones* stories can get an acceptable impression of who may be affected by breast cancer and what it means to suffer from the disease. However, writers and themes on the *Amazones* website do not completely reflect the wider population of young women with breast cancer. Spontaneously published stories on the Internet might be subject to background characteristics of the writers, to main concerns at the time of writing one's story and to the desired degree of self-disclosure. Websites with spontaneously published stories should inform readers about how the set of stories comes into being and that the range of writer characteristics, perspectives and experiences may not fully reflect the wider population.

Acknowledgements

We would like to thank the board of the *Amazones* foundation for giving us permission to conduct this study, Henriët Werij-Wolters and Kristel Schaap for their coding work and Jan Molenaar for retrieving the data from the cancer register.

The Netherlands Organisation for Scientific Research (NWO) financed this study. This study is part of the Narrator project, under the ToKeN2000 umbrella program.

Declaration of Interest

The authors report no conflicts of interest

References

1. Fobair P, Stewart SL, Chang S, D'Onofrio C, Banks PJ, Bloom JR. Body image and sexual problems in young women with breast cancer. Psychooncology 2006;15:579-594.
2. Guinee VF, Olsson H, Moller T, Hess KR, Taylor SH, Fahey T, Gladikov JV, van den Blink JW, Bonichon F, Dische S, Yates JW, Cleton FJ. Effect of pregnancy on prognosis for young women with breast cancer. Lancet 1994;343:1587-1589.
3. Partridge AH, Gelber S, Peppercorn J, Sampson E, Knudsen K, Laufer M, Rosenberg R, Przypyszny M, Rein A, Winer EP. Web-based survey of fertility issues in young women with breast cancer. J Clin Oncol 2004;22:4174-4183.
4. Thewes B, Meiser B, Taylor A, Phillips KA, Pendlebury S, Capp A, Dalley D, Goldstein D, Baber R, Friedlander ML. Fertility- and menopause-related information needs of younger women with a diagnosis of early breast cancer. J Clin Oncol 2005;23:5155-5165.
5. De Sanjose S, Leone M, Berez V, Izquierdo A, Font R, Brunet JM, Louat T, Vilardell L, Borras J, Viladiu P, Xavier Bosch F, Lenoir GM, Sinilnikova OM. Prevalence of BRCA1 and BRCA2 germline mutations in young breast cancer patients: A population-based study. International Journal of Cancer 2003;106: 588-593.
6. Loman N, Johannsson O, Kristoffersson U, Olsson H, Borg A. Family history of breast and ovarian cancers and BRCA1 and BRCA2 mutations in a population-based series of early-onset breast cancer. Journal of the National Cancer Institute 2001;93:1215-1223.
7. Bloom JR, Stewart SL, Chang S, Banks PJ. Then and now: quality of life of young breast cancer survivors. Psychooncology 2004;13:147-160.
8. Walsh SR, Manuel JC, Avis NE. The impact of breast cancer on younger women's relationships with their partner and children. Families, Systems and Health 2005;23:80-93.
9. Smyth JM, Nazarian D, Arigo D. Expressive Writing in the Clinical Context. In: Vingerhoets AJJM, Nyklícek I, Denollet J, editors. Emotion Regulation. Conceptual and Clinical Issues. New York: Springer; 2008. p 215-33.
10. Pennebaker JW, Seagal JD. Forming a story: the health benefits of narrative. J Clin Psychol 1999;55:1243-1254.
11. Sharf BF, Vanderford ML. Illness Narratives and the Social Construction of Health. In: Thompson TL, Dorsey AM, Miller KI, Parrott R, editors. Handbook of Health Communication. London: LEA Publishers; 2003. p 9-34.
12. Healthtalkonline (formerly DIPEx) [Internet]. Breast Cancer in women [cited 2008 May 11]; Available from: http://www.healthtalkonline.org/Cancer/Breast_Cancer (http://www.webcitation.org/5eAfXVpWN)
13. Ziebland S, McPherson A. Making sense of qualitative data analysis: an introduction with illustrations from DIPEx (personal experiences of health and illness). Med Educ 2006;40: 405-414.
14. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A, Thomson R, Barratt A, Barry M, Bernstein S, Butow P, Clarke A, Entwistle V, Feldman-Stewart D, Holmes-Rovner M, Llewellyn-Thomas H, Moumjid N, Mulley A, Ruland C, Sepucha K, Sykes A, Whelan T. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006;333:417-422.
15. Butow P, Fowler J, Ziebland S. Section E: Using personal Stories. In: O'Connor A, Llewellyn-Thomas H, Stacey D, editors. International Patient Decision Aid Standards (IPDAS) Collaboration Background Document. 2005. p 24-27. Available from: http://ipdas.ohri.ca/IPDAS_Background.pdf

16. Overberg RI, Alpay LL, Verhoef J, Zwetsloot-Schonk JHM. Illness stories on the internet: what do breast cancer patients want at the end of treatment? Psychooncology 2007;16: 937-944.
17. De Amazones, jonge vrouwen met borstkanker/The Amazones, young women with breast cancer [Internet]. Ervaringsverhalen/stories [cited 2007 Jan 2]; Available from: http://www.de-amazones.nl/
18. Rozmovits L, Ziebland S. What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs. Patient Educ Couns 2004;53:57-64.
19. KWF Kankerbestrijding/Dutch Cancer Society [Internet]. Videodagboek/video diary [cited 2008 Jun 4]; Available from: http://www.kankerbestrijding.nl (http://www.webcitation.org/5eAgM1yAc)
20. Molenaar S, Sprangers MA, Rutgers EJ, Luiten EJ, Mulder J, Bossuyt PM, van Everdingen JJE, Oosterveld P, de Haes HCJM. Decision support for patients with early-stage breast cancer: effects of an interactive breast cancer CDROM on treatment decision, satisfaction, and quality of life. J Clin Oncol 2001;19:1676-1687.
21. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001;323:1103-1105.
22. Joinson AN. Disinhibition and the Internet. In: Gackenbach J, editor. Psychology and the Internet: Intrapersonal, Interpersonal, and Transpersonal Implications. 2 ed. Burlington/ San Diego/London: Academic Press/Elsevier; 2007. p 75-92.
23. Van Uden-Kraan CF, Drossaert CHC, Taal E, Smit WM, Bernelot Moens HJ, Siesling S, Seydel ER, van de Laar MAFJ. Health-related Internet use by patients with somatic diseases: frequency of use and characteristics of users. Informatics for Health & Social Care 2009;34:18-29.