

Ligt, K.M. de, Rooij, B.H. de, Walraven, I., Heins, M.J., Verloop, J., Siesling, S., Korevaar, J., Poll-Franse, L.V. van de. Varying severities of symptoms underline the relevance of personalized follow-up care in breast cancer survivors: latent class cluster analyses in a cross-sectional cohort. *Supportive Care in Cancer*: 2022, 30(10), p. 7873-7883

---

**Postprint version** : 2.0

**Journal website** : <https://link.springer.com/article/10.1007/s00520-022-07229-6>

**Pubmed link** : <https://pubmed.ncbi.nlm.nih.gov/35727375/>

**DOI** : 10.1007/s00520-022-07229-6

This is a Nivel certified Post Print, more info at [nivel.nl](https://nivel.nl)

## Varying severities of symptoms underline the relevance of personalized follow-up care in breast cancer survivors: latent class cluster analyses in a cross-sectional cohort

Kelly M. de Ligt<sup>1</sup>, Belle H. de Rooij<sup>2,3</sup>, Iris Walraven<sup>1,4</sup>, Marianne J. Heins<sup>5</sup>, Janneke Verloop<sup>2</sup>, Sabine Siesling<sup>2,6</sup>, Joke C. Korevaar<sup>5</sup>, Lonneke V. van de Poll-Franse<sup>1</sup>

<sup>1</sup> Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, PO Box 90203, 1006 BE Amsterdam, The Netherlands

<sup>2</sup> Department of Research and Development, Netherlands Comprehensive Cancer Organisation, Utrecht, The Netherlands

<sup>3</sup> CoRPS - Center of Research On Psychology in Somatic Diseases/Department of Medical and Clinical Psychology, Tilburg University, Tilburg, The Netherlands

<sup>4</sup> Department for Health Evidence, RadboudUMC, Nijmegen, The Netherlands

<sup>5</sup> Nivel, Netherlands Institute of Health Services Research, Utrecht, Netherlands

<sup>6</sup> Department of Health Technology and Services Research, Technical Medical Centre, University of Twente, Enschede, The Netherlands

### Abstract

**Purpose** Insights into the severity of co-existing symptoms can help in identifying breast cancer survivors in need of symptom management. We aimed to identify subgroups of breast cancer survivors based on patterns of symptom severity, and characteristics associated with these subgroups.

**Methods** We selected surgically treated stage I–III breast cancer survivors 1–5 years post-diagnosis from the Netherlands Cancer Registry (N = 876). We assessed experienced severity of fatigue, nausea, pain, dyspnea, insomnia, appetite, constipation, diarrhea, and emotional and cognitive symptoms through the EORTC-QLQ-C30 Quality of Life Questionnaire on a scale of 0–100. We determined subgroups of survivors using latent class cluster analyses (LCA) based on severity of co-existing symptoms and compared their mean severity to the age-matched female reference population to

interpret clinical relevance. We assessed subgroup characteristics by multinomial logistic regression analyses.

**Results** From 404 respondents (46%), three subgroups of survivors with distinct symptom severity were identified: low severity ( $n = 116$ , 28.7%), intermediate severity ( $n = 224$ , 55.4%), and high severity ( $n = 59$ , 14.6%). The low subgroup reported lower symptom severity than the general population; the intermediate subgroup reported a similar symptom severity, although scores for fatigue, insomnia, and cognitive symptoms were worse (small-medium clinical relevance). The high subgroup had worse symptom severity (medium-large clinical relevance). Compared to the intermediate subgroup, one (RRR: 2.75; CI: 1.22–6.19;  $p = 0.015$ ) or more (RRR: 9.19; CI: 3.70–22.8;  $p = < 0.001$ ) comorbidities were significantly associated with the high subgroup. We found no associated treatment characteristics.

**Conclusion** We identified distinct subgroups of breast cancer survivors based on symptom severity, underlining the relevance of further exploring personalized follow-up strategies.

## Introduction

Long-term cancer survivors “remain at risk for (distant) relapse and can potentially experience late treatment-related sequelae” [1]. Many breast cancer survivors suffer from treatment-induced symptoms. Even up to 5 years after diagnosis, reported symptoms include fatigue, pain, lymph edema, cognitive dysfunction, hot flashes, anxiety, depression, insomnia, sexual issues, fear of recurrence, and neuropathy [2–4]. Although the long-term overall health-related quality of life (HRQoL) in breast cancer survivors is as good as that of the general population [5–10], several studies have reported deteriorated functioning and HRQoL in subgroups of survivors [5–10]. Symptoms such as fatigue, depression, pain, and cognitive dysfunction have been associated with these deteriorations [2, 3]. In order to increase HRQoL, symptom management is one of the goals of follow-up care [11, 12]. Current guidelines state however that there is no standardized, evidence-based regimen for breast cancer follow-up [11, 12], and it should thus be based on a tradeoff between patient needs, follow-up costs, and burden [12]. As a result, a variety of follow-up arrangements exist across the world [13]. Evaluations report that many of these not fully meet the needs of breast cancer survivors [1, 11, 14–17], and describe a growing demand for more personalized care planning [1, 11, 14–17].

Survivors rarely report singular symptoms: even several years after diagnosis, multiple burdening symptoms may remain prevalent [2, 3, 6, 10]. These may “relate, hasten, or potentiate one another or contribute to the development of other symptoms through multifaceted underlying factors” in a symptom cluster [20]. Studying the overall symptom severity and patterns of overall symptom severity is expected to better identify subgroups of survivors in need for symptom management [19, 21] and align follow-up arrangements to these needs.

Yet, as reported by an expert panel in 2015, symptoms have mainly been studied as separate, independent items and symptom cluster research was at that time still in its infancy [18, 19]. More recently, several studies have aimed to identify clusters of pre-defined symptoms in breast cancer populations ranging from 240–1500 survivors, from 6 months to 5 years after diagnosis [22–25]. Across studies, methods differed in terms of measurements and type of cluster analysis, and method in general, some of them assessing a broad range of physical and psychosocial symptoms [22, 25], while others establishing clinical relevance of distinguished clusters [23], or determining factors associated with clusters [23, 24], and assessing symptoms over a longer period of time [22].

We aimed to identify subgroups of survivors with different patterns of symptom severity through latent class cluster analyses (LCA) that can be applied to study patterns of severity, burden, or

magnitude of a pre-defined symptom cluster. As survivors reported symptoms even years after diagnosis, we measured a broad range of physical and psychosocial symptoms most relevant to breast cancer survivors [26] up to 5 years after diagnosis. In addition, we aimed to assess the clinical relevance of symptom severity by comparing our sample to the general population. Last, we aimed to identify patient and treatment characteristics that were associated with the defined subgroups of survivors—thereby, combining prominent methodological aspects described above in one study.

## Methods

Utilized data have been collected through a cross-sectional survey study described previously [6, 27]. We selected a random sample of 1000 female breast cancer survivors who had been diagnosed between 2012 and 2016 with non-metastatic breast cancer from the Netherlands Cancer Registry (NCR), a database with national coverage on cancer incidence, diagnosis, and treatment [28]. Survivors had received surgical treatment in one of 20 hospitals committed to our study. In deliberation with these hospitals, we excluded survivors who did not receive follow-up care, who were currently receiving treatment for secondary or recurrent disease, who could not read or write Dutch, or without recent contact information (n = 124). Between September 2017 and March 2018, we invited 876 survivors to complete our survey. Responses were collected until May 2018—at time of completing the survey, survivors were 1 to 5 years after diagnosis. Surveys were administered through the online PROFILES (“Patient-Reported Outcomes Following Initial treatment and Long-term Evaluation of Survivorship”) Registry survey application [29]. Participants gave consent for processing their coded responses and merging these with their clinical data available in the NCR. The use of NCR data was approved by the NCR Privacy Review Board. Formal ethical approval was not required, as the Dutch Medical Research (Human Subjects) Act did not apply for this study.

To better understand the clinical relevance of reported symptom severity, a normative general population sample was included that had completed the EORTC-QLQ-C30. The sample was representative of the Dutch-speaking population in the Netherlands (n = 1369 women, surveyed in 2017) and was provided by CentERdata [30].

## Cross-sectional survey

The survey included three subjects:

1. Symptom severity over the past 4 weeks, measured through the Dutch version of the EORTC-QLQ-C30 Quality of Life Questionnaire for Cancer [26]. This validated questionnaire includes nine symptom scales (fatigue, nausea, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, financial burden). As recent literature deems that the item financial burden is not an actual symptom [31], we left this item out of further analyses. As psychosocial and cognitive symptoms are not represented in the symptom scales but frequently occur in breast cancer survivors [3, 32], we also included the EORTC-QLQ-C30 emotional and cognitive functioning scales. Answer options range from not at all to very much (4-point Likert scale). Scores are transformed to range from 0 to 100, with high scores depicting a high symptom severity (i.e., low HRQoL);
2. Sociodemographic characteristic (age, highest completed level of education);
3. Disease status (current treatment status, presence of comorbidities). Comorbidities at time of survey were assessed through the adapted Self-Administered Comorbidity Questionnaire (SACQ) [33].

## Analyses

We determined mutually exclusive subgroups of survivors using LCA based on severity of co-existing symptoms. All symptom severity scores were included as ordinal variables. As the EORTC-QLQ-C30

includes aspects most relevant to cancer patients [26], we did not exclude symptoms. LCA is a data-driven approach; multiple models with different numbers of subgroups were built. We assessed the fit of each model based on goodness-of-fit statistics (i.e., lowest values of Log likelihood (LL), and LL based on the Bayesian Information Criterion (BICLL), Akaike's Information Criterion (AICLL), and Consistent Akaike Information Criterion (CAICLL) [34, 35]). LCA assumes that observed variables are mutually independent given the class variable, meaning all within-cluster covariance is equal to zero. To check for local dependencies between included variables, bivariate residuals were assessed (bivariate residuals should be 3 or smaller). However, restrictions on local independency are allowed to obtain more parsimonious and stable models: mutual dependency for certain pairs of variables is then assumed within latent classes. The final model was based on the goodness-of-fit statistics (where lowest values indicate a more favorable model fit [35]) and clinically meaningful differences between subgroups. Survivors were assigned to the subgroup with the highest posterior class-membership probability. LCA was performed with Latent GOLD version 5.2.0 (Statistical Innovations Inc., Belmont, USA) [34].

Differences in subgroup characteristics were examined with  $\chi^2$  tests for categorical variables. Multinomial logistic regression analyses were performed to determine factors that were associated with subgroup membership. Variables that were significant in univariate analyses (level of significance,  $p < 0.10$ ) were included in the regression analyses (level of significance,  $p < 0.05$ ). As literature describes the association between age, comorbidities, and HRQoL in long-term breast cancer survivors [36], these variables were included a priori.

Last, the respondents were matched 1:1 to a female normative general population based on age (categories of 5-year age brackets, ranging from  $\leq 25$  to  $> 85$ ). The general population sample included individuals with a history of cancer (4%; based on self-report); we ran the analyses in the cancer-free general population as part of a sensitivity analysis. Mean scores per cluster were compared to the general population mean scores and tested through ANOVA. Clinically relevant differences were described according to the EORTC-QLQ-C30 guidelines for clinical interpretation by Cocks et al. [37].

All analyses other than LCA were performed in Stata/SE 14.2 [38]. Results From 876 administered surveys, 408 were returned, of which 404 (46%) were sufficiently completed and eligible for analyses. Although the survivors in the youngest and oldest age category were slightly underrepresented, the respondent population was deemed representative (previously reported in De Ligt et al. [6, 27]).

Survivors were diagnosed with stage I (46%) or II (43%) disease. Mean time between diagnosis and survey was 3.4 years (range: 0.8–6 years), and 47% reported no comorbid diseases at time of survey. Survivors had received adjuvant radiotherapy (72%), chemotherapy (49%), or anti-hormonal treatment (57%). At time of survey, the latter was still administered in 43% of survivors (Table 1).

### Subgroups of survivors based on symptom severity

Figure 1 shows the identified subgroups. A three-subgroup model had the best goodness-of-fit based on the lowest BICLL and CAICLL (Table 2). Although the AICLL and AIC3LL were lowest for the 6-subgroup and 5-subgroup models, additional subgroups led to less distinctive (i.e., less clinically meaningful) mean scores per subgroup. After relaxing the local independency assumption between cognitive symptoms and nausea/vomiting, and between insomnia and emotional symptoms, all but two bivariate residuals were around 3 or lower (Model 3\*; bivariate residuals for diarrhea/dyspnea: 4.8; cognitive symptoms/emotional symptoms: 4.2). Except for diarrhea, 95% confidence intervals were not overlapping between clusters (Fig. 1).

The first subgroup of survivors had lower mean scores than the general population, meaning their symptom severity was lower (general population, mean scores, fatigue: 19; nausea/vomiting: 0; pain: 3; dyspnea: 8; insomnia: 18; appetite loss: 3; constipation: 5; diarrhea: 3; cognitive symptoms: 10; emotional symptoms: 13). We named this the low severity subgroup ( $n = 116/404$ , 28.7%; mean

scores, fatigue: 5; nausea/vomiting: 0; pain: 0; dyspnea: 1; insomnia: 0; appetite loss: 0; constipation: 1; diarrhea: 1; cognitive symptoms: 4; emotional symptoms: 2). The second subgroup had mean scores similar to those of the general population, although mean scores for fatigue, insomnia, and cognitive symptoms were higher (i.e., more severe, small-medium clinically relevant difference). We named this the intermediate severity subgroup (n = 224/404, 55.4%; mean scores, fatigue: 28; nausea/vomiting: 2; pain: 20; dyspnea: 12; insomnia: 33; appetite loss: 4; constipation: 9; diarrhea: 4; cognitive symptoms: 21; emotional symptoms: 18). The third subgroup had high mean scores: their symptom severity was the worst. Compared to the general population, we found large clinically relevant differences for fatigue, pain, dyspnea, insomnia, appetite loss, and cognitive symptoms, and medium clinically relevant differences for constipation. Mean emotional symptom scores were 28 points worse than in the general population. We named this subgroup the high severity subgroup (n = 59/404, 14.6%; mean scores, fatigue: 66; nausea/vomiting: 13; pain: 47; dyspnea: 34; insomnia: 52; appetite loss: 29; constipation: 21; diarrhea: 9; cognitive symptoms: 46; emotional symptoms: 43).

### [Table 1]

#### Subgroup characteristics

Survivors in the low severity subgroup relatively often had stage I disease (58%, compared to 42% and 32% for the intermediate and high severity subgroup, respectively;  $p = 0.002$ ) and, consequently, were more often treated with breast-conserving surgery (59%;  $p = 0.041$ ). Survivors in the high severity subgroup relatively often had stage III disease (20%, compared to the intermediate (12%) and low severity subgroup (5%);  $p = 0.002$ ) and more often underwent mastectomy (54%;  $p = 0.041$ ). In the low severity subgroup, significantly lower rates of radiotherapy (66%; compared to 78% and 71% for the intermediate and high severity subgroup;  $p = 0.038$ ) and anti-hormonal therapy (49%; compared to 62% and 66% for the intermediate and high severity subgroup;  $p = 0.034$ ) were found.

Age, socioeconomic status, educational level, and time since diagnosis did not differ among the subgroups. Survivors in the high severity subgroup more often reported two or more comorbidities ( $p < 0.001$ ). More specifically, almost 20% of them reported 3–5 comorbidities (3.5% in the intermediate subgroup, 0% in the low subgroup). The most prevalent comorbidities (diseases/impairments in muscles, connective tissue, or joints (32.6%), lung diseases (10.0%), gastrointestinal diseases including liver, gallbladder, and pancreas (8.2%), cardiovascular diseases (7.1%), and complaints in urinary/reproductive system (7.1%)) were significantly more often reported in the high severity subgroup (Table 3).

### [Figure 1] [Table 2]

#### Factors associated with subgroups

Survivors in the low and high severity subgroup were compared to the intermediate severity subgroup (Table 4). Survivors in the low severity subgroup had significantly less comorbidities (one comorbidity: RRR: 0.35, CI: 0.20–0.62;  $p < 0.001$ , two or more comorbidities: RRR: 0.15, CI: 0.05–0.44;  $p < 0.001$ ), more often had completed high vocational education (RRR: 0.51, CI: 0.27–0.95;  $p = 0.033$ ), and were less often treated with chemotherapy (RRR: 0.52, CI: 0.27–1.00;  $p = 0.050$ ). Survivors in the high severity subgroup had significantly more comorbidities (one comorbidity: RRR: 2.75, CI: 1.22–6.19;  $p = 0.015$ ; two or more comorbidities: RRR: 9.19, CI: 3.70–22.8;  $p < 0.001$ ).

Comorbidities especially associated with high symptom severity were diseases/impairments in muscles, connective tissue, or joints (RRR: 3.30, CI: 1.68–6.49;  $p = 0.001$ ), and diseases/impairments of urinary/reproductive system (RRR: 4.14, CI: 1.44–11.9;  $p = 0.008$ ; results not reported in Table 4).

### [Table 3]

## Discussion

Based on patterns of overall symptom severity, our study identified three subgroups of long-term breast cancer survivors, and compared their mean symptom severity to the general population to determine clinical relevance. The varying severities of symptoms in these groups reflect varying needs that may ask for different follow-up arrangements and possible change in clinical practice. Note that this does not include the early detection of loco-regional recurrences or contralateral breast cancer, which is also goal of follow-up arrangements [11, 12].

Fourteen percent of respondents reported a high, clinically relevant symptom severity, which is in range with 13–26% of survivors in recent studies that in similar fashion determined subgroups in comparable populations [23, 39–41]. We found a significant association between number of comorbid diseases and high symptom severity, confirming the literature [23, 41]. A relatively high proportion of respondents in the high severity group reported over three comorbidities. Our results suggest that survivors who already suffered from comorbid diseases will experience more health limitations after breast cancer treatment. This group, which represents one in seven early-stage breast cancer survivors, seems to be in the highest need of supportive health care. We believe these patients may need follow-up arrangements that are sensitive to their more complex, comorbid health status.

The 55% of survivors in the intermediate severity subgroup reported severity scores comparable to the general population, with more severe and clinically relevant scores for fatigue, insomnia, and cognitive symptoms. Fatigue, insomnia, and cognitive symptoms have been reported as part of symptom clusters, however usually with pain or anxiety [23, 42–44]. Although the literature emphasizes the importance of targeted interventions for clustered symptoms, we found only a few, some with limited effects [42, 43]. Even though more research is needed, we expect follow-up arrangements specifically targeted to these symptoms would be more in place for this subgroup.

Last, 29% of survivors reported almost no symptoms. We believe it would be interesting to evaluate if less frequent or “on-demand” follow-up care would serve these survivors, personalized to their needs. We think it is important that, with an increasing prevalence of breast cancer patients [45, 46], alternative and more personalized ways of follow-up are explored. For instance, Kirshbaum et al. [47] evaluated an open-access on-demand follow-up intervention for early-stage breast cancer survivors, in which they could consult the breast cancer clinic when necessary. In terms of HRQoL, women were not disadvantaged by open-access follow-up compared to standard hospital-based follow-up. Note that this only concerned aftercare or symptom management/support, and not the surveillance for recurrent cancer through physical and mammographic examination.

### [Table 4]

Ideally, healthcare providers would want to know early in the care process which follow-up arrangements would probably serve their patients. To understand which factors are associated with symptom severity, we need additional research that supports the prediction of survivors’ need for symptom management. Interestingly, besides the presence of comorbidities, we did not identify any treatment-related factors that were associated with high symptom severity—as well found by Bjerkeset et al. [40]. Contrastingly, three studies reported that chemotherapy was associated with membership to the subgroup with the worst outcomes, but included patients only up to 2 years after diagnosis [23, 39, 41]. The influence of chemotherapy may be more pronounced on shorter term. Furthermore, three studies reported that younger age was associated with membership to the

subgroup with the worst outcomes [23, 40, 41]. The slight underrepresentation of the oldest and youngest survivors in our cohort may explain why we did not.

Furthermore, in clinical practice, a tool is needed that identifies symptom management needs. Survivors could complete questionnaires and a “distress thermometer” at the start of follow-up, as suggested by Iyer et al. [48]. Patient-reported outcome measures (PROMs) could be implemented to structurally identify patient needs during treatment and follow-up [49]. This may even better serve survivorship needs, as women self-reported significantly more symptoms than were registered by the clinical oncologist during clinical consultations [4].

A limitation of our study is the cross-sectional design in post-treatment patients. By pre-treatment assessment of patient characteristics and HRQoL, confounding factors can be measured more correctly. As we measured comorbidity status simultaneously with post-treatment HRQoL, we cannot rule out that all self-reported comorbidities were separate diseases instead of consequences of breast cancer treatment. Cross-referencing between comorbidities and health problems from our previous report [27] demonstrated that survivors who reported diseases/impairments in muscles, connective tissue, or joints significantly more often reported pain and swelling in the breast area. This suggests comorbidity measures may be clouded by treatment-induced lymphedema. We need systematic patient-reported measuring of comorbidities to better understand and adjust for case-mix in cancer populations [50]. Furthermore, prospective longitudinal measurement would better align with the clinical course of post-treatment symptom experience, since symptom burden may change over time, which may differ between subgroups of patients [39, 41].

Comparing our results with other studies is difficult, as there is no universal working definition and assessment method of symptom clusters [19, 21]. This is illustrated by the different methods used in the studies described above, including Markov modeling [39], latent class profile analysis [23, 41], or determining the proportion of survivors who experienced a pre-defined symptom cluster [40]. Furthermore, symptoms were measured through a variety of questionnaires; consensus about which symptoms to measure in identifying symptom clusters is still lacking [19]. We included all EORTC-QLQ-C30 symptom scales, as this questionnaire covers aspects most relevant to cancer patients [26]. We also included (reversely scored) cognitive and emotional functioning: these are commonly prevalent in breast cancer survivors [3, 32], and especially survivors with psychological symptoms report higher symptom severity [39]. The QLQ-C30 includes more functioning scales, and one could argue that all should have been included. We noted no overlapping confidence intervals for the other functioning scales. Still, without a universally approved gold standard for symptom cluster study methods [19, 21], our study may add to the variation of designs and results reported in literature.

### **Clinical implications**

Survivors may be served better by strategies more personalized than current annual hospital-based follow-up. We suggested intensive follow-up sensitive to the more complex, comorbid health status of the high severity subgroup. More research is needed for clustered interventions that target a selection of symptoms that are more severe than in the general population, such as fatigue, insomnia, and cognitive symptoms reported in the intermediate severity subgroup. Furthermore, we need additional research that supports the prediction of survivors’ need for follow-up care and structural prospective and longitudinal assessment by PROMs to measure actual patient needs. Our study did not result in cutoff scores or thresholds, but the means and confidence intervals around these means for each cluster. These can guide future classification of survivors in these subgroups, but are not meant to be used directly as clinical cutoffs. We can however imagine that the availability of cutoff scores would enable the stratification of patients into one of the clusters when assessing symptom scores in clinical practice/during clinical encounters in the future. Only then, alternative follow-up strategies can be set up and evaluated for future clinical implementation.

To conclude, we identified three subgroups of breast cancer survivors based on symptom severity. Our results underline the relevance of further exploring follow-up alternatives suitable for these subgroups. We found that reporting comorbid diseases was associated with higher symptom severity. Yet, treatment factors and time between treatment and survey were not. Future research should longitudinally measure symptoms that are most important for breast cancer patients, including baseline assessment of patient characteristics such as comorbid diseases. This will be useful in clinical practice as well as in future research for determining which survivors require symptom management and through which follow-up strategy.

### **Acknowledgements**

The authors thank all women who participated in completing the survey.

### **Author contribution**

Conceptualization: K.M. de Ligt, B.H. de Rooij, I. Walraven, M.J. Heins, J. Verloop, S. Siesling, J.C. Korevaar, L.V. van de Poll-Franse.

Methodology: K.M. de Ligt, B.H. de Rooij, I. Walraven, M.J. Heins, J. Verloop, S. Siesling, J.C. Korevaar, L.V. van de Poll-Franse.

Formal analyses and investigation: K.M. de Ligt, B.H. de Rooij.

Writing – original draft preparation: K.M. de Ligt.

Writing – review and editing: B.H. de Rooij, I. Walraven, M.J. Heins, J. Verloop, S. Siesling, J.C. Korevaar, L.V. van de Poll-Franse.

All authors have read and approved the final version of the manuscript to be published. The authors ensure that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

### **Data availability**

The data that support the findings of this study are available from the Netherlands Cancer Registry but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the Netherlands Cancer Registry.

### **Code availability**

Codes written in Latent GOLD version 5.2.0; Stata/SE 14.2. Codes are available at the corresponding author upon reasonable request.

### **Declarations**

#### **Ethics approval**

The use of data from the Netherlands Cancer Registry (NCR) was approved by the NCR Privacy Review Board. Formal ethical approval was not required, as the Dutch Medical Research (Human Subjects) Act did not apply for this study.

#### **Consent to participate**

Informed consent was obtained from all individual participants included in the study. Participants gave consent for processing their coded responses and merging these with their clinical data available in the NCR.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

## References

1. Surbone A, Tralongo P (2016) Categorization of cancer survivors: why we need it. *J Clin Oncol* 34(28):3372–3374
2. Wu HS, Harden JK (2015) Symptom burden and quality of life in survivorship: a review of the literature. *Cancer Nurs* 38(1):E29–E54
3. Schmidt ME, Wiskemann J, Steindorf K (2018) Quality of life, problems, and needs of disease-free breast cancer survivors 5 years after diagnosis. *Qual Life Res* 27:2077–2086
4. Bjørklund Ellegaard MB et al (2017) Women with breast cancer report substantially more disease and treatment-related side or late effects than registered by clinical oncologists: a cross-sectional study of a standard follow-up program in an oncological department. *Breast Cancer Res Treat* 164:164
5. Arndt V et al (2017) Quality of life in long-term and very long-term cancer survivors versus population controls in Germany. *Acta Oncol* 56(2):190–197
6. de Ligt KM et al (2019) The impact of health symptoms on health-related quality of life in early-stage breast cancer survivors. *Breast Cancer Res Treat* 178:703–711
7. Doege D et al (2019) Health-related quality of life in long-term disease-free breast cancer survivors versus female population controls in Germany. *Breast Cancer Res Treat* 175:499–510
8. Mols F et al (2005) Quality of life among long-term breast cancer survivors: a systematic review. *Eur J Cancer* 41:2613–2619
9. Peuckmann V et al (2006) Health-related quality of life in long-term breast cancer survivors: nationwide survey in Denmark. *Breast Cancer Res Treat* 104:39–46
10. Koch L et al (2013) Quality of life in long-term breast cancer survivors – a 10-year longitudinal population-based study. *Acta Oncol* 52(6):1119–1128
11. Runowicz CD et al (2016) American Cancer Society/American Society of clinical oncology breast cancer survivorship care guideline. *CA Cancer Soc J Clin Oncol* 66:43–73
12. Cardoso F et al (2019) Early breast cancer: esmo clinical practice guidelines. *Ann Oncol* 30:1194–1220
13. Cruickshank S, Barber M (2019) Breast cancer follow-up after a primary diagnosis: a confused picture. *The Breast* 46:97–100
14. Zorginstituut Nederland (2016) Verbetersignalement Zinnige nacontrole bij vrouwen behandeld voor borstkanker. Diemen. Accessed at: <https://www.zorginstituutnederland.nl/binaries/zinl/documenten/rapport/2016/10/31/zinnige-zorg-verbetersignalement-zinnigenacontrole-bij-vrouwen-behandeld-voor-borstkanker/Verbetersignalement+Zinnige+nacontrole+bij+vrouwen-behandeld+voor+borstkanker.pdf>
15. Fallowfield L, Jenkins V (2014) Psychosocial/survivorship issues in breast cancer: are we doing better? *J Natl Cancer Inst* 107(1):335. <https://doi.org/10.1093/jnci/dju335>
16. Ruban PU et al (2018) Patient evaluation of breast cancer follow-up: a Danish survey. *Patient Educ Couns* 101(1):99–104
17. Mayer DK, Nasso SF, Earp JA (2017) Cancer survivorship in the USA: defining cancer survivors, their needs, and perspectives on survivorship health care in the USA. *Lancet Oncol* 18:e11–18
18. Miaskowski C et al (2007) Conceptual issues in symptom clusters research and their implications for quality-of-life assessment in patients with cancer. *J Natl Cancer Inst Monogr* 37:39–46

19. Miaskowski C et al (2017) Advancing symptom science through symptom cluster research: expert panel proceedings and recommendations. *J Natl Cancer Inst* 109(4):dgw253
20. Brant JM, Beck S, Miaskowski C (2010) Building dynamic models and theories to advance the science of symptom management research. *J Adv Nurs* 66(1):228–240
21. Nguyen J et al (2011) A literature review of symptom clusters in patients with breast cancer. *Expert Rev Pharmacoecon Outcomes Res* 11(5):533–539
22. Nho JH et al (2018) Symptom clusters and quality of life in breast cancer survivors after cancer treatment in a tertiary hospital in Korea. *Eur J Cancer Care (Engl)* 27(6):e12919
23. Lee LJ et al (2020) Symptom clusters in breast cancer survivors: a latent class profile analysis. *Oncol Nurs Forum* 47(1):89–100
24. St Fleur RG et al (2021) A latent profile analysis to assess physical, cognitive and emotional symptom clusters in women with breast cancer. *Psychol Health* 1–17. <https://doi.org/10.1080/08870446.2021.1941960>
25. Berger AM et al (2020) Symptom clusters and quality of life over 1 year in breast cancer patients receiving adjuvant chemotherapy. *Asia Pac J Oncol Nurs* 7(2):134–140
26. Aaronson NK et al (1993) The European Organisation for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 85:365–376
27. de Ligt KM et al (2019) Patient-reported health problems and healthcare use after treatment for early-stage breast cancer. *The Breast* 46:4–11
28. Netherlands Cancer Registry, hosted by Netherlands Comprehensive Cancer Organisation (IKNL) Editor: Utrecht, The Netherlands
29. van de Poll-Franse LV et al (2011) The Patient Reported Outcomes Following Initial treatment and Long term Evaluation of Survivorship registry: scope, rationale and design of an infrastructure for the study of physical and psychosocial outcomes in cancer survivorship cohorts. *Eur J Cancer* 47(14):2188–2194
30. van de Poll-Franse LV et al (2011) Normative data for the EORTC QLQ-C30 and EORTC-sexuality items in the general Dutch population. *Eur J Cancer* 47:667–675
31. Giesinger JM et al (2016) Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust. *J Clin Epidemiol* 69:79–88
32. Kenyon M, Mayer DK, Owens AK (2014) Late and long-term effects of breast cancer treatment and surveillance management for the general practitioner. *J Obstet Gynecol Neonatal Nurs* 43:382–398
33. Sangha O et al (2003) The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. *Arthritis Rheum* 49(2):156–163
34. Vermunt JK (2005) *Latent GOLD 4.0 User's Guide*. Statistical Innovations Inc, Belmont
35. Vermunt JK, Magidson J (2002) Latent class cluster analysis. In: Hageaars J, McCutcheon A (eds) *Applied latent class analysis*. Cambridge University Press, Cambridge, pp 89–106 *Supportive Care in Cancer* 13
36. Chu W et al (2016) Determinants of quality of life among long-term breast cancer survivors. *Qual Life Res* 25(8):1981–1990. <https://doi.org/10.1007/s11136-016-1248-z>
37. Cocks K et al (2011) Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *J Clin Oncol* 29:89–96
38. StataCorp, (2015) *Stata Statistical Software: Release 14*. Stata-Corp LP, College Station
39. Avis NE et al (2017) Longitudinal examination of symptom profiles among breast cancer survivors. *J Pain Symptom Manage* 53(4):703–710

40. Bjerkeset E, Röhrl K, Schou-Bredal I (2020) Symptom cluster of pain, fatigue, and psychological distress in breast cancer survivors: prevalence and characteristics. *Breast Cancer Res Treat* 180:63–71
41. Pinheiro LC et al (2017) Examining health-related quality of life patterns in women with breast cancer. *Qual Life Res* 26:1733–1743
42. Kwekkeboom KL (2016) Cancer symptom cluster management. *Semin Oncol Nurs* 32(4):373–382
43. Barsevick AM (2007) The elusive concept of the symptom cluster. *Oncol Nurs Forum* 34(5):971–980. <https://doi.org/10.1188/07.ONF.971-980>
44. Bender CM et al (2005) Symptom clusters in breast cancer across 3 phases of the disease. *Cancer Nurs* 28(3)
45. Allemani C et al (2018) Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* 391(10125):1023–1075
46. Sung H et al (2021) Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 71(3):209–249
47. Kirshbaum MN et al (2017) Open access follow-up care for early breast cancer: a randomised controlled quality of life analysis. *Eur J Cancer Care* 26:e12577
48. Iyer R, Ring A (2017) Breast cancer survivorship: key issues and priorities of care. *Br J Gen Pract* 67:140–141
49. Greenhalgh J (2009) The applications of PROs in clinical practice: what are they, do they work, and why? *Qual Life Res* 18:115–123
50. Oemrawsingh A et al (2020) Patient-reported morbidity instruments: a systematic review. *Value in Health* 23(6):791–811

**Plain English language summary.**

Many breast cancer survivors suffer from symptoms that may have been caused by the cancer and treatment. To date, these symptoms have been researched as separate symptoms. However, cancer survivors often suffer from multiple symptoms at the same time. We aimed to identify groups of breast cancer survivors based on their overall symptom burden. We found three subgroups of survivors, with either lower, comparable, or higher symptom burden than the general population. We also investigated which characteristics were associated with these subgroups. We found that patients with other diseases besides breast cancer had an increased risk for high symptom burden

## Tables and figures

*Table Respondent characteristics (n = 404)*

	<i>N (404)</i>	<i>%</i>
<i>Tumor characteristics</i>		
Stage		
I	183	45
II	176	44
III	45	11
Hormone-receptor status <sup>x</sup>		
Hormone-receptor-positive	284	70
Hormone-receptor-mixed	55	14
Hormone-receptor-negative	63	16
Tumor grade <sup>x</sup>		
1	95	24
2	176	44
3	97	24
Missing	36	9
<i>Treatment characteristics</i>		
Treatment status at time of survey*		
Treatment completed	180	45
Currently receiving anti-hormonal therapy	173	43
Currently receiving other treatments	29	7
Status unknown	22	5
Surgical treatment		
Breast-conserving surgery	241	60
Mastectomy	163	40
Axillary dissection	86	21
Immediate breast reconstruction	39	10
Adjuvant treatment		
Radiotherapy	297	74
Chemotherapy	204	50
With trastuzumab	50	12
Anti-hormonal therapy	238	60

Table 1 (continued)

	<i>N</i> (404)	%
<i>Patient characteristics</i>		
Age (in years) at time of survey <sup>x</sup>		
Mean (SD; range)	61.5 (11.0; 27–91)	
< 50	56	14
50–59	108	27
60–69	133	33
70+	96	24
Time (in years) between diagnosis and survey <sup>x</sup>		
Mean (SD, range)	3.4 (1.4; 0.8–6.0)	
< 2	82	20
2–3	92	23
3–4	82	20
4–5	88	22
> 5	57	14
Highest completed level of education <sup>x,*</sup>		
Secondary education or lower	122	30
Medium vocational training	170	42
High vocational training	108	27
Number of comorbidities <sup>x,\$</sup>		
0	188	47
1	131	32
≥ 2	61	15
missing	24	6

Figure 1 Mean symptom severity per patient cluster versus reference population, including 95% confidence intervals. CI, confidence interval

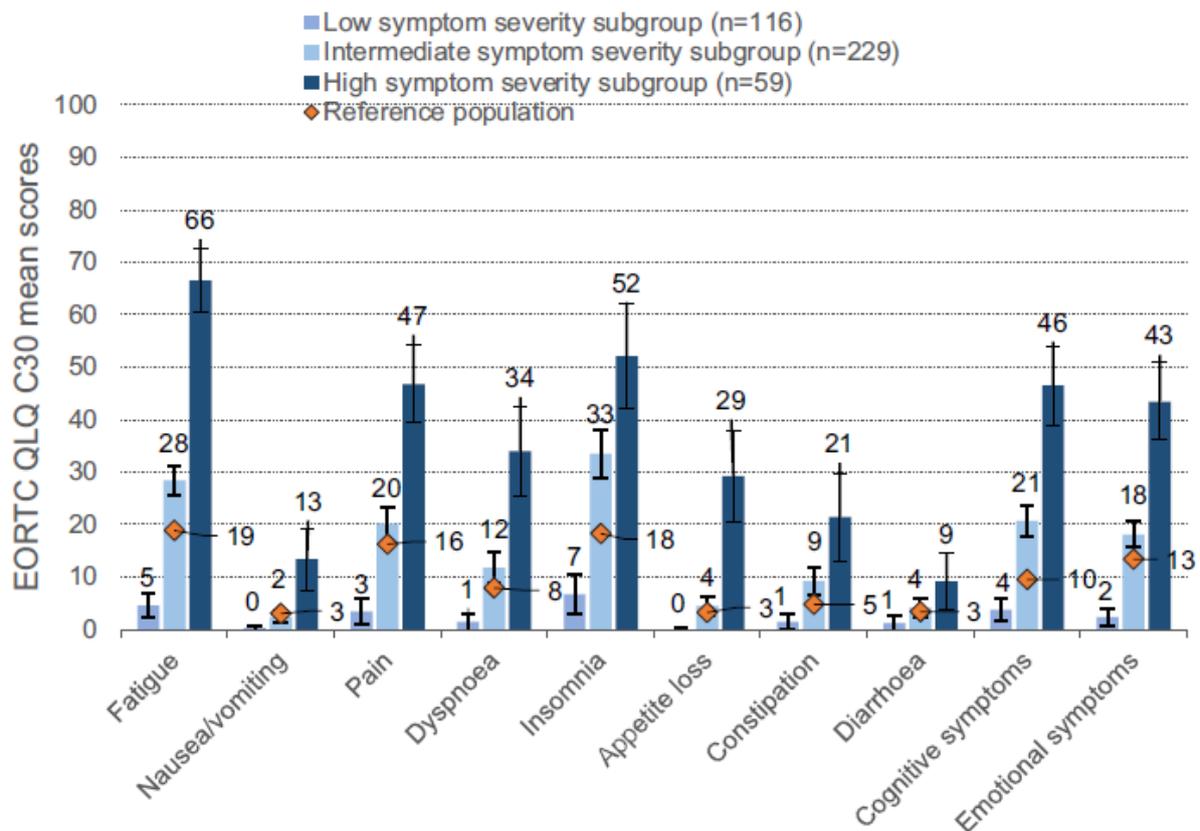


Table 2 LCA model fit statistics

Model	No. of clusters	LL	BIC (LL)	AIC (LL)	(AIC3)LL	CAIC (LL)	Npar	L <sup>2</sup>	df	p value	Class. Err
1	1	4275.0784	8886.2359	8662.1567	8718.1567	8942.2359	56	4351.8431	348	7.8e-682	0.0000
2	2	-3993.7392	8389.5732	8121.4784	8188.4784	8456.5732	67	3789.1649	337	8.5e-576	0.0652
3	3	-3897.5041	<b>8263.1186</b>	7951.0082	8029.0082	<b>8341.1186</b>	78	3596.6946	326	1.7e-543	0.0872
4	4	-3866.0878	8266.3015	7910.1756	7999.1756	8355.3015	89	3533.8620	315	7.8e-537	0.1395
5	5	-3848.4146	8296.9707	7896.8292	<b>7996.8292</b>	8396.9707	100	3498.5157	304	1.2e-535	0.1440
6	6	<b>-3837.1836</b>	8340.5243	<b>7896.3673</b>	8007.3673	8451.5243	111	3476.0537	293	4.6e-537	0.1709
3*	3	-3884.9757	<b>8250.0646</b>	7929.9514	8009.9514	<b>8330.0646</b>	80	3571.6378	324	1.4e-539	0.0814

Description: Model with the lowest model statistic values was deemed the model with the best fit. Lowest values are highlighted in bold

Legend:

\*Model with relaxed local independency between variables reversed “cognitive functioning” and “nausea/vomiting,” and between “insomnia” and reversed “emotional functioning”

Abbreviations: LL, Log likelihood; BIC (LL), Bayesian Information Criterion (based on LL); AIC (LL), Akaike’s Information Criterion (based on LL); (AIC3)LL, modified AIC (based on LL); CAIC (LL), consistent Akaike Information Criterion (based on LL); Npar, number of parameters; L<sup>2</sup>, L-squared; df, degrees of freedom; Class. Err., classification errors

**Table 3 Patient characteristics per cluster of symptom severity**

	Low severity		Intermediate severity		High severity		p*	Total	
	n=116	%	n=229	%	n=59	%		N=404	%
<b>Tumor characteristics</b>									
<b>Stage</b>							<b>0.002</b>		
I	67	58	97	42	19	32		183	45
II	43	37	105	46	28	47		176	44
III	6	5	27	12	12	20		45	11
<b>Hormone-receptor status<sup>x</sup></b>							0.881		
Hormone-receptor-positive	79	69	160	70	45	76		284	70
Hormone-receptor-mixed	16	14	33	14	6	10		55	14
Hormone-receptor-negative	19	17	36	16	8	14		63	16
<b>Tumor grade<sup>x</sup></b>							0.647		
1	21	20	62	30	12	22		95	24
2	55	51	96	46	25	46		176	44
3	31	29	49	24	17	31		97	24
<b>Treatment characteristics</b>									
<b>Treatment status at time of survey</b>							0.062		
Treatment completed	62	56	98	45	20	36		180	45
Currently receiving anti-hormonal therapy	44	40	101	47	28	51		173	43
Currently receiving other treatments	4	4	18	8	7	13		29	7
<b>Surgery<sup>x</sup></b>							<b>0.041</b>		
Breast-conserving surgery	68	59	146	64	27	46		241	60
Mastectomy	48	41	83	36	32	54		163	40
Axillary dissection	18	16	50	22	18	31	0.069	86	21
Immediate breast reconstruction	16	14	20	9	3	5	0.141	39	10
<b>Adjuvant treatment:</b>									
Radiotherapy	76	66	179	78	42	71	<b>0.038</b>	297	74
Chemotherapy	49	42	120	52	35	59	0.069	204	50
Anti-hormonal therapy	57	49	142	62	39	66	<b>0.034</b>	238	60
<b>Survival characteristics</b>									
Died between survey and present day (July 2020)	3	3	7	3	2	3	0.951	12	3
<b>Patient characteristics</b>									
<b>Age (in years) at time of diagnosis (mean, SD)</b>	59.07	11.0	57.8	10.5	59.0	12.1	0.543	58.4	10.9
<b>Age (in years) at time of diagnosis</b>							0.349		
< 50	22	19	47	21	16	27		85	21
50–59	37	32	78	34	13	22		128	32
60–69	40	34	70	31	16	27		126	31
70+	17	15	34	15	14	24		65	16
<b>Time (in years) between diagnosis and survey</b>							0.635		
< 2	20	17	48	21	14	24		82	20
2–3	30	26	51	22	11	19		92	23
3–4	18	16	51	22	13	22		82	20
4–5	27	23	50	22	11	19		88	22
> 5	20	17	27	12	10	17		57	14
<b>Highest completed level of education<sup>x,**</sup></b>							0.332		
Secondary education or lower	39	34	67	30	16	27		122	30
Medium vocational training	54	47	90	40	26	44		170	42
High vocational training	23	20	68	30	17	29		108	27

Table 3 (continued)

	Low severity		Intermediate severity		High severity		<i>p</i> *	Total	
	<i>n</i> = 116	%	<i>n</i> = 229	%	<i>n</i> = 59	%		<i>N</i> = 404	%
<b>Number of comorbidities<sup>x,**</sup></b>	<b>&lt; 0.001</b>								
0	75	70	102	47	11	19		188	47
1	27	25	83	38	21	37		131	32
≥ 2	5	5	31	14	25	44		61	15
<b>Type of comorbidities***</b>									
Lung diseases	3	3	23	11	12	21	<b>0.001</b>	38	9
Cardiovascular diseases	2	2	17	8	8	14	<b>0.012</b>	27	7
gastrointestinal diseases including liver, gallbladder, and pancreas	7	7	14	6	10	18	<b>0.019</b>	31	8
Complaints in urinary/reproductive system	3	3	12	6	12	21	<b>&lt; 0.001</b>	27	7
Diseases or impairments in muscles, connective tissue, or joints	14	13	73	34	37	65	<b>&lt; 0.001</b>	124	31

Table 4 Factors associated with symptom severity following multivariate multinomial logistic regression (relative risk ratios)

	Category	Low symptom severity			Intermediate symptom severity	High symptom severity		
		RRR	95% CI	<i>p</i>		Ref	RRR	95% CI
<b>Treatment characteristics</b>								
Surgery + radiotherapy	BCS + RT	0.59	0.33–1.05	0.075	Ref	0.67	0.32–1.41	0.289
	AMP	Ref			Ref	Ref		
	AMP + RT	0.49	0.20–1.20	0.119	Ref	1.06	0.42–2.69	0.900
Axillary dissection	No	Ref			Ref	Ref		
	Yes	-	-	-	Ref	-	-	-
Chemotherapy	No	Ref			Ref	Ref		
	Yes	0.52	0.27–1.00	0.050	Ref	1.41	0.64–3.10	0.397
	Yes + targeted therapy	1.01	0.44–2.34	0.973	Ref	1.71	0.61–4.78	0.310
Anti-hormonal therapy	No	Ref			Ref	Ref		
	Yes, finished	0.58	0.27–1.26	0.172	Ref	1.16	0.47–2.89	0.748
	Yes, not yet finished	0.72	0.40–1.26	0.249	Ref	1.23	0.60–2.53	0.568
Breast reconstruction	No	Ref			Ref	Ref		
	Yes	-	-	-	Ref	-	-	-
<b>Patient characteristics</b>								
Age <sup>x</sup>	(Continuous)	1.01	0.99–1.04	0.361	Ref	1.00	0.97–1.03	0.934
Education	Secondary school	0.92	0.49–1.74	0.800	Ref	0.65	0.29–1.48	0.308
	Medium vocational training	Ref			Ref	Ref		
	High vocational training	0.51	0.27–0.95	0.033	Ref	1.05	0.50–2.22	0.896
Comorbid diseases*	None	Ref			Ref	Ref		
	One	0.35	0.20–0.62	< 0.001	Ref	2.75	1.22–6.19	0.015
	Two or more	0.15	0.05–0.44	< 0.001	Ref	9.19	3.70–22.8	< 0.001
Time since diagnosis (in years)	Continuous	Ref	-	-	Ref	Ref	-	-

Description: Factors associated with symptom severity following multivariate multinomial logistic regression, presented as relative risk ratios. Statistically significant associated factors ( $p \leq 0.05$ ) are italicized

Legend:

-Not significant in univariate analyses

<sup>x</sup>At time of diagnosis

\*Patient-reported through adjusted Self-Administered Comorbidity Questionnaire at time of survey

Abbreviations: Ref, reference category; RRR, relative risk ratio; BCS, Breast conserving surgery; AMP, amputation; RT, radiotherapy