Data Mining to Reveal Factors Associated with Quality of life among Jordanian Women with Breast Cancer

Najah Al-Shanableh[1], Mazen Al-Zyoud[1], Raya Yousef Al-Husban[2], Noha Al-Shdayfat[3], Ja'far M. Alkhawaldeh[4], Nancy S. Alajarmeh[5], and Sulieman Ibraheem Shelash Al-Hawary[6]

[1]Department of Computer Science, Faculty Of Prince Al-Hussein Bin Abdallah II for IT, Al al-Bayt University, 25113 Mafraq, Jordan
[2]Faculty of Nursing, Zarqa University, 13110 Zarqa, Jordan
[3]Department of Community and Mental Health Nursing, Princess Salma Faculty Of Nursing, Al al-Bayt University, 25113 Mafraq, Jordan
[4]Faculty Of Nursing, Isra University, 11622 Amman, Jordan
[5]Department of Computer Science, College of Information Technology and Communication, Tafila Technical University, 66110 Tafila, Jordan
[6]Department of Business, School of Business, Al al-Bayt University, 25113 Mafraq, Jordan

Abstract: Globally, breast cancer is a major type of cancer among women. This study focused on identifying factors associated with Quality of life among Jordanian women undergoing breast cancer treatment using data mining. Data mining was used to reveal variables by employing several algorithms and the Support Vector Machine to a secondary data set. The results indicated that the factors most strongly related to the physical well-being were pain and fatigue, in the psychological well-being were fear of metastasis, recurrence, and fear from future diagnostic tests, in the social well-being were cancer effect on the women's relationship, and household activity, and in the spiritual well-being were feeling of purpose in life and feeling hopeful. The study signaled the importance of using data mining in health-related studies, and it showed that the dimensions of quality of life of breast cancer women are interrelated and cannot be isolated. Therefore, providing of effective individualized holistic care approach for these women provided by a multidisciplinary team may ease the disease trajectory and improve the quality of life of women with breast cancer

Keywords: Breast cancer, Quality of life, Data mining, Support Vector Machine

1 Introduction

Breast cancer is one of the most prevalent cancers among women worldwide [1,2]. However, recent improvements in early detection strategies and treatment modalities have led to high survival rates [3]. This fact necessitates the need to identify factors that influence the Quality of life (QoL) of women with breast cancer to improve their quality of life. Quality of life is defined by Cella [4] as "patients' appraisal of their current level of functioning compared to what they perceive to be possible or ideal". The dimensions of QOL studied regarding cancer patients tend to be physical, emotional, social, and spiritual.

Factors influencing the QoL has been investigated in many previous studies [3,5,6,7]. Among these depressive disorders, pain, self-blame, and lack of hormonal treatment were related to low QoL among breast cancer survivors [3]. The Pain was also an influential factor in [6] study of 300 women with breast cancer, and they further found poor QoL was associated with paresthesia, chronic pain, and phantom feelings. Chronic pain was influenced by the type of breast cancer treatment they had received, such as mastectomy and radiotherapy, and their stage at diagnosis, age, and time since diagnosis [6]. Body image, disease site, and time since diagnosis were associated with QoL among women diagnosed with metastasis breast cancer [5].

Moreover, the side effects of the disease and its treatment tend to be related to low QOL [8]. One study on Chinese women undergoing breast cancer treatments stressed the importance of managing financial difficulties, fatigue, and treatment side effects because those factors

* Corresponding author e-mail: alzyoud_mazen@aabu.edu.jo

explained about 28–53% of the variance in functional well-being and global QOL [9]. Similarly, Wöckel [10] prospective study found that poor QOL persisted in patients who underwent chemotherapy treatment, mastectomy, or felt extreme fear 8. On the other hand, high educational attainment and high income positively related to QOL in Indonesia. In sum, previous studies have found that physical and psychological states influenced QOL [6]. However, there are cultural and country-level differences.

In 2010, 978 Jordanian women were identified as breast cancer victims, which was a crude incidence rate of 31.8 per 100,000 [11,12] .Still, little is known about the factors that influence the QOL using data mining. Hence, this study will add to the state of knowledge the use of data mining. Data mining for nursing practice has recently expanded to encompass a broad range of healthcare problems and clinical domains, including cancer, diabetes, medication errors, cost estimations, and healthcare applications, among others [13,14]. Data mining has specifically been applied to diagnoses, predicting risk, identifying treatment patterns, and estimating costs [15,16]. Data-mining techniques have been found to be effective and efficient ways to use and manage knowledge [14,17]. All clinical specialties could benefit from data mining [18,19]. Therefore, understanding the factors that influence the Quality of life (QOL) of breast cancer survivors is essential to enhance their Quality of life and decrease their suffering. This study aim is to identify factors affecting the QoL of Jordanian women undergoing breast cancer treatment using data mining techniques.

2 Methodology

This study is a secondary analysis of the previous work [20,21]. Researchers used the existing data that was collected from a convenience sample of 200 women undergoing breast cancer treatment. Therefore the detailed data collection procedure and methodology were explained in the previous work.

2.1 Data Mining

This paper utilized the data mining technique. Data mining is a way to extract information and finding relationships among variables to help develop statistical models [22,23]. Data mining also aids data reduction, description, and model building for revealing and predicting trends [24,25]. Data mining, sometimes referred to as knowledge discovery, extracts important patterns from huge datasets [22,26]. The knowledge discovery in databases (KDD) process aims to find information and build knowledge from massive datasets using mining algorithms.

Data mining's healthcare applications tend to focus on detecting anomalies outside the usual patterns and trends, but it could be used to discover the most important factors related to any disease [15,27]. However, healthcare data tend to be incomplete, incorrect, and stored in nonstandard formats, which considerably impede successful data mining [14]. Nevertheless, several new data preprocessing techniques have emerged to solve these problems and improve results before applying data-mining algorithms.

Several studies have evaluated data mining to explain breast cancer patterns, predictors, and risks using existing datasets [28,29]. Data mining methods also have been used to investigate origins, related factors, and the general behavior of the disease [30,31]. Many predictive algorithms have successfully identified predictive factors for breast cancer. Regarding QOL research, data mining was successful for identifying influences on the health-related QOL of older people with chronic illnesses [32]. Data-mining models mostly outperformed conventional variable selection approaches in the previous literature [33,34,35]. The accuracy of several classifiers was also evaluated on preexisting healthcare datasets. The data-mining methods with the best performance were achieved using decision trees, random forests, and support vector machine methods [33,34,36].

2.2 Design

This study used classification algorithms as the primary data mining task in KDD to identify the factors related to QOL for breast cancer survivors in the sample. The KDD (see Figure 1) takes the following necessary steps to guarantee high performance. (a) Data collection or selection from a previously collected set of data [20];(b) Data pre-processing, including cleaning and integration; (c) Data transformation for building the data-mining model; (d) Data mining using classification or clustering; (e) Pattern evaluation; (f) Knowledge presentation.

2.3 Data Selection

This study analyzed a dataset compiled by King Hussein Medical Center in Amman, Jordan, between June 6, 2017, and October 20, 2017, on women diagnosed with breast cancer and under treatment who had no other physical or mental illnesses. The Quality of life Instrument – Breast Cancer Patient Version (QOL-BC) questionnaire developed by [37] was used to collect the data. Its 46 items comprise four sections focusing on physical, psychological, social, and spiritual QOL [20]. The response options were on a scale of zero = highest QOL to 10 = lowest QOL on a visual analogue scale.
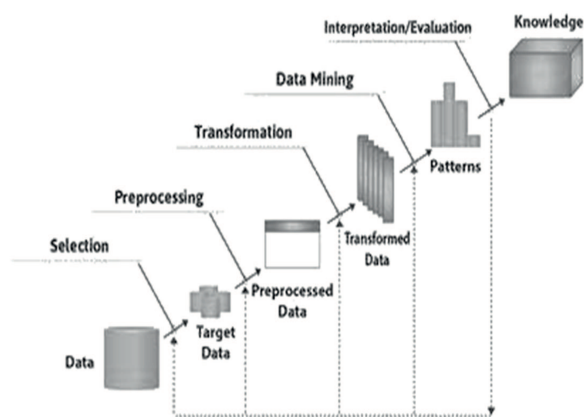
Fig. 1: knowledge discovery in databases (KDD) process

3 Data Analysis

The RapidMiner data-mining tool was used for all of the data cleaning, analysis, and testing of the data-mining model.

3.1 Descriptive Statistics

The available data set was related to a previous study 14 on a sample comprised of 200 women with a mean age of 54 years (SD= 9.14) and about two-thirds of the sample was between ages 41 and 60 years (68%). The distributions of clinical stages of breast cancer are shown in Figure 2. Stage 2 was the most frequent stage (41.5%, n= 83), followed by Stage 3 (31%, n= 62).
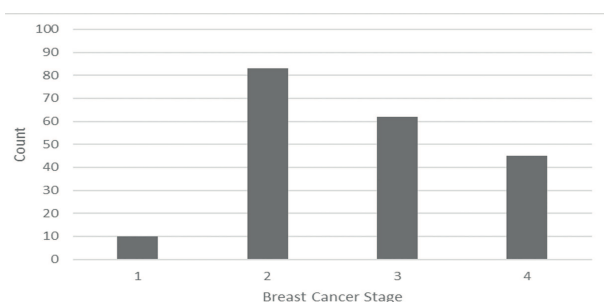

Fig. 2: Distributions of the clinical stages of the participants

3.2 Data Mining Results

The data-mining models were used to identify variables that predict the QOL among women with breast cancer.

Nine classification algorithms were used for outcome prediction, which could be identified through related variables, and Figure 3 lists the accurate classification algorithms used. Based on the accuracy of the models, the Support Vector Machine had the best overall results (80.9% accuracy rate), probably because the dataset was composed of all items.
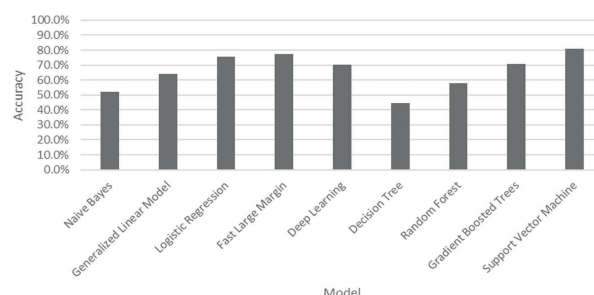

Fig. 3: Data mining models' accuracy

The items that predict the physical well-being were pain, fatigue, appetite and sleep change, and appearance and weight change. In addition, some items from the psychological domain "feeling useful in life, coping with breast cancer treatment, and fear of the results of future diagnostic tests" were strong predictors of the physical well-being (see Table 1).

Table 1: Predictors of physical well-being (N= 200)

Predictor	Weight	Domain
1. Pain	0.7	physical
2. Fatigue	0.6	physical
3. Appetite change	0.6	physical
4. Overall physical health	0.6	physical
5. Weight change	0.6	physical
6. Sleep change	0.6	physical
7. Change in appearance	0.6	physical
8. Menstrual changes or fertility	0.5	physical
9. Arm swelling	0.5	physical
* A higher weight indicates a strong prediction		

The items that predict the **psychological** well-being were fear of metastasis, recurrence, and future diagnostic tests from the psychological domain. Additionally, the "effect of cancer on household activity" from the social domain, and "pain" from the physical domain were associated with **psychological** well-being (see Table 1).

Table 2: Predictors of psychological well-being (N=200)

Predictor	Weight	Domain
1. Fear of cancer metastasis	0.7	Psychological
2. Fear of cancer recurrence	0.7	Psychological
3. Fear of future diagnostic tests	0.7	Psychological
4. Coping with breast cancer	0.7	Psychological
5. Coping with breast cancer treatment	0.7	Psychological
6. Presence of anxiety	0.6	Psychological
7. Effects of cancer on household activity	0.6	Social
8. Change in appearance	0.6	Psychological
9. Over all physical well-being	0.6	physical
10. Feeling of change in self-concept	0.6	Psychological
11. pain	0.6	physical
* A higher weight indicates a strong prediction		

Items from the social domain that predict the **social** well-being were "the effect of cancer on the women's relationship, household activity, and the financial burden. Besides, items that predict the social well-being from the psychological domain were feeling depressed, and fear from cancer recurrence and metastasis (see Table 4).

Table 3: Predictors of social well-being (N= 200)

Predictor	Weight	Domain
1. Effect of cancer on social relationships	0.7	social
2. Amount of financial burden from cancer	0.6	social
3. Effects of cancer on household activity	0.6	social
4. Feeling of isolation	0.6	social
5. Feeling of depression	0.5	Psychological
6. Distress from radiotherapy	0.5	Psychological
7. Fear of cancer metastasis	0.5	Psychological
8. Fear of cancer recurrence	0.5	Psychological
9. Effect of cancer on sexuality	0.5	Sexual
* A higher weight indicates a strong prediction		

The items that predict the **spiritual** well-being were feeling of purpose in life, and feeling hopeful. The Predictors of spiritual well-being are listed in Table 4.

Table 4: Predictors of spiritual well-being (N= 200)

Predictor	Weight	Domain
1. Feeling of purpose or reason for being alive	0.7	Spiritual
2. Feeling of uncertainty	0.7	Spiritual
3. Feeling hopeful after cancer experience	0.7	Spiritual
* A higher weight indicates a strong prediction		

4 Discussion

This study evaluated factors that influencing the dimensions of Quality of life (physical, psychological, social, and spiritual) of women with breast cancer using data mining technique. This study revealed that fatigue, pain, sleep change, and loss of appetite were associated with QoL, which supports [6] results that pain alone or with other symptoms significantly related to poor QoL. McClelland [5] found that women with relatively high fatigue levels reported low physical functionality. Thus, managing fatigue, pain, and other physical symptoms should be the main approach to improving the QoL of

Jordanian women with breast cancer. Furthermore, some psychological factors, such as a change in appearance, which in turn will affect the body image, influenced QoL among the women in the sample, which supports[5]. They found that high scores on body image related to better physical functioning and body image was a significant factor for the QoL of women with breast cancer.

In the psychological domain, findings revealed that fear of cancer metastasis and recurrence in addition to fear of future diagnostic test is a normal response to a life-threatening illness such as cancer. These results are similar to the findings of [37, 38]. However, our results highlighted the role of having a positive outlook and coping with breast cancer treatment for the QoL of women with breast cancer [38]. Therefore, healthcare professionals might focus on strengthening these women's coping strategies through listening, focus groups, consultations, and referrals to psychologists. Nurses might take a central role in this approach and coordinate their efforts with a multidisciplinary healthcare team and support services.

The Factors that affected the social well-being were items about the effect of cancer on Women's relationship and her ability to do the home chores. This can be explained by the fact that breast cancer and its treatment can limit women's physical abilities to socialize due to pain and fatigue. Furthermore, appearance changes such as hair loss can limit social interaction with others. Psychological factors that also can affect social well-being are fear and depression; because the desire to do social interaction is reduced.

Lastly, factors that might affect the spiritual well-being are a feeling of purpose in life and feeling hopeful. This finding is consistent with [39], who found that spirituality was associated with better Quality of life and [40] in which they concluded that hope was strongly related to the mental health of women with breast cancer.

5 Conclusion

In conclusion, Jordanian women undergoing breast cancer experience many problems that might affect their QoL. The study signaled the importance of supporting the use of data mining in health related studies. Accordingly, the physical dimension of QoL for women with breast cancer cannot be isolated from the psychological, social, and spiritual well-being of QoL. Healthcare professionals and nurses, in particular, should focus on regular assessment of patients' QoL to help them devise the most effective individualized holistic approaches possible for these women throughout the disease trajectory. In particular, physical symptom management, psychological support, cognitive therapy, social support, and spiritual enhancement technique including instilling hope must be embraced by all health care team through the use of multidisciplinary team to ease the journey of women with breast cancer.

References

[1] American Cancer Society, Global Cancer Facts & Figures (Third Edition), 1–64 (2022).

[2] S.I. Al-Hawary, The Role of Perceived Quality and Satisfaction in Explaining Customer Brand Loyalty: Mobile Phone Service in Jordan, International Journal of Business Innovation and Research **7**, 393–413 (2013).

[3] P. Brunault, A.L. Champagne, G. Huguet, I. Suzanne, J. Senon, G. Body, V. Camus, Major depressive disorder, personality disorders, and coping strategies are independent risk factors for lower quality of life in non-metastatic breast cancer patients, Psycho-Oncology **25**, 513-520 (2016).

[4] D.F. Cella, Quality of life: Concepts and definition, Journal of Pain and Symptom Management **9**, 186–192 (1994). https://doi.org/10.1016/0885-3924(94)90129-5

[5] S.I. McClelland, K.J. Holland, J.J. Griggs, Quality of life and metastatic breast cancer: the role of body image, disease site, and time since diagnosis, Quality of Life Research **24**, 2939–2943 (2015). https://doi.org/10.1007/s11136-015-1034-3

[6] R. Hamood, H. Hamood, I. Merhasin, L. Keinan-Boker, Chronic pain and other symptoms among breast cancer survivors: prevalence, predictors, and effects on quality of life, Breast cancer research and treatment **167**, 157-16 (2018).

[7] D. Eleimat, M. Ebbini, L. Aryan, S.I. Al-Hawary, The effect of big data on financial reporting quality, International Journal of Data and Network Science **7**, 1775-1780 (2023).

[8] U. Lehto, M. Ojanen, P. Kellokumpu-Lehtinen, Predictors of quality of life in newly diagnosed melanoma and breast cancer patients, Annals of Oncology **16**, 805–816 (2005). https://doi.org/10.1093/annonc/mdi146

[9] J. Xia, Z. Tang, Q. Deng, R. Yang, J. Wang, J. Yu, Predictors of the quality of life in Chinese breast cancer survivors, Breast cancer research and treatment **167**, 537-545 (2018).

[10] A. Wöckel, L. Schwentner, M. Krockenberger, R. Kreienberg, W. Janni, M. Wischnewsky, S. Singer, Predictors of the course of quality of life during therapy in women with primary breast cancer, Quality of Life Research **26**, 2201–2208 (2017). https://doi.org/10.1007/s11136-017-1570-0

[11] H. Setyowibowo, F.D. Purba, J.A.M. Hunfeld, A. Iskandarsyah, S.S. Sadarjoen, J. Passchier, M. Sijbrandij, Quality of life and health status of Indonesian women with breast cancer symptoms before the definitive diagnosis: A comparison with Indonesian women in general, PloS One **13**, e0200966 (2018). https://doi.org/10.1371/journal.pone.0200966

[12] M. Lazenby, J. Khatib, Associations among Patient Characteristics, Health-Related Quality of Life, and Spiritual Well-Being among Arab Muslim Cancer Patients, Journal of Palliative Medicine **15**, 1321–1324 (2012). https://doi.org/10.1089/jpm.2012.0208

[13] M. Alolayyan, S.I. Al-Hawary, A.A. Mohammad, B.A. Al-Nady, Banking Service Quality Provided by Commercial Banks and Customer Satisfaction. A structural Equation Modelling Approaches, International Journal of Productivity and Quality Management **24**, 543–565 (2018).

[14] I. Tăranu, Data mining in healthcare: decision making and precision, Database Systems Journal **5**, 33–40 (2015).

[15] A.E.A. Elrazek, How Can Data Mining Improve Health Care? Applied Mathematics & Information Sciences **11**, 585–588 (2017). https://doi.org/10.18576/amis/110230

[16] M. Al-Azzam, M. Al-Alwan, M. Alqahtani, S.I. Al-Hawary, A. Alserhan, Determinants of behavioral intention to use big data analytics (BDA) on the information and communication technologies (ICT) SMEs in Jordan, Decision Science Letters **12**, 605-616 (2023).

[17] N. Shamaileh, M. Eldahamsheh, S. Alneimat, R. Istaiteyeh, I. Azzam, S.I. Al-Hawary, The effects of smart human resources 4.0 on employee job effectiveness: The mediating role of employee job satisfaction, International Journal of Data and Network Science **7**, 801-808 (2023).

[18] O. Niakšu, O. Kurasova, Data mining applications in healthcare: Research vs practice, CEUR Workshop Proceedings **924**, 58–70 (2012).

[19] H. Al-Shormana, A. AL-Zyadat, M. Khalayleh, A. Al-Quran, M.I. Alhalalmeh, A. Mohammad, s.I. Al-Hawary, Digital Service Quality and Customer Loyalty of Commercial Banks in Jordan: the Mediating Role of Corporate Image, Information science letters **11**, 1887-1896 (2022).

[20] R. Al-Husban, N. Al-Shdayfat, R. Maabreh, A. Alshdefat, N. Elsharkaway, Factors Influencing the Quality of Life among Jordanian Women Undergoing Breast Cancer Treatment: A Cross Sectional Study, World Journal of Medical Sciences **16**, 196-202 (2019).

[21] A. Mohammad, The effect of customer empowerment and customer engagement on marketing performance: the mediating effect of brand community membership, Verslas: Teorija ir praktika/Business: Theory and Practice **21**, 30-38 (2020).

[22] U. Fayyad, D. Haussler, P. Stolorz, Mining scientific data, Communications of the ACM **39**, 51–57 (1996). https://doi.org/10.1145/240455.240471

[23] N.K. Dwijendra, I.G. Arsana, S.I. Al-Hawary, A.S. Prakaash, R.M. Parra, A.T. Jalil, A.T. Hammid, Operation of the Multiple Energy System with Optimal Coordination of the Consumers in Energy Market, Environmental and Climate Technologies **27**, 1-13 (2023).

[24] L. Goodwin, M. VanDyne, S. Lin, S. Talbert, Data mining issues and opportunities for building nursing knowledge, Journal of Biomedical Informatics **36**, 379–388 (2003). https://doi.org/10.1016/j.jbi.2003.09.020

[25] I.R. AlTaweel, S.I. Al-Hawary, The Mediating Role of Innovation Capability on the Relationship between Strategic Agility and Organizational Performance, Sustainability **13**, 7564 (2021).

[26] R. Alshawabkeh, H. AL-Awamleh, M. Alkhawaldeh, R. Kanaan, S.I. Al-Hawary, A. Mohammad, R. Alkhawalda, The mediating role of supply chain management on the relationship between big data and supply chain performance using SCOR model, Uncertain Supply Chain Management **10**, 729-736 (2022).

[27] H. Pallathadka, S.I. Al-Hawary, I. Muda, S. Surahman, A.A. Al-Salami, Z. Nasimova, The study of Islamic teachings in education: With an emphasis on behavioural gentleness, HTS Teologiese Studies/Theological Studies **79**, 8193 (2023).

[28] D.M. Liou, W.P. Chang, *Applying Data Mining for the Analysis of Breast Cancer Data*. In Methods in Molecular Biology (pp. 175–189). Springer: New York.

[29] M.M. Amiruddin, H.S. Alshahrani, N.K. Dwijendra, S.I. Al-Hawary, T.A. Jalil, I. Muda, D. Sunarsi, Religious behaviours and commitment among Muslim healthcare workers in Malaysia, HTS Teologiese Studies/Theological Studies **79**, 8177 (2023).

[30] R.J. Oskouei, N.M. Kor, S.A. Maleki, Data mining and medical world: breast cancers' diagnosis, treatment, prognosis and challenges, American journal of cancer research **7**, 610 (2017).

[31] A. ahamneh, S. Alrawashdeh, A. Bawaneh, Z. Alatyat, A. Mohammad, S.I. Al-Hawary, The effect of digital supply chain on lean manufacturing: A structural equation modelling approach, Uncertain Supply Chain Management **11**, 391-402 (2023).

[32] D.H. Qudsi, M. Kartiwi, N.B. Saleh, Predictive data mining of chronic diseases using decision tree: a case study of health insurance Company in Indonesia, International Journal of Applied Engineering Research **12**, 1334-1339 (2017).

[33] P.A. Idowu, K.O. Williams, J.A. Balogun, A.I. Oluwaranti, Breast Cancer Risk Prediction Using Data Mining Classification Techniques, Transactions on Networks and Communications **3**, 1-11 (2015). https://doi.org/10.14738/tnc.32.662

[34] S. Kulkarni, M. Bhagwat, Predicting Breast Cancer Recurrence using Data Mining Techniques, International Journal of Computer Applications **122**, 26–31 (2015).

[35] M.A. Salem, *Data mining techniques and breast cancer prediction: A case study of Libya*. Doctoral Thesis, Sheffield Hallam University, United Kingdom (2011).

[36] S. Lee, Y. Son, J. Kim, H. Kim, J. Lee, B. Kang, S. Lee, Prediction Model for Health-Related Quality of Life of Elderly with Chronic Diseases using Machine Learning Techniques, Healthcare Informatics Research **20**, 125–134 (2014).

[37] B.R. Ferrell, M. Grant, B. Funk, S. Otis-Green, N. Garcia, Quality of life in breast cancer: Part II: Psychological and spiritual well-being, Cancer nursing **21**, 1-9 (1998).

[38] S. De Aguiar, A. Bergmann, I.E. Mattos, Quality of life as a predictor of overall survival after breast cancer treatment, Quality of Life Research **23**, 627-637 (2014).

[39] S.L. Manne, S. Myers-Virtue, D. Kashy, M. Ozga, D. Kissane, C. Heckman, N. Rosenblum, Resilience, Positive Coping, and Quality of Life Among Women Newly Diagnosed With Gynecological Cancers, Cancer Nursing **38**, 375–382. https://doi.org/10.1097/ncc.0000000000000215

[40] M. Heidari, M. Ghodusi, The relationship between body esteem and hope and mental health in breast cancer patients after mastectomy, Indian journal of palliative care **21**, 198 (2015).