Received: 17 February 2017 | Revised: 4 August 2017 | Accepted: 5 October 2017

DOI: 10.1002/pon.4571

WILEY

PAPER

Cognition, quality-of-life, and symptom clusters in breast cancer: Using Bayesian networks to elucidate complex relationships

Selene Xu[1] | Wesley Thompson[2] | Sonia Ancoli-Israel[2] | Lianqi Liu[2] | Barton Palmer[2,3] | Loki Natarajan[3,4]

[1] Department of Mathematics, University of California, San Diego, CA, USA

[2] Department of Psychiatry, University of California, San Diego, CA, USA

[3] Veterans Affairs, San Diego Health Care System, San Diego, CA, USA

[4] Department of Family Medicine and Public Health, University of California, San Diego, CA, USA

Correspondence
Loki Natarajan, 3855 Health Sciences Drive #0901, La Jolla, CA 92093-090, USA.
Email: lnatarajan@ucsd.edu

Funding information
UCSD Clinical and Translational Research Institute, Grant/Award Number: Pilot grant (2012-2013); Department of Veterans Affairs San Diego Center of Excellence for Stress and Mental Health; UCSD CTRI, Grant/Award Number: UL1RR031980; National Cancer Institute, Grant/Award Number: CA112035 and P30 CA023100

Abstract

Objective: Breast cancer patients frequently complain of cognitive dysfunction during chemotherapy. Patients also report experiencing a cluster of sleep problems, fatigue, and depressive symptoms during chemotherapy. We aimed to understand the complex dynamic interrelationships of depression, fatigue, and sleep to ultimately elucidate their role in cognitive performance and quality of life amongst breast cancer survivors undergoing chemotherapy treatment.

Methods: Our study sample comprised 74 newly diagnosed stage I to III breast cancer patients scheduled to receive chemotherapy. An objective neuropsychological test battery and self-reported fatigue, mood, sleep quality, and quality of life were collected at 3 time points: before the start of chemotherapy (baseline: BL), at the end of cycle 4 chemotherapy (C4), and 1 year after the start of chemotherapy (Y1). We applied novel Bayesian network methods to investigate the role of sleep/fatigue/mood on cognition and quality of life prior to, during, and after chemotherapy.

Results: The fitted network exhibited strong direct and indirect links between symptoms, cognitive performance, and quality of life. The only symptom directly linked to cognitive performance was C4 sleep quality; at C4, fatigue was directly linked to sleep and thus indirectly influenced cognitive performance. Mood strongly influenced concurrent quality of life at C4 and Y1. Regression estimates indicated that worse sleep quality, fatigue, and mood were negatively associated with cognitive performance or quality of life.

Conclusions: The Bayesian network identified local structure (eg, fatigue-mood-QoL or sleep-cognition) and possible intervention targets (eg, a sleep intervention to reduce cognitive complaints during chemotherapy).

KEYWORDS

Bayesian network, breast cancer, cognitive performance, fatigue, mood, oncology, quality of life, sleep

1 | BACKGROUND

Cancer patients report a cluster of fatigue, sleep, and mood problems before and during adjuvant cancer treatment.[1,2] Cancer-related fatigue, one of the most distressing symptoms, is characterized by extreme tiredness.[3] Cancer patients also report sleep problems, eg, difficulties with falling and staying asleep, before, during, and after chemotherapy.[4-6] Insomnia and sleep problems have been associated with fatigue,[4,7,8] depression, and decreased quality of life during treatment.[4,9-11] Depression is common[12,13] with 40% to 82% of patients undergoing chemotherapy reporting clinically significant depressive symptoms.[14] Importantly, our lab has shown that in women with

breast cancer, all symptoms within the symptom cluster, ie, fatigue, sleep complaints, and depressive symptoms, whether or not present before chemotherapy, worsen during cancer treatment.[4,15]

Chemotherapy-related cognitive problems are frequently reported by cancer survivors.[16-18] In a recent meta-analysis, 15 (12 in breast cancer) of 17 studies observed objective cognitive decline in patients treated with chemotherapy.[19] These problems can last for a few weeks, months, or even years after completion of chemotherapy.[20] Imaging studies have found alterations in cerebral activity in cancer patients before and after chemotherapy.[21] It is likely that the cognitive impairment seen in cancer patients might, at least in part, be related to fatigue, sleep problems, and depression.[22]

Unraveling interrelationships between these symptoms and their impact on cognition poses computational challenges. Dimension reduction methods such as principal component analysis or clustering techniques are useful for deriving summary statements regarding association between the symptom cluster as a whole and cognitive symptoms in cancer patients. However, these methods use weighted combinations of suitably normalized symptom cluster variables, rendering results difficult to interpret for individual symptoms. While standard regression modeling cannot disentangle these complex associations, moderated regression methods could be used in this context.[23,24] Yet, alternative novel methods are needed to assess robustness of findings.

Bayesian graphical networks are a powerful approach for examining multivariate relationships. Bayesian networks provide algorithms for discovering and analyzing structure with intuitive graphs for visualizing interrelationships amongst sets of variables. The initial development of Bayesian networks arose in computer science and artificial intelligence,[25] and since then it has become widely used in genomics, and medical applications.[26,27]

Bayesian networks have not been widely used in cancer symptoms research. Herein, we illustrate how to apply Bayesian networks to examine associations amongst symptoms related to chemotherapy treatment, and their role in cognitive dysfunction, and quality of life. Our goal is to demonstrate how this powerful computational method can be used to explore complex interrelationships between variables, and possibly guide design of intervention studies.

2 | BRIEF DESCRIPTION OF *BAYES NETWORK METHODOLOGY*

A Bayesian network is a statistical model that represents multivariate relationships between sets of variables via a graph. Nodes on the graph depict random variables, while edges represent dependencies between variables. Each node has an associated probability function that takes as input a particular set of values for the node's parent variables and gives the probability of the variable represented by the node. The presence of an edge or path between 2 variables indicates a non-zero (partial) correlation between the 2 variables.[25] Specifically, a Bayesian network analysis derives observed probabilistic dependencies and (conditional) independencies between sets of variables. Under certain technical assumptions,[25,28] these observed probabilistic relationships conform to what could have been observed from a hypothetical causal network with the same structure under a controlled experiment, where each variable is manipulated while holding others constant. Thus, a Bayesian network can generate hypotheses that can be tested in future studies.

Fitting a Bayesian network requires learning its structure (ie, which nodes in the graph are connected) and parameters (ie, estimation of conditional probabilities). Specifically, let $\boldsymbol{X}$ comprise the set of variables X_i (eg, X_1 = fatigue, X_2 = sleep quality, X_3 = mood, X_4 = cognition) and M be a Bayesian network on $\boldsymbol{X}$, ie, a graph of edges between variables in $\boldsymbol{X}$, as in Figure 1. The model M encodes conditional independencies that imply a factoring of the joint probability distribution p($\boldsymbol{X}$) of $\boldsymbol{X}$:

$$\Pr(\boldsymbol{X} \mid M) = \prod \Pr(X_i \mid \mathrm{pa}(X_i)), \quad (1)$$

where pa(X_i) denotes variables ("parents") in $\boldsymbol{X}$ with directed edges (ie, arrows) leading to X_i. To learn the structure of the graph M (eg, links, or edges, between fatigue, sleep, mood, and cognition in our application), efficient constraint-based, search-score, and hybrid algorithms can be implemented.[28] For a given graph M, parameters ß of p($\boldsymbol{X}$) can be estimated by regression methods, using multivariate Gaussian (after appropriate transformation if needed) or nonparametric distributions for continuous, and multinomial distributions for categorical variables.[28]

3 | METHODS

3.1 | Study sample and measures

We conducted a secondary analysis of an existing database of a completed NIH-funded study (2004–2010) on the relationship between sleep, fatigue, mood, and cognition in breast cancer patients (PI Ancoli-Israel). Details on the study design and protocol have been previously published.[9] The study was approved by the UCSD Human Research Protections Committee (protocol #s 080120 and 120187) and the UCSD Moores Cancer Center's Protocol Review and Monitoring Committee. We briefly describe the study sample and measures pertinent to our study.

3.1.1 | Study sample

The study recruited 74 newly diagnosed stage I to III breast cancer patients (mean age = 51.8 years) who were scheduled to receive chemotherapy and followed them for 1 year. Data were collected at 3 time points: before the start of chemotherapy (baseline; BL), at the end of cycle 4 chemotherapy (C4), and 1 year after the start of chemotherapy (Y1). In order to reduce confounding by other medical conditions or medications, the study excluded pregnant women, patients with significant anemia, patients currently receiving radiotherapy or bone marrow transplants or treatment for sleep apnea or periodic limb movements in sleep. Also, patients with current diagnosis of major depression, anxiety or psychotic disorder, and patients using medications known to influence sleep for 3 months prior to enrollment, were excluded.

Participant characteristics are provided in Table 1. Briefly, 88% were white, and 51% were college graduates, and close to 70% had

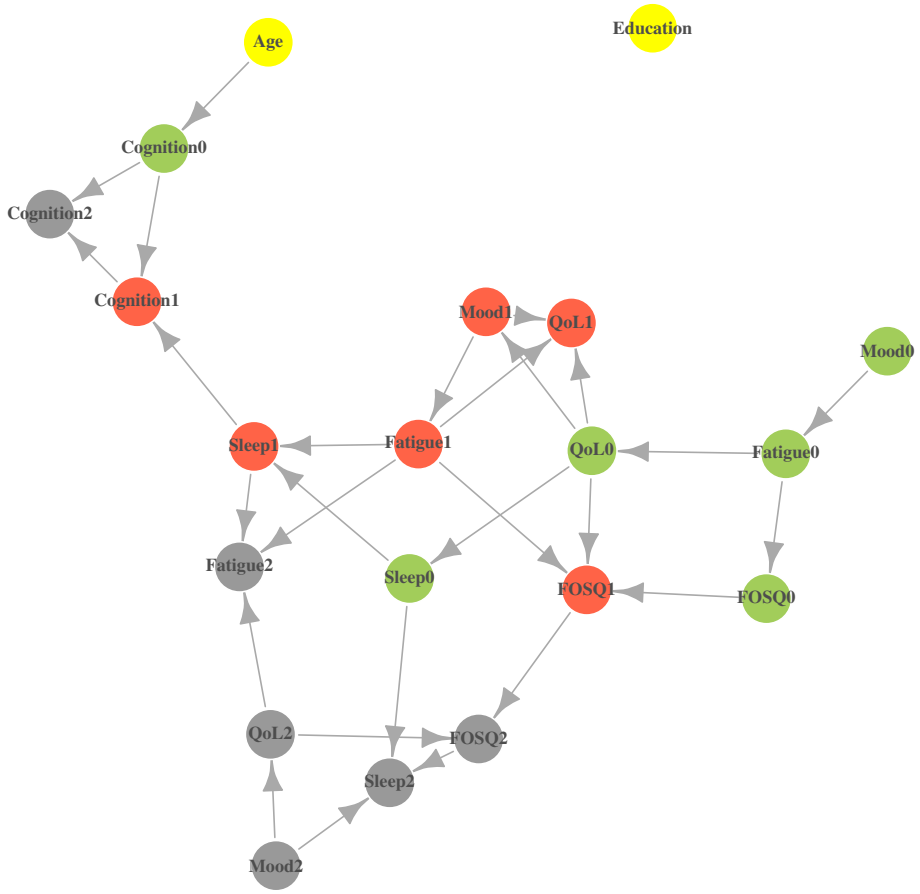

FIGURE 1 Bayesian network of symptoms and outcomes before (BL), during (C4), and after chemotherapy (Y1) amongst breast cancer patients. The circles represent variables, and arrows codify dependencies between variables. The colors and number suffixes represent the 3 time periods: green, 0 = BL; red, 1 = C4; gray, 2: Y1

TABLE 1 Participant characteristics (N = 74, female breast cancer survivors)

	Mean (SD) or Percent
Age (years)	51 (9.5)
Education	
High school	11%
Some college	38%
College graduate	51%
Race	
Caucasian	88%
Asian	7%
African-American	3%
Ethnicity	
Hispanic	8%
Cancer stage	
Stage I	27%
Stage II	41%
Stage III	31%

Stage I or II cancers. The median (25th, 75th percentile) interval between the pre-chemotherapy and end-of-chemotherapy assessments was 79 (64, 84) days, and between the end-of-chemotherapy and the year 1 measures was 341 (317, 409) days.

3.1.2 | Symptom cluster and psychosocial functioning assessment

Symptoms were assessed via validated questionnaires. At BL, C4, and Y1, patients self-reported sleep quality on the Pittsburgh Sleep Quality Index (PSQI)[29]; fatigue on the Multidimensional Fatigue Symptom Inventory-Short Form (MFSI-SF)[30]; mood on the Center of Epidemiological Studies-Depression (CES-D) scale[31]; quality of life (QoL) on the FACT-B scale for breast cancer patients; and functional-outcomes-of-sleepiness (FOS) on the functional-outcomes-of-sleepiness questionnaire (FOSQ).[32] FOS assesses how sleepiness impacts daily functioning.

3.1.3 | Neurophyschological (NP) testing

Cognitive function was assessed with an objective NP test battery, which targeted a number of specific cognitive abilities associated with chemotherapy-related impairment, including episodic learning/memory, attention/working memory, executive functions, and psychomotor/processing speed. Specific component tests in this battery included Digit Span, Digit Symbol, and Symbol Search subtests from the Wechsler Adult Intelligence Scale-Third edition (WAIS-III)[33]; Trail Making Tests A and B[34]; Hopkins Verbal Learning Test-Revised[35]; Wisconsin Card Sorting Task–64 card version (conceptual level of responses)[36]; Stroop Color-Word Interference test (interference

trial)[37]; and the Letter and Category[38] (animals) Fluency test (total words generated).

A summary measure of cognitive ability, a NP composite score, was computed as follows: each component raw test score was converted to a *z*-score by subtracting the baseline mean and dividing by the standard deviation. *Z*-scores were coded so that higher scores represented better functioning, and the composite score was defined as the mean of *z*-scores over the entire battery. As supplementary material, mean (SD) of the individual test scores at each study time point is provided in Table S1.

3.1.4 | Previous analyses of longitudinal symptoms and cognition

We have previously shown[9] that compared with healthy controls, breast cancer patients had worse sleep quality (PSQI), more fatigue (MFSI), worse mood (CESD), worse functional-outcomes-of-sleepiness (FOSQ), and Quality of life prior to chemotherapy. Also, these factors worsened for the patients during chemotherapy, compared with controls. By Y1, symptoms in the patients were not different to their baseline values but were still worse compared with controls.

In summary, our prior analyses examined each symptom individually and demonstrated that, on average, sleep quality, fatigue, mood, and QoL worsened in breast cancer patients during chemotherapy. However, this previous work did not evaluate how these different symptoms influenced each other and cognition over time. The Bayesian network analyses proposed in the next section aims to address this latter question.

4 | STATISTICAL METHODS

We calculated summary statistics of demographic factors, as well as, mean (SD) at BL, C4, and Y1 for the symptom cluster (sleep quality, fatigue, mood), QoL, and cognition.

We fit a Bayesian network to examine multivariate relationships between the symptom cluster, quality of life and cognition, before, during, and after chemotherapy. We also included demographic variables (age, college educated [yes vs no]) in the network. Our network included pre-chemotherapy (BL), post-chemotherapy (C4), and year 1 follow-up measures (Y1), and examined temporal and cross-section relationships amongst variables. In particular, measures at time *t* were allowed to have directed edges to measures at time (*t* + 1), but not vice versa.

We used a score-based hill-climbing algorithm to infer network structure and applied bootstrap resampling to learn a set of 500 network structures. We then averaged these networks in an effort to reduce the impact of locally optimal (but globally suboptimal) networks on learning and inference. The averaged network is a more robust model with better predictive performance than choosing a single, high-scoring network.[28] To quantify stability of inferred edges, we computed *arc strength* and *direction strength*. Arc strength was calculated as the frequency of an edge occurring between 2 variables across the 500 bootstrapped network structures; similarly, directional strength was assessed as the frequency of the observed direction reoccurring in the set of learned network structures. We inferred conditional independencies between variables via the theory of Markov blankets of networks.[28] We applied Bayesian information criteria (BIC) and posterior model probabilities to compare candidate networks.[39] Lower BIC scores indicate better fit; score differences >5 (respectively, between 2.2 and 5) between 2 models strongly (respectively moderately) favor the lower-scoring model; differences <2.2 indicate similar fit for both models. Models were fitted using the R package bnlearn.[28]

5 | RESULTS

Longitudinal scores (Table 2) indicate that, as noted previously,[9] symptoms and QoL worsened during chemotherapy (BL-C4) on average but were generally comparable to BL levels by Y1. Cognitive performance did not change significantly during chemotherapy but was significantly higher at year 1 compared with BL.

5.1 | Bayesian network results

5.1.1 | Decomposition of probability distribution

Using the derived network, we decomposed the joint probability distribution of all 20 variables (2 demographic, and 6 symptoms/outcomes at 3 time points) as a product of conditional distributions. Specifically, letting **X** represent the vector of all the variables, we have the following factorization. The suffix 0 is BL, 1 is C4, and 2 is Y1.

BL-C4-Y1 model:

$$\begin{aligned}P(\mathbf{X}) = {} & P[\text{Mood0}] * P[\text{Mood2}] * P[\text{Age}] * P[\text{Education}] * P[\text{Cognition0}|\text{Age}] * P[\text{Fatigue0}|\text{Mood0}] \\ & * P[\text{QoL2}|\text{Mood2}] * P[\text{QoL0}|\text{Fatigue0}] * P[\text{FOS0}|\text{Fatigue0}] * P[\text{Mood1}|\text{QoL0}] * P[\text{Sleep0}|\text{QoL0}] \\ & * P[\text{Fatigue1}|\text{Mood1}] * P[\text{QoL1}|\text{Mood1} + \text{Fatigue1} + \text{QoL0}] * P[\text{Sleep1}|\text{Fatigue1} + \text{Sleep0}] \\ & * P[\text{FOS1}|\text{Fatigue1} + \text{QoL0} + \text{FOS0}] * P[\text{Cognition1}|\text{Cognition0} + \text{Sleep1}] \\ & * P[\text{Fatigue2}|\text{Fatigue1} + \text{QoL2} + \text{Sleep1}] * P[\text{FOS2}|\text{QoL2} + \text{FOS1}] \\ & * P[\text{Cognition2}|\text{Cognition0} + \text{Cognition1}] * P[\text{Sleep2}|\text{Mood2} + \text{Sleep0} + \text{FOS2}].\end{aligned}$$

P(A|B) denotes the conditional distribution, ie, probability of a variable A, given that we know the value of variable B. Thus, the above decomposition converts the complex model comprising 20 variables

TABLE 2 Symptom scores (mean (SD)) for 74 breast cancer patients before, at completion of, and 1 year after chemotherapy treatment

Symptom	Direction of Better Outcome	Pre-Chemotherapy	End-of Chemotherapy-cycle4	One-Year Post-Chemotherapy
Cognition (NP composite score)	↑	0.062 (0.743)	0.077 (0.691)	0.166 (0.738)
Mood (CESD)	↓	11.5 (10.4)	16.2 (12.9)	10.0 (9.95)
Fatigue (MFSI)	↓	9.66 (18.3)	18.0 (23.9)	7.6 (20.3)
Quality of life (FACT-B)	↑	105 (16.1)	95.0 (23.3)	110 (19.0)
Sleep quality (PSQI)	↓	7.71 (3.87)	9.07 (3.74)	7.49 (4.40)
Functional outcomes of sleepiness (FOS)	↑	18.0 (2.04)	16.0 (2.87)	17.7 (2.19)

into simpler "local" components and highlights subsets of factors that directly influence each variable. In fact, in our network (Figure 1), the maximum number of parents, ie, directed edges pointing to any variable, is 3 (eg, Fatigue at Y1 has parents Fatigue and Sleep at C4, and QoL at Y1), thus substantially fewer than the maximum of 19 possible directed edges. Below, we highlight key findings and describe how to infer (in)dependencies between variables.

5.1.2 | Cognitive functioning and symptoms

The network elicits local structure, so that we can identify parents, namely variables that directly influence any given factor. For example, age was the parent of BL cognition, whereas BL cognition and C4 sleep quality were parents of C4 cognition. The bootstrapped arc strength for age on BL cognition was 0.86 indicating that age was reproducibly associated with cognition. The regression estimate (Table 3) was negative for age indicating that younger age was associated with better cognition at BL.

At C4, cognition was (directly) positively influenced by BL cognition (as might be expected) but negatively influenced by C4 sleep score, indicating that worse sleep quality at the end of chemotherapy was associated with worse cognition. Interestingly, via the Markov property, we infer that after accounting for cognition and C4 sleep, C4 cognitive function was independent of all other variables. We also note that although not directly linked, C4 fatigue affected C4 cognition through C4 sleep quality. Moreover, C4 depression indirectly affected cognition through a direct effect on C4 fatigue and corresponding downstream effects on C4 sleep quality.

At the 1-year follow-up, cognition was directly influenced by both BL cognition and C4 cognition, with, as expected, positive regression coefficients for both variables, indicating that higher BL and C4 cognition scores were associated with better Y1 cognition. Interestingly, no symptoms directly influenced Y1 cognition.

5.1.3 | Symptom clusters, quality of life (QoL), and functional outcomes of sleepiness (FOS)

Focusing on BL Quality of life and functional outcomes of sleepiness, BL fatigue was the only parent of BL QoL (arc-strength = 0.93) and BL FOS (arc-strength = 0.70), with negative regression estimates, indicating that less fatigue was associated with better QoL and FOS (Table 3).

At C4, there were several parents for each of QoL and FOS. C4 mood and fatigue, and BL QoL were all parents of C4 QoL, with C4 mood exhibiting the most consistent effects (arc strength = 0.94, arc direction = 0.89). Also, via the Markov property, we can infer that after conditioning on BL QoL, C4 mood, and C4 fatigue, C4 QoL was independent of all other factors. Factors influencing C4 FOS were C4 fatigue, BL FOS, and QoL, with C4 fatigue and BL FOS exhibiting high consistency (arc strength ≥ 0.93, arc direction ≥ 0.97). Regression estimates (Table 3) indicated that worse C4 mood and/or fatigue (ie, higher score) were associated with worse C4 QoL and FOS.

Factors influencing Y1 QoL and FOS were fewer than at C4. The only predictor for Y1 QoL was Y1 mood, with high consistency (arc strength = 0.96, arc direction = 0.78). C4 FOS and Y1 QoL were directly linked to Y1 FOS but were not stable (arc strength ≤ 0.55).

TABLE 3 Bayesian network structure and associations[a]

Outcome (Child)	Predictors (Parents)	Strength	Direction	Regression Coefficients (SE)
Mood1	QoL0	0.52	0.94	−0.536 (0.095)
Cognition0	Age	0.86	1.00	−0.043 (0.010)
Cognition1	Cognition0 Sleep1	1.00 0.68	1.00 1.00	0.857 (0.062) −0.037 (0.012)
Cognition2	Cognition0 Cognition1	0.95 0.93	1.00 0.90	0.510 (0.113) 0.494 (0.121)
QoL0	Fatigue0	0.93	0.72	−0.706 (0.084)
QoL1	Mood1 Fatigue1 QoL0	0.94 0.64 0.63	0.89 0.58 0.99	−0.980 (0.237) −0.290 (0.121) 0.242 (0.110)
QoL2	Mood2	0.96	0.78	−1.583 (0.171)
FOSQ0	Fatigue0	0.70	0.64	−0.069 (0.014)
FOSQ1	Fatigue1 QoL0 FOSQ0	0.95 0.62 0.93	0.97 1.00 1.00	−0.097 (0.012) −0.047 (0.019) 0.582 (0.131)
FOSQ2	QoL2 FOSQ1	0.55 0.50	0.64 0.99	0.057 (0.013) 0.329 (0.085)
Fatigue0	Mood0	0.99	0.69	1.400 (0.169)
Fatigue1	Mood1	0.87	0.90	1.651 (0.135)
Fatigue2	Fatigue1 QoL2 Sleep1	0.62 0.73 0.67	0.99 0.64 0.97	0.169 (0.112) −0.520 (0.132) 1.439 (0.677)
Sleep0	QoL0	0.96	0.73	−0.158 (0.029)
Sleep1	Fatigue1 Sleep0	0.67 0.52	0.85 0.99	0.080 (0.020) 0.285 (0.125)
Sleep2	Mood2 Sleep0 FOSQ2	0.75 0.76 0.53	0.87 1.00 0.57	0.219 (0.050) 0.353 (0.116) −0.457 (0.227)

[a]Suffixes: 0 = pre-chemotherapy; 1 = end-of-chemotherapy (1); 2 = 1 year after the start of chemotherapy.

Several symptoms showed direct cross-sectional and temporal links. At BL and C4, mood was the (only) parent of fatigue with strong cross-sectional links (arc strength ≥ 0.87); also, regression estimates (Table 3) were positive, indicating that higher CESD scores (ie, worse mood) were associated with higher fatigue scores (ie, worse fatigue). Further, although not exhibiting high consistency, C4 fatigue and sleep quality were parents of Y1 fatigue (arc-strengths ≤ 0.67); Y1 sleep was directly influenced by Y1 mood and BL sleep (arc-strengths ≤ 0.76). Thus, these links suggest a temporal cluster of sleep, mood, and fatigue.

5.1.4 | Comparing networks

Given our focus on cognitive symptoms and quality of life during and after chemotherapy, we conducted sensitivity analyses to test the value of the learned subnetworks for Cognition and QoL. We created a new network in which all edges to- and from- Cognition were removed and refitted this network to the data. The BIC score for this new network was more than 164 points higher than that of the original networks, indicating far superior fit of the original fitted network and providing support for the identified links to cognition. Similarly, a network in which QoL was isolated (ie, all edges to- and from- QoL were deleted) had a 226 point higher BIC score, again strongly favoring the original fitted models.

We also tested the impact of removing a specific edge from the network as follows. If we removed the

- Mood1-QoL1 edge, BIC increased by 12, giving strong evidence for this link
- Fatigue1-QoL1 edge, BIC increased by 2.2, giving weak evidence for this link
- Sleep1-Cognition1 edge, BIC increased by 4.8, giving moderate evidence for this link

Finally, we evaluated the overall fit of our learned network (Figure 1). Our a priori assumption was that network structure might vary during chemotherapy versus during follow-up. Hence, we fit a flexible network in which links amongst symptoms could be different in the chemotherapy treatment (BL-C4) phase compared with during follow-up (C4-Y1). We can quantitatively assess this assumption via BIC scores. Our fitted BL-C4-Y1 network had a BIC score of 4384.4. We then fit a second network in which we constrained the C4-Y1 subnetwork to be identical to the corresponding chemotherapy treatment phase sub-network (the learned BL-C4 sub-network). This constrained network had a BIC score of 4461.6, a 77-point higher score, indicating substantially worse fit for the constrained model compared with the original network.

6 | DISCUSSION

Most patients undergoing chemotherapy complain of symptoms such as fatigue, impaired sleep, and poor mood. Studies of these patients generally focus on average effects and note that mean scores for each of these symptoms usually worsen during chemotherapy. Not much is known regarding how these symptoms influence each other. In the current work, we aimed to address this gap. We applied a powerful Bayesian network approach to discern inter-relationships amongst these symptoms and, furthermore, examined the role of these symptoms on QoL, functional outcomes of sleepiness, and cognitive functioning. Unraveling inter-relationships amongst these many factors is a complex computational problem, and Bayesian networks provide a first glimpse at how we might decompose this large multivariate distribution into a set of lower-dimensional relationships.

6.1 | Clinical implications

There are many potential clinical implications of this work. Understanding cross-sectional and longitudinal inter-relationships amongst symptoms, QoL and cognition could guide the design of effective interventions. For instance, our networks identified sleep quality as the primary symptom influencing cognition. Thus, an intervention aimed at improving sleep during chemotherapy could potentially mitigate some of the neurocognitive symptoms experienced by cancer patients. We emphasize that our goal in the current analysis was not to assess whether cancer patients experienced cognitive dysfunction, a phenomenon that has been well studied, but rather to identify factors that might influence acute cognitive ability for a patient undergoing chemotherapy. Another finding of this work was that mood and fatigue directly influenced QoL and/or FOS in the chemotherapy period, and after accounting for the symptom cluster of sleep, mood, and fatigue, cognition was (conditionally) independent of QoL and FOS. Thus, an intervention aimed at improving this symptom cluster and implemented while patients are undergoing chemotherapy could have numerous benefits.[40]

Using Bayes information criteria, we were able to confirm our hypothesis that inter-relationships between symptoms and outcomes would be different in the chemotherapy treatment phase (BL-C4) as compared with post-chemotherapy (C4-Y1). It is interesting to note similarities and differences between these networks. Post-chemotherapy quality of life and year 1 quality of life were each influenced by mood at the same time point, but post-chemotherapy quality of life was also influenced by concurrent fatigue which was not the case for year 1 quality of life, suggesting that interventions to improve mood during chemotherapy could improve post-chemotherapy quality of life. Similarly, while prior functional outcomes of sleepiness (FOS) score influenced subsequent level at all time points, post-chemotherapy FOS score was strongly influenced by concurrent fatigue, again suggesting that an intervention to reduce fatigue during chemotherapy could improve functional outcomes of sleepiness in breast cancer survivors. Additional differences between the networks are evident in Figure 1, but these differences were not reliable (arc strengths of these differing edges were <0.7).

6.2 | Strengths and limitations

The strengths of this work include a well-characterized cohort of patients undergoing chemotherapy, the use of bootstrap methods, and model averaging, which reduce overfit and improve replicability. There are also limitations. Our sample size is modest and may have impeded our ability to discern important links. Also, our study did not collect self-reported pain, a factor that could have influenced the

observed findings. Furthermore, Bayesian networks are inherently exploratory. Hence, these results need to be confirmed in other cohorts with larger sample sizes that include broad symptom inventories, including pain, and implement alternative computational strategies such as moderated regression.[23,24] We used an established NP battery which affords the opportunity to evaluate objective cognitive performance during chemotherapy. However, self-reported cognitive deficits are commonly noted by cancer patients during treatment. It would be interesting to investigate if networks for self-reported versus objective cognition are similar, and we leave this question and other similar ones (eg, comparing objective sleep assessed via actigraphy to self-reported sleep) for a future study.

7 | CONCLUSIONS

In this article, we have introduced Bayesian networks, a machine learning methodology, to infer networks of symptom cluster and cognitive and psychosocial outcomes for breast cancer patients during and 1 year after undergoing chemotherapy. Our results identified separate pathways and potential links between symptoms, cognitive function and QoL. The network comparison analysis strongly favored the fitted networks, indicating that our findings are robust against alternative network structures. Our work illustrates that Bayesian networks could be a powerful tool in cancer symptoms research; we advocate their use in future studies.

ACKNOWLEDGEMENTS

This study was supported by a UCSD Clinical and Translational Research Institute (CTRI) Pilot grant (2012–2013). The parent study was supported by the National Cancer Institute (CA112035 and P30 CA023100), the UCSD CTRI (UL1RR031980), and the Department of Veterans Affairs San Diego Center of Excellence for Stress and Mental Health (CESAMH).

ORCID

Loki Natarajan http://orcid.org/0000-0001-5719-828X

REFERENCES

1. Sanford SD, Beaumont JL, Butt Z, Sweet JJ, Cella D, Wagner LI. Prospective longitudinal evaluation of a symptom cluster in breast cancer. *J Pain Symptom Manage*. 2014;47(4):721-730.
2. Wu HS, Harden JK. Symptom burden and quality of life in survivorship: a review of the literature. *Cancer Nurs*. 2015;38(1):E29-E54.
3. Guidelines NCCN. Cancer-related fatigue v 1.2013. 2013.
4. Ancoli-Israel S, Liu L, Marler MR, et al. Fatigue, sleep, and circadian rhythms prior to chemotherapy for breast cancer. *Support Care Cancer*. 2006;14(3):201-209.
5. Beck SL, Berger AM, Barsevick AM, Wong B, Stewart KA, Dudley WN. Sleep quality after initial chemotherapy for breast cancer. *Support Care Cancer*. 2010;18(6):679-689.
6. Bower JE. Behavioral symptoms in patients with breast cancer and survivors. *J Clin Oncol*. 2008;26(5):768-777.
7. Berger AM, Mitchell SA. Modifying cancer-related fatigue by optimizing sleep quality. *J Natl Compr Canc Netw*. 2008;6(1):3-13.
8. Wielgus KK, Berger AM, Hertzog M. Predictors of fatigue 30 days after completing anthracycline plus taxane adjuvant chemotherapy for breast cancer. *Oncol Nurs Forum*. 2009;36(1):38-48.
9. Ancoli-Israel S, Liu L, Rissling M, et al. Sleep, fatigue, depression, and circadian activity rhythms in women with breast cancer before and after treatment: a 1-year longitudinal study. *Support Care Cancer*. 2014;22(9):2535-2545.
10. Byar KL, Berger AM, Bakken SL, Cetak MA. Impact of adjuvant breast cancer chemotherapy on fatigue, other symptoms, and quality of life. *Oncol Nurs Forum*. 2006;33(1):E18-E26.
11. O'Donnell JF. Insomnia in cancer patients. *Clin Cornerstone*. 2004;6(Suppl 1D):S6-S14.
12. Avis NE, Levine BJ, Case LD, Naftalis EZ, Van Zee KJ. Trajectories of depressive symptoms following breast cancer diagnosis. *Cancer Epidemiol Biomarkers Prev*. 2015;24(11):1789-1795.
13. So WK, Marsh G, Ling WM, et al. The symptom cluster of fatigue, pain, anxiety, and depression and the effect on the quality of life of women receiving treatment for breast cancer: a multicenter study. *Oncol Nurs Forum*. 2009;36(4):E205-E214.
14. Massie MJ. Prevalence of depression in patients with cancer. *J Natl Cancer Inst Monogr*. 2004;32:57-71.
15. Liu L, Fiorentino L, Natarajan L, et al. Pre-treatment symptom cluster in breast cancer patients is associated with worse sleep, fatigue and depression during chemotherapy. *Psychooncology*. 2009;18(2):187-194.
16. Andreotti C, Root JC, Schagen SB, et al. Reliable change in neuropsychological assessment of breast cancer survivors. *Psychooncology*. 2016;25(1):43-50.
17. Collins B, MacKenzie J, Tasca GA, Scherling C, Smith A. Cognitive effects of chemotherapy in breast cancer patients: a dose-response study. *Psychooncology*. 2013;22(7):1517-1527.
18. Hedayati E, Alinaghizadeh H, Schedin A, Nyman H, Albertsson M. Effects of adjuvant treatment on cognitive function in women with early breast cancer. *Eur J Oncol Nurs*. 2012;16(3):315-322.
19. Dwek MR, Rixon L, Hurt C, Simon A, Newman S. Is there a relationship between objectively measured cognitive changes in patients with solid tumours undergoing chemotherapy treatment and their health-related quality of life outcomes? A systematic review. Psychooncology. 2016.
20. Ganz PA, Kwan L, Castellon SA, et al. Cognitive complaints after breast cancer treatments: examining the relationship with neuropsychological test performance. *J Natl Cancer Inst*. 2013;105(11):791-801.
21. Mo C, Lin H, Fu F, et al. Chemotherapy-induced changes of cerebral activity in resting-state functional magnetic resonance imaging and cerebral white matter in diffusion tensor imaging. *Oncotarget*. 2017;8:81273-81284.
22. Vearncombe KJ, Rolfe M, Wright M, Pachana NA, Andrew B, Beadle G. Predictors of cognitive decline after chemotherapy in breast cancer patients. *J Int Neuropsychol Soc*. 2009;15(6):951-962.
23. Francoeur RB. Using an innovative multiple regression procedure in a cancer population (Part 1): detecting and probing relationships of common interacting symptoms (pain, fatigue/weakness, sleep problems) as a strategy to discover influential symptom pairs and clusters. *Onco Targets Ther*. 2015;8:45-56.
24. Francoeur RB. Using an innovative multiple regression procedure in a cancer population (Part II): fever, depressive affect, and mobility problems clarify an influential symptom pair (pain-fatigue/weakness) and cluster (pain-fatigue/weakness-sleep problems). *Onco Targets Ther*. 2015;8:57-72.
25. Pearl J. *Probabilistic Reasoning in Intelligent systems: Networks of Plausible Inference*. Burlington, MA: Morgan Kaufmann; 2014.
26. Tegner J, Bjorkegren J. Perturbations to uncover gene networks. *Trends Genet*. 2007;23(1):34-41.
27. Verduijn M, Peek N, Rosseel PM, de Jonge E, de Mol BA. Prognostic Bayesian networks I: rationale, learning procedure, and clinical use. *J Biomed Inform*. 2007;40(6):609-618.

28. Nagarajan R, Scutari M, Lèbre S. *Bayesian Networks in R*. New York: Springer-Verlag; 2013.
29. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193-213.
30. Stein KD, Martin SC, Hann DM, Jacobsen PB. A multidimensional measure of fatigue for use with cancer patients. *Cancer Pract*. 1998;6(3):143-152.
31. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psychol Measur*. 1977;1:385-401.
32. Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*. 1997;20(10):835-843.
33. Wechsler D. *Wechsler Adult Intelligence Scale—Third Edition (WAIS-III). Administration and Scoring manual San Antonio*. TX: The Psychological Corporation; 1997.
34. Reitan RM, Wolfson D. *The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation*. 2nd ed. Tucson, AZ: Neuropsychology Press; 1993.
35. Benedict R, Schretlen D, Groninger L, Brandt J. Hopkins Verbal Learning Test-revised: normative data and analysis of inter-form and test-retest reliability. *Clin Neuropsychol*. 1998;12:43-55.
36. Kongs SK, Thompson LL, Iverson GL, Heaton RK. *Wisconsin Card Sorting Test-64 Card Version. Professional Manual*. Odessa, FL: Psychological Assessment Resources, Inc.; 2000.
37. Golden CJ, Freshwater SM. *Stroop Color and Word Test: A Manual for Clinical and Experimental Uses*. 2nd ed. Stoelting Co: Wood Dale, IL; 2002.
38. Gladsjo JA, Schuman CC, Evans JD, Peavy GM, Miller SW, Heaton RK. Norms for letter and category fluency: demographic corrections for age, education, and ethnicity. *Assessment*. 1999;6(2):147-178.
39. Wasserman L. Bayesian model selection and model averaging. *J Math Psychol [Article]*. 2000;44(1):92-107.
40. Ho SY, Rohan KJ, Parent J, Tager FA, McKinley PS. A longitudinal study of depression, fatigue, and sleep disturbances as a symptom cluster in women with breast cancer. *J Pain Symptom Manage*. 2015;49(4):707-715.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Xu S, Thompson W, Ancoli-Israel S, Liu L, Palmer B, Natarajan L. Cognition, quality-of-life, and symptom clusters in breast cancer: Using Bayesian networks to elucidate complex relationships. *Psycho-Oncology*. 2018;27:802–809. https://doi.org/10.1002/pon.4571