

Bayesian Networks and Causal Inference for the Interpretation of Patients' Symptom Experience

Nikolaos Papachristou (✉ nikospapachristou@auth.gr)

Aristotle University of Thessaloniki (A.U.Th.)

Payam Barnaghi

Imperial College London

Sarah J. Allison

University of Surrey

Kord M. Kober

University of California, San Francisco

Christine Miaskowski

University of California, San Francisco

Research Article

Keywords: Bayesian Networks, Causal Inference, Interpretation, Symptom

Posted Date: May 6th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-491010/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Bayesian Networks and Causal Inference for the Interpretation of Patients' Symptom Experience

Nikolaos Papachristou^{1,4}, Payam Barnaghi^{2,3,4}, Sarah J. Allison⁵, Kord M. Kober⁶, and Christine Miaskowski⁶

¹Lab of Medical Physics, School of Medicine, Aristotle University of Thessaloniki (A.U.Th.), Greece

²The Department of Brain Sciences, Imperial College London

³Care Research & Technology Centre, The UK Dementia Research Institute (UK DRI)

⁴Centre for Vision, Speech and Signal Processing (CVSSP), University of Surrey, Guildford, UK

⁵School of Bioscience & Medicine, University of Surrey, Guildford, United Kingdom

⁶University of California, San Francisco, USA

ABSTRACT

Given that oncology patients experience an average of 14 co-occurring symptoms, the identification of sentinel or core symptoms is a critical need for effective symptom management. However, this task is an extremely challenging one. To address this need, we used Bayesian Network Analysis (BNA) approaches on a comprehensive dataset of 38 distinct and co-occurring cancer symptoms from a sample of 1328 cancer patients receiving chemotherapy. We evaluate three classes of algorithms (constrained-based, score-based, hybrid algorithms) on their structural stability and performance with different sample sizes and demographic characteristics (gender and age). We demonstrate their potential clinical use for casual discovery and inference. The hybrid algorithm identified more dense network structures that were clinically relevant, not only for the entire sample but for our case studies of gender and age. Our work demonstrates how to process a comprehensive set of symptom experience data for interactions and provides valuable information on the use of BNA in clinical practice when different algorithms, sample sizes and comparative sub-groups are evaluated.

Introduction

Cancer and its treatments create a significant symptom burden for oncology patients regardless of tumor types, stage of disease, and/or phases of care.^{1,2} Oncology patients undergoing cancer treatment experience an average of 10 to 15 unrelieved symptoms that can be highly variable in their occurrence, severity and distress.³⁻⁵ Research on these symptoms and other patient reported outcomes (PROs) is needed to improve patient' quality of life and in some cases survival.^{1,2,6-11} Until now, studies that evaluated multiple co-occurring symptoms have used analytical techniques such as exploratory factor analysis or cluster analysis.¹²⁻¹⁴ One of the main research hypotheses in this field is that symptoms that co-occur or cluster together may share common underlying mechanisms that are potential targets for therapeutic interventions.¹⁵ Nevertheless, the existing research on multiple co-occurring symptoms and symptom clusters has several limitations.¹² The current analytical approaches are limited to an evaluation of the nature of the relationships between and among individual symptoms and multiple-occurring symptoms. However, these approaches do not permit an examination of how the presence of one or more symptoms may exert an influence on other co-occurring symptoms. The accurate identification of these sentinel or core symptoms may provide target(s) for early therapeutic interventions to decrease patients' symptom burden. In this study, we explore the application of Bayesian Network Analysis (BNA) to evaluate the direction of the relationships between and among multiple co-occurring symptoms in oncology patients receiving chemotherapy (CTX). Being one of the first BNA studies in the domain of cancer symptoms, we explore how BNA can be used to map the pathways that illustrate how symptoms interact and influence each other. In addition, we elaborate on key analytical aspects that are needed to increase the reproducibility of this work as well as for the extension/comparison of clinical findings to different samples of cancer patients.

BNA is an analytical approach that can be used to explore multivariate relationships with intuitive graphs, also known as causal graphs, causal diagrams, or directed acyclic graphs (DAGs). Bayesian networks (BNs) are probabilistic graphical models that represent a set of variables (e.g., cancer symptoms) and their conditional dependencies using a DAG.¹⁶ Using BNA, the multivariate relationships (i.e. joint probability distributions) between these variables are quantified and visualised graphically, with nodes and arcs (see Materials and Methods, Figure 2), making the discovery and analysis of their inter-relatedness more comprehensive. In principle, such causal graphs can be used to answer questions related to: (i) causal discovery (e.g., which symptom(s), among many co-occurring symptoms, can influence another distinct symptom) and (ii) causal inference (e.g., of

what magnitude and causal connection - direct or confounding - these symptoms affect the aforementioned sentinel or core symptom).¹⁷ Because of BNA's ability to handle both observable as well as uncertain information (i.e., latent factors from observational data),^{18–20} BNA has been used in biomedical informatics for more than two decades.

BNs are suitable tools for the probabilistic inference that can aid in clinical decision making for several reasons.^{18,21–23} For example, they provide a visual representation of causality that can be easily understood by a clinician.^{22,24,25} They can integrate knowledge from clinical experts into their modelling procedures while learning the structure and parameters of a network through the available data.^{22,23,26} In addition to their capacity to provide information on observational and causal inferences,^{22,27} BNs can extract parameter estimates through the representation of the joint probability space among the nodes or clinical variables (e.g., symptoms) in a model.²² Likewise, BNs can be used with and perform well with incomplete data.^{22,28–30}

With respect to their healthcare applications, BNs are commonly used in genomics and system biology.^{31–34} The usefulness of their causal diagrams has been explored and applied in epidemiology, public health and health services research.^{35–37} Causal diagrams have been used to control for and identify confounding variables;^{38,39} estimate effect sizes from observational data;¹⁷ select features that are most predictive of a target variable;⁴⁰ and test for external validity,⁴¹ fault diagnosis and health monitoring,⁴² missing data,⁴³ and selection bias.³⁹ Notably, BN applications supported decision making in a wide spectrum of the clinical practice activities (e.g., diagnosis, treatment, prognosis, risk modeling).²¹ In cancer research, BN applications have been used in survival prediction and treatment selection for patients with lung and colon cancer;^{22,28–30} mammography interpretation;⁴⁴ detection, diagnosis and prognosis of breast cancer;^{22,28–30} modelling and predicting the occurrence of brain metastasis from lung cancer⁴⁵ and revealing the genetic and environmental basis of bladder cancer.¹⁶ They have been even applied to model and assess the inter-relationships between life-style, health-related factors and quality of life of cancer patients.^{46,47}

Nevertheless, the use of BNs in cancer symptom management research is extremely limited. To our knowledge, only one longitudinal study used BNA to investigate the effect of sleep disturbance, fatigue, and mood disturbance on cognition and quality of life in breast cancer patients before, during, and after CTX.⁴⁷ While this study illustrates the use of BNs to evaluate the relationships among cancer symptoms, it does not provide an extensive summary of the types of BN algorithms that can be used. In addition, it does not specify how the proposed model can drive data-driven domain knowledge discovery when various symptom structures are unknown. Finally, it does not elaborate on the analytical challenges that such clinical applications may have. While this study provides important information on the relationships among three common co-occurring symptoms and quality of life,⁴⁷ the sample size was small ($n=74$) and included only patients with breast cancer. In addition, only three symptoms were evaluated in the BNA. Given that oncology patients receiving CTX report an average of 14 co-occurring symptoms,³ the goals of our study are multifold and use a broader approach than the previous study. We describe how BNA methods can be used to explore the interrelationships among 38 co-occurring symptoms in a large, clinically representative sample of 1328 cancer patients undergoing CTX. In addition, we show how these methods can be used to build new clinical knowledge by highlighting notable differences in the BNs created using gender and age (i.e. <65 versus >65 years of age based on the definition of an older person from the World Health Organisation).⁴⁸ Finally, we explain how our findings can be used to guide the development and testing of future interventions to improve symptom management in oncology patients.

Our contributions in this paper are two-fold. First, we implemented three different families of BNA algorithms using cross-sectional data on the occurrence of 38 symptoms, as well as on symptom occurrence data dichotomised by gender and age, to assess and compare their structure, performance and stability. Second, we provide the analytical implications, the clinical interpretations and the significance of our findings in relationship to these 38 common co-occurring symptoms in cancer patients receiving CTX. Our work provides the first case study on how to process a comprehensive set of 38 cancer symptoms to assess their interactions and explore potential target symptoms for early interventions. An additional contribution to the field of symptom research lies in the demonstration of how to utilise findings from BNs in clinical practice when using different algorithms, sample sizes and comparative sub-groups.

Results

Sample Characteristics: Of the 1328 ($n_{\text{Complete}}=1328$) patients in this study, 77.8% were female ($n_{\text{Female}}=1033, n_{\text{Male}}=295$) and 72.1% were under the age of 65 ($n_{\text{Age} < 65}=1033, n_{\text{Age} > 65}=295$), with a mean age of 57.3 (± 12.3) years. The majority of the patients had breast (40.2%) or gastrointestinal (30.6%) cancer. The remaining patients had gynaecological cancer (17.5%) and lung cancer (11.7%). These 1328 patients reported an average of 13.9 (± 7.2) symptoms prior to their second or third cycle of CTX. Sample characteristics and symptom occurrence rates are summarised in Appendix A, Table S1 and Online Figures 1 through 5,⁴⁹ respectively.

Table 1. Comparison of Bayesian Network Algorithms

Compared algorithms	Threshold	Hamming Distance (HD)	Structural Hamming Distance (SHD)	Common Arcs found with both algorithms	Arcs found only with the first (1 st) algorithm	Arcs found only with the second (2 nd) algorithm
Grow Shrink (1st) vs Hill-Climbing (2nd)	0.85	27	32	4	6	25
	Statistical Optimisation	59	64	8	14	51
Grow Shrink (1st) vs Max-Min Hill-Climbing (2nd)	0.85	27	32	5	5	24
	Statistical Optimisation	61	66	9	13	53
Hill-Climbing (1st) vs Max-Min Hill-Climbing (2nd)	0.85	0	0	28	1	1
	Statistical Optimisation	2	5	57	2	5

Grow Shrink (GS) Algorithm

To learn the structure of the 38 different symptoms using the GS algorithm, we set the nominal type I error rate to 0.05. We averaged 10,000 BNs and parsed their edges with the two different statistical criteria described in the Methods section.

Both of these networks were relatively sparse, with the majority of the symptoms not connected by any edges. The network that was parsed with a 0.85 threshold identified 10 directed arcs. The network parsed with the optimal threshold (0.4963), described by Scurati and Nagarajan,⁵⁰ identified 22 directed arcs (Appendix A, Table S2). The results are illustrated in Online Figures 6 and 7.⁴⁹

Hill-Climbing (HC) Algorithm

To learn the structure of the 38 different symptoms using the HC algorithm we set its hyperparameters to have 50 random restarts and 30 attempts to randomly insert/remove/reverse an arc on every random restart. We constructed and averaged 10,000 BNs and parsed their edges with the two different statistical criteria described in the Methods section. The results are illustrated in Online Figures 8 and 9.⁴⁹

Compared to the GS algorithm, the BNs learned using the HC algorithm were denser (Online Figures 12 and 13⁴⁹). As illustrated in Online Figure 8,⁴⁹ the network that was parsed with a 0.85 threshold identified 29 directed arcs and seven symptoms (i.e, numbness/tingling in hands/feet, problems with urination, itching, dizziness, difficulty swallowing, constipation, swelling of arms and legs) were not included in the BN. The network that was parsed using the optimal threshold (0.4989), described by Scurati and Nagarajan,⁵⁰ identified 59 directed arcs (Appendix A, Table S2). The only symptom not included within the network was swelling of the arms and legs (Online Figure 9⁴⁹).

Max Min Hill-Climbing (MMHC) Algorithm

To learn the structure of the 38 different symptoms with the MMHC algorithm, we set its hyperparameters to have a nominal type I error rate to 0.05, during the first phase of the algorithm. In addition, we included 50 random restarts and attempted 30 times to randomly insert/remove/reverse an arc on every random restart, during the second phase. We constructed and averaged 10,000 BNs and parsed their edges with different statistical criteria described in the Methods section. The results are illustrated in Appendix A, SF1, SF2, Online Figures 10 and 11.⁴⁹

Similar to the findings with the HC algorithm, the networks learning with the MMHC algorithm were denser in comparison to the networks learned with the GS algorithm and the HC algorithm (Online Figures 14 through 17⁴⁹). As illustrated in Online Figure 10,⁴⁹ the network that was parsed with a 0.85 threshold identified 29 directed arcs and seven symptoms (i.e, numbness/tingling in hands/feet, problems with urination, itching, dizziness, difficulty swallowing, constipation, swelling of arms and legs) were not included in the BN. The network that was parsed using an optimal threshold (0.5010), described by Scurati and Nagarajan,⁵⁰ identified 62 directed arcs (Appendix A, Table S2). The only symptom not included within the network was swelling of the arms and legs (Online Figure 11⁴⁹).

Bayesian Network Algorithms Comparison

To assess the goodness of fit and findings of the networks constructed, we compared GS, HC, and MMHC based on different scoring functions such as k2, BDE, AIC, BIC and logik (Appendix A, Table S2).

Based on these functions, both HC and MMHC performed better than GS. In addition, their performance was similar for all five scoring criteria with MMHC performing better overall when applied to our smaller and larger subgroup analyses (Appendix A, Table S2). Using the complete dataset of all cancer patients, the findings for HC and MMHC were similar (see Table 1). Depending on the threshold used (i.e., 0.85 or optimal threshold), the Hamming Distance (HD) of their networks were 0 and 2 respectively, and their Structural Hamming Distance (SHD) were 0 and 5, respectively. Most of their edges were common as illustrated in Figure 1 (see also Online Figures 12 to 17⁴⁹).

Stability of the Identified BNs

To evaluate the stability of the constructed BNs, we compared the total number of arcs identified in four bootstrap sub-samples ($n_1=332$, $n_2=664$, $n_3=996$, $n_4=1328$). In addition, we compared their similarity based on their common arcs. The averaged BNs were created in a similar way, as explained above. The results are shown in Appendix A, Table S2.

In general, MMHC provided more stable results across the four bootstrap sub-samples ($n_1=332$, $n_2=664$, $n_3=996$, $n_4=1328$). Based on the scoring function, MMHC provided a better fit for the data and it contributed the highest number of common arcs across the procedure (Appendix A, Table S2). Of note, all of the algorithms seemed to be affected by the sample size, with the score and the number of arcs increasing as the sample size increased (Appendix A, Table S2).

Table 2. Comparison of the Bayesian Networks identified using four different subgroups of cancer patients based on gender and age, using the MMHC algorithm.

Group	Sample Size	Threshold	k2	BDE	AIC	BIC	logik	Number of arcs found
Cancer patients below 65	958	0.85	-20522.64	-20524.38	-20391.44	-20547.11	-20327.44	22
	958	0.5033	-19820.67	-19822.95	-19611.51	-19922.86	-19483.51	59
	370	0.85	-20768.43	-20771.37	-20655.08	-20781.56	-20603.08	14
	370	0.5003	-20281.32	-20283.19	-20139.22	-20311.92	-20068.22	27
Cancer patients above 65	370	0.85	-7911.7	-7919.385	-7824.7	-7922.537	-7774.7	12
	370	0.4929	-7525.235	-7534.852	-7397.404	-7589.166	-7299.404	42
Male cancer patients	295	0.85	-5985.751	-6005.247	-5905.206	-5993.694	-5857.206	10
	295	0.5006	-5778.554	-5794.827	-5663.553	-5822.093	-5577.553	37
Female cancer patients	1033	0.85	-22202.87	-22204.15	-22064.96	-22232.93	-21996.96	25
	1033	0.5023	-21452.37	-21454.11	-21240.93	-21557.1	-21112.93	59
	295	0.85	-22675.85	-22677.4	-22564.81	-22688.32	-22514.81	12
	295	0.5012	-22127.84	-22129.48	-21990.69	-22156.19	-21923.69	25
k2: logarithm of the K2 score BDE: logarithm of the Bayesian Dirichlet Equivalent score AIC: Akaike Information Criterion score BIC: Bayesian Information Criterion score logik: multinomial log-likelihood score								

Bayesian Networks for Male and Female Cancer Patients

To demonstrate the potential applications and challenges of BNA in cancer symptom research, we applied the MMHC algorithm to four different subgroups of cancer patients based on their gender and age.

The sample sizes for the two subgroups of male and female cancer patients were not balanced ($n_{\text{male}}=295$, $n_{\text{female}}=1033$). Checking their BNs' similarities provides a view that female cancer patients experience more symptoms than male patients

(Table 2, 3). More specifically, women seem to experience up to 59 different symptom interactions, while men experience up to 37 symptom interactions. Male and female patients experience up to 14 different symptoms in common (Table 3).

However, when we created the BNs using bootstrapped, random samples of female cancer patients of equal size with the male cancer patients (n=295), this view reversed (Table 2). In this case, women seemed to experience up to 25 different symptom interactions compared to 37 in men.

Table 3. Comparison of the Bayesian Networks identified on four different subgroups of cancer patients based on gender and age, using the MMHC algorithm.

Sub-Group 1	Sub-Group 2	Sample size of sub-group 1	Sample size of sub-group 2	Threshold	Hamming Distance (HD)	Structural Hamming Distance (SHD)	Number of common arcs	Number of arcs found only in sub-group 1	Number of arcs found only in sub-group 2
Men	Women	295	1033	0.85	17	20	6	4	19
Men	Women	295	1033	Statistical Optimisation	56	67	14	23	45
Men	Women	295	295	0.85	6	6	6	4	6
Men	Women	295	295	Statistical Optimisation	34	41	8	29	17
Women	Women	1033	295	0.85	13	16	10	15	2
Women	Women	1033	295	Statistical Optimisation	34	48	20	39	5
Age < 65	Age > 65	958	370	0.85	12	15	7	15	5
Age < 65	Age > 65	958	370	Statistical Optimisation	53	62	18	42	24
Age < 65	Age > 65	370	370	0.85	6	6	8	6	4
Age < 65	Age > 65	370	370	Statistical Optimisation	35	45	13	14	29
Age < 65	Age < 65	958	370	0.85	8	10	12	10	2
Age < 65	Age < 65	958	370	Statistical Optimisation	32	48	25	35	2

These results are illustrated in Appendix A, Figures SF3, SF4 and Online Figures 18 through 23.⁴⁹ The differences among the identified BNs between the two aforementioned groups of cancer patients are illustrated in Online Figures 24 through 29.⁴⁹ From these diagrams we can see that when we keep the sub-groups disproportional women seem to experience a more interconnected network of symptoms than men. In contrast, when we balance the gender sub-groups with a similar sample size the findings reverse with men experiencing a more interconnected network of symptoms than women.

Bayesian Network for younger (< 65 years) and older (> 65 years) cancer patients

In a similar fashion, we compared the BNs identified using the MMHC algorithm, between older and younger cancer patients (n_{Age<65}=958, n_{Age>65}=370). The results are illustrated in Appendix A, Figure SF5, SF6 and Online Figures 30 through 35.⁴⁹ The differences among the identified BNs between the two aforementioned groups of cancer patients are shown in Online Figures 36 through 41.⁴⁹

In our initial comparison, keeping the age sub-groups disproportional, younger cancer patients seem to experience a more interconnected network of symptoms than older ones. More specifically, younger people seem to experience up to 59 different symptom interactions, while older ones experience up to 42 symptom interactions. Younger and older cancer patients experience up to 18 different symptoms in common (Table 3).

As in our previous example, when we created the BNs using bootstrapped, random samples of younger cancer patients of equal size with the older cancer patients (n=370), this view reversed (Table 2). The comparison with the balanced sample sizes showed that older patients experience a more interconnected network of symptoms. In this case, younger cancer patients seem to experience up to 27 different symptom interactions compared to 42 interactions in older ones. Younger and older cancer patients experience up to 13 different symptoms in common (Table 3).

Discussion

In this study, we compared the network structures that were learned using constraint-based, score-based, and hybrid algorithms.⁵¹ The hybrid algorithm identified more dense network structures that were clinically relevant, not only for the entire sample but for our case studies of gender and age differences. While our sample size was relatively small (N=1328), and we evaluated the three methods on a constrained set of confounding parameters, our results show that the score-based and hybrid algorithms identified relatively similar structures. The constraint-based algorithm identified less sparse networks which may be the results of the relatively small sample size.⁵² Future analyses, with a broader set of confounding factors (e.g. specific chemotherapy regimens, supportive care interventions) for structure-learning as well as different subgroups of patients (e.g., different cancer types, different cancer treatments) may provide additional insights into the use of BNA in cancer symptom research. In addition, if our findings are confirmed regarding the existence of common core or sentinel symptoms identified in the BNs of different subgroups of cancer patients, this information can be used to develop and test interventions to decrease symptom burden in oncology patients. One can hypothesize that if an intervention is directed at the core/sentinel symptom, co-occurring symptoms associated with this “parent” symptom will decrease as well.

Given the cross-sectional design and the need for replication, one needs to proceed with caution regarding the clinical implications of this study. However, an evaluation of the various edges in Figure SF2, Appendix A, illustrates relationships that are commonly observed in oncology patients. For example cancer patients will often report the co-occurrence of difficulty breathing and shortness of breath.⁵³ In a similar fashion, oncology patients can experience sweats and hot flashes from a variety of therapeutic interventions.^{54,55} Another set of interesting findings from our study is the age and gender differences in the density of BNs when the sample sizes were controlled for in the BNA. While most of the studies of age and gender differences in symptom burden found higher rates and severity of symptoms in younger^{56,57} and female^{56,58} patients, the majority of these studies were cross-sectional and did not account for the inter-relationships among multiple co-occurring symptoms. These findings warrant replication in future studies with cancer types that occur equally in men and women (e.g., gastrointestinal, lung) and in younger and older patients with the same cancer diagnosis who are undergoing similar types of treatments.

We have listed methods and ways in which BNA could inform symptom management research. Among the strengths of this study is the use of an extensive set of cancer symptoms from a comprehensive sample of oncology patients; the complete evaluation of BNA with different algorithms, performance, stability criteria, and contrasting sub-groups of patients; and the use of bootstrap methods and different scoring functions which can reduce overfit and improve replicability. Nevertheless, several limitations warrant consideration. We do not know what the true structure of patients’ cancer symptom experiences. Also, our analyses lack the inclusion of potentially important confounding factors (e.g., specific chemotherapy regimens) that could increase or decrease the occurrence of specific symptoms. Nevertheless, BNs are exploratory tools that are well suited for hypothesis generation. Markov blankets allow us to identify factors that directly influence each symptom and can be regarded a part of a causal (mechanistic) model. Despite the fact that it is very difficult to prove causality from cross-sectional observational data, BNs can provide insights for specific causal models that could be validated in experimental data and/or in longitudinal studies. Therefore, our findings need to be verified in independent cohorts and/or in longitudinal studies.

BNs and network analysis approaches⁵⁹ are promising in detecting patterns among symptoms. The various BNA approaches have the potential to identify possible causal pathways, with predictive value inside the Markov blankets of each network. To take full advantage of this network approach in symptom management research, we need to explore its utility using additional sets of data modelling parameters (e.g., new algorithms, hyper-parameters, statistical criteria). In addition, a need exists to study multiple and diverse sub-groups of patients, based on various demographic, clinical, and molecular characteristics. The application of BNA to a broader set of parameters, similar to those that we present in these analyses, may provide new evidence regarding the mechanisms that underlie multiple co-occurring symptoms in oncology patients undergoing CTX.

Breinan et al.⁶⁰ distinguished two cultures of statistical modelling, namely algorithmic and data modelling. The first one aims to optimise the predictions of the modelling produced from a given set of inputs while treating the process itself as a black box. The second approach focuses on the identification of patterns between all the given inputs and outputs and an explanation of how the specific modelling works. In the context of data modelling for cancer symptoms, algorithmic modelling is crucial for implementing state-of-the-art predictions that can be used in clinical practice.⁶¹ In addition, data modelling can be a useful tool to expand conceptual knowledge about the domain itself.^{3,59} The use of hybrid approaches such as BNA, that combines the strengths of both analytic fields, provides additional tools for both cancer research and care. Its use on temporal data will offer more in-depth insights on the ways in which cancer symptoms interact with each other both in causal and transitional ways.

Material and Methods

Dataset Description

This paper reports on a secondary analysis of data from a longitudinal study of the symptom experience of oncology outpatients receiving CTX. The methods for this study are described in detail in our previous publications.^{62–64} For the BNA experiments, enrollment data from the parent, longitudinal study were analysed ($n_{\text{Complete}}=1328$). Patients were eligible to participate if they: were 18 years of age or older; had a diagnosis of breast, gastrointestinal, gynecological, or lung cancer; had received CTX within the preceding four weeks; were scheduled to receive at least two additional cycles of CTX; were able to read, write, and understand English; and gave written informed consent. Patients were recruited from two Comprehensive Cancer Centers, one Veteran’s Affairs hospital, and four community-based oncology programs. This study was approved by the Committee on Human Research at the University of California, San Francisco, and was implemented in accordance with the Declaration of Helsinki. Written informed consent was obtained from all patients.

Symptom dimensions

To determine the occurrence of 38 symptoms commonly associated with cancer and its treatment, a modified version of the Memorial Symptom Assessment Scale (MSAS)⁶⁵ was used. In addition to the original 32 MSAS symptoms, the following six symptoms were assessed: hot flashes, chest tightness, difficulty breathing, abdominal cramps, increased appetite, and weight gain. Using the MSAS, patients were asked to indicate whether or not they had experienced each symptom in the past week (i.e., symptom occurrence; yes= presence of a symptom, no = absence of symptom). The validity and reliability of the MSAS are well established in studies of oncology inpatients and outpatients.⁶⁵

Bayesian Networks

Bayesian Networks (BNs), also known as a Bayesian Belief Networks (BBNs) or Belief Networks,⁶⁶ are graphical models that represent the relationships among multiple variables and describe the conditional dependencies between them. They comprise a network structure in the form of a DAG and a set of conditional probabilities, one for each variable, that characterize the dependencies represented by the DAG’s edges (see Figure 2).

In BNA terminology, when we have a directed edge (i.e., arc or arrow) going from node X to node Y, node X is considered to be a parent of Y. Accordingly, node Y is considered a child of X. Every variable (i.e., node) inside a DAG, given the values of its parents, is considered to be conditionally independent of the set of all its predecessors in this graph.^{16,67} Based on the Markov Condition when two nodes are not connected directly through an arrow, we consider these two variables to be independent. Taking advantage of the aforementioned graphical structure and the chain rule of probability calculus the total joint probability distribution of the group of variables represented inside a BN can be expressed as the product of their conditional probabilities:¹⁶

$$p(x|\theta) = \prod_{i=1}^n p(x_i | pa(x_i), \theta_i) \quad (1)$$

where $x = \{x_1, \dots, x_n\}$ are the variables (i.e., the nodes inside a BN) and $\theta = \{\theta_1, \dots, \theta_n\}$ are the BN’s parameters. Each i is the set of parameters necessary to specify the distribution of the variable x_i given its parents $pa(x_i)$.¹⁶

The structure mentioned above can support researchers and clinicians to understand the main cause(s) of a certain problem (e.g., symptom). It can also be used to describe the probabilities of different effects (e.g., increased occurrence rates of other symptoms) given an action (i.e., presence of a core or sentinel symptom). To predict the value or behaviour of a specific node inside a DAG, we need its Markov Blanket. The Markov Blanket of a node (i.e., variable or symptom) is a set of nodes, consisting of its parent nodes, child nodes, and all of the other nodes that share a child with it. For example, in Figure 2, nodes

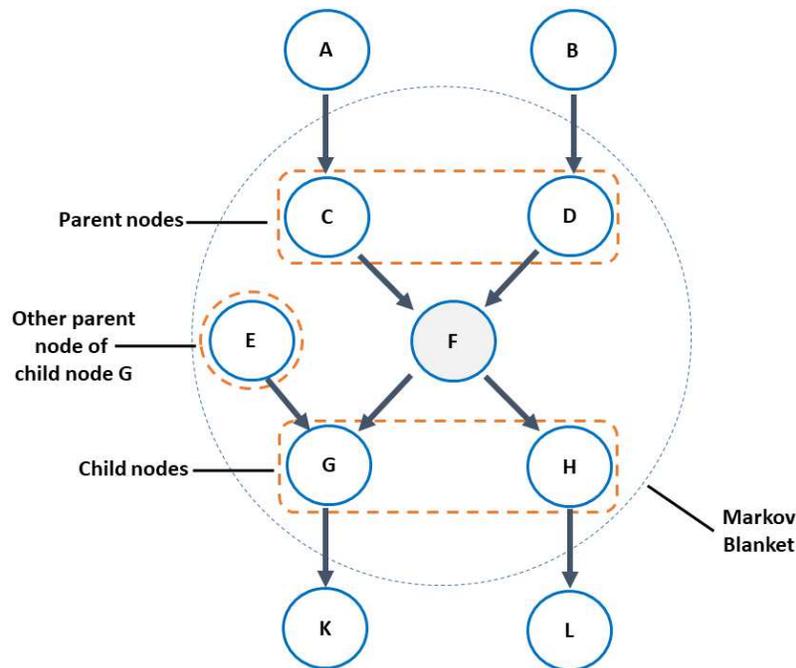


Figure 2. Bayesian Network: parents, child nodes and Markov Blanket of node A in a directed graph

C and D are the parent nodes and nodes G and H are the child nodes of node F. Node E is a parent node of G, sharing the same child node with node F. Thus, the Markov Blanket of variable F is the group of all the aforementioned variables or nodes (i.e., C, D, F, E, G, H). The minimum set of nodes that are necessary to predict the behaviour of variable F are the ones included in this Markov Blanket (see Appendix B).⁶⁸

Bayesian Network Learning Algorithms

The process of identifying and understanding the structure of a BN from the available data is called Learning, and it requires two steps.⁶⁹ The first step, called Network Structure or Structure Learning (i.e., identification of the topology of the BN),⁶⁹ consists of finding the graph structure that encodes the conditional independencies present in the data. The second step, called Parameter Learning, deals with the estimation of the parameters of the global distribution given the identified network topology.⁶⁹

Regarding Network Structure Learning, three different types of algorithms exist for constructing BNs from observational data, namely, constraint-based, score-based, and hybrid algorithms.⁵¹ In Parameter Learning, two main categories of algorithmic approaches exist, one that is focused on complete data (e.g., Maximum Likelihood Estimate) and the other that is focused on incomplete data (e.g., Expectation-maximization, Monte-Carlo, Gaussian approximation).⁷⁰ In this study, we focus on network structure learning.

Structural Learning Algorithms

In this study, we use three well known structural learning algorithmic implementations, one for each of the aforementioned approaches. These algorithms are the Grow-shrink (GS, constraint-based), the Hill-Climbing (HC, score-based) and the Max-Min Hill-Climbing (MMHC, hybrid).⁵¹

Constraint-based algorithms focus on recognising the conditional independence (i.e., Markov Condition) of all of the observed variables between them.⁶⁹ They learn the network structure with conditional independence tests, such as χ^2 , to determine the absence of edges between variables, and then construct a graph, adding directions to edges that satisfy the d-separation criterion. Given a causal graph, d-separation is a criterion for deciding whether a set X of variables is independent of another set Y, given a third set Z. More specifically, if X, Y, Z are three disjoint sets of variables in a BN, “Z is said to d-separate X from Y if and only if Z blocks every path from a node in X to a node in Y”.⁷¹ Thus, **GS** is based on the recovery of the optimal Markov Blanket of every variable X, based on pairwise independence tests.⁷² Grow-shrink’s name refers to its two different phases, the growing and the shrinking one. Initially, the GS algorithm starts with a variable X and an empty set S.^{16,72,73} The growing phase adds variables to S as long as they are dependent on X, conditional on the rest of the variables in S. In the

subsequent shrinking phase, variables that violate the Markov blanket property of X are removed from S . Then, the algorithm identifies the local neighbourhood (direct parents and children) of each variable in the network within the Markov blanket to recover the exact structure around each variable. Edge directions are determined by examining the triples of variables using the d -separation criterion.

Score-based methods calculate a number of possible BN structures and allocate a score to each that measures how well it explains the observed set of data.¹⁶ These algorithms select the BN structure that maximizes this score. Because the number of potential structures is exponential to the number of nodes, even datasets with few variables have too many possible BN structures to allow for an exhaustive search. That is the reason why most score-based algorithms have employed heuristic search techniques, such as HC or simulated annealing.^{16,51} More specifically, a **HC** search starts from either an empty, full or possibly random network and existing prior knowledge can be used to seed the initial candidate network. The main loop consists of attempting every possible single-edge addition, removal, or reversal in a candidate network. The structure that increases the score, the most becomes the current candidate. The process iterates until a change in a single-edge does not increase the score.

Finally, **hybrid algorithms** were developed that combine the constraint- and score-based methods to maximise their advantages. They consist of two steps, called restrict and maximise.⁵¹ During the restrict step, a constraint-based algorithm is implemented to identify the skeleton of the network using conditional independence tests. During the maximise step, a score-based method is employed to maximise the score function of this network and identify the best set of edge orientations. In this manner, **MMHC** first learns the skeleton of a BN using a local discovery algorithm called Max-Min Parents and Children (MMPC). MMPC consists of a forward phase and a backward phase. In the forward phase, all variables with an edge to and from a variable are selected using a heuristic function. In the backward phase, all false positives chosen in the first phase are removed. After MMPC identifies the skeleton of the BN, a greedy Bayesian-scoring HC search is used to orient the edges. The algorithm has the advantages of reliably scaling up to thousands of variables in a reasonable amount of time.

Computation, Performance, Stability, Reproducibility and Structure Assessment

We implemented each of the analyses in this study using the *bnlearn* R package.⁷⁴ We implemented the visualisations using the *visNetwork* R package⁷⁵. The algorithms (i.e., GS, HC, MMHC) used were assessed for their performance (i.e., the goodness of fit with respect to particular score functions), stability (i.e., performance and topology across different sample sizes) and structure (i.e., discrepancies between the different algorithms).

Computation

To ensure the reliability of the identified DAGs, we bootstrapped 10.000 BNs, for each analysis, and averaged them to obtain the final, resultant network. To identify the possible true, high-confidence arcs (i.e., directed edges), we assessed the frequency of the presence of an arc (i.e., directed edge) across the aforementioned bootstrapped networks. The final visualised BNs contained the arcs that had a predefined frequency/inclusion threshold. These thresholds were based on both an ad-hoc approach and an 85% frequency, according to Sachs et al.⁷⁶ or a statistically motivated approach and statistically optimised frequency, according to Scutari et al.⁵⁰. In the first case, for example, for the final averaged BNs we kept only the edges that appeared in at least 85%⁷⁶ of the 10.000 bootstrapped networks (i.e., at least 8.500).

In order to keep only the edges that were significant in a specific direction, we kept the ones that pointed from symptom X to symptom Y in at least 51% of the 10.000 bootstrapped networks (i.e., at least 5.100). Finally, the selected arcs made a different contribution to the the goodness of fit of the final BNs. This contribution was represented by their width. Arcs which characterized the structure of the data with more confidence were visualised as thicker than others. The thicker an arc, the more damaging it would be for the model fit to remove it from the network.⁷⁷

You can see the full list of these, visualised networks in Online Figures 6 through 41.⁴⁹

Performance

To evaluate the averaged networks learned, on all of the combinations of the different algorithms, thresholds, and samples sizes used, we compared their post-hoc scores on a set of different scoring functions (Appendix A, Table S2). The selected functions were suitable for assessing the discrete BNs made from our categorical, symptom occurrence variables (i.e., Yes/No).^{74,78} These metrics were the logarithm of the K2 score (k_2), the logarithm of the Bayesian Dirichlet Equivalent score (BDE), the Bayesian Information Criterion score (BIC), the Akaike Information Criterion score (AIC) and the multinomial log-likelihood (loglik) score. The values for these scoring metrics reflect their goodness of fit, which the algorithm attempts to maximize. In Table S2, Appendix A, higher values reflect better goodness of fit.

Many of these scoring metrics are in the form of a penalised log-likelihood (LL) function.⁷⁹ The LL is the log probability of the dataset D , used, given a learnt network structure B . With the assumption that the data are independent and identically distributed, the likelihood of the data D given a structure B can be calculated as:⁷⁹

$$LL(D|B) = \sum_j^N \log P(D_j|B) = \sum_i^n \sum_j^N \log P(D_{ij}|PA_{ij}) \quad (2)$$

where D_{ij} is the instantiation of X_i in data point D_j , and PA_{ij} is the instantiation of X_i 's parents in D_j .

Adding an arc to a network increases the likelihood of the network. If this arc does not provide more information it is ignored. Nevertheless, extra arcs may lead to overfitting of the training data. To address the overfitting problem these penalized LL functions add a penalty term which penalizes the complex networks. In this manner, despite the fact that the complex networks may have a very good LL score, the high penalty term will reduce their score to be below that of a less complex network.

Stability and Reproducibility

Our study examines the data-driven, exploratory usefulness of BNA for cancer symptom data. Until now, very few network analyses studies were done on this domain. Furthermore, no consensus or evidence exists on the true network structure of how co-occurring symptoms' interact with each other.⁵⁹ Finally, an inherent problem in BNA is obtaining this "true" network structure.⁸⁰ The identified structures could be sensitive to the sample size, the dataset used, and/or the specific algorithms used. To assess the stability and reproducibility of our findings we tested their consistency among four equally sized, bootstrapped and randomly assigned subsets ($n_1=332, n_2=664, n_3=996, n_4=1328$). We ran 10,000 BNs for each bootstrapped and randomly assigned subset.

In Appendix A, Table S2 provides the full list of the results of these analyses. These analyses showed the stability of the identified networks as well as the reproducibility of each BNA algorithm based on our cancer symptoms' dataset.

Comparing Networks

To facilitate comparison between the BNs in this study, we compared them on their structural similarities and distances with the Hamming Distance (HD) and the Structural Hamming Distance (SHD).^{74, 80–82} HD and SHD describe the changes that have to be made to a network for it to turn into the one that it is being used for comparison. HD simply counts the number of edges that are different between the skeleton of the two networks, ignoring the direction of their arcs. SHD counts how these networks differ, taking into account the direction of their arcs. It sums the counts of the edges that are present in the original structure but are missing in the comparator structure, the counts of the edges that are found in the comparator structure but are not present in the original structure, and the counts of directed edges in the comparator structure that are oriented differently.

We provide a visual side-by-side comparison of the networks, identified with all three algorithms (i.e., GS, HC, MMHC) on the complete dataset ($n_{\text{Complete}}=1328$), and the two sub-groups based on gender ($n_{\text{Male}}=295, n_{\text{Female}}=1033$) and age ($n_{\text{Age} < 65}=958, n_{\text{Age} > 65}=370$). Lacking a "gold standard" network structure for cancer symptoms' data, in order to compare the performance and stability of the GS, HC, MMHC algorithms across the four different sub-samples ($n_1=332, n_2=664, n_3=996, n_4=1328$) we used the networks obtained from the complete, unrandomised data set ($n_{\text{Complete}}=1328$) as the reference network. Online Figures 12 through 17, 24 through 29, and 36 through 41,⁴⁹ provide a visual overview of the structural differences between the networks identified in this study. More precisely, by taking one network as a reference network, we computed the true positive, false positive and false negative arcs in the other network. In addition, we created an interactive visualisation of these differences.

Data availability

The data used in this study will be available upon request and subject to ethics approval. All data requests should be sent to Christine Miaskowski (chris.miaskowski@ucsf.edu).

References

1. Bubis, L. D. *et al.* Symptom burden in the first year after cancer diagnosis: an analysis of patient-reported outcomes. *J. Clin. Oncol.* **36**, 1103–1111 (2018).
2. Dantzer, R., Meagher, M. W. & Cleeland, C. S. Translational approaches to treatment-induced symptoms in cancer patients. *Nat. reviews Clin. oncology* **9**, 414 (2012).

3. Papachristou, N. *et al.* Congruence between latent class and k-modes analyses in the identification of oncology patients with distinct symptom experiences. *J Pain Symptom Manag.* **55**, 318–333 (2018).
4. Miaskowski, C. *et al.* Latent class analysis reveals distinct subgroups of patients based on symptom occurrence and demographic and clinical characteristics. *J Pain Symptom Manag.* **50**, 28–37 (2015).
5. Esther Kim, J. E., Dodd, M. J., Aouizerat, B. E., Jahan, T. & Miaskowski, C. A review of the prevalence and impact of multiple symptoms in oncology patients. *J Pain Symptom Manag.* **37**, 715–736 (2009).
6. Jaffee, E. M. *et al.* Future cancer research priorities in the usa: a lancet oncology commission. *The Lancet Oncol.* **18**, e653–e706 (2017).
7. Penedo, F. J. *et al.* The increasing value of ehealth in the delivery of patient-centred cancer care. *The Lancet Oncol.* **21**, e240–e251 (2020).
8. Lindner, O. C., Velikova, G. & Stark, D. P. Digitally enabled patient-reported outcome measures in cancer care. *The Lancet Oncol.* **20**, e2 (2019).
9. Heathcote, L. C. *et al.* Digitally enabled patient-reported outcome measures in cancer care—authors’ reply. *The Lancet Oncol.* **20**, e3 (2019).
10. Basch, E. Patient-reported outcomes—harnessing patients’ voices to improve clinical care. *New Engl. J. Medicine* **376**, 105–108 (2017).
11. Buiting, H. M. & Olthuis, G. Importance of Quality-of-Life Measurement Throughout the Disease Course. *JAMA Netw. Open* **3**, e200388–e200388 (2020).
12. Miaskowski, C. *et al.* Advancing symptom science through symptom cluster research: Expert panel proceedings and recommendations. *J. Natl. Cancer Inst.* **109** (2017).
13. Miaskowski, C. Future directions in symptom cluster research. *Semin Oncol Nurs.* **32**, 405–415 (2016).
14. Barsevick, A. Defining the symptom cluster: How far have we come? *Semin Oncol Nurs.* **32**, 334–350 (2016).
15. Cleeland, C. S., Fisch, M. J. & Dunn, A. J. *Cancer symptom science: measurement, mechanisms, and management* (Cambridge University Press, 2010).
16. Su, C., Andrew, A., Karagas, M. R. & Borsuk, M. E. Using bayesian networks to discover relations between genes, environment, and disease. *BioData mining* **6**, 6 (2013).
17. Guo, R., Cheng, L., Li, J., Hahn, P. R. & Liu, H. A survey of learning causality with data: Problems and methods. *arXiv preprint arXiv:1809.09337* (2018).
18. Chao, Y.-S. *et al.* A network perspective of engaging patients in specialist and chronic illness care: The 2014 international health policy survey. *PloS one* **13**, e0201355 (2018).
19. Lazić, N., Bishop, C. & Winn, J. Structural expectation propagation (sep): Bayesian structure learning for networks with latent variables. In *Artificial Intelligence and Statistics*, 379–387 (2013).
20. Gao, T. & Ji, Q. Constrained local latent variable discovery. In *IJCAI*, 1490–1496 (2016).
21. Kyrimi, E. *et al.* A comprehensive scoping review of bayesian networks in healthcare: Past, present and future. *arXiv preprint arXiv:2002.08627* (2020).
22. Sesen, M. B., Nicholson, A. E., Banares-Alcantara, R., Kadir, T. & Brady, M. Bayesian networks for clinical decision support in lung cancer care. *PloS one* **8**, e82349 (2013).
23. Yet, B., Perkins, Z. B., Rasmussen, T. E., Tai, N. R. & Marsh, D. W. R. Combining data and meta-analysis to build bayesian networks for clinical decision support. *J. biomedical informatics* **52**, 373–385 (2014).
24. Barber, D. *Bayesian reasoning and machine learning* (Cambridge University Press, 2012).
25. Puga, J. L., Krzywinski, M. & Altman, N. Points of significance: Bayesian networks (2015).
26. Flores, M. J., Nicholson, A. E., Brunskill, A., Korb, K. B. & Mascaro, S. Incorporating expert knowledge when learning bayesian network structure: a medical case study. *Artif. intelligence medicine* **53**, 181–204 (2011).
27. Korb, K. B. & Nicholson, A. E. *Bayesian artificial intelligence* (CRC press, 2010).
28. Jayasurya, K. *et al.* Comparison of bayesian network and support vector machine models for two-year survival prediction in lung cancer patients treated with radiotherapy. *Med. physics* **37**, 1401–1407 (2010).

29. Stojadinovic, A. *et al.* Clinical decision support and individualized prediction of survival in colon cancer: Bayesian belief network model. *Annals surgical oncology* **20**, 161–174 (2013).
30. Dekker, A. *et al.* Survival prediction in lung cancer treated with radiotherapy: Bayesian networks vs. support vector machines in handling missing data. In *2009 International Conference on Machine Learning and Applications*, 494–497 (IEEE, 2009).
31. Tegner, J. & Bjorkegren, J. Perturbations to uncover gene networks. *Trends Genet.* **23**, 34–41 (2007).
32. Markowitz, F. & Spang, R. Inferring cellular networks—a review. *BMC bioinformatics* **8**, S5 (2007).
33. Friedman, N. Inferring cellular networks using probabilistic graphical models. *Sci.* **303**, 799–805 (2004).
34. Yu, J., Smith, V. A., Wang, P. P., Hartemink, A. J. & Jarvis, E. D. Advances to bayesian network inference for generating causal networks from observational biological data. *Bioinforma.* **20**, 3594–3603 (2004).
35. Acid, S. *et al.* A comparison of learning algorithms for bayesian networks: a case study based on data from an emergency medical service. *Artif. intelligence medicine* **30**, 215–232 (2004).
36. Maglogiannis, I., Zafiroopoulos, E., Platis, A. & Lambrinouidakis, C. Risk analysis of a patient monitoring system using bayesian network modeling. *J. Biomed. Informatics* **39**, 637–647 (2006).
37. Sadeghi, S., Barzi, A., Sadeghi, N. & King, B. A bayesian model for triage decision support. *Int. journal medical informatics* **75**, 403–411 (2006).
38. Lederer, D. J. *et al.* Control of confounding and reporting of results in causal inference studies. guidance for authors from editors of respiratory, sleep, and critical care journals. *Annals Am. Thorac. Soc.* **16**, 22–28 (2019).
39. Williams, T. C., Bach, C. C., Matthiesen, N. B., Henriksen, T. B. & Gagliardi, L. Directed acyclic graphs: a tool for causal studies in paediatrics. (2018).
40. Kuschner, K. W. *et al.* A bayesian network approach to feature selection in mass spectrometry data. *BMC bioinformatics* **11**, 177 (2010).
41. Pearl, J., Bareinboim, E. *et al.* External validity: From do-calculus to transportability across populations. *Stat. Sci.* **29**, 579–595 (2014).
42. Tidriri, K., Chatti, N., Verron, S. & Tiplica, T. Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges. *Annu. Rev. Control.* **42**, 63–81 (2016).
43. Moreno-Betancur, M. *et al.* Canonical causal diagrams to guide the treatment of missing data in epidemiologic studies. *Am. journal epidemiology* **187**, 2705–2715 (2018).
44. Rubin, D. L., Burnside, E. S. & Shachter, R. A bayesian network to assist mammography interpretation. In *Operations Research and Health Care*, 695–720 (Springer, 2005).
45. Wang, K.-J., Makond, B. & Wang, K.-M. Modeling and predicting the occurrence of brain metastasis from lung cancer by bayesian network: a case study of taiwan. *Comput. biology medicine* **47**, 147–160 (2014).
46. Xu, S. *et al.* Modeling interrelationships between health behaviors in overweight breast cancer survivors: Applying bayesian networks. *PloS one* **13** (2018).
47. Xu, S. *et al.* Cognition, quality-of-life, and symptom clusters in breast cancer: Using bayesian networks to elucidate complex relationships. *Psycho-oncology* **27**, 802–809 (2018).
48. Kowal, P. & Dowd, J. E. Definition of an older person. proposed working definition of an older person in africa for the mds project. *World Heal. Organ. Geneva, doi* **10**, 5188–9286 (2001).
49. Papachristou, N., Barnaghi, P., Allison, S. J., Kober, K. M. & Miaskowski, C. Online Figures of the paper titled Bayesian Networks and Causal Inference for the Interpretation of Patients’ Symptom Experience. (2020). URL <https://doi.org/10.5281/zenodo.4001527>. DOI 10.5281/zenodo.4001527.
50. Scutari, M. & Nagarajan, R. Identifying significant edges in graphical models of molecular networks. *Artif. intelligence medicine* **57**, 207–217 (2013).
51. Scutari, M., Graafland, C. E. & Gutiérrez, J. M. Who learns better bayesian network structures: Constraint-based, score-based or hybrid algorithms? *arXiv preprint arXiv:1805.11908* (2018).
52. Cowell, R. G. Conditions under which conditional independence and scoring methods lead to identical selection of bayesian network models. *arXiv preprint arXiv:1301.2262* (2013).

53. Russell, J. *et al.* Stability of Symptom Clusters in Patients With Lung Cancer Receiving Chemotherapy. *J Pain Symptom Manag.* **57**, 909–922 (2019).
54. Sullivan, C. W. *et al.* Stability of Symptom Clusters in Patients With Breast Cancer Receiving Chemotherapy. *J Pain Symptom Manag.* **55**, 39–55 (2018).
55. Savard, M. H., Savard, J., Quesnel, C. & Ivers, H. The influence of breast cancer treatment on the occurrence of hot flashes. *J Pain Symptom Manag.* **37**, 687–697 (2009).
56. Tejada, M. *et al.* Identification of subgroups of chemotherapy patients with distinct sleep disturbance profiles and associated co-occurring symptoms. *Sleep* **42** (2019).
57. Wright, F. *et al.* Morning Fatigue Severity Profiles in Oncology Outpatients Receiving Chemotherapy. *Cancer Nurs* **42**, 355–364 (2019).
58. Abid, H. *et al.* Common and Distinct Characteristics Associated With Trajectories of Morning and Evening Energy in Oncology Patients Receiving Chemotherapy. *J Pain Symptom Manag.* **53**, 887–900 (2017).
59. Papachristou, N. *et al.* Network Analysis of the Multidimensional Symptom Experience of Oncology. *Sci Rep* **9**, 2258 (2019).
60. Breiman, L. *et al.* Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Stat. science* **16**, 199–231 (2001).
61. Papachristou, N. *et al.* Learning from data to predict future symptoms of oncology patients. *PLoS ONE* **13**, e0208808 (2018).
62. Kober, K. M. *et al.* Subgroups of chemotherapy patients with distinct morning and evening fatigue trajectories. *Support. Care Cancer* **24**, 1473–1485 (2016).
63. Miaskowski, C. *et al.* The symptom phenotype of oncology outpatients remains relatively stable from prior to through 1 week following chemotherapy. *Eur J Cancer Care (Engl)* **26** (2017).
64. Wright, F. *et al.* Inflammatory pathway genes associated with inter-individual variability in the trajectories of morning and evening fatigue in patients receiving chemotherapy. *Cytokine* **91**, 187–210 (2017).
65. Portenoy, R. K. *et al.* The memorial symptom assessment scale: an instrument for the evaluation of symptom prevalence, characteristics and distress. *Eur. J. Cancer* **30**, 1326–1336 (1994).
66. Pearl, J. *Probabilistic reasoning in intelligent systems: networks of plausible inference* (Elsevier, 2014).
67. Neapolitan, R. E. *et al.* *Learning bayesian networks*, vol. 38 (Pearson Prentice Hall Upper Saddle River, NJ, 2004).
68. Darwiche, A. *Modeling and reasoning with Bayesian networks* (Cambridge university press, 2009).
69. Beretta, S., Castelli, M., Gonçalves, I., Henriques, R. & Ramazzotti, D. Learning the structure of bayesian networks: A quantitative assessment of the effect of different algorithmic schemes. *Complex.* **2018** (2018).
70. Ji, Z., Xia, Q. & Meng, G. A review of parameter learning methods in bayesian network. In *International Conference on Intelligent Computing*, 3–12 (Springer, 2015).
71. Pearl, J. *Causality: models, reasoning and inference*, vol. 29 (Springer, 2000).
72. Margaritis, D. Learning bayesian network model structure from data. Tech. Rep., Carnegie-Mellon Univ Pittsburgh Pa School of Computer Science (2003).
73. Margaritis, D. & Thrun, S. Bayesian network induction via local neighborhoods. In *Advances in neural information processing systems*, 505–511 (2000).
74. Scutari, M. Learning bayesian networks with the bnlearn r package. *arXiv preprint arXiv:0908.3817* (2009).
75. Almende, B., Thieurmel, B. & Robert, T. visnetwork: Network visualization using “vis. js” library. *R package version 1* (2016).
76. Sachs, K., Perez, O., Pe’er, D., Lauffenburger, D. A. & Nolan, G. P. Causal protein-signaling networks derived from multiparameter single-cell data. *Sci.* **308**, 523–529 (2005).
77. McNally, R. J., Heeren, A. & Robinaugh, D. J. A bayesian network analysis of posttraumatic stress disorder symptoms in adults reporting childhood sexual abuse. *Eur. J. Psychotraumatology* **8**, 1341276 (2017).
78. Nagarajan, R., Scutari, M. & Lèbre, S. Bayesian networks in r. *Springer* **122**, 125–127 (2013).

79. Liu, Z., Malone, B. & Yuan, C. Empirical evaluation of scoring functions for bayesian network model selection. In *BMC bioinformatics*, vol. 13, S14 (Springer, 2012).
80. Liu, Z., Malone, B. & Yuan, C. Empirical evaluation of scoring functions for Bayesian network model selection. *BMC Bioinforma.* **13 Suppl 15**, S14 (2012).
81. Tsamardinos, I., Brown, L. E. & Aliferis, C. F. The max-min hill-climbing bayesian network structure learning algorithm. *Mach. learning* **65**, 31–78 (2006).
82. de Jongh, M. & Druzdzel, M. J. A comparison of structural distance measures for causal bayesian network models. *Recent Adv. Intell. Inf. Syst. Challenging Probl. Sci. Comput. Sci. series* 443–456 (2009).

Acknowledgements

This project received funding from the European Union’s Horizon 2020 research and innovation ACTIVAGE project under grant agreement No 732679, the European Union’s Horizon 2020 research and innovation LifeChamps project under grant agreement No 875329 and the Care Research and Technology Centre at the UK Dementia Research Institute (UK DRI). Part of this study was also funded by the National Cancer Institute (CA134900).

Author contributions statement

Christine Miaskowski was responsible for the design and execution of the clinical study, including patient recruitment and retention and data collection. Nikolaos Papachristou, Payam Barnaghi and Christine Miaskowski conceived the experiment. Nikolaos Papachristou conducted and analysed the results. Nikolaos Papachristou, Payam Barnaghi, Christine Miaskowski, Sarah J. Allison and Kord M. Kober contributed to the interpretation of the study findings and participated in the editing and revision of the final version of the manuscript.

Competing Interests

We declare that we have no commercial or financial relationship with any company that can inappropriately influence our work. All authors declare that we have no conflicting or competing interests.

Corresponding authors

Correspondence to Nikolaos Papachristou or Payam Barnaghi.

Figures

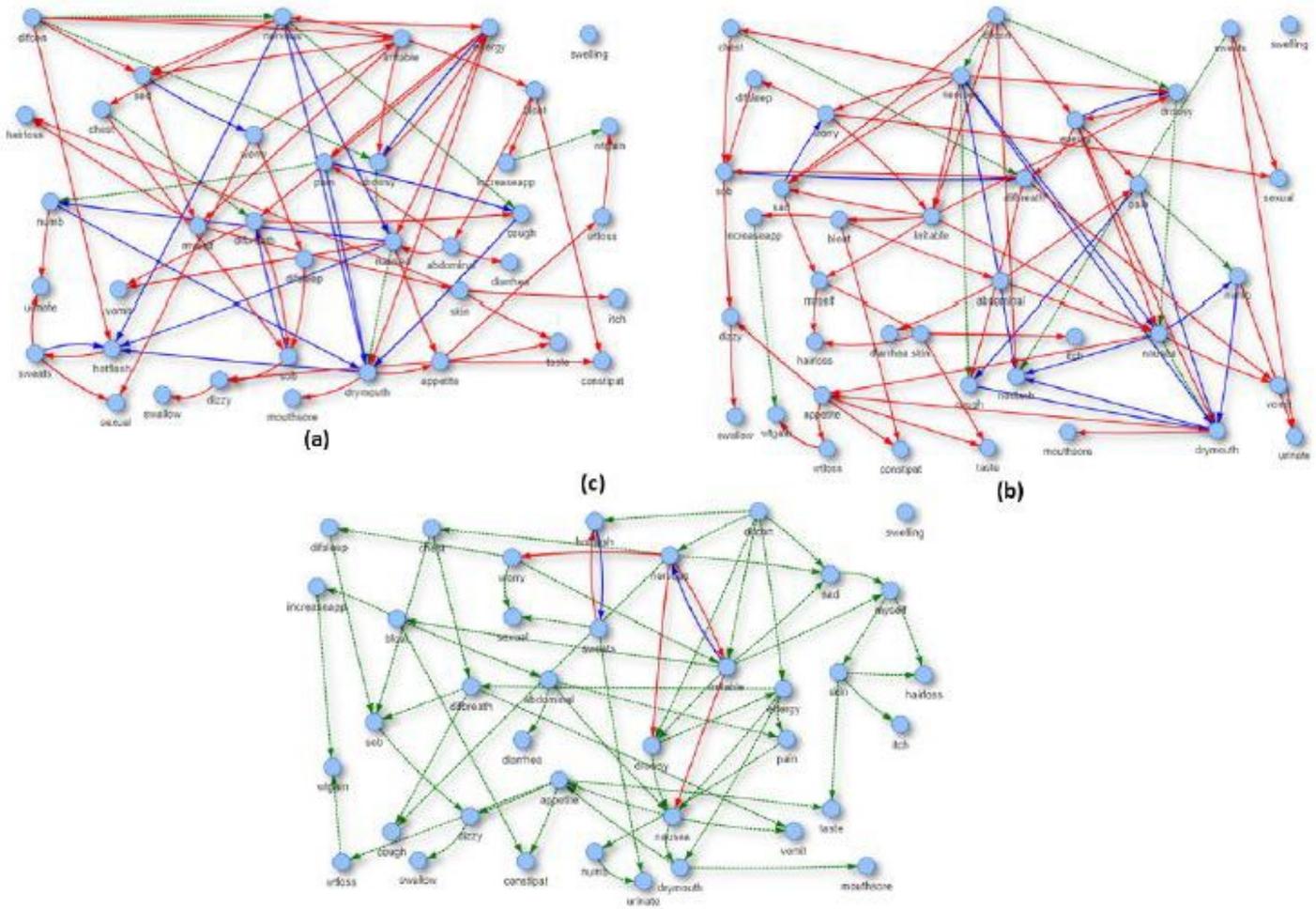


Figure 1

Bayesian Networks Comparison: a) Grow-Shrink vs Hill-Climbing, b) Max-Min Hill-Climbing, c) Hill-Climbing vs Max-Min Hill-Climbing. The green dotted lines represent common arcs between the compared networks. The blue and red arcs represent different arcs that belong to each of the networks that were compared. For a better view of the diagrams see Online Figures 12 to 17.49

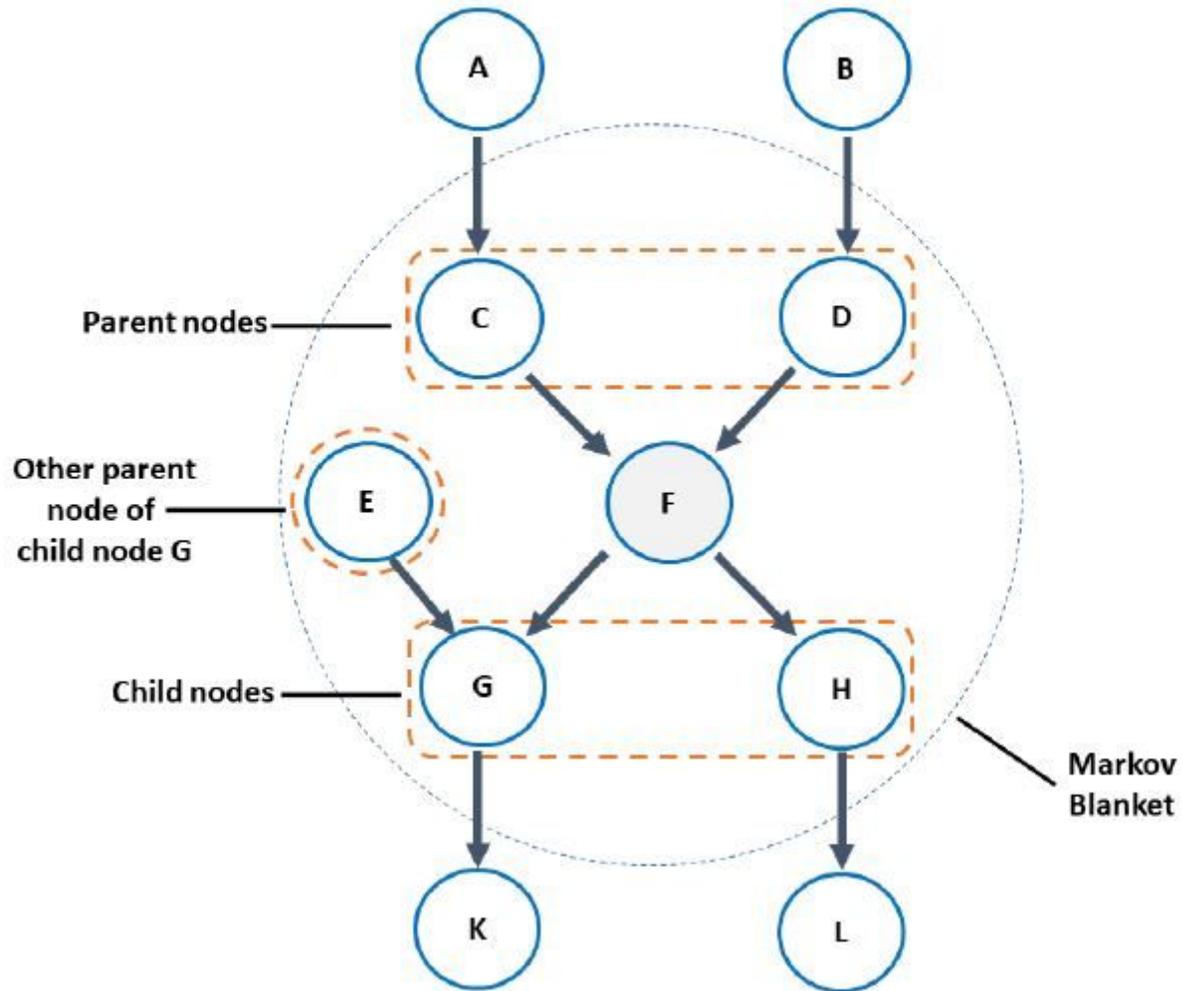


Figure 2

Bayesian Network: parents, child nodes and Markov Blanket of node A in a directed graph C and D are the parent nodes and nodes G and H are the child nodes of node F. Node E is a parent node of G, sharing the same child node with node F. Thus, the Markov Blanket of variable F is the group of all the aforementioned variables or nodes (i.e., C, D, F, E, G, H). The minimum set of nodes that are necessary to predict the behaviour of variable F are the ones included in this Markov Blanket (see Appendix B).68

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AppendixA.pdf](#)
- [AppendixB.pdf](#)