

Supplementary information Appendix A

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

Supplemental Tables S1-S2

Table S1. Demographic and Clinical Characteristics of the Sample (n=1328)

Characteristic	Mean (SD)
Age (years)	57.3 (12.3)
Education (years)	16.2 (3.0)
Body mass index (kg/m ²)	26.2 (5.7)
Karnofsky Performance Status score	80.1 (12.4)
Self-administered Comorbidity Questionnaire score	5.5 (3.2)
Time since diagnosis (years)	2.0 (3.9)
Time since diagnosis (years, median)	0.4
Number of prior cancer treatments	1.6 (1.5)
Number of metastatic sites including lymph node involvement ^a	1.2 (1.2)
Number of metastatic sites excluding lymph node involvement	0.8 (1.0)
Average number of symptoms reported prior to their next dose of CTX	13.9 (7.2)
	% (n)
Gender (% female)	77.8 (1033)
Self-reported ethnicity	
White	69.9 (916)
Non-white	30.1 (412)
Married or partnered (% yes)	64.4 (843)
Lives alone (% yes)	21.5 (282)
Currently employed (% yes)	35.2 (462)
Annual household income	
Less than \$30,000	18.4 (219)
\$30,000 to \$70,000	21.2 (252)
\$70,000 to \$100,000	16.7 (198)
Greater than \$100,000	43.7 (520)
Child care responsibilities (% yes)	22.0 (286)
Elder care responsibilities (% yes)	7.9 (95)
Exercise on a regular basis (% yes)	71.0 (922)
Cancer diagnosis	
Breast cancer	40.2 (534)
Gastrointestinal cancer	30.6 (407)
Gynecological cancer	17.5 (232)
Lung cancer	11.7 (155)
Prior cancer treatment	
No prior treatment	25.0 (323)
Only surgery, CTX, or RT	42.0 (542)
Surgery and CTX, or surgery and RT, or CTX and RT	19.9 (257)
Surgery and CTX and RT	13.1 (169)
Metastatic sites	
No metastasis	32.3 (423)

Continued on next page

Table S1 – continued from previous page

	% (n)
Only lymph node metastasis	22.1 (289)
Only metastatic disease in other sites	21.1 (277)
Metastatic disease in lymph nodes and other sites	24.5 (321)
Cycle length	
14 day cycle	42.1 (558)
21 day cycle	50.5 (670)
28 day cycle	7.3 (97)

^a Total number of metastatic sites evaluated was 9.

Abbreviations: CTX = chemotherapy, kg = kilograms, m² = meters squared, RT = radiation therapy, SD = standard deviation

Table S2. Performance and Stability of Bayesian Network Algorithms

Algorithm	Sample size	Threshold	k2	BDE	AIC	BIC	logik	Arcs Number	Number of Common Arcs*	Hamming Distance*	Structural Hamming Distance*
GS	332	0.8500	-29818.64	-29818.74	-29718.37	-29677.37	-29824.81	3	3	7	8
	332	0.4930	-29237.59	-29239.17	-29125.53	-29247.55	-29078.53	8	5	16	18
	664	0.8500	-29541.37	-29541.20	-29436.34	-29550.57	-29392.34	5	5	5	5
	664	0.4903	-29131.40	-29134.36	-29012.75	-29147.75	-28960.75	12	10	12	17
	996	0.8500	-29471.19	-29470.84	-29362.68	-29482.1	-29316.68	7	7	3	4
	996	0.4906	-29104.73	-29110.37	-28971.68	-29143.02	-28905.68	17	13	11	13
	1328	0.8500	-29451.03	-29462.67	-29331.35	-29487.11	-29271.35	10	10	0	0
	1328	0.4985	-29130.00	-29179.53	-28975.69	-29240.49	-28873.69	22	22	1	1
	1328**	0.8500	-29451.03	-29462.67	-29331.35	-29487.11	-29271.35	10	-	-	-
	1328**	0.4963	-29130.00	-29179.53	-28975.69	-29240.49	-28873.69	22	-	-	-
HC	332	0.8500	-28722.65	-28726.57	-28603.20	-28735.61	-28552.20	13	9	16	20
	332	0.4987	-27999.91	-28003.56	-27853.40	-28027.34	-27786.40	25	21	34	52
	664	0.8500	-28492.51	-28495.74	-28360.98	-28511.55	-28302.98	18	14	11	14
	664	0.4878	-27612.57	-27619.04	-27429.49	-27676.12	-27334.49	42	35	18	44
	996	0.8500	-28177.69	-28179.24	-28033.37	-28202.12	-27968.37	24	21	5	11
	996	0.5000	-27446.14	-27449.48	-27246.63	-27519.22	-27141.63	49	43	11	44
	1328	0.8500	-28060.89	-28061.89	-27904.77	-28094.28	-27831.77	29	29	0	0
	1328	0.5034	-27197.85	-27202.55	-26967.37	-27310.05	-26835.37	60	57	0	3
	1328**	0.8500	-28060.89	-28061.89	-27904.77	-28094.28	-27831.77	29	-	-	-
	1328**	0.4989	-27164.21	-27165.58	-26932.69	-27270.18	-26802.69	59	-	-	-
MMHC	332	0.8500	-28722.65	-28726.57	-28603.20	-28735.61	-28552.20	13	10	16	20
	332	0.4984	-27940.17	-27943.85	-27791.88	-27968.42	-27723.88	26	22	36	55
	664	0.8500	-28533.85	-28535.16	-28400.16	-28553.33	-28341.16	18	16	11	12
	664	0.5092	-27603.43	-27607.41	-27419.52	-27660.96	-27326.52	42	38	21	52
	996	0.8500	-28177.75	-28179.24	-28033.37	-28202.12	-27968.37	24	23	5	11
	996	0.4951	-27461.06	-27470.1	-27256.78	-27550.14	-27143.78	50	48	13	25
	1328	0.8500	-28060.88	-28061.89	-27904.77	-28094.28	-27831.77	29	29	0	0
	1328	0.4992	-27206.61	-27219.05	-26966.79	-27343.22	-26821.79	62	62	1	1
	1328**	0.8500	-28060.88	-28061.89	-27904.77	-28094.28	-27831.77	29	-	-	-
	1328**	0.5010	-27206.61	-27219.05	-26966.79	-27343.22	-26821.79	62	-	-	-

*between the subsample and the complete, non-randomised dataset

**complete, non-randomised dataset

Abbreviations in Table S2 stand for the following: GS = Grow-Shrink algorithm, HC = Hill Climbing algorithm, MMHC = Max-Min Hill Climbing algorithm. Better scores have green shading. Worst scores have red shading.

Supplemental Figures S1-S6

The 38 cancer symptoms represented in the graphs below are coded in the following fashion: difcon: Difficulty Concentrating, pain: Pain, energy: Lack of Energy, cough: Cough, nervous: Feeling Nervous, hotflash: Hot Flashes, drymouth: Dry Mouth, nausea: Nausea, drowsy: Feeling Drowsy, numb: Numbness or Tingling in Hands or Feet, chest: Chest Tightness, difbreath: Difficulty Breathing, difsleep: Difficulty Sleeping, bloat: Feeling Bloated, urinate: Problems with Urination, vomit: Vomiting, sob: Shortness of Breath, diarrhea: Diarrhea, sad: Feeling Sad, sweats: Sweats, sexual: Problems with Sexual Interest or Activity, worry: Worrying, itch: Itching, appetite: Lack of Appetite, abdominal: Abdominal Cramps, increaseapp: Increased Appetite, wtgain: Weight Gain, dizzy: Dizziness, swallow: Difficulty Swallowing, irritable: Feeling Irritable, mouthsore: Mouth Sore, wtloss: Weight Loss, hairloss: Hair Loss, constipat: Constipation, swelling: Swelling, taste: Change in the Way Food Tastes, myself: I Do Not Look Like Myself, skin: Changes in Skin.

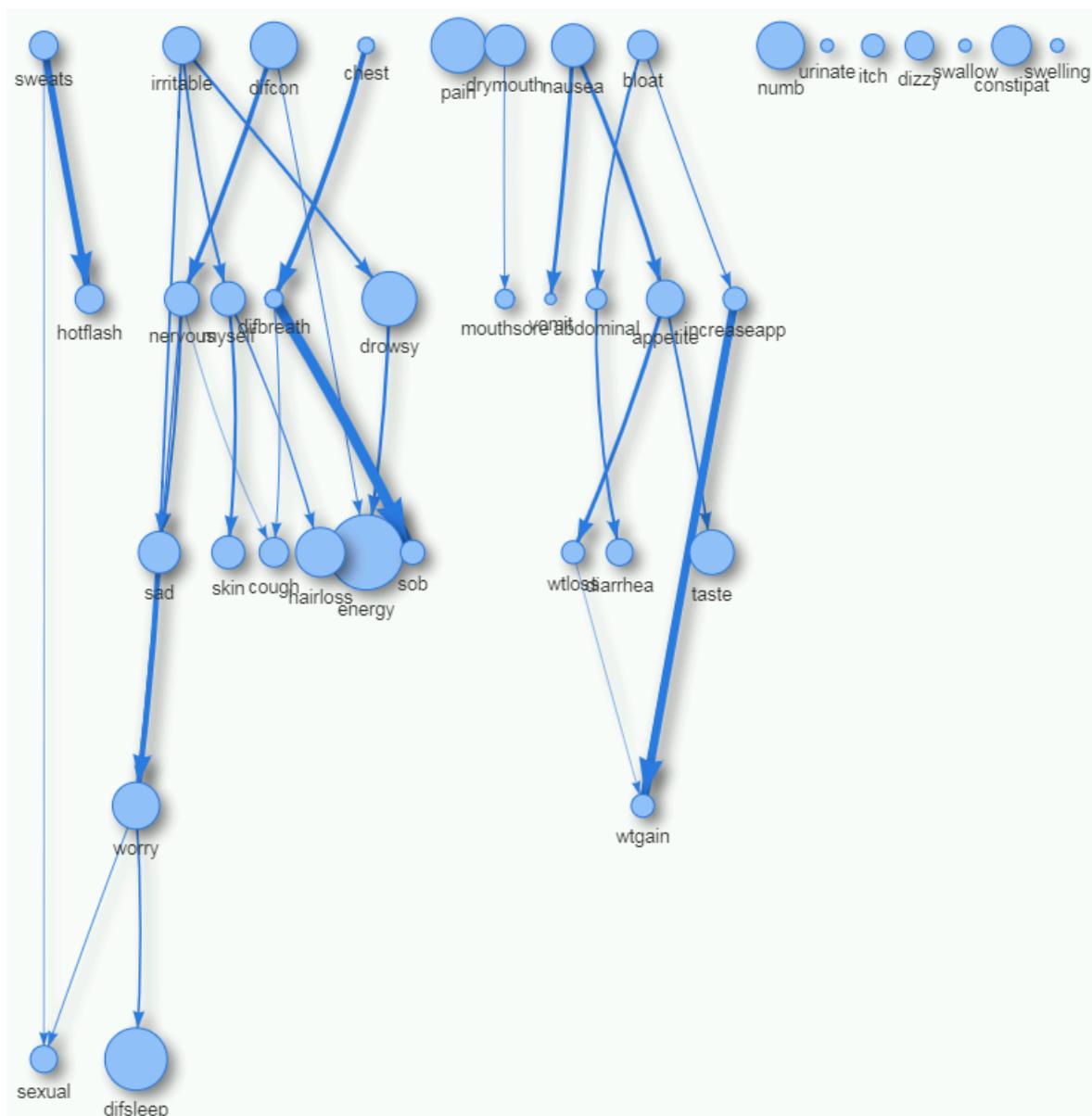


Figure S1. Network on symptom data for all the cancer patients (n=1328), made with the Max-Min Hill Climbing algorithm and 0.85 threshold (Sachs et al., 2005). To view the interactive diagram of this network see Online Figure 10¹

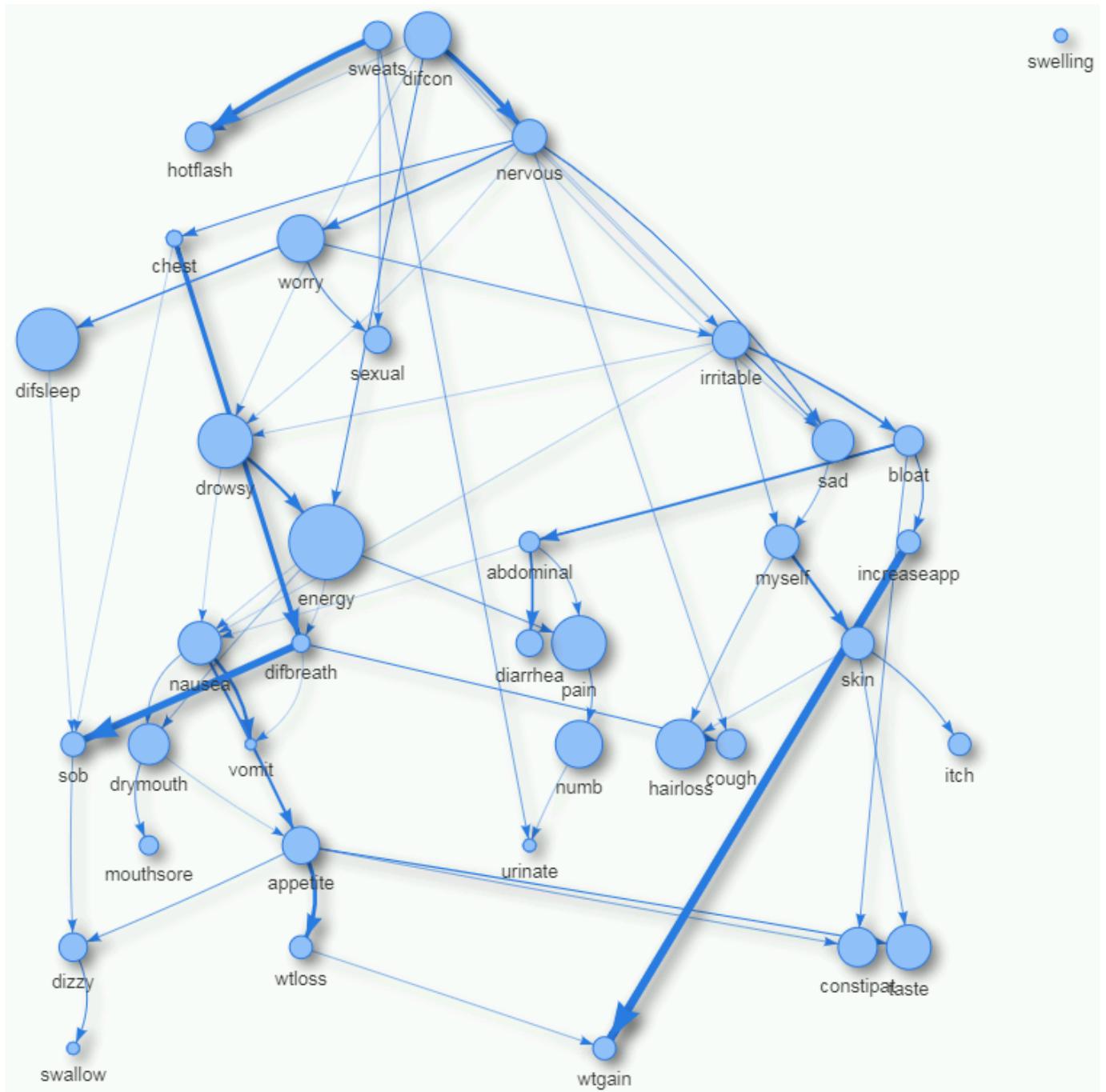


Figure S2. Network on symptom data for all the cancer patients (n=1328), made with the Max-Min Hill Climbing algorithm and optimal threshold according to Scurati Nagarajan, 2013. To view the interactive diagram of this network see Online Figure 11¹

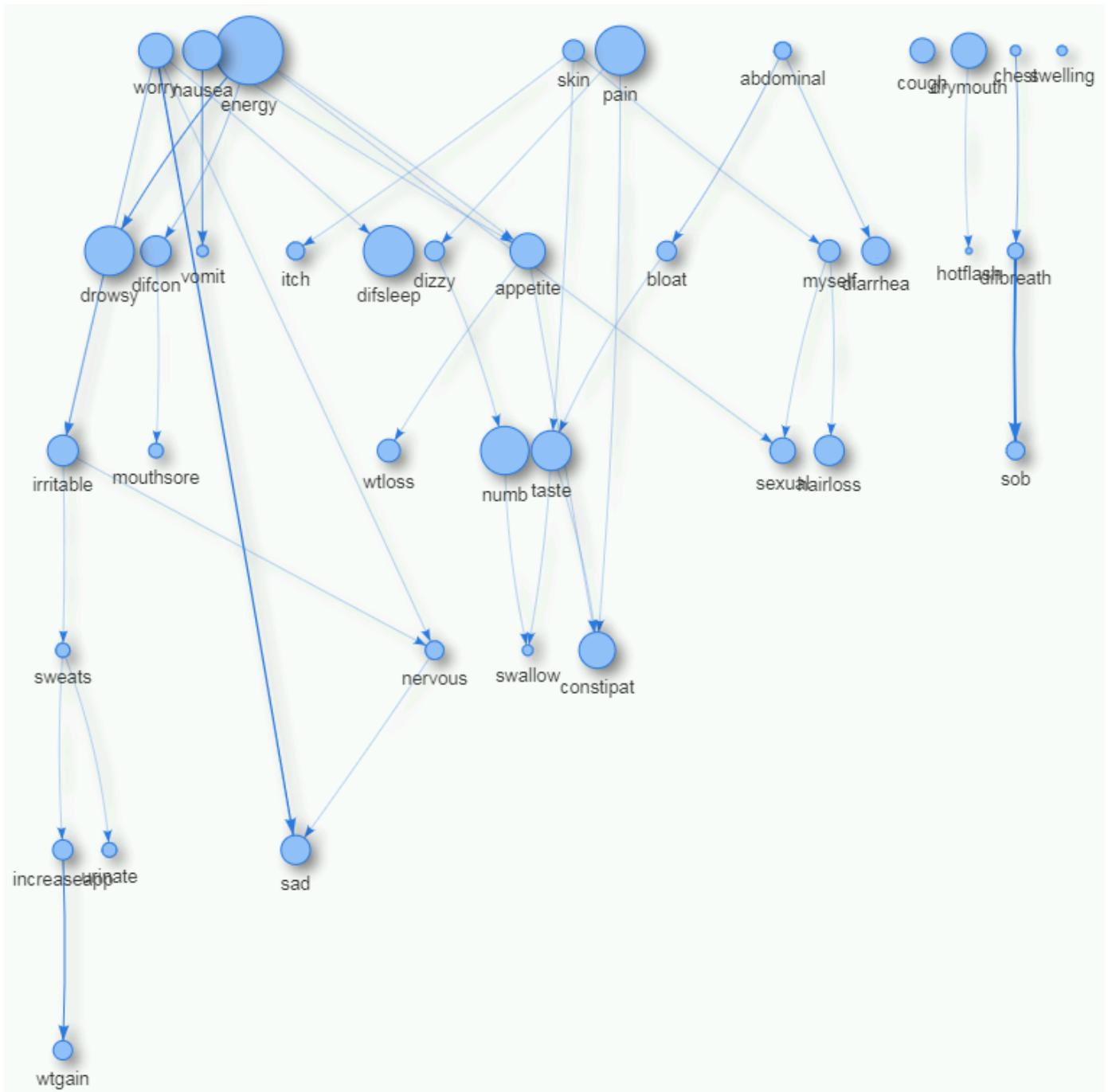


Figure S3. Network on symptom data for the male cancer patients (n=295), made with the Max-Min Hill Climbing algorithm and the optimal threshold according to Scurati Nagarajan, 2013. To view the interactive diagram of this network see Online Figure 19¹

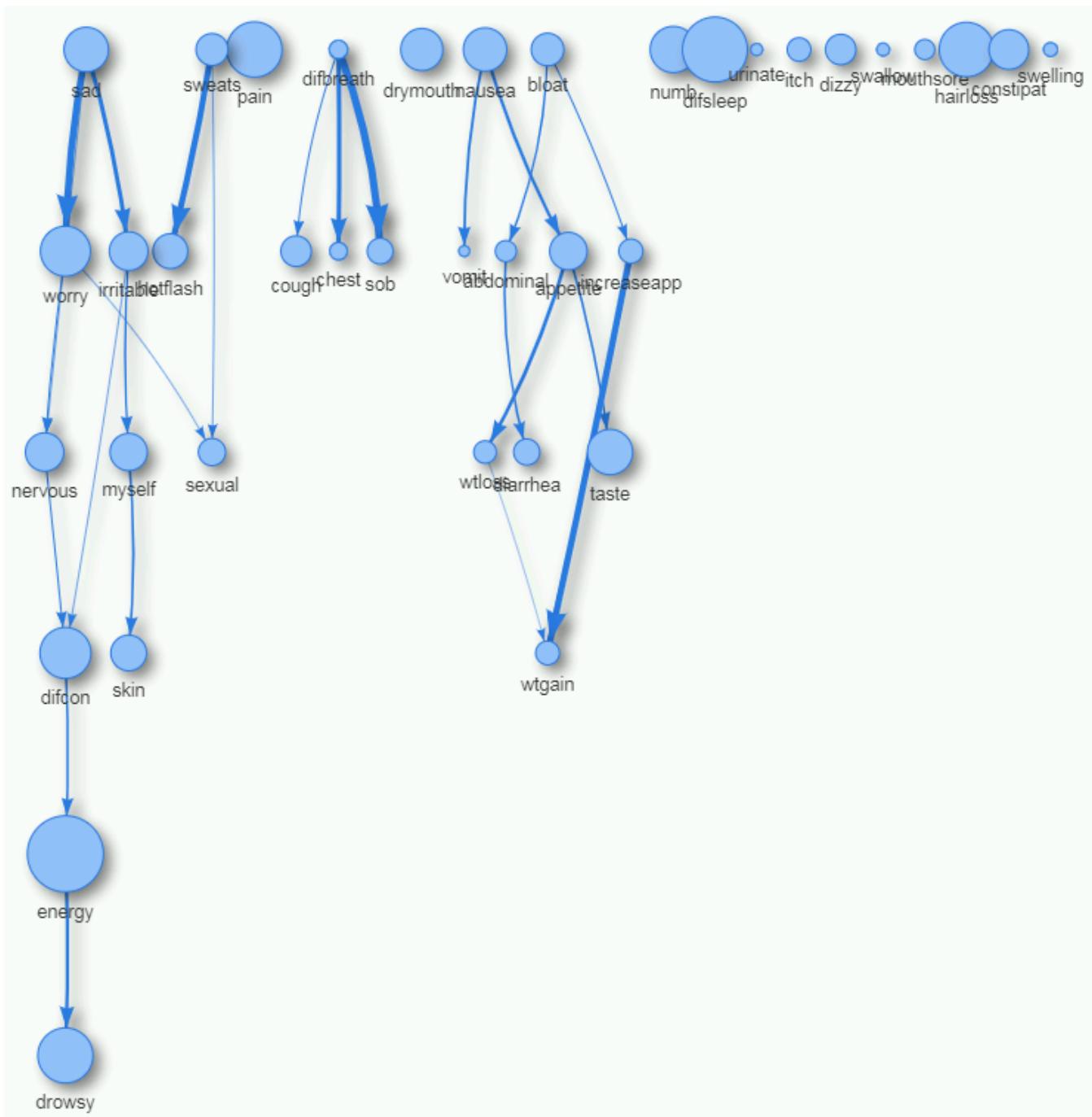


Figure S4. Network on symptom data for the female cancer patients (n=295, bootstrapped sampling), made with the Max-Min Hill Climbing algorithm and the optimal threshold according to Scurati Nagarajan, 2013. To view the interactive diagram of this network see Online Figure 23¹

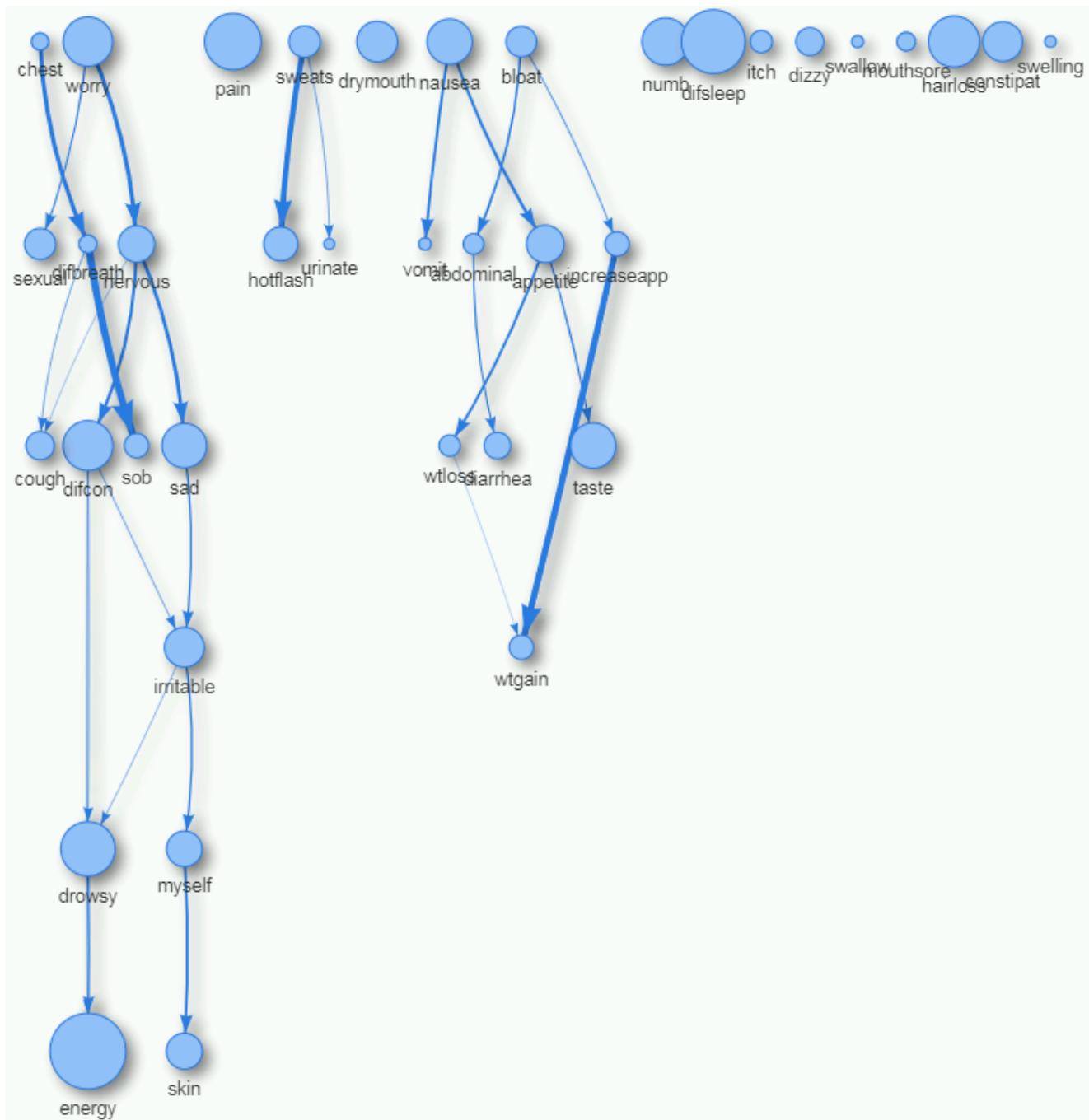


Figure S5. Network on symptom data for the cancer patients with age below 65 (n=370, bootstrapped sampling), made with the Max-Min Hill Climbing algorithm and the optimal threshold according to Scurati Nagarajan, 2013. To view the interactive diagram of this network see Online Figure 33¹

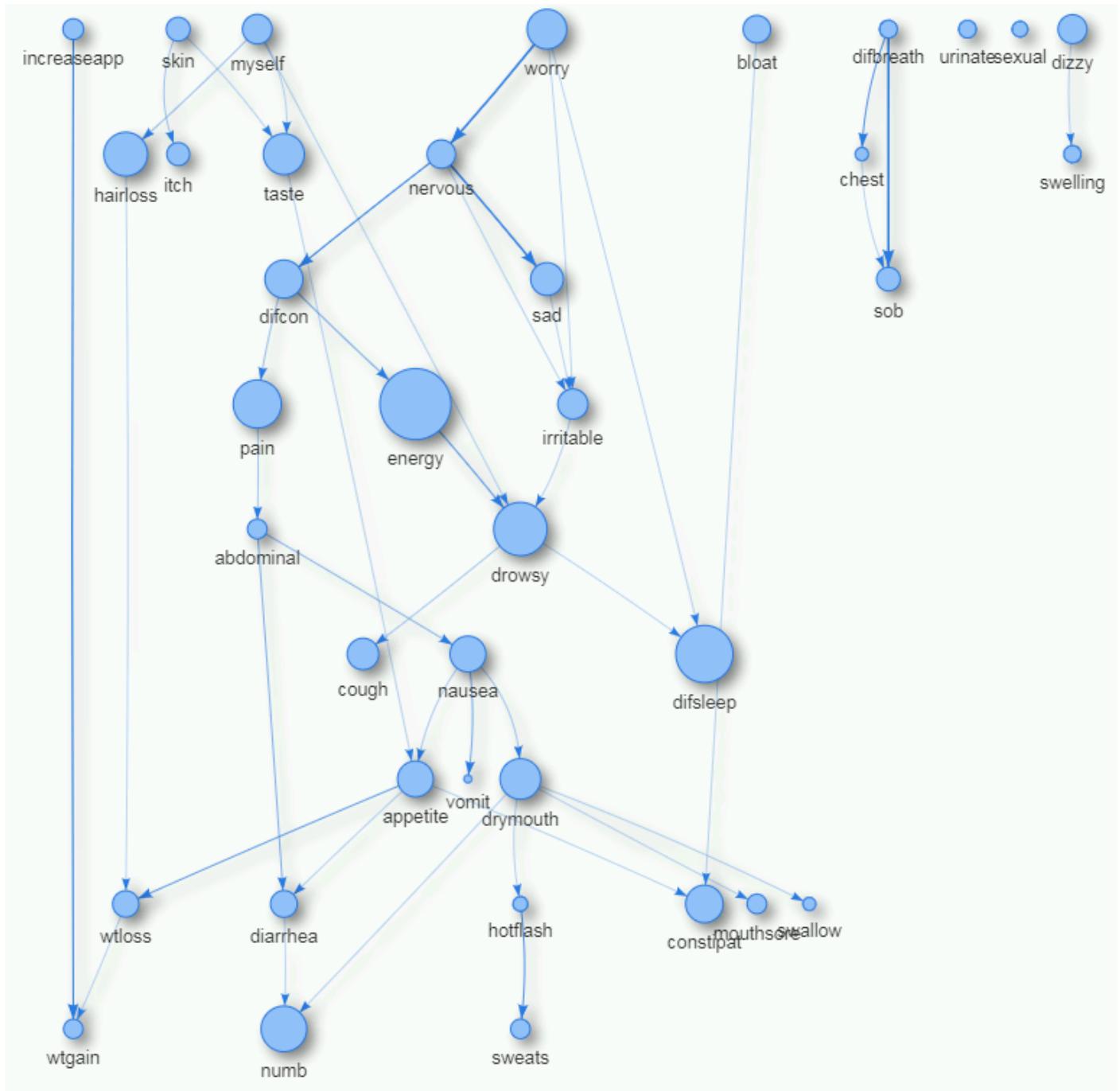


Figure S6. Network on symptom data for the cancer patients with age above 65 (n=370), made with the Max-Min Hill Climbing algorithm and the optimal threshold according to Scurati Nagarajan, 2013. To view the interactive diagram of this network see Online Figure 35¹

References

1. 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.