Chicken or the Egg: Identifying Nature vs. Nurture as the Decisive Predecessor for Obesity

Anonymous

Vrije Universiteit Amsterdam

Abstract. Obesity is one of the most critical illnesses of modern society. Several interconnected factors influence the susceptibility of obesity. Moreover, influential factors can be divided into being based on nature vs. nurture. To shed light on this critical topic, especially since it has been overshadowed by the COVID-19 pandemic, a Bayesian Network (BN) was set up to give new and interesting insights into the interconnections between the different influencing factors and point out which predecessors are especially decisive. The results of this research were inexpressive and did not characterize a great distinction between genetic and environmental influences. This could be based on a too simple representation of a very complex and highly interactive problem scenario. Additionally, the used BN potentially missed important influential factors of obesity.

1 Introduction

Our modern society is suffering from an increasingly growing epidemic. Obesity is one of the major health risks in the 21st century and is characterized as having a Body Mass Index (BMI) of 30 or higher [19]. After a spiking popularity of fast food chains, and a generally high consumption of hypercaloric and sugary food as well as less physical activity, the BMI of the global population has been growing steadily. According to the WHO, in 2016 a total of 1.9 billion adults worldwide were overweight, with 650 million being obese. For example, 36.2 % of the adults in the United States and 20.4 % of the Netherlands were obese [9]. The ongoing COVID-19 pandemic even intensified this already dire situation, as lockdowns decreased physical activity, resulting in gain weight for every second person[7].

Furthermore, the consequences of this silent epidemic can be severe. Obesity was associated with many physical illnesses as well as psychological consequences. For example, obesity leads to an increased likelihood of having a serious course of COVID-19 [3], developing type 2 diabetes, cardiovascular diseases, various cancers, and musculoskeletal disorders [18]. Hence, obesity can be characterized as a risk for premature death or substantial disability.

Fighting against obesity can be extremely hard on an individual and on a societal level. As for most illnesses, putting down precise predecessors can be difficult given the various influences of the environment and genetic dispositions. Nurture influential factors can be mainly traced back to an unhealthy change in diet. High calorie and sugary food, big meals and in-activeness promote an unhealthy lifestyle with all its physical consequences [6]. Moreover, socioeconomic status was also associated with obesity [6], which raises the subject to a sociopolitical level. In terms of nature, twin studies showed that genetic factors played a major role on BMI variations as well [6]. One factor that was repeatedly mentioned in this context was parental obesity [13][10].

Given that both, nature and nurture factors, and the interaction of these explains variance, this research aims to get an understanding about some specific predecessors of obesity and analyze these variables in terms of their decisive influence. Hence, a (small) literature research was conducted to find crucial nature and nurture influences and their associations with obesity and each other. Given the time and money constraints of this project, the interactive structure of the factors is kept simple and variables are limited to influential factors mentioned frequently in previous research. In order to display the network of identified predecessors and to meet the complexity of the interconnected structure of environmental and genetic factors, a Bayesian Network (BN) was used to display the problem statement and associations between the variables graphically.

Taken together, this research aims to shed light on the critical topic of obesity, especially since it has been overshadowed by the COVID-19 pandemic. Therefore, a BN was set up to give new and interesting insights into the interconnections between the different influencing factors and to point out which predecessors are decisive ones. Results of this paper could be used for improving prevention strategies.

As a research question, this exploratory research investigates whether nature, nurture or the interaction of the two variables is the more decisive influential factor for obesity. Therefore, an a-priori marginal query, a posterior marginal query and one MAP and one MEP query were conducted with different nature and nurture factors given as evidence.

2 Theoretical Background

Due to the topic's criticality, obesity and associated factors have been analyzed in previous research. However, the focus of investigated factors is different in each paper. To conduct a BN with popular related factors and probabilities that underline the precise relation between the variables, a literature research was conducted. Around 30 papers considering obesity, different influential factors of obesity, and connection amongst these factors were examined. In total, the results of 13 papers were used to find the relations of the different factors. The connections found are displayed in a BN, which can be found in Figure 2 the Appendix. The exact definition of each variable as well as the corresponding scientific grounding can be found in the Methods section.

3 Methods

3.1 Bayesian Network

Bayesian Networks serve as a graphical modelling tool to illustrate the probabilistic relationships between events. As such they constitute a suitable method to investigate questions concerning obesity and its contributing factors. A BN is uniquely defined by its network structure and its parameterisation. The former is depicted as a directed acyclic graph (DAG) which shows the directional relations between events, and the latter is manifested by conditional probability tables (CPTs) which contain the conditional dependencies between variables. For the purpose of this study, functions needed to conduct inference queries on BNs, that is marginal joint and posterior distributions, and most likely explanation (MPE) and maximum a posteriori (MAP), were implemented. Furthermore, to reduce the complexity of such queries, d-separation, ordering, and network pruning is used.

3.2 Inference Engine

D-Separation D-separation is a method used to investigate whether two sets of variables (for example X and Y) are independent from each other given another set (for example Z). The simple idea is that these sets are either connected or independent in the graph of the network, therefore translated into d-separation and d-connectedness. Two variables are d-connected if there is an unblocked path between them, while two sets are d-separated when Z is blocking all connecting paths.

For the d-separation, an algorithm published by the MIT was used [15]. The algorithm splits into 4 easy steps. First, the ancestral graph is specified of the sets X and Y. Second, for each pair of variables in the ancestral graph connect the parents with an edge. This is called the moralization process. Third, the graph gets "disoriented" by replacing all directed edges with undirected edges. Forth and last, the given variables (set Z) and their edges get deleted. If X and Y are connected with edges they are d-connected, while, if they are not, they are d-separated. A visual display of the algorithm found in [15] can be found in Figure 1 in the Appendix. In the example, the sets A and B are not d-separated given D and F.

Ordering When BN variables are summed out or maximised out, the order of such an elimination process can affect computational budget. While there is no general rule of thumb, several heuristics exist, like the min-degree heuristic and the min-fill heuristic. Both of these were implemented, and the pseudo-code may be found in the Appendix.

Network Pruning When querying the BN, not necessarily all variables of the network need to be involved in the computation. Therefore, pruning redundant nodes and edges, resulting in a reduced number of variables in some CPTs, may save computational budget. On the one hand, node pruning entails the deletion of leaf nodes in case these are not part of the variables involved in a query. On the other hand, edge pruning can possibly be conducted if a query contains evidence, and consists of removing outgoing edges from the evidence. Subsequently, the CPTs of the disconnected child nodes are updated given the evidence. In general, pruning is an iterative process. The pseudo-code for the implementation of both pruning algorithms can be found in the Appendix.

Marginal Joint and Posterior Distributions Joint marginals and posterior marginals are queries of the form $\Pr(\mathrm{X})$ and $\Pr(\mathrm{X}|, e)$ respectively, where $\mathbf{X}$ represents the intersection of a set of variables and $\mathbf{e}$ is a set of instantiated variables (i.e. evidence). The implementation of both these algorithms are included in the Appendix.

MAP BNs may not only be queried for the probabilities of predefined instantiations, but also for the most likely instantiations of a set of variables X given a set of evidence e. Both, node and edge pruning can be applied before to reduce computational complexity. The Appendix contains the pseudo-code for the implementation of the MAP-algorithm.

MPE MPE may be seen as a special type of MAP queries. Instead of looking for the most likely instantiation of a subset of variables given some evidence e, MPE returns the most likely instantiation of all other variables in the BN. Consequently, node pruning is omitted for MPE. The pseudo-code for the full algorithm is contained in the Appendix.

3.3 Evaluation of BN Implementation

The implemented methods were tested on 7 BNs of different size regarding their applicability and quality of performance. The networks were generated with a self-built BN generator. The the amount of variables for the BNs is 5, 10, 15, 20, 25, 30, and 35 nodes. To evaluate the quality, the running time for the methods MAP and MPE with the heuristics min-degree, min-fill and random order were assessed with one random variable being set to True as evidence. Given the randomness of the random heuristic, the average running time of 50 runs was taken for this method.

The results of the evaluation are displayed in Figure 4 and 5 in the Appendix. For the MPE method, as can be seen in the plot, the running time of the two heuristics min-degree and min-fill are quite similar and increase with heightened amount of nodes. The running time of the random heuristic increases exponentially with a heightened amount of nodes. These results highlight the quality of the implemented heuristics. As can be seen, deciding smartly on a variable elimination order saves much computational budget in terms of running time and memory.

The heuristics display a different and interesting picture for the MPA method. All methods increase in running time with a heightened amount of nodes. However, the fastest ordering method for MPA is the random order, not the min-fill or min-degree heuristic. We could not find an explanation for this surprising finding. However, the quality of all methods and the MPA function was tested on different examples and compared to results highlighted in the course Knowledge Representation and, apart from the running time, seemed to work perfectly. Therefore we continued with this implementation for the use case analysis.

3.4 BN Implementation

In order to test the stated above hypothesis, a BN was implemented and tested with the described methods. The BN has three root nodes, namely parental education, age, and parental obesity. The root nodes are connected to several factors that were associated with obesity in previous literature. Moreover, the factors are connected amongst each other. As influential factors on obesity, this research takes physical activity, healthy eating habits, TV consumption, smoking, snacking habits, and alcohol consumption next to the mentioned root nodes into account. A visualization of the network can be found in Figure 2 in the Appendix.

The variable selection was based on research on influential factors of obesity. Various studies and meta analyses emphasize the interconnections between the variables and their relation to obesity. All data concerning the probabilities of the connections between the variables were based on associations found in the literature search or estimations based on these stated effects. The following variable definitions were derived from the literature and included in this research.

Parental Education Parental education is a factor that, amongst other things, influences the level of physical activity, healthy eating habits, and smoking habits of their children. Parental education was operationalized as a dichotomous variable, by differentiating between parents that have a university degree or higher (referred to as True in the BN) and parents who do not have a university degree (referred to as False in the BN). Studies that reported effects refarding the connection to the influenced variables were [11], [1], [4], [2], and [5].

Age Age influences the amount of physical activity a person participates in. Age is operationalized as a dichotomous variable, by differentiating between people above 50 (referred to as True in the BN) and people below that age (referred to as False in the BN). Studies that reported effects regarding the connection to the influence variables were [16], and [12].

Parental Obesity Parental Obesity refers to at least one parent having a BMI of 30 or above (referred to as True in the BN) and is treated as a dichotomous variable (with BMI ¡ 30 referred to as False in the BN). Parental obesity influences Healthy Eating and obesity of their children [13], [10], [4], and [19].

Obesity Obesity is characterized as having a BMI above 30 (referred to as True in the BN while a BMI ¡ 30 is referred to as False) [19]. Since this research is focusing on the influencing factors on obesity, the consequences of obesity were neglected in this paper.

Smoking Previous research highlighted smoking as one predecessor of obesity [8] and [2]. Smoking was also operationalized in a dichotomous manner with characterizing oneself as a smoker being referred to as True in the BN, while characterizing oneself as a non-smoker is being referred to as False.

Physical Activity An influencing factor for TV consumption and smoking, which in turn were associated with other predecessors of obesity, is physical activity [8], [5] and [17]. Physical activity was defined as being active at least 60 min. per day (referred to as True in the BN) or not (referred to as False in the BN).

Healthy Eating Habits Healthy Eating habits can be defined by eating more than two portions of fruit and/or vegetables per day [1](referred to as True in the BN; less fruit intake is being referred to as False in the BN). Healthy eating influences, amongst other things, alcohol consumption and TV consumption [10][1].

Snacking Snacking habits can be classified separately next to healthy eating habits as food intake that influences obesity. It is classified as sweet and high calorie food on more than five days a week [5] (referred to as True in the BN; False otherwise). It was identified as a direct predecessor of obesity [5].

Alcohol consumption Alcohol consumption was defined as consuming alcohol at least one day a week. Moreover, it had a direct influence on the amount of snacking [14][5].

TV Consumption Last not but not least, high TV consumption was identified as an interconnected influence on obesity. High TV consumption was defined as watching four or more hours of TV a day [14]. Previous research associated high TV consumption with heightened alcohol and snacking intake [14].

As aforementioned, the nodes edges and probabilities of the BN and according variables were derived from the stated above research. In order to work with the network it was converted into a .BIFXML file. The CPTs for each node can be found in the Appendix.

4 Experimental Set-up

In order to investigate the the influences of nature vs. nurture factors on obesity, an a-priori marginal query, a posterior marginal query and one MAP and one MEP query were conducted with different nature and nurture variables as evidence. Min-fill was used as an order heuristic. Parental obesity was used to represent nature, while parental education and healthy lifestyle were used as nurture variables. For this experiment, the variables physical activity (True), healthy eating (True), and alcohol consumption (False) were clustered into the group healthy lifestyle.

First, with the a-priori marginal function, the following probabilities were investigated:

- Pr(Obesity=True, Parental Obesity=True)
- Pr(Obesity=True, Parental Obesity=False)
- Pr(Obesity=True, Healthy Lifestyle=True) and
- Pr(Obesity=True, Healthy Lifestyle=False)

Second, with the posterior marginal query, the following probabilities were investigated:

- Pr(Obesity=True|Parental Obesity=True, Healthy Lifestyle=False),
- Pr(Obesity=True|Parental Obesity=False, Healthy Lifestyle=True),
- Pr(Obesity=True|Parental education=True) and
- Pr(Obesity=True|Parental education=False)

Third, the following probabilities were investigated with the MPE method:

- Pr(Obesity=True, Parental Obesity=False)
- Pr(Obesity=True, Parental Obesity=True)
- Pr(Obesity=True, Healthy Lifestyle=True)
- Pr(Obesity=True, Healthy Lifestyle=False)

forth, the following probabilities were investigated with the MAP method:

- Pr(Obesity|Parental Obesity=True)
- Pr(Obesity|Parental Obesity=False)
- Pr(Obesity|Healthy Lifestyle=True)
- Pr(Obesity|Healthy Lifestyle=False)
- Pr(Obesity|Parental Obesity=True, Healthy Lifestyle=True)
- Pr(Obesity|Parental Obesity=False, Healthy Lifestyle=False)
- Pr(Obesity|Parental Obesity=True, Healthy Lifestyle=False)
- Pr(Obesity|Parental Obesity=False, Healthy Lifestyle=True)

5 Results

5.1 Prior Marginals

The probability for all stated above conditions can be found in the tables below. The probabilities for the prior marginals are all low despite the probability for parental obesity=True and obesity=False, which is 0.77. Therefore, it is generally unlikely to be obese, despite parental obesity. Additionally, the probability for obesity given that the parental obesity is False is especially small (p=0.01). The high probability of obesity=True and obesity=False in this research was surprising and will be discussed in the according section.

Table 1: Pr(Obesity, Parental Obesity)

Parental Obesity	**Obesity**	**p**
False	False	0.115438
False	True	0.014562
True	False	0.773886
True	True	0.115254

5.2 Posterior Marginals

As can be seen in Table 2, the results for the posterior marginals show very similar probabilities. Despite the different variables given as evidence, the probability for obesity=False was around p=0.88 for all four conditions while obesity=True was p=0.12.

Obesity	p
False	0.870293
True	0.129707

Obesity	p
False	0.892249
True	0.107751

Obesity	p
False	0.889324
True	0.129816

Obesity	p
False	0.889324
True	0.129816

Table 2:
∗I: Pr(Obesity |Parental Obesity = True, Healthy Lifestyle = False)
∗II: Pr(Obesity |Parental Obesity = False, Healthy Lifestyle = True)
∗III: Pr(Obesity |Parental Education = True)
∗IV: Pr(Obesity |Parental Education = False)

5.3 MPE

The tables depicting the results for the MPE method highlight the most likely instantiation of the all variables with the according probability. All probabilities for the set conditions were close to zero. Furthermore, the assignment of the most likely instantiations was similar for all four evidence conditions, excluding the instantiations based on the evidence of course. Most variables were assigned to True across all four conditions. However, some minor changes did occur as stated in the tables.

Table 3: MPE Evidence: { Obesity = True, Parental Obesity = True }

True	False	p
Parental Education	Physical Activity	0.004131
Age	-	-
Parental Obesity	-	-
Smoking	-	-
Alcohol Consumption	-	-
Healthy Eating	-	-
Snacking	-	-
TV Consumption	-	-
Obesity	-	-

Table 4: MPE Evidence: { Obesity = True, Parental Obesity = False }

True	False	p
Parental Education	Physical Activity	0.00058
Age	Snacking	-
Parental Obesity	-	-
Smoking	-	-
Alcohol Consumption	-	-
Healthy Eating	-	-
TV Consumption	-	-
Obesity	-	-

5.4 MAP

The table below displays the most likely instantiation of obesity given the different evidences and corresponding probabilities. The assigned instantiation for obesity was False in all cases and all probabilities were around 0.1.

Table 5: MPE Evidence: { Obesity = True, Healthy Lifestyle = True }

True	False	p
Parental Education	Alcohol Consumption	0.000447
Age	Smoking	-
Parental Obesity	Snacking	-
Healthy Eating	-	-
TV Consumption	-	-
Obesity	-	-
Physical Activity	-	-

Table 6: MPE Evidence: { Obesity = True, Healthy Lifestyle = False }

True	False	p
Age	Snacking	0.00394
Obesity	Healthy Eating	-
Parental Obesity	Parental Education	-
Smoking	Physical Activity	-
TV Consumption	-	-
Parental Obesity	-	-

Obesity			
Parental Obesity	Healthy Lifestyle	Most likely Instantiation	Probability
False	False	False	0.085
False	True	False	0.0815
True	False	False	0.0919
True	True	False	0.093
False	-	False	0.103
True	-	False	0.078
-	False	False	0.0902
-	True	False	0.079

Table 7: Table depicting the most likely instantiation for obesity using the maximum a posteriori(MAP) estimation. The two left columns depict the given variable for these cases. The right column depicts the most likely instantiation for obesity and corresponding probability.

6 Discussion

This research aimed to investigate influencing factors on obesity with respect to nature vs. nurture. The results highlighted that generally all probabilities were very small. Moreover, the most likely instantiations did not expose a significant contrast between the different conditions. We explain the considerably small probabilities regarding obesity with the overall low percentage of people that are obese. Despite obesity being a growing concern in modern society, most people still have an BMI below 30. The overall low percentage of people that have obesity also explains the assigned instantiation of False in each condition for the MAP result. Moreover, this research could not gain insights into the differentiation between nature vs. nurture factors. Despite comparing nature and nurture evidences and their effect on obesity, we could not identify a pattern highlighting a decisive origin of the issue. As a possible reason, we suggest that environmental and genetic influences are so highly intertwined that it is not possible to identify corresponding main and interaction effects, especially due to the relatively simple study design and low baseline probability to have obesity. Second, another reason for these vague results could be a too simple BN. The time and budget constraints of this research limited the complexity of our BN and therefore significant connections and interactions between the variables as well as unconsidered important factors could be missing.

6.1 Limitations and Future Research

This research contains several limitations. First, the data acquisition was incomplete given the novelty of a nature vs. nurture analysis regarding obesity with a use of a BN and, therefore, finding specific data on that subject was not possible. Moreover, the used literature included papers from the last 30 years as well as different age groups. The different time and age background diminishes the validity of this research. The surprising result in Table 1, assigning a probability of 0.77 to parental obesity=True and obesity=False, could be traced back to this critical issue. Last not but least, given the multi-factorial influences obesity inherits, the BN used in this research missed many influential factors. This could have distorted the results. For future research, we suggest to take a larger research body with precise probability values and all influential factors into account.

7 Conclusion

The conducted research aimed to bring insight into the causality of nature vs. nurture factors on obesity. Our research could not bring new understandings about the impact of genetic and environmental factors as well as their interaction, nonetheless we hope that our research sheds a light on the danger of obesity despite the low probabilities found in this study. In conclusion, our research proposes a high interplay between nature and nurture factors that are not easy to unravel. Furthermore, we encourage colleagues to not get discouraged by our result as well as to take our study a step further and investigate the interesting and complex network of obesity with taking all its predecessors into account.

References

1. van Ansem, W.J., Schrijvers, C.T., Rodenburg, G., van de Mheen, D.: Maternal educational level and children's healthy eating behaviour: role of the home food environment (cross-sectional results from the inpact study). Int J Behav Nutr Phys Act **11**, 113 (2014). https://doi.org/10.1186/s12966-014-0113-0
2. Assari, S., Boyce, S., Caldwell, C.H., Bazargan, M.: Parent education and future transition to cigarette smoking: Latinos' diminished returns. Frontiers in pediatrics **8**, 457 (2020). https://doi.org/10.3389/fped.2020.00457
3. Chu, Y., Yang, J., Shi, J., Zhang, P., Wang, X.: Obesity is associated with increased severity of disease in COVID-19 pneumonia: a systematic review and meta-analysis. European Journal of Medical Research **25**(1) (2020). https://doi.org/10.1186/s40001-020-00464-9
4. Chung, W., Lim, S.: Factors contributing to educational differences in obesity among women: evidence from south korea. BMC Public Health **20**, 1136 (2020). https://doi.org/10.1186/s12889-020-09221-3
5. Gregori, D., Foltran, F., Ghidina, M., Berchialla, P.: Understanding the influence of the snack definition on the association between snacking and obesity: a review. International Journal of Food Sciences and Nutrition **62**(3), 270–275 (2011). https://doi.org/10.3109/09637486.2010.530597, https://doi.org/10.3109/09637486.2010.530597
6. Jackson, S.E., Llewellyn, C.H., Smith, L.: The obesity epidemic – Nature via nurture: A narrative review of high-income countries. SAGE Open Medicine **8**, 205031212091826 (2020). https://doi.org/10.1177/2050312120918265
7. München, T.U.: Studie zu ernährung und bewegung seit beginn der corona-pandemie vorgestellt. "corona befeuert eine andere pandemie" (2021)
8. Nduaguba, S.O., Ford, K.H., Rascati, K.: The role of physical activity in the association between smoking status and quality of life. Nicotine tobacco research : official journal of the Society for Research on Nicotine and Tobacco **21**, 1065–1071 (2019). https://doi.org/10.1093/ntr/nty052
9. ProCon.org: Global obesity levels (2018)
10. Statista: Would you say that you eat a healthy and balanced diet? (2019)
11. voor de Statistiek, C.B.: No gender gap in higher education. CBS-Statistics Nederland (2019)
12. voor de Statistiek, C.B.: Half of dutch population meet physical activity guidelines (2020)
13. Tang, D., Bu, T., Dong, X.: Are parental dietary patterns associated with children's overweight and obesity in china? BMC Pediatrics **20**(1) (2020). https://doi.org/10.1186/s12887-020-1910-z
14. Tebar, W.R., Christofaro, D.G.D., Diniz, T.A., Lofrano-Prado, M.C., Botero, J.P., Correia, M.d.A., Cucato, G.G., Ritti-Dias, R.M., Prado, W.L.d.: Increased screen time is associated with alcohol desire and sweetened foods consumption during the covid-19 pandemic. Frontiers in Nutrition **8** (2021). https://doi.org/10.3389/fnut.2021.630586
15. of Technology, M.I.: d-separation: How to determine which variables are independent in a bayes net (2015)
16. government UK, S.: Age groups (2018)
17. Vale, S., Ricardo, N., Soares-Miranda, L., Santos, R., Moreira, C., Mota, J.: Parental education and physical activity in pre-school children. Child: Care, Health and Development **40**(3), 446–452 (2013). https://doi.org/10.1111/cch.12041

18. WHO: Obesity: Health consequences of being overweight (2013)
19. WHO: Obesity and overweight (2020)

8 Appendix

8.1 Appendix for the Methods Section

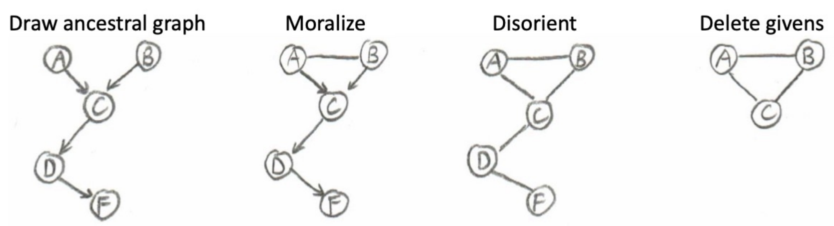

Fig. 1: D-separation algorithm testing whether A and B are conditionally independent, given D and F

8.2 Pseudocode

Algorithm 1 MinDegree(**N, X**)

Input:
N: Bayesian network
X: Variables in network **N**
Output:
an ordering π of variables **X**
Main:

$G \leftarrow$ interaction graph of the CPTs in network **N**
for $i = 1$ to number of variables in **X** **do**
    $\pi(i) \leftarrow$ variable in **X** with smallest number of neighbors in G
    add an edge between every pair of non-adjacent neighbors of $\pi(i)$ in G
    delete variable π from G and from **X**
end for
return π

Algorithm 2 MinFill(**N, X**)

Input:
N: Bayesian network
X: Variables in network **N**
Output:
an ordering π of variables **X**
Main:

```
1: G ← interaction graph of the CPTs in network N
2: for i = 1 to number of variables in X do
3:     π(i) ← variable in X with that adds the smallest number of edges on Line 4
4:     add an edge between every pair of non-adjacent neighbors of π(i) in G
5:     delete variable π from G and from X
6: end for
return π
```

Algorithm 3 PriorMarginal(N, Q, π)

Input:
N: Bayesian network
Q: Variables in network **N**
π: Ordering of network variables not in **Q**
Output:
The prior marginal Pr(**Q**)
Main:

$S \leftarrow$ CPTs of network **N**
for $i = 1$ to length of order π **do**
    $f \leftarrow \prod_k f_k$, where f_k belongs to S and mentions variable $\pi(i)$
    $f_1 \leftarrow \sum_{\pi(i)} f$
    replace all factors f_k in S by factor f_i
end for
return $\prod_{f \in S} f$

Algorithm 4 PosteriorMarginal(**N, Q, e,** π)

Input:
N: Bayesian network
Q: Variables in network **N**
e: Instantiation of some variables in network **N**
π: Ordering of network variables not in **Q**
Output:
The posterior marginal $\Pr(Q \mid e)$
Main:
$S \leftarrow \{f^e : f$ is a CPT of network **N**$\}$
for $i = 1$ to length of network **N** **do**
    $f \leftarrow \prod_k f_k$, where f_k belongs to S and mentions variable $\pi(i)$
    $f_1 \leftarrow \sum_{\pi(i)} f$
    replace all factors f_k in S by factor f_i
end for
divide marginal posterior distribution values by Pr(**e**)
return $\prod_{f \in S} f$

Algorithm 5 MAP(**N, M, e**)

Input:
N: Bayesian network
M: Some variables in network **N**
e: evidence ($E \cap M = \emptyset$)
Output:
Trivial factor containing the MAP probability $MAP_P(M, e)$
Main:
$N' \leftarrow$ pruneNetwork(N, M, e) ▷ Prune nodes and edges
$\pi \leftarrow$ a variable elimination order for N' in which variables M appear last
$S \leftarrow \{f^e : f$ is a CPT of network **N'**$\}$
for $i = 1$ to length of order π **do**
    $f \leftarrow \prod_k f_k$, where f_k belongs to S and mentions variable $\pi(i)$
    if $\pi(i) \in M$ **then**
        $f_i \leftarrow max_{\pi(i)} f$
    else
        $f_1 \leftarrow \sum_{\pi(i)} f$
    end if
    replace all factors f_k in S by factor f_i
end for
return trivial factor $\prod_{f \in S} f$

Algorithm 6 MPE($\boldsymbol{N}$, $\mathbf{e}$)

input:
$\boldsymbol{N}$: a Bayesian Network
$\mathbf{e}$: evidence
output: trivial factor f, where $f(\top)$ is the MPE probability in evidence $\mathbf{e}$
main:

$\boldsymbol{N}\leftarrow$prune Edges($\boldsymbol{N}$,$\mathbf{e}$)$\mathbf{Q}\leftarrow$variables in network $\boldsymbol{N'}$
$\pi\leftarrow$elimination order of variables $\mathbf{Q}$
$S\leftarrow\{f^{\mathrm{e}}$: f is a CPT in network $\boldsymbol{N'}\}$
for $i = 1$ to $|\mathbf{Q}|$ **do**
    $f\leftarrow\prod_{\mathrm{k}} f_{\mathrm{k}}$, where f_{k} belongs to S and mentions variable $\pi(i)$
    $f_i\leftarrow\max_{\pi(i)} f$
    replace all factors f_{k} in S by factor f_i
end for
return trivial factor $\prod_{f\in S} f$ =0

8.3 Bayesian Network

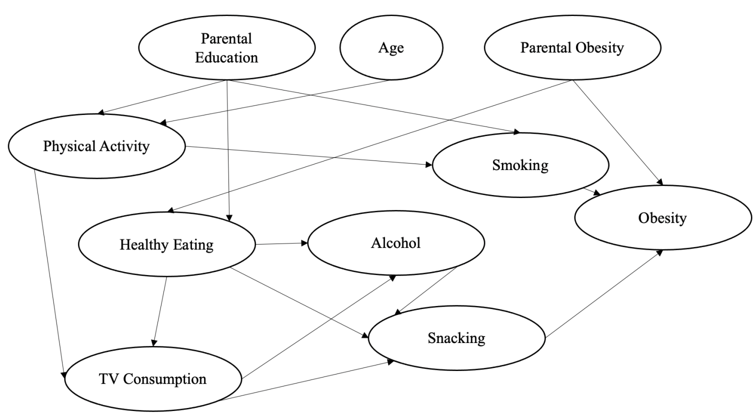

Fig. 2: Influential Factors of Obesity

Table 8

Healthy Eating	TV Consumption	Alcohol	p
False	False	False	0.22
False	False	True	0.78
False	True	False	0.20
False	True	True	0.80
True	False	False	0.30
True	False	True	0.70
True	True	False	0.28
True	True	True	0.72

Table 9

Parental Education	Parental Obesity	Healthy Eating	p
False	False	False	0.50
False	False	True	0.50
False	True	False	0.75
False	True	True	0.25
True	False	False	0.33
True	False	True	0.67
True	True	False	0.39
True	True	True	0.61

Table 10

Parental Obesity	Snacking	Smoking	Obesity	p
False	False	False	False	0.92
False	False	False	True	0.08
False	False	True	False	0.87
False	False	True	True	0.13
False	True	False	False	0.90
False	True	False	True	0.10
False	True	True	False	0.87
False	True	True	True	0.13
True	False	False	False	0.89
True	False	False	True	0.11
True	False	True	False	0.88
True	False	True	True	0.12
True	True	False	False	0.85
True	True	False	True	0.15
True	True	True	False	0.85
True	True	True	True	0.15

Table 11

Age	**p**
False	0.22
True	0.78

Table 12

Parental Education	**p**
False	0.3
True	0.7

Table 13

Parental Obesity	**p**
False	0.13
True	0.87

Table 14

Physical Activity	**Parental Education**	**Smoking**	**p**
False	False	False	0.70
False	False	True	0.30
False	True	False	0.76
False	True	True	0.24
True	False	False	0.88
True	False	True	0.22
True	True	False	0.80
True	True	True	0.20

Table 15

Healthy Eating	**Parental Education**	**Smoking**	**p**
False	False	False	0.30
False	False	True	0.70
False	True	False	0.45
False	True	True	0.55
True	False	False	0.40
True	False	True	0.60
True	True	False	0.55
True	True	True	0.45

Table 16

Healthy Eating	**TV Consumption**	**Alcohol**	**Snacking**	**p**
False	False	False	False	0.73
False	False	False	True	0.27
False	False	True	False	0.60
False	False	True	True	0.40
False	True	False	False	0.65
False	True	False	True	0.35
False	True	True	False	0.60
False	True	True	True	0.40
True	False	False	False	0.70
True	False	False	True	0.30
True	False	True	False	0.60
True	False	True	True	0.40
True	True	False	False	0.65
True	True	False	True	0.35
True	True	True	False	0.50
True	True	True	True	0.50

Table 17

Physical Activity	**Healthy Eating**	**TV Consumption**	**p**
False	False	False	0.2
False	False	True	0.8
False	True	False	0.4
False	True	True	0.6
True	False	False	0.5
True	False	True	0.5
True	True	False	0.3
True	True	True	0.7

8.4 Appendix for the Evaluation of MAP and MPE

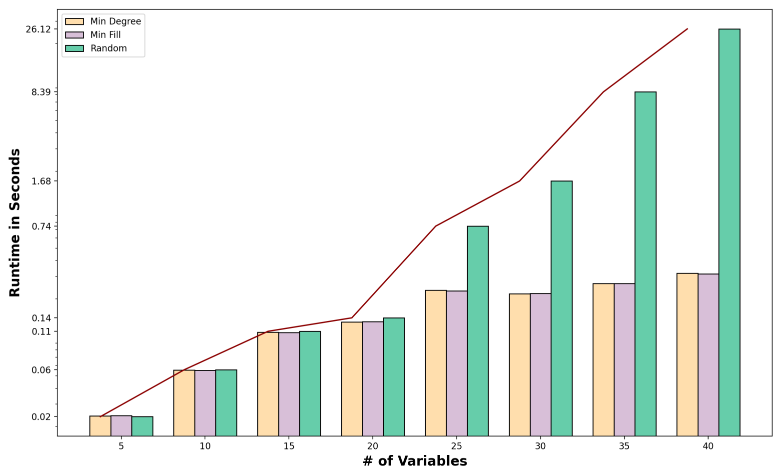

Fig. 3: Results of the evaluation of MPE

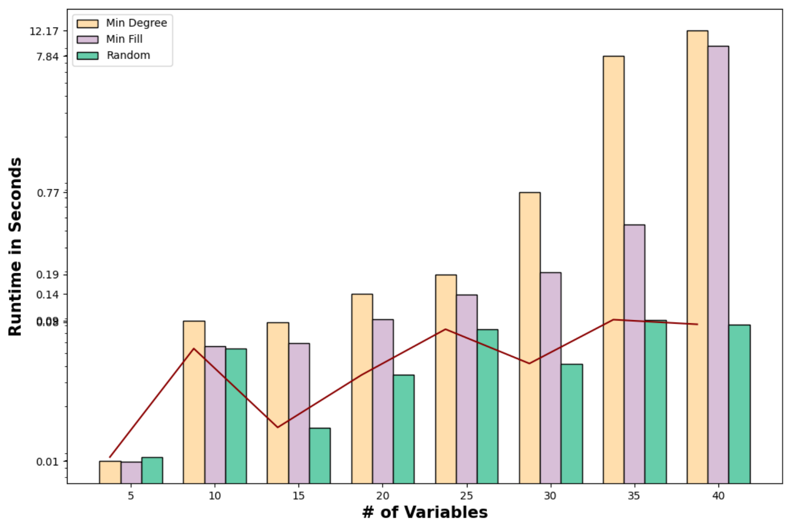

Fig. 4: Results of the evaluation of MAP