Jurnal Sisfotek Global
ISSN (Online): 2721 - 3161, ISSN (Print): 2088 – 1762
DOI: http://dx.doi.org/10.38101/sisfotek.v15i2.15992
Vol. 15, No. 2, September 2025, pp. 74-80

Causal Relationship of Factors Influencing Heart Disease and Diabetes Using the GES Algorithm

Nurhaeka Tou[1*], Putri Mentari Endraswari[2], & Iski Zaliman[3]

[1,2,3] Teknologi Informasi, Universitas Bangka Belitung, Bangka, Indonesia, 33172
E-mail: [1]nurhaeka@ubb.ac.id, [2]putrimentari@ubb.ac.id, [3]iski.zaliman@ubb.ac.id

ARTICLE HISTORY

Received : August 16, 2025
Revised : September 18, 2025
Accepted : Septembet 27, 2025

KEYWORDS

Causality
Diabetes
Greedy Equivalence Search
Heart
Modelling

ABSTRACT

Diabetes and heart disease are leading causes of global death, and if not properly managed, can be fatal. Previous studies have explored the factors influencing these two diseases through correlation analysis, multivariate analysis, machine learning, and deep learning. However, these approaches generally only identify associations without being able to predict causal relationships. This study aims to model the causal relationships between factors in two clinical datasets: heart disease (13 parameters) and diabetes (9 parameters), in order to support early diagnosis and prevention. The Greedy Equivalence Search (GES) algorithm is used to determine the direction of the causal relationship between parameters. The results show that heart disease exhibits three directional relationships: between blood pressure and age, between Maximum Heart Rate (MHR) and age, and between age and cholesterol. Then, diabetes exhibits two bidirectional relationships: blood pressure and BMI, then BMI and Diabetes Pedigree Function. In addition, diabetes also exhibits three directional relationships: Diabetes Pedigree Function and Glucose, BMI and Glucose, and Blood Pressure and Glucose. Thus, it can be concluded that the algorithm can identify the causal relationship between diabetes and heart disease.

1. Introduction

Machine learning and deep learning are two fields of artificial intelligence that are being increasingly applied in clinical domains. Both fields are used to build automated systems that can discover various models from a set of data. Learning techniques in these two fields are categorized into two main types: supervised learning and unsupervised learning [1][2]. Supervised learning relates to classification and prediction techniques, such as Naïve Bayes classification, Support Vector Machine (SVM), and artificial neural networks. Meanwhile, unsupervised learning is related to clustering and component analysis, such as K-Means, Independent Component Analysis, and authenticators[3]. However, both machine learning and deep learning only target the association between input variables (X) and output variables (Y), namely P(Y|X). Here, both variables X and Y are observed domains. This learning process is made to project the outcome Y based on the given input value X [4]. For example, studying whether a person has diabetes (output) or not, based on the results of a clinical examination (input) obtained from the hospital. This process is called classification. Although the resulting classification model is fairly accurate, the learning process results in a large number of trained model parameters that are difficult to explain [4]. This then becomes a drawback of machine learning and deep learning [5].

A primary goal in various scientific fields, including clinical science, is to create generative models. These models are expected to show how variables in a data set obtain their values. However, this goal cannot be achieved directly using machine learning and deep learning models, as these two fields generally focus only on the relationship between input (X) and output (Y) variables, P(Y|X) [6].

Therefore, causal modeling is an alternative method that can be implemented to solve generative models. Causal models essentially study how the value of variable Y changes based on variable X. Variable X is often obtained from an intervention, such as medical therapy. Unlike P(X|Y) in machine learning and deep learning, in this causal model we will look at the mechanism generated by P(Y| do (x); z), that is, what will be Y, if we do X = x, and observe the variable Z. This mechanism can help us to determine the model obtained based on the data we have. In the clinical field, causal models are useful for developing better therapy and treatment models, because we will focus on variables that have a large impact on the disease[7] [8]. Causal modeling is generally divided into two stages: the score-based stage and the constraint-based stage[9]. The constraint-based stage uses conditional independence tests and orientation rules to derive a causal model. Meanwhile, the score-based stage utilizes scores to maximize the causal model, which will form the basis of the final causal model[10] [1].

This study refers to several previous studies that also used the GES algorithm for causal modeling, such as the first study entitled "Causal Modeling of Factors in Stunting Using the Peter Clark and Greedy Equivalence Search Algorithms" [11]. The study conducted causal modeling of eight parameters causing stunting in toddlers. The study showed that the GES algorithm could identify a causal relationship of 0.66 based on a comparison of directional density performance.

The second study was entitled "Applying PC Algorithm and GES to Three Clinical Data Sets: Heart Disease, Diabetes, and Hepatitis"[12]. The results of the study showed that the PC and GES algorithms were able to represent causal models of the three data sets.

Therefore, in this case study, we aim to apply causal modeling to estimate generative models on two clinical datasets: heart disease and diabetes mellitus. Both conditions are among the top five causes of death globally, making understanding the underlying causal mechanisms of each disease crucial. This is especially true in recent years, as the medical world has faced challenges in developing more efficient and effective treatments and care.Technically, this study will employ a score-based method known as GES (Greedy Equivalent Search). The authors acknowledge that this causal method is not the most advanced. However, GES represents the foundation and simplicity of score-based approaches in general. On the other hand, newer causal methods actually build upon this approach, adding more complex assumptions, steps, or algorithms. Since the primary objective of this study is to introduce causal modeling as an alternative approach, the authors chose to use a simpler method, such as the GES algorithm.

2. Research Methodology

This study uses the R programming language to analyze the data. Specifically, the authors used the pclag package to implement the GES algorithm. For each data set, the authors deliberately selected several continuous and interesting variables. Since the main objective of this study is to introduce causal modeling as an alternative method, we will conduct a simple application, rather than a comparison. Here, we apply the GES algorithm to two types of diseases: diabetes and heart disease.

2.1 Heart Disease Data

The heart disease dataset was obtained from kaggle.com, which contains observational data from 1,000 respondents suspected of having heart disease. Each respondent has 13 attributes related to their health condition. In this study, several specific variables were selected, such as resting blood pressure (Trestbps), cholesterol levels, fasting blood sugar (fbs), resting electrocardiogram results (restecg), maximum heart rate (thalach), exercise-induced ST depression (oldpeak), number of major blood vessels (ca), and the patient's thalassemia category (thal).

2.2 Diabetes Disease Data

The Pima Indian diabetes data set was obtained from the National Institute of Diabetes, Digestive, and Kidney Diseases and also through kaggle.com. This data set consists of 768 respondents with nine parameters, consisting of continuous attributes and one outcome attribute. In this case, the researchers selected the attributes Glucose, Blood Pressure, Skin Thickness, Insulin, BMI, Diabetes Pedigree Function, Age, and target.

2.3 Method

This research will use the R programming language to analyze the data. Specifically, the researchers used the pcalg package to implement the Greedy Equivalence Search (GES) algorithm.

The causal algorithm is a causal method based on score values for comparing models based on the fit of the resulting models. The GES algorithm in causal science can be used to construct causal structures in the form of directed graphs. In the GES algorithm, each iteration selects changes that can improve the causal model based on statistical criteria, such as the Bayesian score. The concept of the GES algorithm can be seen in The Concept of Greedy Equivalence Search AlgorithmFigure 1.

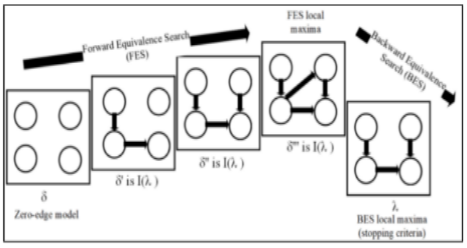


Figure 1.The Concept of Greedy Equivalence Search Algorithm

Referring to Figure 1, the GES algorithm utilizes the Bayesian Information Criterion (BIC) score calculated from Gaussian or multinomial data, especially for data with a higher match rate. This algorithm consists of two main stages: the forward equivalence search (FES) stage (described in Algorithm 2) and the backward equivalence search stage (described in Algorithm 3).

The GES algorithm first starts with a graph model with no edges. Next, each single edge obtained through a progressive method is added to the DAG graph. This addition stage is called the forward equivalence search (FES) stage. However, if the FES stage has not yet achieved an improvement in the score, then the FES stage is advanced using the BIC stage. The BIC stage will improve the score and eliminate or add edges to the temporary DAG causal model. Technically, the GES algorithm has two modeling stages: a search stage that begins with an empty graph or a very simple graph.

Then, the GES algorithm will add a forward or elimination phase (backward phase) process to build variable relationships into a complex graph and find a causal model capable of representing the values in the data (see Figure 2 and Figure 3). The second stage in the GES algorithm is the Greedy stage, which is used to increase changes in building more complex graphs (see Figure 4).

Algorithm 2
Add Operation (Forward Phase)

Input *A* dataset *D* and a CPDAG G
Output *A* CPDAG *H*
Let *M* be a set of CPDAGs obtainable from *G* by ***adding*** one edge
Let *K* be a CPDAG from *M* with the highest score *S(D,K)*

if *S(D,G)) < S(D,K)* ***then***
H ← *K*
else
H ← *G*
end if
return H

Figure 2. Algorithm 2 GES

Figure 2 above is the process of adding one edge to improve the complete partial directed acyclic graph (CPDAG) framework. The edge addition stage begins with graph G. The algorithm searches for a new set of graph M by adding edges to graph G. Next, in graph M, a set K with a high score will be selected, named S(D,K) and used to measure its suitability with the set in data D. If graph K has a higher score than graph G, then graph H will be modified into graph K. However, if not, graph H remains G. The process aims to improve the graph by adding edges that can increase its similarity to the data set. Next, the Backward Phase stage can be seen in Figure 3.

Algorithm 3
Delete Operation (Backward Phase)

Input *A* dataset *D* and a CPDAG G
Output *A* CPDAG *H*
Let *M* be a set of CPDAGs obtainable from *G* by ***deleting*** one edge
Let *K* be a CPDAG from *M* with the highest score *S(D,K)*

if *S(D,G)) < S(D,K)* ***then***
H ← *K*
else
H ← *G*
end if
return H

Figure 3. Algorithm 3 GES

Figure 3 above is an algorithm used to remove one edge in the CPDAG graph, so that the framework can be improved. At that stage, if one edge of graph G has been removed, the algorithm will create a set of graphs M. Then, graph K, which has a high score S(D,K) will be used to measure the effectiveness of the model obtained based on the selected data D. If removing an edge results in a high score on graph G, then graph H is modified into graph K. However, if removing an edge does not increase the score, then graph H remains graph G. Removing edges that do not provide substantial contributions to the model will improve the structure of the causal framework. Furthermore, when the forward phase has been completed and there are no more edges that can improve the assessment score, the GES algorithm will move to the backward phase. Technically, the GES algorithm can be seen in Figure 4.

Algorithm 4
Greedy Equivalence Search Algorithm

Input *A* dataset *D*
Output *A* CPDAG *G*

while *S(D, Add(D,G)) > S(D,G)* ***do***
G ← *Add(D,G)*
end while
while *S(D, Del(D,G)) > S(D,G) do*
G ← *Del(D,G)*
end while
return G

Figure 4. Algorithm 4 GES

2.4 Research Stages

This research was conducted in several stages: literature review, data pre-processing, causal modeling analysis, evaluation, and dissemination. Details of the research stages can be seen in Figure 5.

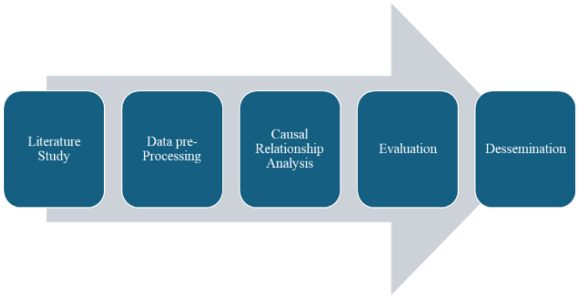


Figure 5. Research Stages

This research stage begins with a literature study stage to review the findings of several previous studies that are relevant to this research [13],[14],[15],[16] with regard to the incidence of diabetes and heart disease.The second stage is data pre-processing, which aims to check the completeness of the data set for any missing data values and to examine the distribution of the data. At this stage, if there is missing data, it can be resolved by cleaning it using the code NewData <- Data[complete.case(data)]. The third stage of this research is applying the GES algorithm to the data set that is ready for analysis. Model computation is performed using the R programming language and packages called "pcalg" and "GaussL0penObsScore". The output of this stage is a causal model of the factors that cause diabetes and heart disease. The goal of causal discovery is to identify causal structures, namely

the patterns of influence between predictor variables on each other or on the desired response variable. The process then continues with model evaluation and dissemination of the resulting causal model.

3. Results and Discussion

3.1 Results

The results of this study are described to obtain a causal model using the Greedy Equivalence Search (GES) method with the following data details.

```
> str(kausal)
'data.frame':   1025 obs. of  14 variables:
 $ age     : int  52 53 70 61 62 58 58 55 46 54 ...
 $ sex     : int  1 1 1 1 0 0 1 1 1 1 ...
 $ cp      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ trestbps: int  125 140 145 148 138 100 114 160 120 122 ...
 $ chol    : int  212 203 174 203 294 248 318 289 249 286 ...
 $ fbs     : int  0 1 0 0 1 0 0 0 0 0 ...
 $ restecg : int  1 0 1 1 1 0 2 0 0 0 ...
 $ thalach : int  168 155 125 161 106 122 140 145 144 116 ...
 $ exang   : int  0 1 1 0 0 0 0 1 0 1 ...
 $ oldpeak : num  1 3.1 2.6 0 1.9 1 4.4 0.8 0.8 3.2 ...
 $ slope   : int  2 0 0 2 1 1 0 1 2 1 ...
 $ ca      : int  2 0 0 1 3 0 3 1 0 2 ...
 $ thal    : int  3 3 3 3 2 2 1 3 3 2 ...
 $ target  : int  0 0 0 0 0 1 0 0 0 0 ...
```

Figure 6. Heart disease dataset details

```
> str(kausal)
'data.frame':   768 obs. of  9 variables:
 $ Pregnancies             : int  6 1 8 1 0 5 3 10 2 8 ...
 $ Glucose                 : int  148 85 183 89 137 116 78 115 197 125 ...
 $ BloodPressure           : int  72 66 64 66 40 74 50 0 70 96 ...
 $ SkinThickness           : int  35 29 0 23 35 0 32 0 45 0 ...
 $ Insulin                 : int  0 0 0 94 168 0 88 0 543 0 ...
 $ BMI                     : num  33.6 26.6 23.3 28.1 43.1 25.6 31 35.3 30.5 0 ...
 $ DiabetesPedigreeFunction: num  0.627 0.351 0.672 0.167 2.288 ...
 $ Age                     : int  50 31 32 21 33 30 26 29 53 54 ...
 $ Outcome                 : int  1 0 1 0 1 0 1 0 1 1 ...
```

Figure 7. Diabetes disease dataset details

Figure 6 dan Figure 7 above detail the variables in the diabetes and heart disease datasets. Next, the researchers performed data preprocessing to identify missing values before performing calculations. Missing data was checked using the code New Data <- Data[complete.cases(keeps)], specifying the variables to be calculated, and a Gaussian conditional independence test using 0.05. Then, identification was performed based on the correlation values generated for each variable.

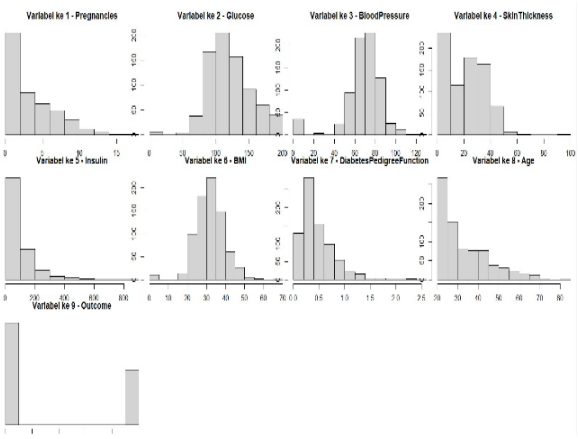


Figure 8. Diabetes Dataset Distribution

```
> cor(dataku[,1:9])
                          Pregnancies     Glucose BloodPressure SkinThickness
Pregnancies                1.00000000 0.12945867    0.14128198   -0.08167177
Glucose                    0.12945867 1.00000000    0.15258959    0.05732789
BloodPressure              0.14128198 0.15258959    1.00000000    0.20737054
SkinThickness             -0.08167177 0.05732789    0.20737054    1.00000000
Insulin                   -0.07353461 0.33135711    0.08893338    0.43678257
BMI                        0.01768309 0.22107107    0.28180529    0.39257320
DiabetesPedigreeFunction -0.03352267 0.13733730    0.04126495    0.18392757
Age                        0.54434123 0.26351432    0.23952795   -0.11397026
Outcome                    0.22189815 0.46658140    0.06506836    0.07475223
                            Insulin       BMI DiabetesPedigreeFunction
Pregnancies               -0.07353461 0.01768309              -0.03352267
Glucose                    0.33135711 0.22107107               0.13733730
BloodPressure              0.08893338 0.28180529               0.04126495
SkinThickness              0.43678257 0.39257320               0.18392757
Insulin                    1.00000000 0.19785906               0.18507093
BMI                        0.19785906 1.00000000               0.14064695
DiabetesPedigreeFunction   0.18507093 0.14064695               1.00000000
Age                       -0.04216295 0.03624187               0.03356131
Outcome                    0.13054795 0.29269466               0.17384407
                                 Age    Outcome
Pregnancies                0.54434123 0.22189815
Glucose                    0.26351432 0.46658140
BloodPressure              0.23952795 0.06506836
SkinThickness             -0.11397026 0.07475223
Insulin                   -0.04216295 0.13054795
BMI                        0.03624187 0.29269466
DiabetesPedigreeFunction   0.03356131 0.17384407
Age                        1.00000000 0.23835598
Outcome                    0.23835598 1.00000000
```

Figure 9. Correlation matrix dataset

Figure 8 above shows that of the 9 variables in the diabetes data set, only 4 variables are normally distributed: variable 2 (glucose), variable 3 (Blood Pressure), variable 6 (BMI), and variable 7 (DiabetesPredigreeFunction). Furthermore, Figure 9 above shows the distribution of correlation matrix values from the diabetes data set.

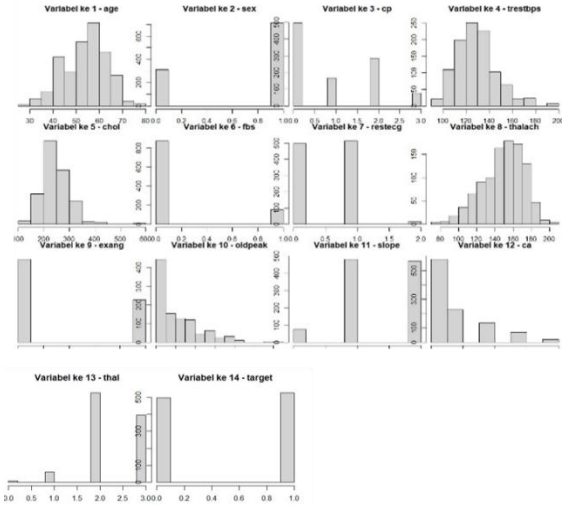


Figure 10. Distribution of heart disease datasets

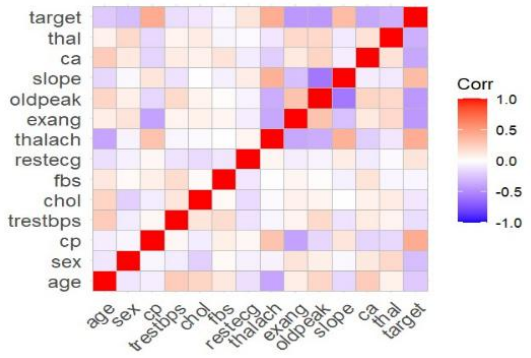
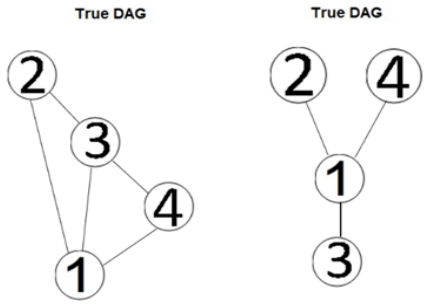


Figure 11. Correlation matrix of heart disease data set

Figure 10 above shows that of the 14 variables in the diabetes dataset, only four are normally distributed: variable 1 (age), variable 4 (trestbps), variable 5 (cholesterol), and variable 8 (talach). Furthermore, Figure 11 above shows the distribution of correlation matrix values in the diabetes dataset.

These variables will then serve as the primary reference for identifying causal relationships. Figure 12 (a) shows the results of the causal relationship for diabetes, and Figure 12 (b) shows the results of the causal relationship for heart disease.



(a) (b)

Figure 12. Visualization of causal models using the GES algorithm: (a) True DAG of diabetes, (b) True DAG of heart disease.

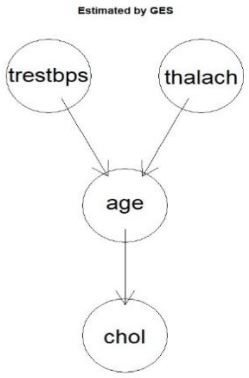


Figure 13. Visualization of causal models of heart disease using the GES algorithm

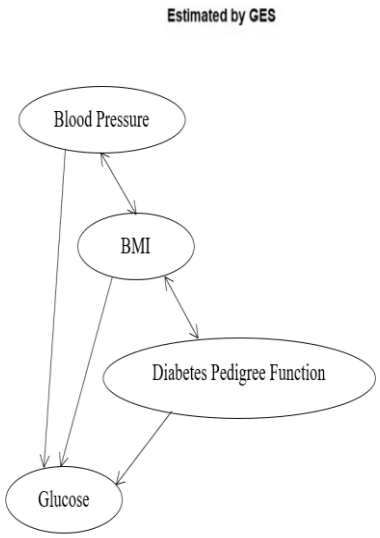


Figure 14. Visualization of a case model of diabetes using the GES algorithm

Visualization of the causal model using the GES method applied to two datasets, namely diabetes and heart disease, shows that four variables in heart disease produce three relationships and four variables in diabetes produce five relationships. Figure 12 is a causal model framework called "True DAG", which shows the existence of a causal relationship but does not yet display its direction. True DAG is a model generated by eliminating the direction of unnecessary nodes to obtain a good estimate of the model structure. Figure 12 part (a) shows node 1 is Glucose, Node 2 is Blood Pressure, node 3 is BMI, and node 4 is Diabetes Pedigree Function. Furthermore, Figure 12 part (b) shows node 1 is age, node 2 is trestbps, node 3 is chol, and node 4 is thalach.

Furthermore, because a true DAG is not yet able to show the direction of the causal relationship between variables, the GES algorithm is used to clarify the causal relationship that occurs in each variable. Figure 13 shows three directional relationships from the heart disease data set: trestbps and age, thalach and age, and age and chol. Then, Figure 14 for diabetes shows two bidirectional relationships: Blood Pressure and BMI, and then BMI and Diabetes Pedigree Function. Furthermore, Figure 14 also shows three directional relationships: Diabetes Pedigree Function and Glucose, BMI and Glucose, and Blood Pressure and Glucose.

3.2 Discussion

The results of the study above show that the variable of resting blood pressure is related to age. As age increases, blood pressure tends to increase, especially systolic pressure. This is in line with research conducted [17] and other studies, which found a very close relationship between resting blood pressure and age. Systolic blood pressure tends to increase, especially at the age of 40 years. Increasing age and resting blood pressure are closely related to

heart disease. Age is one of the non-modifiable risk factors in heart disease. As age increases, the elasticity of blood vessels decreases, hardening of the arteries increases, so that the function of the heart and blood vessels physiologically decreases [18]. Furthermore, high resting blood pressure is a major indicator of hypertension, which directly increases the risk of coronary heart disease, heart failure, and stroke [17]. So the risk of heart disease increases if a person gets older and has high blood pressure.

Another finding in this study was the relationship between age and thalach (maximum heart rate). Age is clearly a major factor in heart disease. A high resting heart rate is an independent marker of heart disease risk, especially if it occurs at a relatively young age [18].

Another relationship was also found in the heart disease data set, namely the relationship between age and cholesterol variables. Increasing age is associated with increased cholesterol levels. This is in line with research conducted [19] which states that high cholesterol levels can increase the risk of heart disease in middle age and old age, especially LDL levels. With increasing age, cholesterol levels, especially in women, increase. Another study stated that increased cholesterol levels due to increasing age can increase cardiovascular risk factors [20].

Furthermore, the causal relationship produced in this study is the relationship between the variables Blood Pressure and Glucose in diabetes. Diabetes is a decrease in insulin receptors that are able to absorb insulin. This results in insulin resistance and causes hyperglycemia. This hyperglycemia then increases high blood pressure. Too much glucose in the blood can cause red blood cells to stick to blood vessel walls. The more that stick, the more clots or plaques will form, which can cause thickening or narrowing of the blood vessels, which can lead to increased blood pressure [21].

Another relationship found between diabetes is blood pressure and BMI. High BMI is strongly associated with high blood pressure in type 2 diabetes patients. This is in line with research conducted [22] which states that the combination of high BMI and hypertension can significantly increase the incidence of diabetes and cardiometabolic complications. The higher the BMI, the greater the volume of blood needed to supply oxygen and nutrients to body tissues, thus increasing high blood pressure[23].

Another causal relationship found in this study is that BMI is related to glucose. A high BMI, or obesity, can worsen insulin resistance through chronic inflammation, which can lead to hyperglycemia. Research [24] suggests that BMI management is key to improving blood sugar control and preventing diabetes complications.

The next relationship is between the BMI variable and Diabetes Pedigree Function. People with a high BMI and a history of diabetes in their family environment have a higher risk of developing type 2 diabetes. Diabetes is a non-communicable disease, but it can be inherited genetically; family members with a history of diabetes will be at greater risk of developing diabetes than family members who do not have a history of diabetes [25].

The final finding in this study is the relationship between glucose variables and diabetes pedigree function. Genetic factors can affect beta cells and change their ability to recognize and distribute insulin secretory stimuli so that a person with a history of diabetes has a 15% chance of developing diabetes [25].

4. Conclusion

In this study, the authors applied the GES algorithm to the incidence of heart disease and diabetes. This study found a simple model representing the causal relationship between factors influencing the incidence of heart disease and diabetes, supported by previous research relevant to this study. The resulting model consists of three directional relationships from the heart disease dataset: blood pressure and age, thalach and age, and age and cholesterol. Furthermore, for diabetes, a two-way relationship model was generated: blood pressure and BMI, BMI and Diabetes Pedigree Function, and three directional relationships: Diabetes Pedigree Function and Glucose, BMI and Glucose, and Blood Pressure and Glucose. This model not only finds correlation values but can also establish causal relationships between factors that contribute to the incidence of diabetes and heart disease. Causal models enable us to reason at a higher level by understanding what happens under different circumstances.

This causal model is expected to serve as a scientific reference for healthcare professionals, such as doctors, nurses, nutritionists, researchers, and others involved in diabetes and heart disease. For further research, this model can be applied to real-world datasets, and other causal algorithms can be used to generate more causal models that can provide broader insights into addressing these issues.

References

[1] N. Afrianto *et al.*, "Applying PC Algorithm and GES to Three Clinical Data Sets : Heart Disease , Diabetes , and Hepatitis," in *IOP Conference Series: Materials Science and Engineering*, 2021, pp. 1–8.

[2] M. Bakator and D. Radosav, "Deep Learning and Medical Diagnosis : A Review of Literature," *Multimodal Technol. Interact.*, vol. 02, no. 03, pp. 1–12, 2018.

[3] M. Fatima and M. Pasha, "Survey of Machine Learning Algorithms for Disease Diagnostic," *J. Intell. Learn. Syst. Appl.*, vol. 9, no. 1, pp. 1–

16, 2017.
[4] B. Y. J. Pearl, “The Seven Tools of Causal Inference , with Reflections on Machine Learning,” *Commun. ACM*, vol. 62, no. 3, pp. 54–60, 2019.
[5] S. Gupta, W. Zhang, and F. Wang, “Model Accuracy and Runtime Tradeoff in Distributed Deep Learning: A Systematic Study IEEE 16 th,” in *Int. Conf. on Data Mining (ICDM)*, 2016, no. 171–180.
[6] R. Rahmadi, P. Groot, M. Heins, H. Knoop, T. Heskes, and T. Optimistic, “Causality on Cross-Sectional Data : Stable Specification Search in Constrained Structural Equation Modeling,” *Appl. Soft Comput.*, vol. 52, pp. 687–698, 2017.
[7] X. Chen, “Transmission Line Outage Probability Prediction Under Extreme Events Using Peter-Clark Bayesian Structural Learning.”
[8] N. Tou, C. Effendy, and R. Rahmadi, “Causal Relations of Factors Representing the Elderly Independence in Doing Activities of Daily Livings Using S3C-Latent Algorithm,” *Int. J. Artif. Intell. Res.*, vol. 5, no. 1, pp. 65–77, 2021.
[9] M. Pasquato and B. L. Davis, “Causa prima : cosmology meets causal discovery for the first time,” *Macnwsjfjf*, 2023.
[10] M. Z. Naser, “Discovering causal models for structural , construction and defense- related engineering phenomena,” *Def. Technol.*, vol. 43, pp. 60–79, 2025.
[11] Y. Setiya, R. Nur, A. Sa, D. Aldo, and B. Masulah, “Causal Modeling of Factors in Stunting using the Peter-Clark and Greedy Equivalence Search Algorithms,” *J. Ilmu Pengetah. dan Teknol. Komput.*, vol. 10, no. 3, pp. 523–533, 2025.
[12] N. Afrianto *et al.*, “Applying PC Algorithm and GES to Three Clinical Data Sets: Heart Disease, Diabetes, and Hepatitis,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1077, no. 1, p. 012067, 2021.
[13] G. Nursa, Y. Fauzi, and J. Habibi, “Faktor-Faktor Yang Mempengaruhi Kejadian Diabetes Melitus Di Puskesmas Bintuhan Kabupaten Kaur Tahun 2022,” *J. Hygeia Public Heal.*, vol. 1, no. 1, pp. 1–6, 2022.
[14] S. A. Muna Lubis, T. N. F. Aminah, S. Pangestuty, R. Atika, S. P. Sembiring, and Z. Aidha, “Faktor-Faktor yang Berhubungan dengan Kejadian Diabetes Melitus (DM) di Desa Kubah Sentang Kecamatan Pantai Labu,” *J. Ilm. Univ. Batanghari Jambi*, vol. 23, no. 2, p. 2061, 2023.
[15] R. D. Christy, H. Lukman, and H. Karnirius, “Faktor-Faktor Yang Mempengaruhi Kejadian Penyakit Jantung Koroner Di Rsud Rantau Prapat Tahun 2020,” *PREPOTIF J. Kesehat. Masy.*, vol. 5, no. 2, 2021.
[16] V. B. Valentine and I. A. Rakhmawati, “Analisis Faktor-Faktor Yang Mempengaruhi Kejadian Penyakit Jantung Koroner Pada Anggoita POLRI MAPOLDA Jawa Timur Tahun 2024,” *J. Manaj. dan Adm. Rumah Sakit Indones.*, vol. 8, no. 2, pp. 224–232, 2024.
[17] P. K. Whelton *et al.*, “Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults,” *J. Am. Coll. Cardiol.*, vol. 71, no. 19, pp. e127–e248, 2018.
[18] E. J. Benjamin *et al.*, *Heart Disease and Stroke Statistics-2019 Update*, vol. 139, no. 10. 2019.
[19] D. A. Saputri and A. Novitasari, “Hubungan Usia Dengan Kadar Kolesterol Masyarakat Di Kota Bandar Lampung,” *BIOEDUKASI (Jurnal Pendidik. Biol.*, vol. 12, no. 2, p. 238, 2021.
[20] M. Bertolotti, G. Lancellotti, and C. Mussi, “Changes in cholesterol homeostasis associated with aging and with age-related conditions : pathophysiological and clinical implications,” *Res. Gerontol. Geriatr.*, pp. 32–42, 2024.
[21] A. A. Lestari, E. Haryanto, Museyaroh, and A. Handayati, “No Title,” *Anal. Kesehat. Sains*, vol. 13, no. 1, pp. 20–24, 2024.
[22] T. P. Utami, “Hubungan antar Indeks Massa Tubuh dan Tekanan Darah pada Pasien Diabetes Melitus Tipe 2,” *J. Arch. Pharm.*, vol. 1, no. 1, pp. 19–22, 2019.
[23] R. Between, B. Mass, and I. N. T. H. E. Elderly, “Relationship Between Body Massa Index and Blood Pressure in the Elderly,” *J. Kesehat. Manarang*, vol. 10, no. 3, pp. 329–335, 2024.
[24] S. F. Jannah, Wantonoro, and Widiastuti, “Korelasi indeks massa tubuh dengan gula darah pada pasien diabetes melitus,” in *Prosiding Seminar Nasional Penelitian dan Pengabdian kepada Masyarakat LPPM Universitas Aisyiyah Yogyakarata*, 2025, vol. 3, pp. 807–814.
[25] C. K. Lariwu, C. P. Sarayar, L. Pondaag, G. Merentek, and M. E. Lontaan, “Indeks Masa Tubuh , Riwayat Keluarga dan Kebiasaan Konsumsi Gula : Faktor Dominan Penyebab Diabetes Melitus Tipe 2 Pada Lanjut Usia Di Kota Tomohon Abstrak,” *J. Ilmu Pendidik. Nonform.*, vol. 10, no. January, pp. 379–386, 2024.