JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)
VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6184

# CAUSAL MODELING OF FACTORS IN STUNTING USING THE PETER-CLARK AND GREEDY EQUIVALENCE SEARCH ALGORITHMS

**Yohani Setiya Rafika Nur [1*]; Aminatus Sa'adah[2]; Dasril Aldo[3]; Bidayatul Masulah[4]**

Informatics Engineering[1, 2, 3]
Telkom University, Purwokerto, Indonesia[1, 2, 3]
info@telkomuniversity.ac.id[1]
yohanin@telkomuniversity.ac.id[1*], aminatuss@telkomuniversity.ac.id[2],
dasrilaldo@telkomuniversity.ac.id[3]

Mathematical Institute[4]
Universiteit Leiden, Netherlands[4]
studielijn@leidenuniv.nl[4]
bidayatul.masulah@umail.leidenuniv.nl[4]

(*) Corresponding Author
(Responsible for the Quality of Paper Content)



***Abstract***— *Stunting is one of the nutritional problems that can hinder the growth and development process in toddlers. Untreated stunting can lead to fatal outcomes. Previous research on the factors that exist in the incidence of stunting mostly used multivariate analysis. Previous research on stunting factors has primarily used multivariate or correlation analyses. However, this study uniquely focuses on establishing causal relationships between these factors, a crucial step in improving early diagnosis for stunting prevention and treatment. The data used in this research was 83 data on stunting incidents and consisted of eight parameters. The purpose of this study is to model the causal relationship between factors that represent the incidence of stunting. This study uses two simple causal approaches, namely the Peter-Clark (PC) algorithm to obtain the initial concept of a graph model of the relationship between variables and the Greedy Equivalence Search (GES) algorithm to refine the model by obtaining the direction of the causal relationship. There are six bi-directed relationships that have been found, namely from food variables to support; maternal knowledge with sanitation; Height/Age and Weight/Age with Child Nutrition; height/age with weight/age and stunting. In addition, both algorithms in this study have successfully obtained a causal model, by comparing performance using directional and causal densities that the GES algorithm was able to identify a relationship of 0.66 compared to the PC algorithm.*

***Keywords***: *causality, greedy equivalence search, peter-clark, stunting.*

**Intisari**— *Stunting merupakan salah satu masalah gizi yang dapat menghambat proses tumbuh kembang balita. Stunting yang tidak ditangani dapat berakibat fatal. Penelitian terdahulu mengenai faktor-faktor yang ada pada kejadian stunting sebagian besar menggunakan analisis multivariat. Penelitian sebelumnya mengenai faktor-faktor stunting lebih banyak menggunakan analisis multivariat atau korelasi. Namun, penelitian ini secara unik berfokus pada pembentukan hubungan sebab akibat antara faktor-faktor tersebut, yang merupakan langkah penting dalam meningkatkan diagnosis dini untuk pencegahan dan penanganan stunting. Data yang digunakan dalam penelitian ini sebanyak 83 data kejadian stunting dan terdiri dari delapan parameter. Tujuan dari penelitian ini adalah untuk memodelkan hubungan sebab akibat antara faktor-faktor yang mewakili kejadian stunting. Penelitian ini menggunakan dua pendekatan kausal sederhana, yaitu algoritma Peter-Clark (PC) untuk mendapatkan konsep awal model grafik hubungan antar variabel dan algoritma Greedy Equivalence Search (GES) untuk menyempurnakan model dengan mendapatkan arah hubungan kausal. Hasil model pada penelitian ini mendapatkan enam arah hubungan* *secara dua arah yang telah ditemukan, yaitu dari variabel pangan dengan dukungan; pengetahuan ibu dengan sanitasi; Tinggi Badan/Umur dan Berat Badan/Umur dengan Gizi Anak; tinggi badan/umur dengan berat badan/umur dan stunting. Kedua algoritma dalam penelitian ini telah berhasil mendapatkan model kausal, dengan melakukan perbandingan kinerja menggunakan densitas terarah dan kausal bahwa algoritma GES mampu mengidentifikasi hubungan sebesar 0.66 dibandingkan dengan algoritma PC.*

***Kata Kunci****: kausal, greedy equivalence search, peter-clark, stunting.*

## INTRODUCTION

Stunting is a condition of failure to grow and develop in children. Stunting in babies and toddlers is brought on by either a persistent lack of nutrients or food consumption that falls short of recommended daily allowances [1]. In toddlers or children under five, stunting can lead to decreased growth and development, poor learning outcomes, and an increased risk of infection [2]. The consequences that stunting sufferers have are direct and can even have a long-term impact, and if not treated, will result in death. Zero Hunger has a goal in sustainable development with key indicators and is included in the 17 main targets of Global Nutrition in reducing stunting rates, especially for 2030 [2]. The incidence of stunting in Indonesia has a high prevalence rate. The prevalence of stunting incidence in 2018, was 31.91%, with a population vulnerable to stunting of 17,591,715 people [3]. The cases are distributed in Sumatra, Kalimantan, Sulawesi, Java, and Papua. This incidence over time decreased in 2020 compared to 2019 by 22% [4]. However, the decline did not provide a significant change. The occurrence of stunting in toddlers is often not realized, and after two years, it is seen that the toddler is short. Chronic nutritional problems in toddlers are caused by insufficient nutritional intake for a long time due to parents/families not knowing or not being aware of providing food that is in accordance with their child's nutritional needs [5], [6].

Applications in artificial intelligence technologies such as machine learning [7], [8]; Expert System [9], and so on, which is a very useful innovation in helping in the health sector. Stunting data can be processed using a variety of computational techniques, but there are still some drawbacks. For example, typical machine learning algorithms may need lengthy training times and comprehensive computing, which makes them ineffective when utilized in resource-constrained environments. However, leveraging ML and DL models cannot achieve this goal directly. This approach largely or fundamentally does not look deeper into the causal relationship between the input and output variables $P(Y|X)$ [10], [11]. In contrast, to address the above generative models, a technique known as causal modeling or causal discovery can be used. Basically, the causal model studies how the variable Y (causation) can only change its value depending on the variable X (causation) [12]. A primary objective across many sciences, particularly those in the therapeutic domain, is to find generative models. Such a model makes clear how the values of the variables in the data are obtained [10].

A novel approach to cause and effect research of a variable is called causation [13]. The aim of causal discovery is to find the causal structure, or the method under which one predictor variable causes another predictor variable and/or the desired result [13]. Causal relationships are generally divided into four types of edges, namely 1.) directed edges to represent the causal relationship of the parent/causal variable to the inheritance of the effect, 2.) undirected edges where between each connected variable cannot be defined the causal relationship, 3.) bidirectional edges represent whether the variables can be the cause or effect of each other, and finally 4.) edges that represent whether there are latent variables. In the case of cancer, we are embarking on a study that delves into the (probable) cause-and-effect relationship between quality of life in the sufferer [11]. For example, in the case of cancer patients with low quality of life status, they are very sensitive to their emotions and the emotions that arise cause poor sleep quality (insomnia). Thus, unstable emotional impacts have an impact on other cognitive functions. Thus, in order to address risks and develop solutions, it is essential to understand these cause-and-effect relationships.

Previous research has been conducted on stunting, such as determining the correlation between symptom factors and the incidence of stunting [7], identifying causal to the incidence factor of stunting disease [14], prediction on stunting prevalence with linear model [15], understanding determinants of childhood stunting reduction applied using the standardized mixed-methods framework [16], and diagnostic predictive model for determining child stunting [17]. Some studies perform identified causality, such as quality of life factors in cancer [11], in three clinical datasets [18]. However, the study only looks at the



Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6184

results of the correlation relationship and needs to understand and look more deeply at the relationship between cause and effect. The causal method is technically divided into two: the constraint-based approach and the score-based approach [11]. There are several approaches or algorithms in identifying causality, namely, Peter-Clark (PC), Greedy Equivalence Search (GES), Fast Causal Inference (FCI), Stable Specification Search (S3C)- Latent [11].

Based on the explanation above, there have been many studies related to the incidence of stunting, both with correlation, deep learning, and machine learning approaches. However, the statement by Spirtes et al. cited in studies [11] that do not answer fundamental questions about causal relationships. Therefore, understanding a causal relationship will help in understanding the problem and thus designing intervention models to help treatments and therapies that can reduce the incidence of stunting. Another research conducted in [18] has implemented simple causal algorithms namely Peter-Clark (PC) and Greedy Equivalence Search (GES), but used 3 clinical datasets namely "heart disease, diabetes, and hepatitis". However, the study only implemented one algorithm for one clinical dataset case. This research aims to designing interventions such as treatment and therapy that can reduce the incidence of stunting. Other research that has been conducted in [18] has applied simple causal algorithms, namely Peter-Clark (PC) and Greedy Equivalence Search (GES), but using 3 clinical datasets such as "heart disease, diabetes and hepatitis". However, the study only implemented one algorithm for one clinical dataset case. So, in this study, the two algorithms will be re-implemented with the clinical dataset "incidence in stunting" to obtain a model of causal relationships. Based on the explanation above, this study will identify the causal relationship in the stunting incidence dataset using two approaches, namely Peter-Clark (PC) and Greedy Equivalence Search (GES), and also to determine the performance of each method in obtaining a causal relationship model. Hopefully, this research can provide an overview and recommendations to broaden our understanding of this clinical problem.

## MATERIALS AND METHODS

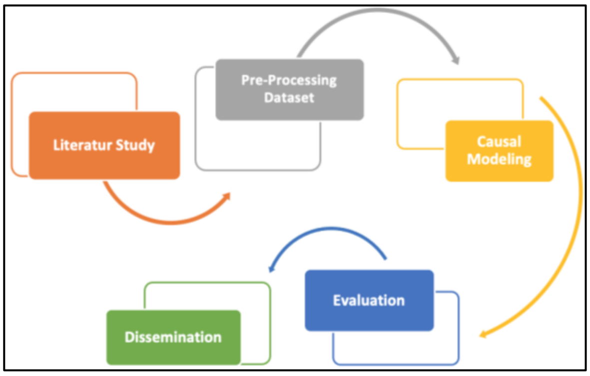
The data set used in this study is sourced from Kaggle [19] contains 9 parameters and 83 respondents. Figure 1 shows the dataset details. These parameters are related to the recording of respondent data, age, weight (BB), height (TB), height/age (TB/U), weight/age (BB/U), child nutrition, support, maternal knowledge, sanitation, culture, and stunting.

```
> str(hapus)
'data.frame':   83 obs. of  9 variables:
 $ TB.U           : num  -0.5 -2.567 -2.512 0.591 -1.027 ...
 $ BB.U           : num  -0.857 -0.5 -0.789 0.778 -1 ...
 $ GIZI.ANAK      : int  0 0 1 0 0 0 0 0 -2 -2 ...
 $ PENGETAHUAN.IBU: int  6 5 7 6 5 5 5 5 6 5 ...
 $ SANITASI       : int  12 11 11 9 12 10 13 11 10 10 ...
 $ PANGAN         : int  37 44 30 30 45 45 34 45 45 45 ...
 $ BUDAYA         : int  10 0 8 8 3 2 7 3 8 7 ...
 $ DUKUNGAN       : int  2 1 2 1 2 1 2 1 1 1 ...
 $ STUNTING       : int  0 1 1 0 0 0 1 0 0 0 ...
```

Source : (Research Results, 2024)

Figure 1. Dataset Details

To see the relationship between each of these parameters, this study will use the Peter-Clark Algorithm and the GES algorithm to produce a model with visualization using a causal graph simulated with related datasets. The present study is conducted with following stages, i.e., literature study, data pre-processing, causal modeling, evaluation, and dissemination. For brief methodology, see Figure 2.




Source : (Research Results, 2024)

Figure 2. Research Stages

The research phase begins by conducting a literature study to review relevant previous research findings [7], [8], [9], [20] with the incidence of stunting in children. The second stage is data pre-processing; this stage aims to check the details of the data set if there are any missing values, as well as check the distribution of the data. At that stage, the missing value problem can be solved by cleaning up using code NewData ← Data[complete.cases(Data),]. The third step is to apply the Peter-Clark (PC) and Greedy Equivalence Search (GES) algorithms to the prepared dataset. Model computation is done using an R package named "pcalg" and "GaussL0penObsScore" [18]. The output of this stage is a causal model between stunting incidence factors.

In the science of examining cause-and-effect correlations between variables, causation is a novel model [18]. Finding the causal structure, or the pattern of impact between one predictor variable and another predictor variable or the intended response, is the aim of causal discovery [21]. A number of methods or algorithms are employed to determine causality, such as S3C (Stable Specification Search) Latent [11], PC (Peter-Clark), GES (greedy equivalency search, FCI (rapid causal inference)

One of the widely used techniques in causal structure identification for determining cause-and-effect linkages between variables is the PC (Peter-Clark) algorithm. It generates a graph or network structure that illustrates the causal linkages between variables using conditional independence (CI) tests to find causal relationships. in the data and generating a graph or network structure that shows the causal relationship between variables. There are two primary steps in the PC algorithm method [22]. The procedure begins by assuming that all variables are related to one another, which is followed by an edge removal step. The algorithm then eliminates edges that do not represent causal linkages and methodically runs conditional independence checks on every pair of variables. The algorithm then moves on to the edge orientation stage, which uses causality rules (such as the d-separation rule) to determine the direction of the relationship until only pertinent edges are left. A directed graph illustrating the pattern of causal interaction between the variables is the final product.

**Algorithm 1**
Peter-Clark Algorithm

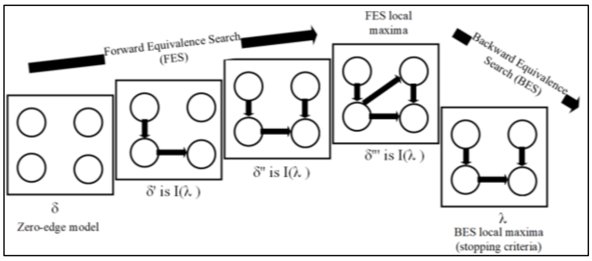
```
Input Dataset D with a set of variables V and significant level α
Output The undirected graph G with a set of edges E
Assume all nodes are connected initially
Let depth d = 0

  repeat
    for each ordered pair of adjacent vertices X and Y in G do
      if |adj(X,G) \ {Y}| ≥ d then
        for each subset Z ⊂ adj(X,G)\{Y} and |Z| = d do
          Test I(X, Y | Z)
          If I(X, Y|Z) then
              Remove edge between X and Y
              Save Z as the separating set of (X,Y)
              Update G and E
              break
          end if
        end for
      end if
    end for
    Let d = d+1
  until |adj(X, G)\{Y}| < for every pair of adjacent vertices in G
```

Source: (Afrianto, 2021) [18]
Figure 3. Peter-Clark Algorithm

According to the aforementioned technique on Figure 3 above, the edge between two variables, X and Y, is eliminated if they are shown to be independent of one another without taking other variables into account. The variables that set them apart are then saved as a distinct set at each iteration. This is accomplished by increasing the depth d and looking at a bigger selection of variables. A graph showing the causal structure between the variables is produced once the procedure is complete, at which point no more pairs of variables may be separated [23]. A causal model with directional and non-directional edges in each graph is concluded by the PC algorithm's ultimate output, which is partially represented by a completed directed acyclic graph (CPDAG) with complete side criteria [12]. These rules mainly rely on the notion of axilality and the fact that selection variables produce unique independent structures that allow for the recovery of some v-structures.

This is different from the greedy equivalence search (GES) algorithm. Chickering has developed a method for studying a graph structure known as greedy equivalence search (GES). The learning structure is by studying a relationship between variables by looking at cause and effect. The GES algorithm is a score-based causal method that compares models based on the size of the match of a model produced [24], [25]. The GES (Greedy Equivalence Search) algorithm for causality is used to discover the causal structure between variables in the data by mapping the causal relationships in the form of a directed graph. At each iteration, the algorithm selects changes that improve the quality of the causality model based on certain statistical criteria, such as Bayesian scores. The concept of the GES algorithm can be seen in Figure 4.




Source: (Chakraborti, 2023) [25]
Figure 4. The Concept of the Greedy Equivalence Search Algorithm

Based on the above concept, the GES algorithm uses Bayesian Information Criterion (BIC) scores calculated from Gaussian or multinomial data for data with larger match sizes. GES has two phases, namely forward equivalence search (FES; can be seen in Algorithm 2) and backward equivalence search (BES; can be seen in Algorithm 3). This algorithm originally started as a zero-side graph model. Furthermore, any possible

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6184

single side of the progressive way will be added to the DAG until the algorithm reaches the local maximum. The addition phase is called forward equivalence search (FES) or maximum local FES. After the FES phase has reached the local maximum or has not experienced an increase in score, the FES technique is improving using the BIC. This BIC mechanism will repeatedly maximize the increase in scores and consider the elimination or addition of the addition of the causal relationship side in the temporary DAG model . The GES algorithm works in two main stages. The first stage is Search. In this stage, the algorithm starts with an empty or very simple graph. Then, gradually, the algorithm adds operation (forward phase) or delete operation (backward phase) the relationships between variables to build a more complex graph with the aim to find the causal structure that best fits the data (Can see Figure 5 and Figure 6). The second stage is the Greedy Step. In this stage, the algorithm selectively chooses increasing changes to build a more complex graph (can see Figure 7).

**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
```

Source: (Nur, 2024) [14]
Figure 5. Forward Phase

In the algorithm pseudocode above, the CPDAG (Completed Partially Directed Acyclic Graph) structure will be improved by adding one edge. First starting with a graph G, the algorithm searches for a new graph set M by adding one edge to G. From the set M, the graph K with the highest score is selected, called S(D,K), which measures the graph's match with the dataset D. If graph K has a higher score than G, then H is updated to K, and if not, then H remains G. The algorithm's goal is to improve the graph by adding edges that improve the match with the data. The algorithm of the backward phase can be seen in Figure 6.

**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
```

Source: (Nur, 2024) [14]
Figure 6. Forward Phase

This pseudocode shows how to remove one edge from a CPDAG (Completed Partially Directed Acyclic Graph) to enhance its structure. After eliminating one edge from graph G, the method creates a collection of graphs M. The graph K with the greatest score S (D, K), which measures how effectively the model fits dataset D, is then selected. If the edge removal produces a higher score than G, the graph H is modified to K. The graph H stays G else. By eliminating edges that do not make a substantial contribution to the model, this approach seeks to enhance the causal structure. When the forward phase is finished and there are no more edges that might raise the scoring criteria, the algorithm can at anytime go to the backward phase. Technically, the overall algorithm structure of GES can be seen in Figure 7.

**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
```

Source: (Nur, 2024) [14]
Figure 7. Forward Phase

## RESULTS AND DISCUSSION

### Results

The results of this study are described to obtain a causal model from the application of the Peter-Clark (PC) and Greedy Equivalence Search (GES) algorithms and to see the direction of the relationship between the variables in the data used. The demographic characteristics of the research subjects were described based on gender, age (months), Height (TB), Body Weight (BB),

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v10i3.6184

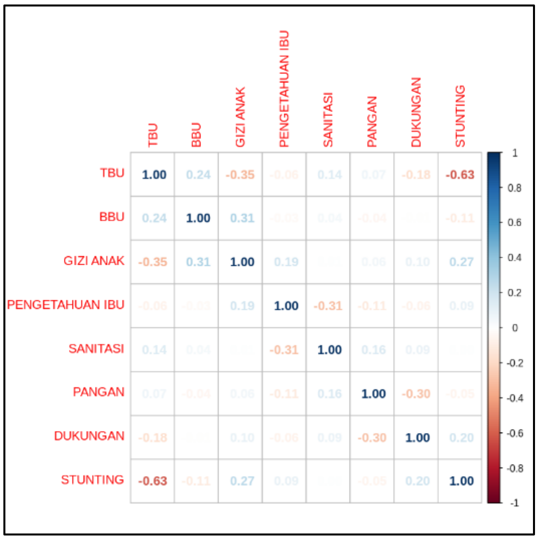

Height/Age (TB/U), Weight/Age (BB/U), Weight/Height (BB/TB), Child Nutrition, Maternal Knowledge, Food, Culture, Support, Stunting.

Table 1. Demographics Of Respondents (N = 83)

| Parameter | Indicator | Frequency (N) | Percent (%) |
|---|---|---|---|
| Genders | Male | 49 | 59 |
| | Female | 34 | 40.9 |
| Ages (month) | 0-30 | 41 | 49.4 |
| | 31-59 | 42 | 50.6 |
| Hight for Age (Z score) | < -2 to ≥ -3 | 8 | 9.6 |
| | ≥ - 2 | 65 | 78.3 |
| | < -3 | 8 | 9.6 |
| Weight for Age (Z score) | < -2 to ≥ -3 | 73 | 87.9 |
| | ≥ - 2 | 8 | 9.6 |
| | < -3 | 2 | 2.4 |
| Stunting | Yes | 18 | 21.7 |
| | No | 65 | 78.3 |

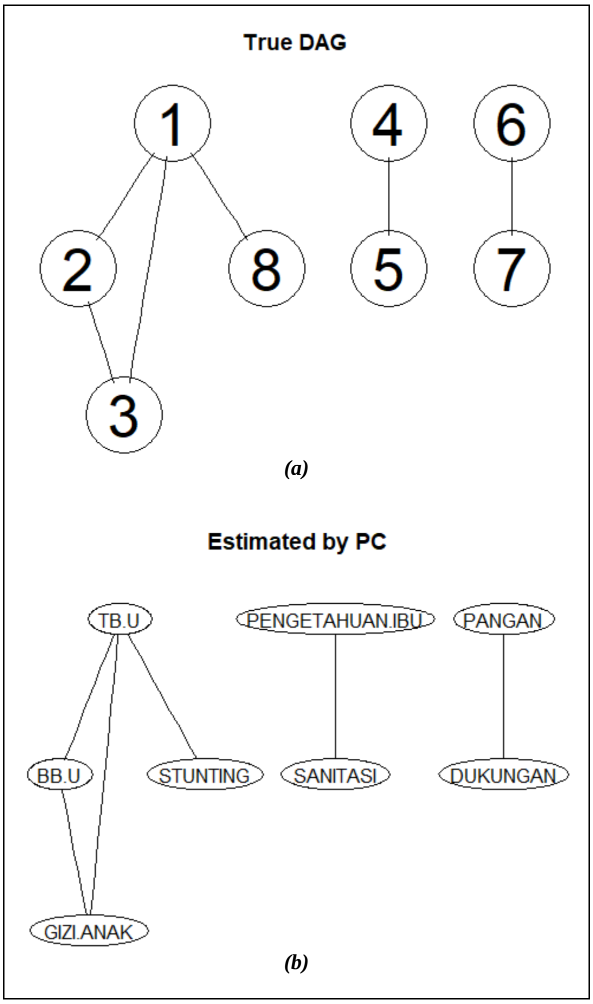
Source : (Research Results, 2024)

Table 1 shows that most of the respondents in this study are men (59%), most are 31-59 years old (42%), the Z-score of TB/Age is mostly ≥ - 2 (78.3%) while BB/Age < -2 to ≥ -3 (87.9%), and the incidence of stunting is mostly in the status of No (78.3%). Technically, the research process conducts pre-processing by checking the missing value first, before then the calculation is carried out. Checking for missing datasets using the NewData←Data[complete.cases(keeps),] code, setting the variables to be computed and Gaussian conditional independence test for the PC algorithm using 0.05. It then performs identification based on the resulting correlation of each variable (can be seen in Figure 8).




Source: (Research Results, 2024)
Figure 8. Variable's Identification Based on Correlation

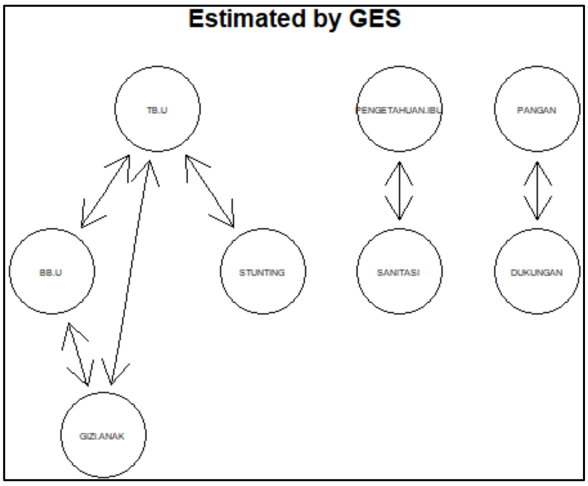
Based on Figure 8, there are results with positive correlation variables >0.30 and negative correlation variables <-0.30, i.e., TB.U and Child Nutrition, BB.U, Stunting; Maternal Knowledge and Sanitation, Food Needs and Support. In this case these parameters will be the main reference for producing a causal model. After identification and performing the above process, the results of the causal model with the PC algorithm can be seen in Figure 9: (a) True DAG, which is the initial result of the causal model framework; and (b) PC Model Estimation which is the final model with CPDAG graph visualization. In addition, the model results of the GES algorithm can be seen in Figure 10.




Source: (Research Results, 2024)
Figure 9. Visualization of causal models of stunting incidence based on algorithm PC: (a) True DAG; (b) Estimated by PC


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6184




Source: (Research Results, 2024)

Figure 10. Visualization of causal model of stunting incidence based on GES algorithm

The causal model visualization of the two algorithms, namely PC and GES above, shows that the eight variables produce six relationships that are interconnected with each other. Technically, the initial results of the PC algorithm will produce a causal model framework called the "True DAG", which represents the results of the identification of causal relationships between variables but not their direction. The results of the true DAG produced do not differ from the final estimation results with the visualization of the CPDAG graph, which is referred to as "Estimation by PC". True DAG is the result of a model that has repeatedly omitted directions on nodes that are not needed to get the estimation model framework. Node 1 is the Height/Age variable (TB. U); node 2 is the Weight/Age variable (BB. U); node 3 is a variable of Child Nutrition; node 4 is Mother Knowledge; node 5 is Sanitation; node 6 is Food; node 7 is Support; node 8 is the status of stunting events.

Furthermore, because the results of the PC algorithm only produce relationships between variables in other words, it does not show the direction of the causal relationship. So, the GES algorithm produces a causal model that has a relationship and shows the direction of cause and effect. Based on Figure 9 above, there are six bi-directional relationships, namely, food variables and support; maternal knowledge with sanitation; TB. U, BB.U with child nutrition; TB. U with BB. U; and TB. U with stunting. The assessment metrics of the two algorithms are carried out using Directed Density and Causal Density identification, which are based on the model estimate outcomes of the PC and GES algorithms. Whereas causal density (CD) is a directed edge, directed density (DD) represents directed edges with the total number of edges in the model. For example for DD metric in DAG with three known links, of which two are directed and one is not. becomes 2/3 = 0.67 [26].

Table 2. Evaluation of Directed Density (DD) and Causal Density (CD) Matrics (%)

| Matrics | PC Algorithm | GES ALgorithm |
|---|---|---|
| DD | 0/6 (0) | 0/6 (0) |
| CD | 1/6 (0.16) | 4/6 (0.66) |

Source: (Research Results, 2024)

Based on the evaluation results, the performance of the two algorithms based on DD and CD can be concluded that the GES algorithm is able to identify causal relationships using the CD matric.

**Discussion**

The results of the above study show that maternal knowledge causes sanitation. This finding is in line with the study conducted in [27] and other studies, which found a very close relationship between the variables of maternal knowledge and sanitary behaviour [28]. Mother's knowledge and sanitary behaviour are very closely related to stunting prevention [27]. Poor access to water, sanitation, and hygiene is correlated with the substantial global burden of disease and disability caused by subsequent famine. Improving access to water, sanitation, and hygiene, for instance, may reduce infectious disorders like diarrheal illnesses, which are associated with an increased risk of stunting [28], [29]. A mother needs to pay attention to the cleanliness of both herself and the environment where she lives, such as the management of drinking water, clean food, and household waste. The wider the mother's knowledge, the better she can show a good attitude and do something positive.

Another relationship obtained from this study, namely food and family support, is also very necessary to prevent stunting. These results are in line with research in [30]. Family activities are crucial to ensuring that children under five have a healthy diet. Additionally, the satisfaction of all demands, including the provision of sufficient nourishment, determines the quality of family life. Quality of life and physical health are improving. According to reports, other family members—particularly grandmothers who live with mothers—have a significant influence on decisions about infant feeding [31]. This relationship is also strengthened by the results of the study in [32], which states that the support of the family, as well as knowledge in fulfilling nutritional needs, greatly determines the health status of children. The provision and fulfilment of the nutritional quality of babies and mothers is highly dependent on the availability of good and nutritious foodstuffs, so family support is also needed to meet these needs.

In addition, the relationship between TB. U has a relationship with the nutritional status of children. This relationship is supported by the results of the study in [32] also show that the support of the family, as well as knowledge in fulfilling nutritional needs, greatly determines the health status of children. People with conditions living in coastal areas can affect the nutritional status of children because these communities have different socio-cultural conditions from other regions, especially in parenting, patterns and food needs for children, as well as locations with very minimal access to cleanliness, sanitation, water, and food [32].

The next relationship is the BB.U with child nutrition has a relationship with being strengthened by the results of the study in [33] that children with a history of underweight will experience a 10 times greater risk of poor nutritional status when compared to children with normal weight. Children with a history of BBLR have poor digestion, which can lead to poor absorption of nutrients [34].Weight loss slows or stops linear development until weight is recovered and illness is addressed. Stated differently, during stress, nutritional weight (lean and fat mass) can be maintained or restored at the price of linear growth (height/length) [35].

The correlation between TB.U (height/length) and BB.U (body weight) is as follows: children with a low weight status and a low height increase at the beginning of the period, while children with a weight gain and a lower height increase more than children with a normal weight for their height [36]. Babies with low birth weight experience delayed intrauterine growth since they are in the womb. This will continue until the baby is born when they do not receive support with proper parenting and nutritional intake. Ultimately, this baby will fail to reach the growth rate that should be achieved at his age [37].

Furthermore, the relationship between TB. U with stunting. Research [38] explained that the birth length of children with low status has a 6.29 times greater risk of stunting than children with normal birth length. The incidence of low-height of children is also influenced by other factors, namely the child's weight status. Children with low birth weight are 6.16 times more likely to be stunted than children enrolled in the BBLR category, or 60.9% more likely to be stunted [39]. Mothers who have children with stunting conditions have a habit of delaying and providing food intake without paying attention to their nutritional needs. This condition causes the food intake of toddlers to be less good in terms of quality and quantity, so toddlers are prone to stunting [40].

The PC and GES algorithms in this study have succeeded in obtaining causal models in identifying stunting events, but the GES algorithm is superior in identifying directional relationships. Based on the evaluation of the model performance of the two algorithms, it is also in line with the research Naser [26] that the GES algorithm has the best performance evaluation results in identifying directional edges, i.e. capturing conditional independence in existing data, regardless of its practical and actual nature compared to the PC algorithm.

## CONCLUSION

In this study, we applied simple causal methods, namely the PC algorithm and GES algorithm, to stunting incidence data. The findings of this study are a simple model to represent a causal relationship supported by the results of relevant previous studies, as well as the opinions of experts in the field. The model shows that there are six relationships in bi-direction, namely food variables ↔ backing; Mother's Knowledge ↔ Sanitation; Height/Age and Weight/Age ↔ Child nutrition; Height/Age ↔ weight/age and stunting. Based on the results obtained from the many variables, it does not only produce and see correlation or classification values, but can answer a fundamental question about the cause-and-effect relationship between variables, especially variables that trigger stunting. Causal models enable us to adopt a higher degree of reasoning (understanding what might happen under different conditions) thanks to causal models [26].

The model obtained is expected to be a scientific reference for those who are struggling, especially in this field, such as doctors, nurses, midwives, researchers, nutritionists, and others who focus on handling stunting events. This study also aims to add a reference to an alternative model in the clinical domain, namely the causal model, which can be useful to understand the problem well and then propose a solution for its solution. For future research, it can be implemented with real datasets with other demographic information, as well as use other algorithms to get more causal relationships between variables so as to gain broader insights into this problem.


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6184

## REFERENCE

[1] S. H. Quamme and P. O. Iversen, "Prevalence of child stunting in Sub-Saharan Africa and its risk factors," *Clinical Nutrition Open Science*, vol. 42, pp. 49–61, Apr. 2022, doi: 10.1016/j.nutos.2022.01.009.

[2] N. Fauziah, M. A. Ar-Rizqi, S. Hana, N. M. Patahuddin, and A. Diptyanusa, "Stunting as a Risk Factor of Soil-Transmitted Helminthiasis in Children: A Literature Review," *Interdiscip Perspect Infect Dis*, vol. 2022, pp. 1–14, 2022, doi: 10.1155/2022/8929025.

[3] K. D. Negeri, "Dashboard Peta Sebaran Kerentanan Stunting." [Online]. Available: https://gis.dukcapil.kemendagri.go.id/arcgis/apps/webappviewer/index.html?id=b832cb42d23b4fe197e278ceca893ad7

[4] N. A. Latifah, F. Fajrini, N. Romdhona, D. Herdiansyah, Ernyasih, and Suherman, "Systematic Literature Review: Stunting pada Balita di Indonesia," *Jurnal Kedokteran dan Kesehatan*, vol. 20, no. 1, pp. 55–73, 2024, [Online]. Available: https://jurnal.umj.ac.id/index.php/JKK

[5] D. Handayani, E. Kusuma, H. Puspitasari, and A. D. Nastiti, "The Factors Affecting Stunting on Toddlers in Coastal Areas," *Jurnal Aisyah : Jurnal Ilmu Kesehatan*, vol. 7, no. 3, pp. 755–764, Aug. 2022, doi: 10.30604/jika.v7i3.967.

[6] N. K. Dewi, H. A. R. Kusumasari, S. Andarini, and I. W. A. Indrawan, "Nutritional Factors Affecting Stunting Among Toddlers," *Amerta Nutrition*, vol. 7, no. 1SP, pp. 25–29, May 2023, doi: 10.20473/amnt.v7i1SP.2023.25-29.

[7] O. N. Chilyabanyama *et al.*, "Performance of Machine Learning Classifiers in Classifying Stunting among Under-Five Children in Zambia," *MDPI Journal*, vol. 9, no. 7, Jul. 2022, doi: 10.3390/children9071082.

[8] S. M. J. Rahman *et al.*, "Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach," *PLoS One*, vol. 16, no. 6 June 2021, Jun. 2021, doi: 10.1371/journal.pone.0253172.

[9] B. Riswanto, W. Setiawan, and S. C. E. Sahputro, "Sistem Pakar Diagnosa Stunting pada Balita Berbasis Website Menggunakan Metode Forward Chaining dan Metode Waterfall," *Digital Transformation Technology*, vol. 3, no. 2, pp. 468–477, Sep. 2023, doi: 10.47709/digitech.v3i2.2881.

[10] P. Spirtes, "Introduction to Causal Inference," *Journal of Machine Learning Research*, vol. 11, pp. 1643–1662, 2010.

[11] Y. S. R. Nur, R. Rahmadi, and C. Effendy, "Causal Modeling Between Factors on Quality of Life in Cancer Patients Using S3C-Latent Algorithm," *Journal of Information Systems Engineering and Business Intelligence*, vol. 7, no. 1, p. 74, Apr. 2021, doi: 10.20473/jisebi.7.1.74-83.

[12] J. J. Kossakowski, L. J. Waldorp, and H. L. J. van der Maas, "The Search for Causality: A Comparison of Different Techniques for Causal Inference Graphs," *Psychol Methods*, vol. 26, no. 6, pp. 719–742, 2021, doi: 10.1037/met0000390.

[13] M. Z. Naser, "Discovering causal models for structural, construction and defense-related engineering phenomena," *Defence Technology*, Apr. 2024, doi: 10.1016/j.dt.2024.04.007.

[14] Y. S. R. Nur, D. Aldo, A. Sa'adah, and Faizah, "Implementation of PC Algorithm to the Incidence Factor of Stunting Disease," in *Proceeding - 2024 International Conference on Information Technology Research and Innovation, ICITRI 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 93–98. doi: 10.1109/ICITRI62858.2024.10699213.

[15] Mambang, F. D. Marleny, and M. Zulfadhilah, "Prediction of linear model on stunting prevalence with machine learning approach," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 1, pp. 483–492, Feb. 2023, doi: 10.11591/eei.v12i1.4028.

[16] N. Akseer, T. Vaivada, O. Rothschild, K. Ho, and Z. A. Bhutta, "Understanding multifactorial drivers of child stunting reduction in exemplar countries: A mixed-methods approach," *American Journal of Clinical Nutrition*, vol. 112, pp. 792S-805S, Jul. 2020, doi: 10.1093/ajcn/nqaa152.

[17] J. Mkungudza, H. S. Twabi, and S. O. M. Manda, "Development of a diagnostic predictive model for determining child stunting in Malawi: a comparative analysis of variable selection approaches," *BMC Med Res Methodol*, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12874-024-02283-6.

[18] 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, Feb. 2021, doi: 10.1088/1757-899x/1077/1/012067.

[19] "IMERI Stunting." Accessed: May 20, 2024. [Online]. Available: https://www.kaggle.com/datasets/dwitarahmadini/imeri-stunting?select=YES.csv
[20] H. Shen, H. Zhao, and Y. Jiang, "Machine Learning Algorithms for Predicting Stunting among Under-Five Children in Papua New Guinea," *Children*, vol. 10, no. 10, Oct. 2023, doi: 10.3390/children10101638.
[21] X. Chen, Q. Huang, and Y. Zhou, "Transmission Line Outage Probability Prediction Under Extreme Events Using Peter-Clark Bayesian Structural Learning," *arXiv preprint*, Nov. 2024, [Online]. Available: http://arxiv.org/abs/2411.11980
[22] A. H. Petersen, M. Osler, and C. T. Ekstrøm, "Data-Driven Model Building for Life-Course Epidemiology," *Am J Epidemiol*, vol. 190, no. 9, pp. 1898–1907, Sep. 2021, doi: 10.1093/aje/kwab087.
[23] M. Pasquato, Z. Jin, P. Lemos, B. L. Davis, and A. V. Macciò, "Causa prima: cosmology meets causal discovery for the first time," Nov. 2023, [Online]. Available: http://arxiv.org/abs/2311.15160
[24] U. Hasan and M. O. Gani, "Optimizing Data-driven Causal Discovery Using Knowledge-guided Search," Apr. 2023, [Online]. Available: http://arxiv.org/abs/2304.05493
[25] A. Chakraborti, H. Vainio, K. T. Koskinen, and J. Lammi, "A Graph-Based Model Reduction Method for Digital Twins," *Machines*, vol. 11, no. 7, Jul. 2023, doi: 10.3390/machines11070733.
[26] M. Z. Naser, "Discovering causal models for structural, construction and defense-related engineering phenomena," *Defence Technology*, 2024, doi: 10.1016/j.dt.2024.04.007.
[27] R. Wiratama *et al.*, "Hubungan Pengetahuan Terhadap Perilaku Sanitasi Dalam Pencegahan Stunting Pada Ibu Balita," *Community Dev J*, vol. 4, no. 2, pp. 4083–4088, 2023.
[28] S. Huo *et al.*, "Influence of Maternal Exposure to Mass Media on Growth Stunting Among Children Under Five: Mediation Analysis Through the Water, Sanitation, and Hygiene Program," *JMIR Public Health Surveill*, vol. 8, no. 4, Apr. 2022, doi: 10.2196/33394.
[29] B. Banu, S. Haque, S. A. Shammi, and M. A. Hossain, "Maternal Nutritional Knowledge And Determinants Of The Child Nutritional Status In The Northern Region Of Bangladesh," *Bangladesh Journal of Multidisciplinary Scientific Research*, vol. 7, no. 1, pp. 11–21, Apr. 2023, doi: 10.46281/bjmsr.v7i1.2018.
[30] T. Susanto, H. Rasni, and L. A. Susumaningrum, "Prevalence of malnutrition and stunting among under-five children: A cross-sectional study family of quality of life in agricultural areas of Indonesia," *Med J Nutrition Metab*, vol. 14, no. 2, pp. 147–161, 2021, doi: 10.3233/MNM-200492.
[31] D. C. Rahmadiyah, J. Sahar, Widyatuti, R. A. D. Sartika, and H. Hassan, "Family Resilience With Stunted Children Aged Below 5 Years: A Qualitative Study in Depok City, Indonesia," *Glob Qual Nurs Res*, vol. 11, Jan. 2024, doi: 10.1177/23333936231221753.
[32] P. F. Wiliyanarti, Y. Wulandari, and D. Nasrullah, "Behavior in fulfilling nutritional needs for Indonesian children with stunting: Related culture, family support, and mother's knowledge," *J Public Health Res*, vol. 11, no. 4, Oct. 2022, doi: 10.1177/22799036221139938.
[33] P. T. Andreanetta, Q. Santosa, V. Indriani, K. Arifah, and W. Fatchurohmah, "Hubungan Berat Badan Lahir dengan Status Gizi dan Perkembangan Anak Usia 6 – 60 Bulan," *JURNAL MEDIKA UDAYANA*, vol. 11, no. 9, pp. 34–38, 2022.
[34] R. Windra, W. Program, S. Gizi, J. Ilmu, and K. Masyarakat, "Hubungan Riwayat BBLR, Asupan Protein, Kalsium, Dan Seng Dengan Kejadian Stunting Pada Balita," *Nutrition Research and Development Journal*, vol. 1, no. 2, pp. 1–12, 2021, [Online]. Available: https://journal.unnes.ac.id/sju/index.php/nutrizione/
[35] S. Thurstans *et al.*, "The relationship between wasting and stunting in young children: A systematic review," Jan. 01, 2022, *John Wiley and Sons Inc*. doi: 10.1111/mcn.13246.
[36] K. Sadler *et al.*, "How Can Nutrition Research Better Reflect the Relationship Between Wasting and Stunting in Children? Learnings from the Wasting and Stunting Project," *Journal of Nutrition*, vol. 152, no. 12, pp. 2645–2651, Dec. 2022, doi: 10.1093/jn/nxac091.
[37] Risnawati, H. Mas'ud, and Hendrayati, "Kejadian BBLR Dan Ispa Pada Batita 12-36 Bulan Dengan Kejadian Stunting Di Puskesmas Maros Baru," *Media Gizi Pangan*, vol. 28, no. 2, pp. 21–31, 2021.


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6184

[38] N. Hidayati, Z. Munawarah, P. P. Utami, and A. Ulmiaprilani, "Birth Weight and Length of Stunting Toddlers in Working Area of SetelukPublic Health Center in West Sumbawa in 2019: A Case Study," *Internatio nal Journal of Nursing and Health Services (IJNHS)*, vol. 5, no. 5, 2022, doi: 10.35654/ijnhs.v5i5.632.

[39] Ardan, R. Tosepu, and D. S. Effendy, "Identification of Stunting Determinants in Toddlers in Prevention of Nutritional Problems in the North Buton Regency," *KnE Life Sciences*, Sep. 2022, doi: 10.18502/kls.v0i0.11787.

[40] E. M. Mahfouz, E. S. Mohammed, S. F. Alkilany, and T. A. A. Rahman, "The relationship between dietary intake and stunting among pre-school children in Upper Egypt," *Public Health Nutr*, vol. 25, no. 8, pp. 2179–2187, Aug. 2022, doi: 10.1017/S136898002100389X.