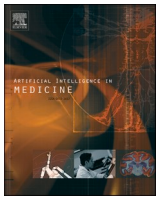
Artificial Intelligence In Medicine 128 (2022) 102310

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed



Medical checkup data analysis method based on LiNGAM and its application to nonalcoholic fatty liver disease

Tsuyoshi Uchida [a], Koichi Fujiwara [b,*], Kenichi Nishioji [c], Masao Kobayashi [c], Manabu Kano [a], Yuya Seko [d], Kanji Yamaguchi [d], Yoshito Itoh [d], Hiroshi Kadotani [e]

[a] Kyoto University, Japan
[b] Nagoya University, Japan
[c] Japanese Red Cross Kyoto Daini Hospital, Japan
[d] Kyoto Prefectural University of Medicine, Japan
[e] Shiga University of Medical Science, Japan

ARTICLE INFO

Keywords:
Nonalcoholic fatty liver disease
Medical checkup data
Linear non-Gaussian acyclic model
Collaborative filtering
Causal analysis

ABSTRACT

Although medical checkup data would be useful for identifying unknown factors of disease progression, a causal relationship between checkup items should be taken into account for precise analysis. Missing values in medical checkup data must be appropriately imputed because checkup items vary from person to person, and items that have not been tested include missing values. In addition, the patients with target diseases or disorders are small in comparison with the total number of persons recorded in the data, which means medical checkup data is an imbalanced data analysis. We propose a new method for analyzing the causal relationship in medical checkup data to discover disease progression factors based on a linear non-Gaussian acyclic model (LiNGAM), a machine learning technique for causal inference. In the proposed method, specific regression coefficients calculated through LiNGAM were compared to estimate the causal strength of the checkup items on disease progression, which is referred to as LiNGAM-beta. We also propose an analysis framework consisting of LiNGAM-beta, collaborative filtering (CF), and a sampling approach for causal inference of medical checkup data. CF and the sampling approach are useful for missing value imputation and balancing of the data distribution. We applied the proposed analysis framework to medical checkup data for identifying factors of Nonalcoholic fatty liver disease (NAFLD) development. The checkup items related to metabolic syndrome and age showed high causal effects on NAFLD severity. The level of blood urea nitrogen (BUN) would have a negative effect on NAFLD severity. Snoring frequency, which is associated with obstructive sleep apnea, affected NAFLD severity, particularly in the male group. Sleep duration also affected NAFLD severity in persons over fifty years old. These analysis results are consistent with previous reports about the causes of NAFLD; for example, NAFLD and metabolic syndrome are mutual and bi-directionally related, and BUN has a negative effect on NAFLD progression. Thus, our analysis result is plausible. The proposed analysis framework including LiNGAM-beta can be applied to various medical checkup data and will contribute to discovering unknown disease factors.

1. Introduction

Electronic health record (EHR), in particular, medical checkup data, would be useful for discovering unknown factors of disease progression [1–3]. Most persons have a medical checkup once a year in Japan. Various checkup items can be collected simultaneously, such as height, body weight, waist circumference, blood components, electrocardiography, and lifestyle habits through questionnaires and interviews. That is, we can easily collect and utilize big medical data for analyzing factors of disease progression.

Many studies have adopted statistical tests for analyzing disease progression [4,5]; however, they are not satisfactory. Statistical tests cannot take into consideration the causal relationship between checkup items. Causal inference is of importance because causes and results need to be distinguished when disease progression factors are investigated.

A linear non-Gaussian acyclic model (LiNGAM) is a statistical causal inference method used in machine learning, which can distinguish between correlation and causality based on a structure equation model [6].

* Corresponding author at: The Department of Material Process Engineering, Nagoya University, Nagoya 464-8601, Japan.
E-mail address: fujiwara.koichi@hps.material.nagoya-u.ac.jp (K. Fujiwara).

https://doi.org/10.1016/j.artmed.2022.102310
Received 16 February 2021; Received in revised form 24 March 2022; Accepted 17 April 2022
Available online 22 April 2022

Some applications of LiNGAM to medical data have been attempted.

Helajärvi et al. evaluated the relationships among changes in TV watching time, waist circumference, and body mass index (BMI) in participants using LiNGAM. Their analysis result suggested that TV watching time is antecedent to increases in the waist circumference and BMI, and reduction of TV watching time may be effective for long- or short-term weight management [7]. Ogawa et al. applied LiNGAM to functional magnetic resonance imaging (fMRI) data for analyzing brain activities. They attempted to distinguish frontoparietal motor-related networks based on the effective connectivity in the motor imagery (MI) and the motor execution (ME) conditions. LiNGAM was applied to identify the effective connectivity among the frontoparietal motor-related brain regions [8]. These studies show that LiNGAM is a promising machine learning technique for causal inference.

In the present work, we propose a framework for discovering unknown causal factors of disease progression based on LiNGAM. The purpose of the original LiNGAM is to estimate a directed graph representing a causal structure among variables, which is expressed as the regression coefficient matrix $\boldsymbol{\beta}$. The proposed method, referred to as LiNGAM-beta, focuses on the relationship between the target variable and other variables by extracting the specific row in the regression coefficient matrix $\boldsymbol{\beta}$ that corresponds to the target variable. By using the proposed LiNGAM-beta, we can quantitatively analyze the causal directions and strengths of health checkup data, which is useful for the development of a new therapy in the future.

However, there are many missing values in medical checkup data, which makes analysis difficult. Checkup items vary from person to person because he/she can select the medical checkup menu, or his/her social health insurance may specify the menu. Some tests may not be performed due to economic constraints. Thus, items that have not been tested result in missing values, and such missing values should be appropriately imputed. In addition, the patients with target diseases or disorders are limited in comparison with the total number of persons recorded in the data, which makes analysis difficult. Hence, we adopted collaborative filtering (CF) [9] and a sampling approach [10] for dealing with these problems before the application of LiNGAM-beta.

CF is a machine learning technique for interpolating missing values in a table data format, which is mainly used for recommender systems [11]. Gräßer et al. presented two methods for data-driven therapy decision support based on CF. Both algorithms aim to predict the individual response to different therapy options using diverse patient data and recommend the therapy that is assumed to provide the best outcome for a specific patient and time [12]. Zeng et al. applied CF to the prediction of pathogenic human genes. They proposed a probability-based collaborative filtering model (PCFM) to predict pathogenic human genes. They compared the results of PCFM with the results of four state-of-the-art approaches, and the results show that PCFM performs better than other advanced approaches [13]. We used (CF) for statistical imputation of missing values in the medical checkup data.

The sampling approaches are widely used in machine learning when the data distribution is imbalanced. RUSBoost is a well-known imbalance data algorithm that combines random under-sampling (RUS) and boosting [10]. In addition, Fujiwara et al. developed an imbalance data algorithm that uses under-sampling and over-sampling, simultaneously [14]. These approaches to sampling can improve the accuracy of statistical models for the imbalanced data by tuning the data distribution compared to ordinal classification algorithms when the imbalanced data is analyzed. Thus, the sampling approach is useful for medical checkup data analysis.

In this study, we also propose an analysis framework consisting of LiNGAM-beta, CF, and the sampling approach for causal inference of medical checkup data. The usefulness of the proposed analysis framework is demonstrated through disease progression factor identification from real-world medical checkup data.

The target disease in this study is nonalcoholic fatty liver disease (NAFLD), which is considered to be associated with metabolic syndrome bi-directionally and could lead to various diseases such as cardiovascular diseases [15]. NAFLD is the most common cause of chronic liver disease and results in serious public health problems in many countries worldwide. NAFLD encompasses a wide spectrum of liver pathology, ranging from non-alcoholic fatty liver (NAFL), which is usually benign, to non-alcoholic steatohepatitis (NASH), which is characterized by steatosis, lobular inflammation, and hepatocellular injury, and may progress to liver cirrhosis, hepatic failure, and hepatocellular carcinoma in the absence of significant alcohol consumption [16–19]. Besides metabolic syndrome, NAFLD is reported to be associated with sleep physiology, cholesterol and uric acid metabolism [20–22].

2. Material & methods

In this section, we propose a new medical checkup data analysis method based on LiNGAM [6] and introduce collaborative filtering (CF) for missing value imputation in the medical checkup data [9]. In addition, the details of the medical checkup data analyzed in this study are described.

2.1. LiNGAM-beta

LiNGAM is a model expressing the causal structure among variables [6] and is designed to be applied to data containing confounders. An example of a causal structure is shown in Fig. 1. The vertices represent variables, and the directed edges express causal dependencies among the variables. For example, there is a directed edge from vertex x_1 to x_2, which means x_1 has a causal effect on x_2.

LiNGAM assumes that each variable is generated as the linear combination of the causal antecedent variables and an exogenous variable. In the LiNGAM assumption, the causal structure is expressed as a directed acyclic graph (DAG), which is a directed graph without a cycle. An example of a LiNGAM model in Fig. 1 is shown below:

$$x_1 = e_1 \tag{1}$$

$$x_2 = b_{21}x_1 + e_2 \tag{2}$$

$$x_3 = b_{31}x_1 + b_{32}x_2 + e_3 \tag{3}$$

where x_i and $e_i (i=1,2,3)$ are observed variables and exogenous noise, respectively. b_{21}, b_{31}, and b_{32} are the coefficients.

When the data include p variables, LiNGAM is expressed as a linear equation as follows:

$$\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{e} \tag{4}$$

where $\boldsymbol{x} \in \Re^p$ is the variable vector, $\boldsymbol{e} \in \Re^p$ is the exogenous noise vector, and $\boldsymbol{B} \in \Re^{p \times p}$ is the coefficient matrix, which must be a lower triangular matrix whose diagonal components are zero due to the causal assumption. The goal of the causal discovery with LiNGAM is to estimate the

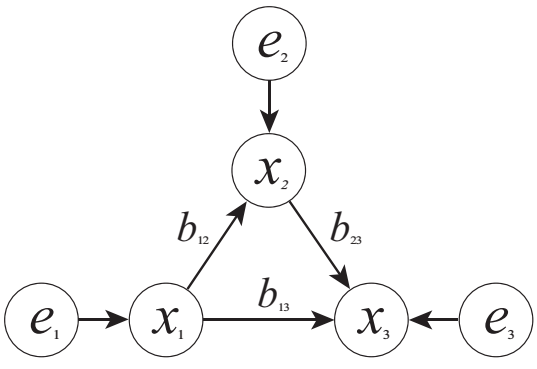


Fig. 1. Causal structure.

coefficient matrix $\boldsymbol{B}$, which describes the causal relationships among the variables. Although there are some assumptions in the LiNGAM model, one of the biggest assumptions is that all variables in the data need to be continuous.

The coefficients in the row vector in $\boldsymbol{B}$ that associate with the severity of the target disease express the causal effects of the test items in the medical checkup on disease severity. Their absolute values and positive/negative signs indicate the strength and the direction of the causal effects, respectively. Thus, a specific row vector $\boldsymbol{\beta}$ in $\boldsymbol{B}$ is extracted. The absolute values and positive/negative signs of the extracted regression coefficients are compared with each other to estimate the causal effects on target disease progression. This method is referred to as LiNGAM-beta.

Several LiNGAM algorithms have been proposed. ICA-LiNGAM algorithm is the first proposed LiNGAM algorithm that is based on independent component analysis (ICA). Since ICA-LiNGAM requires some additional information, including the initial guess and the convergence criteria [6], it may get stuck in local optima and not converge to a reasonable solution if the initial guess is not appropriately chosen. DirectLiNGAM is a method for estimating a causal ordering of variables with no prior knowledge of the structure and is guaranteed to converge to the right solution within a small fixed number of steps if the data strictly follows the LiNGAM assumptions and the sample size is infinite. DirectLiNGAM enables a more accurate estimation of a causal order of the variables in a disambiguated and direct procedure [23]. Pairwise LiNGAM is a modification of DirectLiNGAM that uses simple first-order approximations of the likelihood ratio for reducing computational burden [24]. Thus, we adopted Pairwise LiNGAM in this study, taking its advantages into account.

2.2. Collaborative filtering

In general, medical checkup items vary from person to person. For example, a person undergoes a brain examination using MRI, and another person does not undergo MRI due to its economic burden. In such a case, items that have not been tested include missing values, and such missing values must be imputed appropriately before analysis.

CF [9] is a machine learning technique for interpolating missing values in a table data format, which is mainly used for recommender systems [11]. The key idea behind user-based CF is that a rating of a user u for a new item i is likely to be similar to that of a user v if users u and v rated items similarly. Likewise, the user u is likely to rate two items i and j in a similar fashion if other users gave similar ratings to these two items [25]. The rating of the user u for the item i is expressed as r_{ui}. Users similar to the user u are called the nearest neighbors of u.

The aim of user-based CF is to interpolate the missing value of r_{ui} using ratings given to i by u's nearest neighbors. r_{ui} is predicted as

$$\widehat{r}_{ui} = \overline{r}_u + \sigma_u \frac{\sum_{v \in N_i(u)} \mathrm{sim}\left(u, v\right) \cdot \frac{r_{vi} - \overline{r}_v}{\sigma_v}}{\sum_{v \in N_i(u)} \mathrm{sim}\left(u, v\right)} \tag{5}$$

where $\widehat{r}_{ui}$ is the predicted rating of the item i by the user u, and r_{vi} is the known rating of i by v. $\overline{r}_u$, $\overline{r}_v$, σ_u and σ_v represent mean values and standard deviations of the ratings by users u and v, respectively. $\mathrm{sim}(u, v)$ denotes the similarity between users u and v. $N_i(u)$ is the set of the nearest neighbors of the user u who already rated the item i. The Pearson correlation was adopted as the similarity $\mathrm{sim}(u, v)$ in this study, which is expressed as follows:

$$\mathrm{sim}(u, v) = \frac{\sum_{i \in I_{uv}} (r_{ui} - \overline{r}_u)(r_{vi} - \overline{r}_v)}{\sqrt{\sum_{i \in I_{uv}} (r_{ui} - \overline{r}_u)^2} \sqrt{\sum_{i \in I_{uv}} (r_{vi} - \overline{r}_v)^2}} \tag{6}$$

where I_{uv} is the set of items rated by both u and v.

2.3. Data description

The medical checkups were carried out between April 2014 and March 2015 in the health care division of the Japanese Red Cross Kyoto Daini Hospital. The data contain records on 3511 persons (1920 males, 23–93 y.o. mean: 57.7 y.o.; 1591 females, 27–87 y.o. mean: 57.4 y.o.) who received a medical checkup, including physical and physiological examinations, questionnaires about lifestyle such as sleep duration, and a blood screening examination. In addition, the presence of fatty liver was diagnosed by means of abdominal ultrasonography [26].

It has been reported that obstructive sleep apnea (OSA) is associated with NAFLD severity [27]. In order to diagnose OSA, polysomnography (PSG) or portable sleep monitoring devices are required; however, they are hard to perform in standard medical checkups. Instead of these diagnostics, we checked snoring frequency, which is strongly associated with OSA [28,29]. The snoring frequency was collected by means of a questionnaire and was rated on a four-point scale: "0: almost never", "1: sometimes", "2: often", and "3: almost every day". In addition, there was another choice, "Don't know", which was treated as a missing value in this work.

Persons were excluded if they had any of the following: (1) hepatitis virus infection (i.e., hepatitis B surface antigen-positive or hepatitis C virus antibody-positive); (2) excessive use of alcohol ($\geq$30 g/day for males and $\geq$20 g/day for females); (3) autoimmune liver disease (e.g., autoimmune hepatitis, primary biliary cirrhosis); (4) uncontrolled biliary disease (e.g., choledocholithiasis or cirrhosis); (5) ongoing treatment for autoimmune disease (e.g., rheumatoid arthritis or systemic lupus erythematosus); (6) ongoing hormone therapy for cancer (i.e., breast cancer or prostate cancer); (7) ongoing treatment for thyroid disease [26].

The abdominal ultrasonography (US) protocol and the definition of fatty liver adopted in this study were as follows: all eligible persons received the abdominal US to assess their liver. The abdominal US was performed by expert sonographers with a 3.5 MHz convex array probe (Xario unit, Toshiba Medical Systems, Tokyo, Japan). The certified gastroenterologists reviewed the images and made fatty liver diagnosis according to four known criteria (i.e., hepatorenal echo contrast, bright liver, deep attenuation, and vascular blurring) [26].

The test items in the medical checkup, their abbreviations, and summaries of each variable are listed in Table 1. In the summary column, the mean and the standard deviation are shown if a variable is continuous, and the number in each category if categorical. Table 2 indicates the number of missing values and the missing rate of items in the medical checkup data. The overall missing rate was 0.77%, and the item 'snoring frequency' had the largest missing rate of 18.1%.

The study was approved by the institutional review board of the Japanese Red Cross Kyoto Daini Hospital (S25–15), and the study was conducted in accordance with the Declaration of Helsinki. The health checkup data will be made available by the corresponding author after approval by the institutional review board of the Japanese Red Cross Kyoto Daini Hospital to colleagues who propose a reasonable scientific request.

2.4. NAFLD severity

In the present study, NAFLD severity was scored with a four-point discrete variable. This score was determined based on the results of abdominal ultrasonography and FIB4 [30], which is an index indicating the degree of fibrosis advancement, expressed as follows:

$$\mathrm{FIB4} = \frac{\mathrm{Age\,[years]} \times \mathrm{AST\,[U/L]}}{\mathrm{PLT}\left[10^9/\mathrm{L}\right] \times \sqrt{\mathrm{ALT\,[U/L]}}} \tag{7}$$

Shah et al. investigated the optimal cut-off values of FIB4 for predicting the existence of advanced fibrosis. Fibrosis risk is low if FIB4$\leq$1.30, modest if 1.30<FIB4<2.67, and high if FIB4$\geq$2.67 [31].

Table 1
Medical checkup items.

Item	Unit	Summary
Abdominal ultrasonography	–	NAFLD: 893, Healthy: 2618
Sex	–	Male: 1920, Female: 1591
Age	years	58±12
Sleep duration	h	6.4±0.9
Snoring frequency*	–	0:907, 1:1104, 2:466, 3: 398, missing: 636
Systolic blood pressure (SBP)	mmHg	120±17
Diastolic blood pressure (DBP)	mmHg	71±11
Body fat percentage (BFP)	%	22.8±5.5
Body mass index (BMI)	–	22.2±3.3
Waist circumference (WC)	cm	81.1±9.3
C-reactive protein (CRP)	mg/dL	0.09±0.26
Total protein (TP)	g/dL	7.2±0.4
Serum albumin (ALB)	g/dL	4.11±0.26
Alkaline phosphatase (ALP)	U/L	200±57
Gamma-glutamyl transpeptidase (γ-GTP)	U/L	30±32
Aspartate aminotransferase (AST)	U/L	23±9
Alanine aminotransferase (ALT)	U/L	21±14
Lactate dehydrogenase (LDH)	U/L	171±31
Cholinesterase (ChE)	U/L	328±72
Blood urea nitrogen (BUN)	mg/dL	14.5±3.8
estimated glomerular filtration rate (eGFR)	mL/min	70±12
Uric acid (UA)	mg/dL	5.4±1.3
Triglyceride (TG)	mg/dL	96±60
High-density lipoprotein cholesterol (HDL)	mg/dL	67±19
Low-density lipoprotein cholesterol (LDL)	mg/dL	123±29
White blood cell count (WBC)	/μL	52±14
Hemoglobin count (Hb)	g/dL	13.9±1.3
Platelet count (PLT)	/μL	23.0±5.2
Fasting plasma glucose (FPG)	mg/dL	100±15
Hemoglobin A1c (HbA1c)	%	5.9±0.6

* Snoring frequency: "0: almost never", "1: sometimes", "2: often", and "3: almost every day", and missing means the answer "Don't know."

Table 2
Missing values in dataset.

Item	#Missing	Missing rate [%]
CRP	137	3.90
ALB	113	3.22
ALP	113	3.22
LDH	113	3.22
ChE	137	3.90
BUN	137	3.90
snoring frequency	636	18.1

Since FIB4 is a screening tool for NASH, the confirmed diagnosis of NASH requires a liver biopsy, and very few patients with a high risk of NASH have a liver biopsy due to its invasiveness. It is difficult to define appropriate NAFLD severity based on liver biopsy because a small proportion of patients suspected of moderate to severe liver fibrosis are recommended to receive liver biopsy in daily clinical practice.

Before the study, we have confirmed the applicability of FIB4 by evaluating the severity of hepatic fibrosis of a Japanese NAFLD cohort in 2012 [32]. In a large cohort study of NAFLD in the USA, Kanwal et al. described that "we calculated FIB-4 score to define liver fibrosis severity and found a stepwise increase in the risk of HCC with increasing FIB-4 score" [33]. These reports justify the use of FIB4 for evaluating the degree of hepatic fibrosis for NAFLD studies in the general population because liver biopsy could not be performed in all the participants. Therefore, it is concluded that our definition of NAFLD severity in this study is appropriate.

In this work, the severity of NAFLD was defined according to the result of abdominal ultrasonography and FIB4, as shown in Table 3. Persons without NAFLD were classified into class 0. Persons with NAFLD were classified into classes 1–3 according to FIB4 as follows; the low-risk group was class 1, the modest-risk group was class 2, and the high-risk group was class 3.

Table 3
Classification of NAFLD severity.

NAFLD severity	0	1	2	3
NAFLD	no	yes	yes	yes
FIB4	–	≤1.30	1.3<, <2.67	2.67≤

The numbers of persons belonging to each of the NAFLD severity classes are listed in Table 4, which shows that NAFLD severity was unevenly distributed and that 66% of males and 85% of females belonged to class 0 (healthy).

2.5. Analysis flow

We analyzed the medical checkup data by dividing them into two gender-specific groups because the prevalence of NAFLD is different between male and female [34].

As preprocessing, we eliminated, from the target medical checkup data, binary medical checkup items like questionnaires about lifestyle since LiNGAM can only handle continuous variables. Each checkup item was standardized with zero mean and a standard deviation of one. We performed under-sampling so that the numbers of persons in each class became balanced since NAFLD severity was unevenly distributed, as shown in Table 4.

We applied CF and the proposed LiNGAM-beta to the preprocessed data in order to estimate the causality of the medical checkup data. Finally, the coefficients in the row vector in $\boldsymbol{B}$ that are associated with NAFLD severity were extracted. The above procedure was repeated 1000 times, and the mean was calculated for precise evaluation because we randomly under-sampled records.

Females are more likely to store subcutaneous fat before menopause and visceral fat after menopause than males. The prevalence of NAFLD is elevated in postmenopausal females [35]. Thus, we performed an additional analysis by dividing the gender-specific groups into four groups according to age: young male, old male, young female, and old female. The age boundary was set to fifty years old, in accordance with the average age at which menopause begins.

In addition, we performed the t-test for validating the significance of each regression coefficient estimated by the proposed LiNGAM-beta. $p<.01$ was adopted as the significance level.

We used the Python library Surprise for performing CF [36], and MATLAB codes published by Shimizu and Hyvärinen [24] were used here for executing Pairwise LiNGAM.

3. Results

Fig. 2 shows the causal effects of each medical checkup item on NAFLD severity estimated by LiNGAM-beta in the male and the female data, respectively. Bars express the directions and the strengths of the causal effects of items on NAFLD severity. The items with positive values would increase NAFLD severity, and those with negative values would decrease NAFLD severity. The absolute values denote the strengths of their causal effects. Asterisks denote items whose regression coefficient was significant. Although almost all items were significant factors of NAFLD according to the t-test, in particular, items related to metabolic syndromes such as WC, BFP, BMI, and age were highly ranked in both

Table 4
Number of samples grouped by NAFLD severity*.

NAFLD severity	0	1	2	3	total
Male	651 (1269)	400	235	16	1302 (1920)
Female	242 (1349)	144	93	5	484 (1591)

* Numbers in parentheses are numbers of samples before under-sampling.

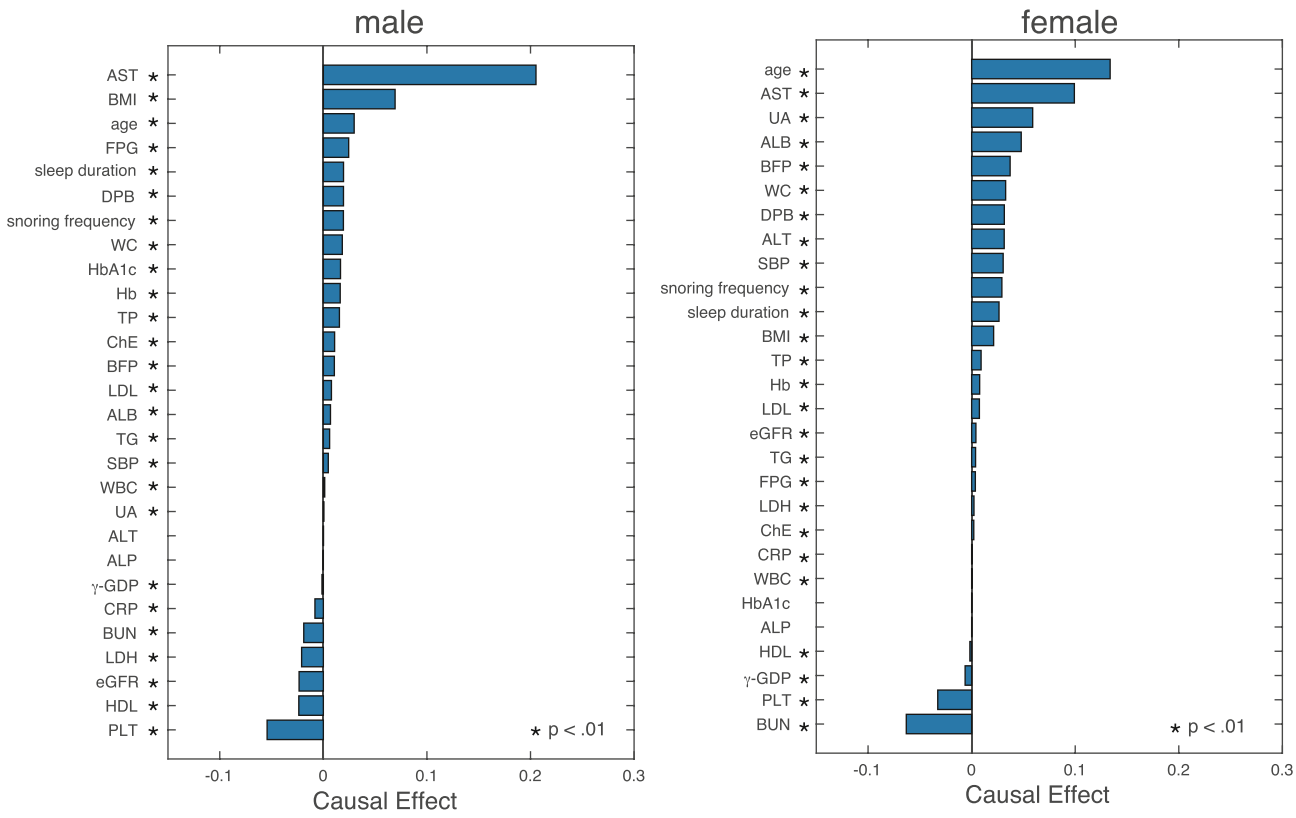


Fig. 2. Causal effects on NAFLD severity in male (left) and female (right).

sexes. Age had a strong positive effect on the female result, and BUN indicated a relatively strong negative effect on NAFLD severity in both sexes.

According to Fig. 2, snoring frequency positively affected NAFLD severity in both sexes, which indicated that OSA was associated with NAFLD severity in our analysis as well; however, Fig. 3 suggests that snoring frequency has a negative effect on NAFLD severity only in the young female group, and a positive effect in the other groups. In addition, sleep duration had a negative effect on NAFLD severity in the young groups according to Fig. 3. Sleep condition improvement may contribute to NAFLD prevention.

In this work, we adopt CF for analyzing medical checkup data that contain many missing values before applying the proposed LiNGAM-beta; however, other imputation methodologies have been proposed. The simplest method is mean imputation, which fills missing values with the mean values of the item. In addition, the hot-deck imputation is also a well-known method that replaces a missing value with an observed value of the most similar sample [38]. We performed additional experiments to validate the imputation efficacy of CF through comparison with the mean imputation and the hot-deck imputation. We generated some artificial missing values in 'snoring frequency,' originally having the largest missing rate, such that the missing rates would become 20%, 40%, 60%, and 80%. In addition, the missing rates of other items were determined so that the ratio of the number of missing values in the item to 'snoring frequency' was consistent with that of the original data.

The artificial missing values were imputed using three methods, and their imputation errors were compared. In the hot-deck method, the Euclidean distance was used as the similarity. The root means squared error (RMSE) between the original values before the missing and the imputed values were used for evaluation. This procedure was repeated 100 times for precise evaluation. Table 5 shows the comparison result, in which the bold fonts indicate the best performance. The mean imputation became slightly worse as the missing rate increased, and the performance of the hot-deck method was worst. In contrast, CF achieved the best imputation performance of the three methods, and it did not deteriorate even when the missing rate increased. This experiment clearly illustrates that CF is effective for missing value imputation. The mean imputation always substitutes the constant value for the missing values even if the variance is large. The hot-deck imputation uses information from one sample only. In contrast, CF uses multiple samples weighted by the similarity as written in Eq. (5). Hence, CF can adequately impute missing values.

4. Discussion

According to the analysis results, age and AST had a positive effect on NAFLD severity while PLT was negative in both sexes, which is to be expected because age and AST are in the numerator, and PLT is in the denominator in the FIB4 definition Eq. (7). Thus, it is concluded that our analysis result by the proposed LiNGAM-beta is appropriate.

Usually, in machine learning problems, the trained model performance is of interest, and the receiver operating characteristic (ROC) curve is frequently used for performance evaluation. On the other hand, the aim of causal inference is not prediction or classification, and we cannot evaluate the model performance with a ROC curve. Instead, we assessed the validity of our analysis results by comparing them with previous reports on NAFLD progression.

Hamaguchi et al. carried out a longitudinal survey using medical checkup record data of healthy Japanese subjects to investigate the relationship between metabolic syndrome and NAFLD progression and concluded that metabolic syndrome is a strong predictor of NAFLD [39]. Since the association between metabolic syndrome factors such as WC, BFP, BMI, and NAFLD has been reported [26,40], our results are consistent with previous reports according to Fig. 2. The results also showed that age has the most significant effect on increasing NAFLD severity in the female data. Eguchi et al. reported that the prevalence of NAFLD gradually increases with age in females [41], which is consistent with our analysis result of the female data.

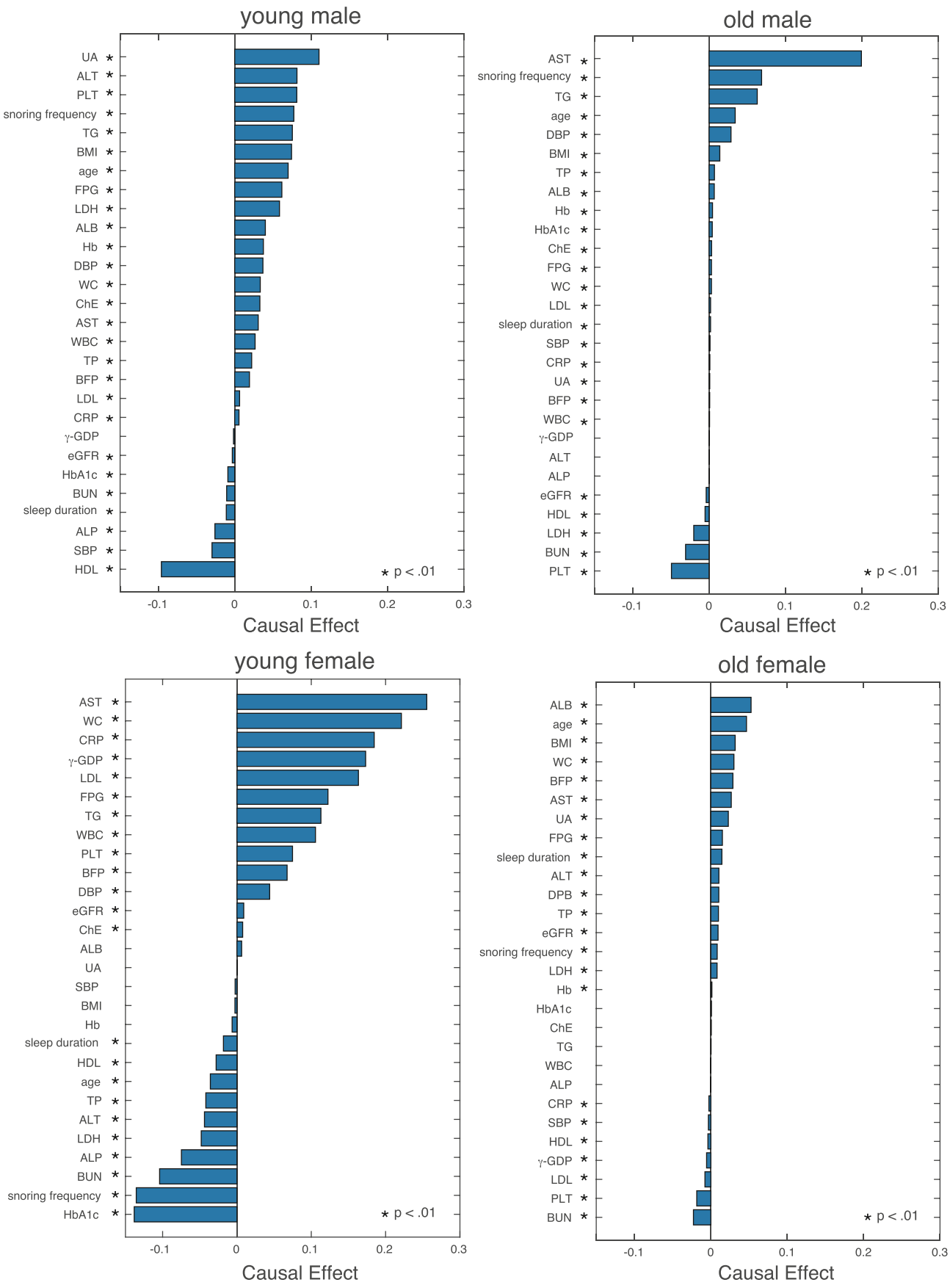


Fig. 3. Causal effects on NAFLD severity in young male (top left), old male (top right), young female (bottom left), and old female (bottom right).

Table 5
RMSE by three imputation methods.

Method	Missing ratio of "snoring frequency"			
	20%	40%	60%	80%
CF	**0.23±0.01**	**0.23±0.00**	**0.23±0.00**	**0.23±0.00**
Mean	0.24±0.01	0.24±0.00	0.25±0.00	0.25±0.00
Hot-deck	0.31±0.01	0.31±0.01	0.32±0.00	0.32±0.01

In our analysis, blood urea nitrogen (BUN) showed a strong negative effect on NAFLD severity in both sexes, particularly in females. The association between BUN and NAFLD has been investigated in some studies. Thomsen et al. studied ureagenesis, which is an essential metabolic liver function for whole-body nitrogen homeostasis, in a rodent model of diet-induced NASH, and suggested that NASH progress leads to impairment of the ability of the liver to adjust to changes in demand of nitrogen excretion, and to the removal of potentially neurotoxic nitrogenous substances by means of the synthesis of urea [42]. Kani et al. compared the dietary intake of NAFLD patients and that of healthy persons. They showed that BUN levels of the former were higher than those of the latter [43]. Erçin et al. investigated the difference of BUN levels between patients with NAFLD and healthy persons and concluded that there was no significant difference between the two groups; however, all of the participants in their study were male, and patients with severe NAFLD were not included [44].

Some studies focused on the relationship between NAFLD and urea cycle enzymes (UCEs) related to BUN. Eriksen et al. analyzed liver mRNA expression in the gene pathway that governs hepatic nitrogen conversion in NAFLD patients and healthy individuals. They found that gene expressions of most UCEs were downregulated in NAFLD patients in comparison with healthy individuals and concluded that downregulation of urea cycle flux-generating carbamoyl phosphate synthetase (CPS1) is associated with the functional capacity loss for ureagenesis in NASH [45]. Chiara et al. measured ammonia concentrations in rats and humans with NASH as well as the gene and protein expression of ornithine transcarbamylase (OTC) and CPS1. They found that NASH is associated with a reduction in the gene and protein expression in UCEs, which results in hyperammonemia [46]. UCE genes hypermethylation and urea synthesis impairment would lead to scar tissue development and the progression of NAFLD. Thus, while these findings support our analysis result that BUN has a negative effect on NAFLD severity, additional experiments and collection of clinical data are needed to confirm our result.

Lonardo et al. reported that both insulin resistance and uric acid were independent predictors of NAFLD [47], and several studies including two meta-analytic reviews support these findings [48,49]. Furthermore, although insulin resistance and serum uric acid concentrations are associated, both parameters are independently associated with NASH [50]. This result suggests that each factor contributes to the development of progression of NAFLD independently. The result of this study may support these previous studies.

Taketani et al. investigated the relationship between gastroesophageal reflux disease (GERD) and insomnia in subjects with biopsy-proven NAFLD. They concluded that about 30% of Japanese patients with NAFLD had insomnia [51]. In addition, it was reported that sleep duration is related to NAFLD [52]. In our analysis, sleep duration generally contributed to NAFLD severity, which is consistent with a previous report [52]. The association between sleep duration and NAFLD has been under discussion. The effect of sleep duration varied by age group according to Fig. 3: negative effect in the young group and a positive effect in the old group.

The directions of the causal effect of eGFR on NAFLD severity were negative in males and positive in females, respectively, according to Fig. 2. There was relatively weak correlation between eGFR and NAFLD in females, and it is assumed that the BUN level in NAFLD may reflect the ureagenesis and urea cycle enzymes rather than renal function. The result regarding eGFR in females was not consistent with a systematic review of NASH and NAFLD, which reported that advanced fibrosis is associated with low eGFR [53]. According to Table 4, the number of women with severe NAFLD was very small, which suggests it might be an unreliable result. In addition, the causal effect of HbA1c on NAFLD severity in young females was different from the other groups, which also may be due to the small proportion of NAFLD patients compared to the other groups. In addition, NAFLD is a sexually dimorphic condition [54–56]. The results of this study reinforce these previous reports of sexual differences in NAFLD.

Some reports indicated that snoring frequency is related to fat deposition. Whittle et al. analyzed the volume of soft tissue in the neck using MRI and found no significant sexual differences in fat deposition in the regions prone to collapse during sleep [57]. O'Donnell et al. suggested that circulating levels of leptin are higher in females than males because subcutaneous fat produces more leptin than visceral fat and that leptin or some other obesity-related factors may play a role in sex-related differences in breathing during sleep [58]. On the other hand, there are some previous reports suggesting sexual differences in snoring severity. Mohsenin found that obese females had significantly smaller pharyngeal airways than obese males; however, the severity of SAS was influenced by the pharynx size in males only [59]. Our analysis result suggested that snoring frequency may have a positive effect on NAFLD severity in groups other than young females. However, the number of female persons with frequent snoring was small, as shown in Fig. 4. It has been reported that females tend to underestimate the loudness of their snoring [60]. That is, our data on female snoring might not reflect facts. It was difficult to estimate the causal effect of snoring on NAFLD severity with our data. Exact female snoring data measured by a snoring recording system like PSG is necessary for further analysis.

In this study, we adopted under-sampling in order to tune the imbalanced data distribution of NAFLD severity; the number of persons in class 0 was under-sampled randomly from 1269 to 651 in the male data and from 1349 to 242 in the female data, respectively, as shown in Table 4. Although the sampling operation, such as over-sampling and under-sampling altered the data distribution, the sampling approaches are common in machine learning when the data distribution is imbalanced. Over-sampling requires a certain number of samples, and the number of patients with severe NAFLD was very small in this data. Thus, only under-sampling was used in this study because the distribution of NAFLD severity was highly-imbalanced.

Although LiNGAM-beta estimates the causal directions and their strengths, other machine learning methods can calculate the contribution of input variables to an output variable. Partial least squares (PLS) is a well-known regression method with which it is possible to avoid the multicollinearity problem and construct a highly accurate prediction model by using the latent variables [61,62]. We tried to use PLS for causal effect estimation on NAFLD severity. We applied CF and PLS to the medical checkup data following the same procedure as Section 2.5 and obtained regression coefficients shown in Fig. 5. The casual directions of LDL and ALP were different between the proposed LiNGAM-beta and PLS both in male and female.

We were able to identify missing LDL as a factor with a positive causal direction by means of the proposed LiNGAM-beta in this study, whereas the PLS could not. A large cohort study in Japan reported that hyperlipidemia including high LDL cholesterol level and low HDL cholesterol level were major complications of NAFLD [63]. Sun et al. classified NAFLD patients into four groups based on their LDL levels in medical checkup data and found that NAFLD incidence became high as the LDL level became high [64]. Min et al. reported a direct relationship between the hepatic expression of HMGCoA reductase (HMGCR) and the severity of the NAFLD activity score, suggesting the greater levels of LDL in patients with NAFLD [65]. Although the causal strength of ALP is minimal in the LiNGAM-beta study, a previous report by Rafiq et al. reporting that an increased ALP level was an independent predictor of liver-related mortality of NASH patients reinforces our results [66].

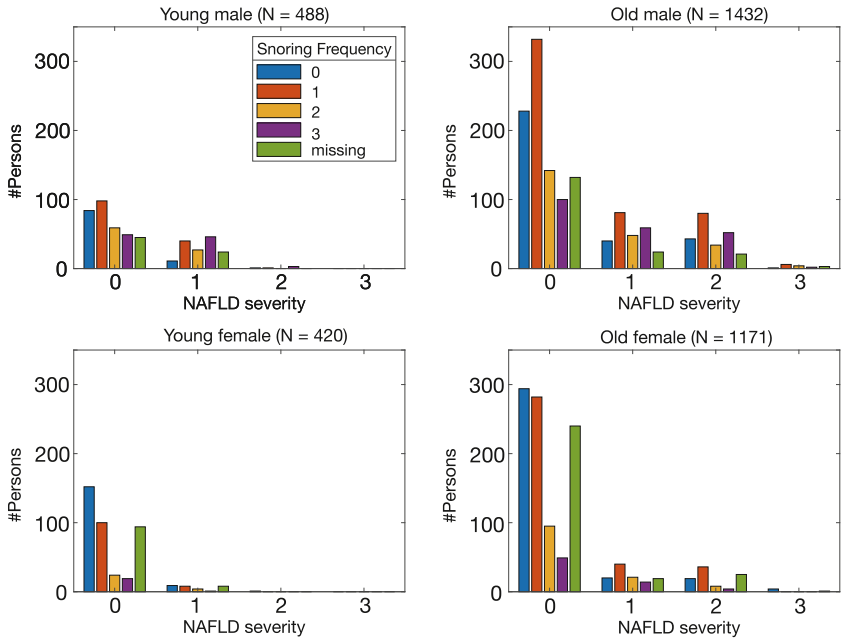


Fig. 4. Number of persons grouped by NAFLD severity and snoring frequency of young male (top left), old male (top right), young female (bottom left), and old female (bottom right).

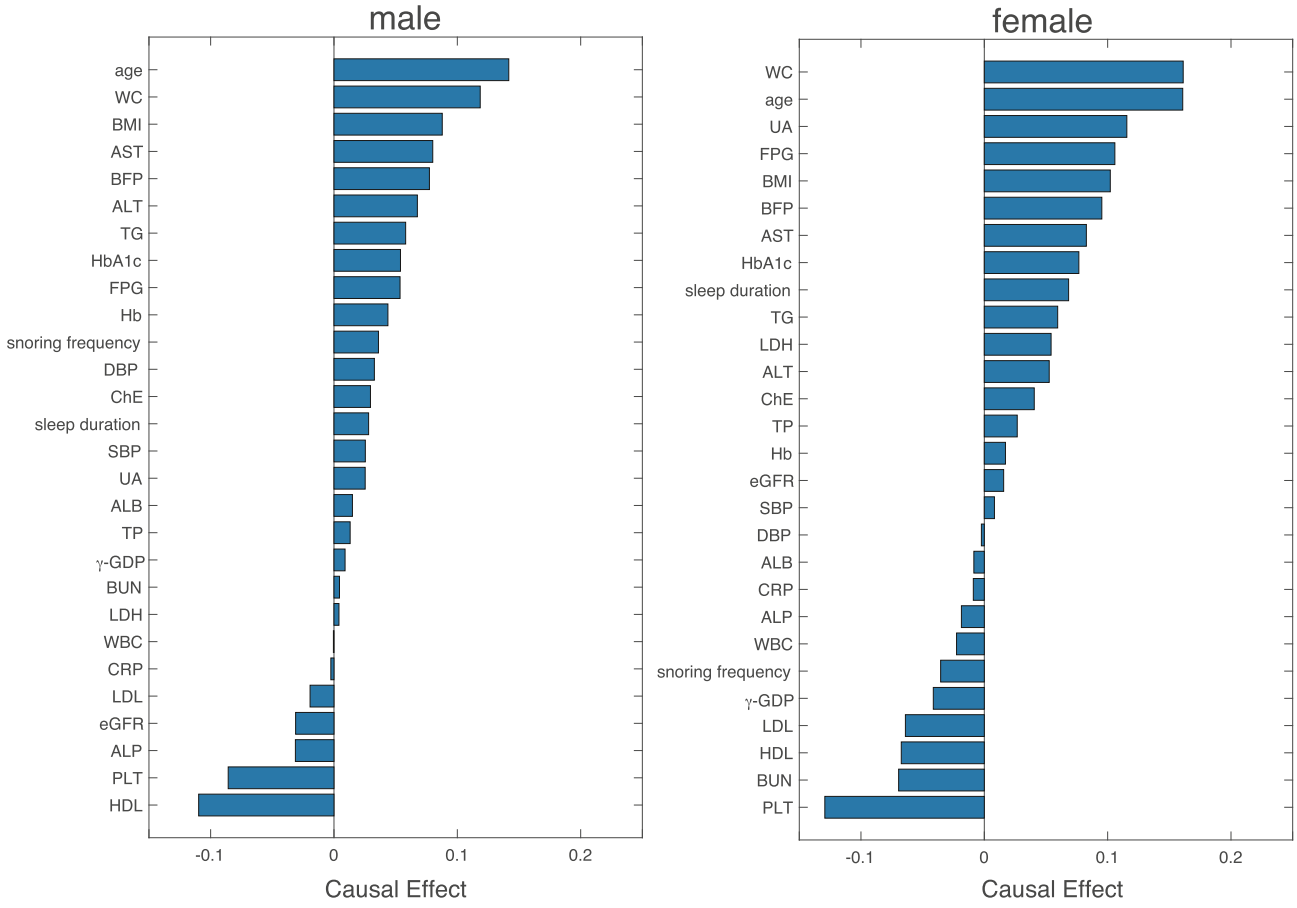


Fig. 5. Causal effects on NAFLD severity estimated with PLS in male (left) and female (right).

Thus, the proposed LiNGAM-beta was plausible because its estimation result had more agreement with previous studies than PLS. Although random forest (RF) also can compute the variable importance [67,68], it takes only a positive value and does not indicate the causal direction. We did not consider RF in this study.

The purpose of regression methods, including PLS, is prediction.

Some information not related to prediction such as confounders should be ignored to maximize the prediction performance, which may cause a reversal of the causal directions, as was the case with LDL and ALP in this experiment. On the other hand, LiNGAM assumes the causal structure as a directed acyclic graph that includes confounders, and confounders are decomposed as exogenous variables and endogenous variables. We can quantitatively discuss the causal directions and strengths among health checkup data by considering the sign and magnitude of the estimated coefficients in the proposed LiNGAM-beta. Thus, the proposed LiNGAM-beta can estimate the first and independent effects on disease progression.

Since our analysis results were consistent with existing reports on NAFLD severity in many respects, it is concluded that the proposed LiNGAM-beta is effective in medical checkup data analysis.

The limitations of this study include the medical checkup data, such as the fact that the population was limited to Japanese. In addition, we did not analyze binary items because of the constraint of LiNGAM.

5. Conclusion

In this study, we proposed a new medical checkup data analysis method based on LiNGAM. In the proposed LiNGAM-beta, we can quantitatively analyze the causal directions and strengths among health checkup items. We adopted CF for missing value imputation of the medical checkup data. We performed a causal effect analysis to identify NAFLD severity factors from the medical checkup data based on the proposed LiNGAM-beta. Since the causal relationships estimated by LiNGAM-beta were consistent with previous reports on NAFLD progression, it is appropriate to use the proposed method for causal analysis of medical checkup data. In particular, our analysis indicated that BUN is a candidate factor of NAFLD progression, although additional experiments and collection of clinical data are needed to confirm our result.

In future works, we will try to analyze medical checkup data including binary or discrete variables, such as answers to questionnaires. Since there has been an attempt to extend LiNGAM so that it can deal with discrete variables [69], we will adopt such a method. We will apply the proposed method to other types of EHR in order to identify unknown factors of various diseases. As for clinical data, we will apply the proposed LiNGAM-beta with appropriate feature extraction and formatting, and compare it with other existing methods.

Funding

This work was supported in part by JST PRESTO #JPMJPR1859.

Data availability statement

The health examination data will be made available by the corresponding author to colleagues who propose a reasonable scientific request after approval by the institutional review board of the Japanese Red Cross Kyoto Daini Hospital.

CRediT authorship contribution statement

K. F is with Quadlytics Inc. as well as Nagoya University. M. K is with Quadlytics Inc. as well as Kyoto University. H. K's laboratory is supported by donations from Fukuda Lifetech Co., Ltd., Fukuda Life Tech Keiji Co., Ltd., Tanaka Sleep Clinic, Akita Sleep Clinic, and Ai Ai Care Co., Ltd., made to the Shiga University of Medical Science. Other authors declare that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

References

[1] Pivovarov R, Perotte AJ, Grave E, Angiolillo J, Wiggins CH, Elhadad N. Learning probabilistic phenotypes from heterogeneous EHR data. J Biomed Inform 2015;58:156–65.
[2] Perotte A, Ranganath R, Hirsch JS, Blei D, Elhadad N. Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis. J Am Med Inform Assoc 2015;22(4):872–80.
[3] Jin B, Che C, Liu Z, Zhang S, Yin X, Wei X. Predicting the risk of heart failure with EHR sequential data modeling. IEEE Access 2018;6:9256–61.
[4] Canapari CA, Hoppin AG, Kinane TB, Thomas BJ, Torriani M, Katz ES. Relationship between sleep apnea, fat distribution, and insulin resistance in obese children. J Clin Sleep Med 2011;7(3):268–73.
[5] Nagata T, Hata J, Sakata S, Oishi E, Honda T, Furuta Y, Ohara T, Yoshida D, Hirakawa Y, Shibata M, Ide T, Kitazono T, Tsutsui H, Ninomiya T. Serum N-terminal pro-B-type natriuretic peptide as a predictor for future development of atrial fibrillation in a general population: the Hisayama Study. Int J Cardiol 2020;320:90–6.
[6] Shimizu S, Hoyer PO, Hyvärinen AKA. A linear non-Gaussian acyclic model for causal discovery. J Mach Learn Res 2006;7:2003–30.
[7] Helajärvi H, Rösenstrom T, Pahkala K, Kähönen M, Lehtimäki T, Heinonen OJ, Oikonen M, Tammelin T, Viikari JS, Raitakari OT. Exploring causality between TV viewing and weight change in young and middle-aged adults. The Cardiovascular Risk in Young Finns study. PLoS ONE 2014;9(7):e101860.
[8] Ogawa T, Shimobayashi H, Hirayama J-I, Kawanabe M. Asymmetric directed functional connectivity within the frontoparietal motor network during motor imagery and execution. Neuroimage 2022;247:118794.
[9] Goldberg D, Nichols D, Oki BM, Terry D. Using collaborative filtering to weave an information tapestry. Commun ACM 1992;35(12):61–70.
[10] Seiffert C, et al. RUSBoost: a hybrid approach to alleviating class imbalance. IEEE Trans Syst Man Cybern A Syst Hum 2010;40(1):185–97.
[11] Resnick P, Varian HR. Recommender systems. Commun ACM 1997;40(3):56–8.
[12] Gräßer F, Beckert S, Küster D, Schmitt J, Abraham S, Malberg H, Zaunseder S. Therapy decision support based on recommender system methods. J Healthc Eng 2017;2017:8659460.
[13] Zeng X, Ding N, Rodríguez-Patón A, Zou Q. Probability-based collaborative filtering model for predicting gene–disease associations. BMC Medical Genom. 2017;10(5):45–53. https://doi.org/10.1186/s12920-017-0313-y.
[14] Fujiwara K, Huang Y, Hori K, Nishioji K, Kobayashi M, Kamaguchi M, Kano M. Over- and under-sampling approach for extremely imbalanced and small minority data problem in health record analysis. Front Public Health 2020;8:178.
[15] Lonardo A, Leoni S, Alswat KA, Fouad Y. History of nonalcoholic fatty liver disease. Int J Mol Sci 2020;21(16):5888.
[16] Stefan N, Häring H-U, Cusi K. Non-alcoholic fatty liver disease: causes, diagnosis, cardiometabolic consequences, and treatment strategies. Lancet Diabetes Endocrinol 2019;7(4):313–24.
[17] Diehl AM, Day C. Cause, pathogenesis, and treatment of nonalcoholic steatohepatitis. N Engl J Med 2017;377(21):2063–72.
[18] European Association for the Study of the Liver (EASL), European Association for the Study of Diabetes (EASD), European Association for the Study of Obesity (EASO). EASL–EASD–EASO clinical practice guidelines for the management of non-alcoholic fatty liver disease. J Hepatol 2016;64(6):1388–402.
[19] Rinella ME. Nonalcoholic fatty liver disease: a systematic review. JAMA 2015;313(22):2263–73.
[20] Bedogni G, Tamini S, Caroli D, Cicolini S, Domenicali M, Sartorio A. Development and internal validation of fatty liver prediction models in obese children and adolescents. J Clin Med 2021;10(7):1470.
[21] Deprince A, Haas JT, Staels B. Dysregulated lipid metabolism links NAFLD to cardiovascular disease. Mol Metab 2020;42:101092.
[22] Orr WC, Fass R, Sundaram SS, Scheimann AO. The effect of sleep on gastrointestinal functioning in common digestive diseases. Lancet Gastroenterol Hepatol 2020;5(6):616–24.
[23] Shimizu S, Inazumi T, Sogawa Y, Hyvärinen A, Kawahara Y, Washio T, et al. DirectLiNGAM: a direct method for learning a linear non-Gaussian structural equation model. J Mach Learn Res 2011;12:1225–48.
[24] Hyvärinen A, Smith SM. Pairwise likelihood ratios for estimation of non-Gaussian structural equation models. J Mach Learn Res 2013;14(1):111–52.
[25] Desrosiers C, Karypis G. A comprehensive survey of neighborhood-based recommendation methods. In: Recommender systems handbook. NY: Springer; 2011.
[26] Nishioji K, Sumida Y, Kamaguchi M, Mochizuki N, Kobayashi M, Nishimura T, Yamaguchi K, Itoh Y. Prevalence of and risk factors for non-alcoholic fatty liver disease in a non-obese Japanese population, 2011–2012. J Gastroenterol 2015;50(1):95–108.
[27] Musso G, Cassader M, Olivetti C, Rosina F, Carbone G, Gambino R. Association of obstructive sleep apnoea with the presence and severity of non-alcoholic fatty liver disease.A systematic review and meta-analysis. Obes Rev 2013;14(5):417–31.
[28] Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: a population health perspective. Am J Respir Crit Care Med 2002;165(9):1217–39.
[29] Takegami M, Hayashino Y, Chin K, Sokejima S, Kadotani H, Akashiba T, Kimura H, Ohi M, Fukuhara S. Simple four-variable screening tool for identification of patients with sleep-disordered breathing. Sleep 2009;32(7):939–48.
[30] Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, Sulkowski MS, Torriani FJ, Dieterich DT, Thomas DL, Messinger D, Nelson M. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006;43(6):1317–25.

[31] Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2009;7(10):1104–12.
[32] Sumida Y, Yoneda M, Hyogo H, Itoh Y, Ono M, Fujii H, Eguchi Y, Suzuki Y, Aoki N, Kanemasa K, Fujita K, Chayama K, Saibara T, Kawada N, Fujimoto K, Kohgo Y, Yoshikawa T, Okanoue T. Validation of the FIB4 index in a Japanese nonalcoholic fatty liver disease population. BMC Gastroenterol 2012;12:2.
[33] Kanwal F, Kramer JR, Mapakshi S, Natarajan Y, Chayanupatkul M, Richardson PA, Li L, Desiderio R, Thrift AP, Asch SM, Chu J, El-Serag HB. Risk of hepatocellular cancer in patients with non-alcoholic fatty liver disease. Gastroenterology 2018;155(6):1828–37.
[34] Hassan K, Bhalla V, Regal MEEl, A-Kader HH. Nonalcoholic fatty liver disease: a comprehensive review of a growing epidemic. World J Gastroenterol 2014;20(34):12082–101.
[35] Florentino GS, Cotrim HP, Vilar CP, Florentino AV, Guimaraes GM, Barreto VS. Nonalcoholic fatty liver disease in menopausal women. Arq Gastroenterol 2013;50(3):180–5.
[36] Hug N. Surprise: a Python library for recommender systems. JOpen Source Softw 2020;5(52):2174. https://doi.org/10.21105/joss.02174. URL doi:10.21105/joss.02174.
[38] Andridge RR, Little RJ. A review of hot deck imputation for survey non-response. Int Stat Rev 2010;78(1):40–64.
[39] Hamaguchi M, Kojima T, Takeda N, Nakagawa T, Taniguchi H, Fujii K, Omatsu T, Nakajima T, Sarui H, Shimazaki M, et al. The metabolic syndrome as a predictor of nonalcoholic fatty liver disease. Ann Intern Med 2005;143(10):722–8.
[40] Eguchi Y, Eguchi T, Mizuta T, Ide Y, Yasutake T, Iwakiri R, Hisatomi A, Ozaki I, Yamamoto K, Kitajima Y, Kawaguchi Y, Kuroki S, Ono N. Visceral fat accumulation and insulin resistance are important factors in nonalcoholic fatty liver disease. J Gastroenterol 2006;41(5):462–9.
[41] Eguchi Y, Hyogo H, Ono M, Mizuta T, Ono N, Fujimoto K, et al. Prevalence and associated metabolic factors of nonalcoholic fatty liver disease in the general population from 2009 to 2010 in Japan: a multicenter large retrospective study. J Gastroenterol 2012;47(5):586–95.
[42] Thomsen KL, Grønbæk H, Glavind E, Hebbard L, Jessen N, Clouston A, George J, Vilstrup H. Experimental nonalcoholic steatohepatitis compromises ureagenesis, an essential hepatic metabolic function. Am J Physiol Gastrointest Liver Physiol 2014;307(3):295–301.
[43] Hashemi Kani A, Alavian SM, Esmaillzadeh A, Adibi P, Azadbakht L. Dietary quality indices and biochemical parameters among patients with non alcoholic fatty liver disease (NAFLD). Hepat Mon 2013;13(7):e10943.
[44] Erçin CN, Doğru T, Çelebi G, Gürel H, Genç H, Sertoğlu E, Bağci S. The relationship between blood urea nitrogen levels and metabolic, biochemical, and histopathologic findings of nondiabetic, nonhypertensive patients with nonalcoholic fatty liver disease. Turk J Med Sci 2016;46(4):985–91.
[45] Eriksen PL, Vilstrup H, Rigbolt K, Suppli MP, Sørensen M, Heebøll S, Veidal SS, Knop FK, Thomsen KL. Non-alcoholic fatty liver disease alters expression of genes governing hepatic nitrogen conversion. Liver Int 2019;39(11):2094–101.
[46] De Chiara F, Heebøll S, Marrone G, Montoliu C, Hamilton-Dutoit S, Ferrandez A, Andreola F, Rombouts K, Grønbæk H, Felipo V, Gracia-Sancho J, Mookerjee RP, Vilstrup H, Jalan R, Thomsen KL. Urea cycle dysregulation in non-alcoholic fatty liver disease. J Hepatol 2018;69(4):905–15.
[47] Lonardo A, Loria P, Leonardi F, Borsatti A, Neri P, Pulvirenti M, Verrone AM, Bagni A, Bertolotti M, Ganazzi D, Carulli N, Group PS. Fasting insulin and uric acid levels but not indices of iron metabolism are independent predictors of non-alcoholic fatty liver disease. a case-control study. Dig Liver Dis 2002;34(1):204–11.
[48] Zhou Y, Wei F, Fan Y. High serum uric acid and risk of nonalcoholic fatty liver disease: a systematic review and meta-analysis. Clin Biochem 2016;49(7):636–42.
[49] Jaruvongvanich V, Ahuja W, Wirunsawanya K, Wijarnpreecha K, Ungprasert P. Hyperuricemia is associated with nonalcoholic fatty liver disease activity score in patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol 2017;29(9):1031–5.
[50] Mosca A, Nobili V, De Vito R, Crudele A, Scorletti E, Villani A, et al. Serum uric acid concentrations and fructose consumption are independently associated with NASH in children and adolescents. J Hepatol 2017;66(5):1031–6.
[51] Taketani H, Sumida Y, Tanaka S, Imajo K, Yoneda M, Hyogo H, Ono M, Fujii H, Eguchi Y, Kanemasa K, et al. The association of insomnia with gastroesophageal reflux symptoms in biopsy-proven nonalcoholic fatty liver disease. J Gastroenterol 2014;49(7):1163–74.
[52] Kim CW, Yun KE, Jung HS, Chang Y, Choi ES, Kwon MJ, Lee EH, Woo EJ, Kim NH, Shin H, Ryu S. Sleep duration and quality in relation to non-alcoholic fatty liver disease in middle-aged workers and their spouses. J Hepatol 2013;59(2):351–7.
[53] Musso G, Gambino R, Tabibian JH, Ekstedt M, Kechagias S, Hamaguchi M, Hultcrantz R, Hagström H, Yoon SK, Charatcharoenwitthaya P, George J, Barrera F, Haflioadóttir S, Björnsson ES, Armstrong MJ, Hopkins LJ, Gao X, Francque S, Verrijken A, Yilmaz Y, Lindor KD, Charlton M, Haring R, Lerch MM, Rettig R, Völzke H, Ryu S, Li G, Wong LL, Machado M, Cortez-Pinto H, Yasui K, Cassader M. Association of non-alcoholic fatty liver disease with chronic kidney disease: a systematic review and meta-analysis. PLoS Med 2014;11(7):e1001680.
[54] Balakrishnan M, Patel P, Dunn-Valadez S, Dao C, Khan V, Ali H, El-Serag L, Hernaez R, Sisson A, Thrift AP, Liu Y, El-Serag HB, Kanwal F. Women have a lower risk of nonalcoholic fatty liver disease but a higher risk of progression vs men: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2021;19(1):61–71.e15.
[55] Lonardo A, Suzuki A. Sexual dimorphism of NAFLD in adults. Focus on clinical aspects and implications for practice and translational research. J Clin Med 2020;9(5):1278.
[56] Lonardo A, Nascimbeni F, Ballestri S, Fairweather D, Win S, Than TA, Abdelmalek MF, Suzuki A. Sex differences in nonalcoholic fatty liver disease: state of the art and identification of research gaps. Hepatology 2019;70(4):1457–69.
[57] Whittle AT, Marshall I, Mortimore IL, Wraith PK, Sellar RJ, Douglas NJ. Neck soft tissue and fat distribution: comparison between normal men and women by magnetic resonance imaging. Thorax 1999;54(4):323–8.
[58] O'Donnell CP, Schwartz AR, Smith PL. Upper airway collapsibility: the importance of gender and adiposity. Am J Respir Crit Care Med 2000;162(5):1606–7.
[59] Mohsenin V. Gender differences in the expression of sleep-disordered breathing: role of upper airway dimensions. Chest 2001;120(5):1442–7.
[60] Westreich R, Gozlan-Talmor A, et al. The presence of snoring as well as its intensity is underreported by women. J Clin Sleep Med 2019;15(3):471–6.
[61] Mejdell T, Skogestad S. Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression. Ind Eng Chem Res 1991;30(12):2543–55.
[62] Kresta JV, Marlin TE, MacGregor JF. Development of inferential process models using PLS. Comput Chem Eng 1994;18(7):597–611.
[63] Nakahara T, Hyogo H, Yoneda M, Sumida Y, Eguchi Y, Fujii H, Ono M, Kawaguchi T, Imajo K, Aikata H, Tanaka S, Kanemasa K, Fujimoto K, Anzai K, Saibara T, Sata M, Nakajima A, Itoh Y, Chayama K, Okanoue T, Japan Study Group of Nonalcoholic Fatty Liver Disease (JSG-NAFLD). Type 2 diabetes mellitus is associated with the fibrosis severity in patients with nonalcoholic fatty liver disease in a large retrospective cohort of Japanese patients. J Gastroenterol 2014;49(11):1477–84.
[64] Sun DQ, Liu WY, Wu SJ, Zhu GQ, Braddock M, Zhang DC, Shi KQ, Song D, Zheng MH. Increased levels of low-density lipoprotein cholesterol within the normal range as a risk factor for nonalcoholic fatty liver disease. Oncotarget 2016;7(5):5728–37.
[65] Min HK, Kapoor A, Fuchs M, Mirshahi F, Zhou H, Maher J, Kellum J, Warnick R, Contos MJ, Sanyal AJ. Increased hepatic synthesis and dysregulation of cholesterol metabolism is associated with the severity of nonalcoholic fatty liver disease. Cell Metab 2012;15(5):665–74.
[66] Rafiq N, Bai C, Fang Y, Srishord M, McCullough A, Gramlich T, Younossi ZM. Long-term follow-up of patients with nonalcoholic fatty liver. Clin Gastroenterol Hepatol 2009;7(2):234–8.
[67] Ho TK. The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell 1998;20(8):832–44.
[68] Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
[69] Inazumi T, Washio T, Shimizu S, Suzuki J, Yamamoto A, Kawahara Y. Discovering causal structures in binary exclusive-or skew acyclic models. 2012. arXiv:1202.3736.