

Article

Another Look at Obesity Paradox in Acute Ischemic Stroke: Association Rule Mining

Pum-Jun Kim ¹, Chulho Kim ^{1,2,*}, Sang-Hwa Lee ^{1,2}, Jong-Hee Shon ^{1,2}, Youngsuk Kwon ^{1,3}, Jong-Ho Kim ^{1,3}, Dong-Kyu Kim ^{1,4}, Hyunjae Yu ^{1,5}, Hyo-Jeong Ahn ^{1,5}, Jin-Pyeong Jeon ^{1,6}, Youngmi Kim ¹ and Jae-Jun Lee ^{1,3}

- ¹ Institute of New Frontier Research Team, Hallym University College of Medicine, Chuncheon 24252, Korea; pumjun4093@gmail.com (P.-J.K.); bleulsh@naver.com (S.-H.L.); deepfoci@hallym.or.kr (J.-H.S.); gettys@hallym.or.kr (Y.K.); poik99@hallym.or.kr (J.-H.K.); doctordk@naver.com (D.-K.K.); yunow@hallym.or.kr (H.Y.); 21hobs@gmail.com (H.-J.A.); jjs6553@daum.net (J.-P.J.); kym8389@hanmail.net (Y.K.); iloveu59@hallym.or.kr (J.-J.L.)
- ² Department of Neurology, Chuncheon Sacred Heart Hospital, Chuncheon 24253, Korea
- ³ Department of Anesthesiology and Pain Medicine, Chuncheon Sacred Heart Hospital, Chuncheon 24253, Korea
- ⁴ Department of Otorhinolaryngology-Head and Neck Surgery, Chuncheon Sacred Heart Hospital, Chuncheon 24253, Korea
- ⁵ Chuncheon Artificial Intelligence Center, Chuncheon Sacred Heart Hospital, Chuncheon 24253, Korea
- ⁶ Department of Neurosurgery, Chuncheon Sacred Heart Hospital, Chuncheon 24253, Korea
- * Correspondence: gomdol52@hallym.or.kr; Tel.: +82-33-240-5255; Fax: +82-33-255-6244



Citation: Kim, P.-J.; Kim, C.; Lee, S.-H.; Shon, J.-H.; Kwon, Y.; Kim, J.-H.; Kim, D.-K.; Yu, H.; Ahn, H.-J.; Jeon, J.-P.; et al. Another Look at Obesity Paradox in Acute Ischemic Stroke: Association Rule Mining. *J. Pers. Med.* **2022**, *12*, 16. <https://doi.org/10.3390/jpm12010016>

Academic Editor:
Jasminka Ilich-Ernst

Received: 19 October 2021
Accepted: 14 December 2021
Published: 29 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Though obesity is generally associated with the development of cardiovascular disease (CVD) risk factors, previous reports have also reported that obesity has a beneficial effect on CVD outcomes. We aimed to verify the existing obesity paradox through binary logistic regression (BLR) and clarify the paradox via association rule mining (ARM). Patients with acute ischemic stroke (AIS) were assessed for their 3-month functional outcome using the modified Rankin Scale (mRS) score. Predictors for poor outcome (mRS 3–6) were analyzed through BLR, and ARM was performed to find out which combination of risk factors was concurrently associated with good outcomes using maximal support, confidence, and lift values. Among 2580 patients with AIS, being obese (OR [odds ratio], 0.78; 95% CI, 0.62–0.99) had beneficial effects on the outcome at 3 months in BLR analysis. In addition, the ARM algorithm showed obese patients with good outcomes were also associated with an age less than 55 years and mild stroke severity. While BLR analysis showed a beneficial effect of obesity on stroke outcome, in ARM analysis, obese patients had a relatively good combination of risk factor profiles compared to normal BMI patients. These results may partially explain the obesity paradox phenomenon in AIS patients.

Keywords: infarction; risk factors in epidemiology; outcome research; association rule mining; body mass index

1. Introduction

Stroke is a leading cause of death and disability globally [1]. Obesity is a known risk factor for cardiovascular disease (CVD) [2]. Reports also suggest that obesity may be protective for CVD severity or outcome [3–5]. Though most population-based cohort studies have shown that obesity increases the risk of stroke [6–9], obesity has been associated with better outcomes among patients with acute ischemic stroke (AIS) [10–16]. Khan et al. evaluated associations between BMI and lifetime CVD risk and mortality among a population-based cohort without established CVD at baseline [17]. Observing more than 3 million person-years, they found that being overweight was associated with earlier development of CVD, and obesity was associated with reduced longevity and cardiovascular survival. In addition, obesity increased the risk of hypertension, diabetes, and dyslipidemia.

There is still insufficient explanation to determine whether the relationship between obesity and CVD outcome analyzed using logistic regression is beneficial or not.

The association rule mining (ARM) algorithm was introduced for market basket analysis by Agrawal et al. and has identified significant association patterns within a variety of settings [18–23]. ARM is used for explanatory data visualization to summarize certain concurrent combinations of dependent factors (important association rules) with a specific condition in a non-hierarchical fashion [19,24]. Moreover, ARM helps to discover important relationships among complex phenomena [18,19]. Therefore, we aimed to verify the obesity paradox with logistic regression and provide a new perspective through the ARM method on the relationship between obesity and ischemic stroke outcome.

2. Materials and Methods

2.1. Participants

We obtained data from the prospective stroke registry of our hospital, the detailed information of which has been described previously [25]. Briefly, patients were admitted to the hospital within 7 days after the onset of focal neurologic deficits (2012–2019). Patients included in the study were identified as having relevant acute lesions on diffusion MRI. Patients with cerebral hemorrhages were excluded from the study. During the data capture period, procedures affecting the prognosis of patients, such as extending the time window of intravenous thrombolysis and endovascular treatment, were reflected in all included patients (Figure S1). We collected demographic data and accessed data for clinical and laboratory findings for patients admitted with AIS using a standardized web server, which included central coordinator requests for regular audits and amendments of the data.

2.1.1. Measurement of Body Mass Index

Height and weight were recorded by stroke nurses immediately after hospitalization for stroke. BMI was calculated as weight in kilograms divided by the square of height in meters (kg/m^2). Patients were grouped as underweight ($<18.5 \text{ kg}/\text{m}^2$), normal weight ($18.5\text{--}22.9 \text{ kg}/\text{m}^2$), overweight ($23\text{--}24.9 \text{ kg}/\text{m}^2$), and obese ($\geq 25 \text{ kg}/\text{m}^2$) according to the Asian Pacific World Health Organization criteria [26].

2.1.2. Covariates

Vascular risk factors were defined based on our previous report [25]. ARM generally assesses possible associations between discrete variables, and the algorithm can be easily applied to categorical variables. Therefore, we used the following categorical variables for the elucidation of risk profiles and the underlying mechanisms of stroke:

- Hypertension was defined as blood pressure $\geq 140/90$ mmHg in more than two consecutive readings or taking antihypertensive agents;
- Diabetes was defined as fasting blood glucose ≥ 126 mg/dL, random blood glucose readings ≥ 200 mg/dL with relevant diabetic symptoms or glycated hemoglobin $\geq 6.5\%$ [27];
- Hyperlipidemia was defined as total cholesterol ≥ 240 mg/dL or taking lipid-lowering agents [27];
- Current smoking was defined as smoking within 6 months prior to the study;
- Stroke subtypes were defined as cardioembolism and non-cardioembolism;
- Stroke severity was categorized as a score of the National Institute of Health Stroke Scale score (mild: 0–5, moderate: 6–14, and severe: >14) [28];
- Thrombolysis was defined as patients receiving intravenous or intra-arterial thrombolytic agents, or mechanical thrombectomy;
- Patient's Age (years) was categorized as years < 54 , $55 \leq$ years, < 65 , and years ≥ 65 [28].

2.1.3. Study Outcomes

Stroke physicians or certified nurses prospectively assessed patients' modified Rankin Scale (mRS) score as a 3-month functional outcome measure when the patients visited the outpatient clinic or via a telephonic interview. The poor functional outcome was defined as an mRS score of 3–6.

2.2. Statistical Analyses

We compared the baseline differences for independent variables between good and poor outcome groups using the χ^2 test or Student's *t*-test, as appropriate. Univariate and multivariate binary logistic analyses were performed to assess predictors for poor outcomes at 3 months. Independent variables with a two-sided *p*-value < 0.05 in univariate analyses were assessed by multivariate analyses.

2.3. Association Rule Mining

The frequent pattern growth (fp-growth) is one of the ARM algorithms used to evaluate associations for each variable, and to find important linkage patterns using the quantitative parameters of support, confidence, and lift [29]. The definition and formula of these quantitative parameters are stated in Table 1. To understand these parameters, let *X* and *Y* as events in the real world, then *p*(*X*) and *p*(*Y*) are the probabilities that these events will occur. Here, the sup (*X*→*Y*) means the probability that *X* and *Y* occur at the same time. The conf (*X*→*Y*) is a conditional probability such that the likelihood of *Y* occurs given event *X*. The last parameter lift (*X*→*Y*) is defined as a weighted probability of confidence, *p*(*Y* | *X*)/*p*(*Y*), to verify whether event *X* and *Y* are mutually dependent (when the lift is equal to one, then *X* and *Y* are independent).

Table 1. Definition of formula and explanation of support, confidence, and lift.

	Formula	Definition & Meaning
Support	$\text{sup}(X \rightarrow Y) = p(X \cap Y)$	The value of support means how frequent this rule is appearing in the data.
Confidence	$\text{conf}(X \rightarrow Y) = p(Y X)$	The confidence indicates how much the rule is accurate.
Lift	$\text{lift}(X \rightarrow Y) = p(Y X)/p(Y)$	The lift measures the dependency between the predictor and the response. The value of lift close to zero indicates independence.

Let *X* as a subset of predictors and *Y* as a response. *p* for probability of an association.

The fp-growth algorithm is emphasized by its efficiency of a reduced number of calculations compared to the apriori algorithm, due to its tree-like structure [29]. For illustration of this algorithm, let *T* denote set of transactions {*t*₁, *t*₂, . . . , *t*_{*n*}} and list of all items, *I* = {*i*₁, *i*₂, . . . , *i*_{*m*}} (Table 2). First, reorder items in each *t*_{*j*} according to the support (*i*_{*k*}) in descending order for all *k* in [1, *m*]. In these ordered transactions, the less frequent items which are smaller than user-defined minimum support are excluded for fewer calculations later. As a second step to grow an fp-tree, position and connect the items in the initial transaction, *t*₁, starting with a root node (null node), as in Figure 1. Each time, when items are added to the tree, increment the count of these item nodes by one. Third, for the following transaction, if the part of the sequence from the first item overlaps with one of the antecedent transactions, then the overlapping path of these items has to be merged, and the rest of the items are connected after the merged node. However, if there is no identical order of prefix, the new transaction follows the second step by connecting a sequence of items to a null node. Last, since the unit item is also a candidate for being a frequent pattern, the dashed lines in the fp-tree are drawn to denote the same unit items in the different branches of the tree. With a fully-grown tree by those steps, the candidates of frequent patterns are considered through minimum support and confidence. For example, in Figure 1, when the minimum support τ equals two, the sequence {null, *i*₁, *i*₃} (equality with {*i*₁, *i*₃}) is the only

candidate satisfying the lower bound, τ , which means this sequence is a frequent pattern (rule) of the transaction data. Other candidate patterns through the fp-growth algorithm can also be considered through less conservative minimum supports.

Table 2. Example of ordinal and reordered transaction table.

Original Items	Reordered Frequent Items
$t_1 = \{i_1, i_3, i_5, i_6, i_8\}$	$t_1 = \{i_1: 3, i_3: 2, i_6: 2, i_8: 2\}$
$t_2 = \{i_1, i_7, i_9\}$	$t_1 = \{i_1: 3, i_7: 2\}$
$t_3 = \{i_6, i_7, i_8\}$	$t_1 = \{i_6: 2, i_7: 2, i_8: 2\}$
$t_4 = \{i_1, i_2, i_3, i_4\}$	$t_1 = \{i_1: 3, i_3: 2\}$

: denotes the frequency of items in transaction table. t, transactions; i, item.

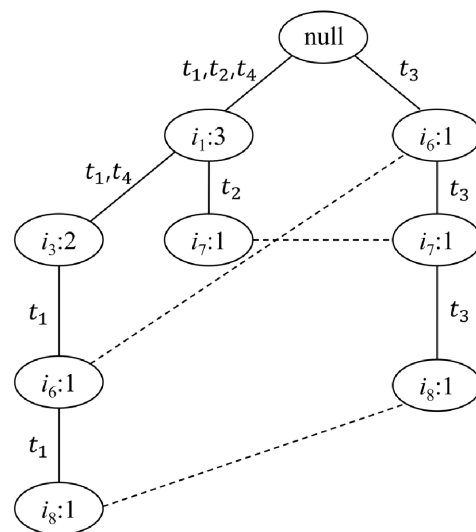


Figure 1. Schematic representation of the frequent pattern growth algorithm with an illustration of fully grown frequent pattern tree using transaction data. t, transactions; i, item.

In the resulting rules in this paper, pruning step, a rule XY is redundant if $\exists X^* \subset X$ and $\text{confidence}(X^* \rightarrow Y) \geq \text{confidence}(X \rightarrow Y)$, is applied to drive a more generalized frequent pattern without redundant rules [30]. Since the goal of ARM is to find the rule of associations between categorical variables, the age variable was grouped as age < 55, 55 ≤, age < 65, and age ≥ 65 in our analysis. The moonBook (version 0.2.3), arules (version 1.6.4), and rCBA (version 0.4.3) packages for R software [R Foundation for Statistical Computing (version 3.6.3)] were used to perform the binary logistic regression and ARM algorithm. The exact R code for ARM using the fp-growth algorithm was shown in Supplemental Data.

3. Results

In total, 2580 AIS patients (mean age 68.1 ± 13.0 years, 59.1% males) were included in this study (Table 3). The proportions of underweight, normal weight, overweight, and obese patients were 4.6%, 34.8%, 26.4%, and 34.1%, respectively. The proportions of overweight and obese patients were higher in the good outcome group compared to the poor outcome group (p for χ^2 trend <0.001). Participants with a poor outcome were more likely to be older, have severe stroke severity, to present with hypertension and diabetes compared to those with a good outcome. Current smoking was higher in patients with a good outcome than among those with a poor outcome.

Table 3. Comparison of baseline characteristics of the participants.

	Good Outcome (N = 1671)	Poor Outcome (N = 909)	p
Age, years			<0.001
18–54	441 (24.6%)	120 (13.2%)	
55–64	401 (24.0%)	70 (7.8%)	
≥65	859 (51.4%)	719 (79.1%)	
BMI, kg/m ²			<0.001
Underweight (<18.5)	47 (2.8%)	71 (7.8%)	
Normal weight (18.5–22.9)	540 (32.3%)	359 (39.5%)	
Overweight (23.0–24.9)	463 (27.7%)	220 (24.2%)	
Obese (≥25)	621 (37.2%)	259 (28.5%)	
Stroke Severity, NIHSS			<0.001
Mild (0–5)	1534 (91.8%)	479 (52.7%)	
Moderate (6–14)	100 (6.0%)	226 (24.9%)	
Severe (>14)	37 (2.2%)	204 (22.4%)	
Men	1068 (63.9%)	457 (50.3%)	<0.001
Hypertension	947 (56.7%)	624 (68.6%)	<0.001
Diabetes	446 (26.7%)	319 (35.1%)	<0.001
Current Smoking	578 (34.6%)	184 (20.2%)	<0.001
Cardioembolism	292 (17.5%)	245 (27.0%)	<0.001
Thrombolysis	138 (8.3%)	135 (14.9%)	<0.001
Hyperlipidemia	293 (17.5%)	156 (17.2%)	0.854

BMI, body mass index; NIHSS, National Institute of Health Stroke Scale.

Table 4 shows differences in clinical characteristics for patients according to BMI. We found that age was inversely associated with BMI and that male sex was more prevalent in the overweight group than in the other BMI groups. Hyperlipidemia and current smoking were less frequent, and cardioembolism was more frequent in the underweight group than in the normal weight, overweight, and obese groups.

Table 4. Differences in clinical characteristics of the patients stratified by body mass index.

	Underweight (N = 118)	Normal Weight (N = 899)	Overweight (N = 683)	Obese (N = 880)	p
Age, years					<0.001 [†]
18–54 years	13 (11.0%)	142 (15.8%)	110 (16.1%)	206 (23.4%)	
55–64 years	12 (10.2%)	162 (18.0%)	169 (24.7%)	188 (21.4%)	
≥65 years	93 (78.8%)	595 (66.2%)	404 (59.2%)	486 (55.2%)	
Stroke Severity, NIHSS					<0.001
Mild (0–5)	70 (59.3%)	658 (73.2%)	550 (80.5%)	735 (83.5%)	
Moderate (6–14)	25 (21.2%)	131 (14.6%)	76 (11.1%)	94 (10.7%)	
Severe (>14)	23 (19.5%)	110 (12.2%)	57 (8.3%)	51 (5.8%)	
Stroke Mechanism					<0.001
CE	39 (33.1%)	220 (24.5%)	130 (19.0%)	148 (16.8%)	
Non-CE	79 (66.9%)	679 (75.5%)	553 (81.0%)	732 (83.2%)	
Outcome at 3 months					<0.001
Good (mRS score 0–2)	47 (39.8%)	540 (60.1%)	463 (67.8%)	621 (70.6%)	
Poor (mRS score 3–6)	71 (60.2%)	359 (39.9%)	220 (32.2%)	259 (29.4%)	
Men	49 (41.5%)	495 (55.1%)	432 (63.3%)	549 (62.4%)	<0.001
Hypertension	67 (56.8%)	492 (54.7%)	200 (29.3%)	281 (31.9%)	<0.001
Hyperlipidemia	9 (7.6%)	146 (16.2%)	113 (16.5%)	181 (20.6%)	0.002
Current Smoking	19 (16.1%)	268 (29.8%)	201 (29.4%)	274 (31.1%)	0.010
Diabetes	37 (31.4%)	247 (27.5%)	200 (29.3%)	281 (31.9%)	0.217
Thrombolysis	12 (10.2%)	101 (11.2%)	70 (10.2%)	90 (10.2%)	0.891

[†] p for analysis of variances. Statistical significance of cate variables is represented with the p-values of the χ^2 test. NIHSS, National Institute of Health Stroke Scale; CE, cardioembolic; mRS, modified Rankin Scale score.

Table 5 shows the results of BLR analyses for predictors of a poor outcome at 3 months. Being underweight was a significant predictor for poor outcomes in multivariate analysis. However, obesity had a beneficial effect on the outcome at 3 months after adjusting for multiple covariates. Moreover, addition logistic regression analysis considering age*BMI interaction and entering age and BMI as a continuous variable were not different from the original BLR analysis (Tables S1 and S2).

Table 5. Results of binary logistic regression analysis for predictors of poor functional outcome at 3 months in patients with acute ischemic stroke.

	Univariate OR (95% CI)	<i>p</i>	Multivariate OR (95% CI)	<i>p</i>
Age, years				
18–54	0.29 (0.24–0.36)	0.001	0.63 (0.44–0.91)	0.013
55–64	1.00 (reference)	-	1.00 (reference)	-
≥65	2.87 (2.29–3.59)	<0.001	2.26 (1.74–2.93)	<0.001
BMI, kg/m ²				
Underweight (<18.5)	2.27 (1.54–3.36)	<0.001	1.69 (1.07–2.66)	0.024
Normal weight (18.5–22.9)	1.00 (reference)	-	1.00 (reference)	-
Overweight (23.0–24.9)	0.71 (0.58–0.88)	0.001	0.82 (0.64–1.05)	0.119
Obese (≥25)	0.63 (0.52–0.76)	<0.001	0.78 (0.62–0.99)	0.041
Stroke Severity, NIHSS				
Mild (0–5)	0.14 (0.11–0.18)	<0.001	0.11 (0.08–0.15)	<0.001
Moderate (6–14)	1.00 (reference)	-	1.00 (reference)	-
Severe (>14)	2.44 (1.60–3.72)	<0.001	2.37 (1.51–3.70)	<0.001
Diabetes	1.49 (1.25–1.77)	<0.001	1.61 (1.30–1.99)	<0.001
Current Smoking	0.48 (0.40–0.58)	<0.001	0.67 (0.52–0.86)	0.001
Thrombolysis	1.94 (1.51–2.49)	<0.001	0.64 (0.45–0.91)	0.012
Cardioembolism	1.74 (1.44–2.11)	<0.001	0.75 (0.58–0.96)	0.024
Hypertension	1.67 (1.41–1.98)	<0.001	1.14 (0.92–1.41)	0.228
Hyperlipidemia	0.97 (0.79–1.21)	0.811	0.90 (0.70–1.17)	0.444
Men	0.57 (0.48–0.67)	<0.001	0.92 (0.75–1.14)	0.854

OR, odds ratio; CI, confidence interval; BMI, body mass index; NIHSS, National Institute of Health Stroke Scale.

For fp-growth algorithms, we set minimum support bound of 0.04 and a confidence bound of 0.8, according to the average of all possible association rules that could be generated in our dataset. With this parameter setting, 329 rules were generated in total. After a pruning procedure for reducing redundant association rules, five association rules satisfied the support and confidence limits (Table 6). In the top five rules with high confidence & lift values, obesity, age less than 55 years, and mild stroke severity were presented to have a relationship with a good outcome. Figure 2A depicts the relationship of these association rules according to support, confidence, and lift values. Obese was associated with a good outcome, in which no diabetes, mild stroke severity, no hyperlipidemia, male, smoking, and younger age (<55 years) were concurrently observed. Figure 2B shows the coordination plot of significant association rules in our fp-growth algorithm. The rule of being obese, younger, and presenting with no diabetes, no hyperlipidemia, and mild stroke severity was associated with a good outcome with the highest lift value in our rules. Table S3 shows the result of the five most frequent association rules in each BMI subgroup. Underweight patients with severe stroke severity and older age were especially associated with poor outcomes.

Table 6. Results of the frequent pattern growth algorithm of independent and dependent variables in patients with acute ischemic stroke.

	LHS	RHS	Support	Confidence	Lift	Count
1	{Hyperlipidemia = No, Diabetes = No, Severity = mild, Age = 18–54 years, BMI = Obese}	Good	0.0488	0.9767	1.5080	126
2	{Severity = mild, Age = 18–54 years, Smoking = Yes, BMI = Obese}	Good	0.0434	0.9739	1.5037	112
3	{Diabetes = No, Severity = mild, Age = 18–54 years, BMI = Obese}	Good	0.0577	0.9738	1.5036	149
4	{Hyperlipidemia = No, Severity = mild, Gender = Male, Age = 18–54 years, BMI = Obese}	Good	0.0472	0.9682	1.4949	122
5	{Hyperlipidemia = No, Severity = mild, Age = 18–54 years, BMI = Obese}	Good	0.0573	0.9673	1.4935	148

LHS, left hand side; RHS, right hand side; BMI, body mass index; CE, cardioembolism.

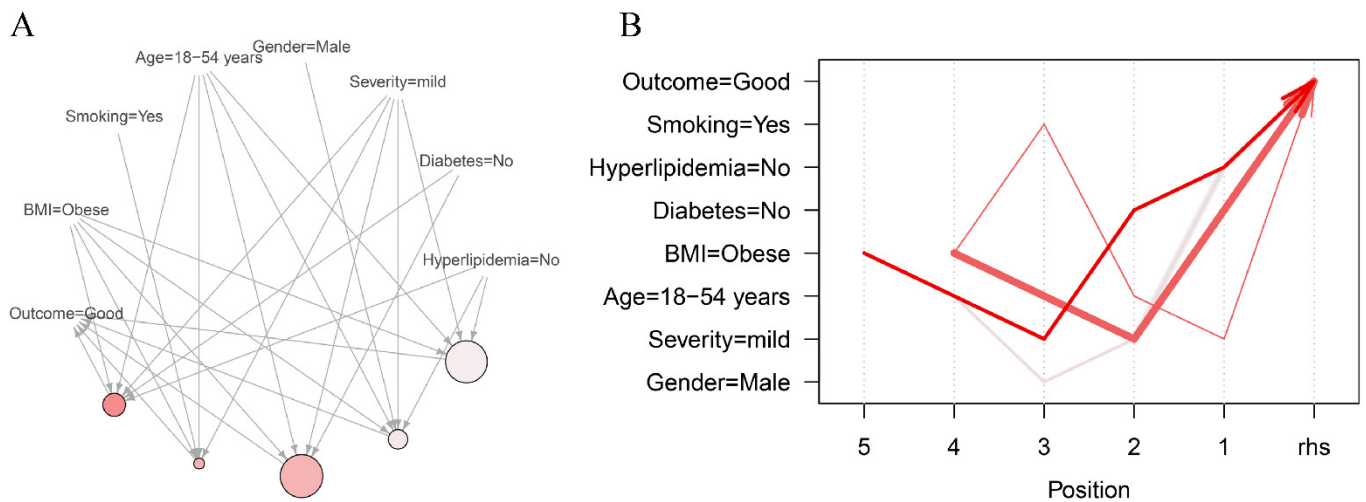


Figure 2. Visualization of significant association rules in patients with acute ischemic stroke, via fp-growth algorithm with minimum support (0.04) and confidence (0.8). (A) Network of significant association rules. (B) Parallel coordination plot for significant association rules. In the Figure 2A,B, thickness and redness of the lines means support and confidence of the association rules. BMI, body mass index; CE, cardioembolism; rhs, right hand side.

4. Discussion

In the present study, logistic regression analyses in the established ischemic stroke population revealed that being obese was associated with a good outcome. However, obese patients had a lower age, male gender, and were less likely to present with cardioembolism. In other words, the obesity paradox could be partially explained by the fact that obese patients may be younger than patients with normal BMI or may be more likely to have a stroke that is associated with a non-cardioembolism mechanism.

Welton et al. conducted a prospective investigation of 8528 patients with diabetes and reported that mortality among obese patients was less than among patients with a normal BMI [31]. However, this tendency decreased when the fitness level was high in obese patients, and the authors suggested that the fitness level measured by the metabolic equivalent of the task may be a moderator of the association between obesity and mortality. Meanwhile, age was different for each BMI group in the present study, and notably, the mean age was lowest in the obese group (BMI > 30 kg/m²). Bhaskaran et al. analyzed the data of 3.6 million participants from the UK general population for associations between BMI and mortality using prospective survival analysis, and they suggested that mortality was higher in subjects with low (<18.5 kg/m²) or high BMI (≥25 kg/m²) [32]. The average age was 37 years in their study, and the overweight or obese group was found to be 10 years older than the normal BMI group. Thus, age can be an important factor for analyzing obesity and stroke outcomes. Khan et al. analyzed the relationship between obesity and CVD using individual patient data from ten population-based cohorts [17]. The increase in BMI at cohort enrollment accelerated CVD onset, which eventually increased the longevity of patients with CVD. In addition, high BMI increased CVD mortality compared to non-CVD death. This study followed 19,000 patients for over 50 years overall and was stratified according to age, sex, and BMI status. This approach may have minimized the selection bias, such as lead-time bias or survivorship bias.

After a report regarding the obesity paradox in stroke patients, there has not been a clear explanation for the relationship between positive stroke outcomes and high BMI status [10]. Results of prospective cohort studies within the general population consistently show that high BMI increases stroke risk [6–9]. However, the contradicting results found within the obesity paradox for stroke have been reported in studies with established stroke cohorts [10–16,33–37]. Several studies have also reported that the strength of correlation between high BMI and low stroke mortality was attenuated by age [10,13]. An additional

consideration is that those with a high BMI group in the established stroke population are consistently reported to be about 3–7 years younger than those with normal BMI in these observational stroke cohorts. According to the Danish Stroke Register report, higher BMI is associated with accelerated ischemic stroke occurrence [38]. Andersen et al. reported that obesity was associated with lower mortality and risk of readmission for recurrent stroke than the normal weight group after applying Cox proportional hazard models [14]. However, concerns regarding the proportional hazard assumption and age discrepancies between each BMI group have not been discussed in detail. Results from the Atherosclerosis Risk in Communities study, which adjusted for time-varying covariates using the G-estimation method, reported that the association between stroke outcome and BMI status could vary due to mishandling time-varying confounders [9]. Hence, age is the most important variable for explaining the relationship between BMI and stroke outcome. In addition, from the results of the logistic regression model, it can be suggested that high BMI is related to a good outcome, but residual confounding factors related to age should be considered when interpreting these associations.

The ARM algorithm was initially designed to determine the specific purchasing patterns of people in the market and to analyze customers' behavior [18]. Since then, studies using ARM to find valuable linkages among various factors have been reported in the medical field. Szalkai et al. analyzed factors affecting cognitive decline below 15 points in the Mini-Mental Status Examination score within more than 5000 Alzheimer disease databases [39]. Factors such as high aspartate aminotransferase and high serum sodium were identified as significant rules showing a strong association with cognitive decline. In addition, ARM can specify which non-hierarchical phenotype has been achieved by effective clustering of multiple diseases [40–42]. However, ARM usually deals with Boolean data, so it is difficult to analyze associations among the numerical data. Further, the antecedents and consequents in an ARM analysis are not a way to depict a potential causal relationship because variables with high proportions are usually applied as antecedents.

Logistic regression analysis represents the individual risk as an odds ratio; hence, it focuses on individual risk [43]. However, ARM evaluates the entire dataset on a microscopic level and can find important patterns in the groups of interest, though the effect sizes are small [44]. To be specific, when logistic regression and ARM analyze the data, these methods can reach a similar result. However, when it comes to the interpretation of the associations between the predictors and response, logistic regression predicts a probability of dependent variable using all the independent variables and significance of a single predictor. ARM methodology finds association rules between the subset of predictors and response with its frequency and significance. As in the Supplementary Table S2, for clarification of the result, an interaction term among the age and BMI is considered in the logistic regression to account for an additional effect between predictors on a dependent variable, but it presented no significant relationship with the response. In addition, ARM does not include prior hypotheses for statistical testing and needs to consider interactions among the variables of interest [45]. In other words, ARM is complementary to the conventional logistic regression model. Therefore, if we implement ARM in addition to logistic regression analysis in evaluating the relationship between disease and risk factors, we can make a more explanatory hypothesis by checking the macroscopic and microscopic associations.

The present study used the data of a representative stroke population. However, our study has a few limitations. First, the results did not include numerical laboratory data such as low-density lipoprotein or glucose levels because of the innate disadvantages of the ARM algorithm. Second, a single-center, retrospective observation is prone to have a selection bias, so the results may not be generalizable to the entire stroke population. Third, we used BMI values at the time of admission for evaluating the degree of obesity; however, BMI is well-known to be a time-varying covariate. Therefore, our study does not provide results on how changes in BMI after stroke affect the outcome at 3 months. Finally, we used Asian Pacific World Health Organization criteria for the obesity categorization. Because BMI depends on age, gender, and ethnicity, it is difficult to generalize our findings

to all ischemic stroke patients. Therefore, for an international comparison of the impact of BMI on the prognosis of stroke, we should consider which criteria were used for the BMI categorization. To accurately understand the impact of BMI on CVD in AIS, we need later prospective studies considering age stratification and temporal variation of BMI in the research design and analytic phase.

5. Conclusions

Our findings using logistic regression analysis suggest that obesity is associated with a good outcome after stroke. However, the results of ARM analysis revealed that being obese was associated with good outcomes by way of younger age at the onset and mild stroke severity. This suggests that the good outcome in AIS patients was not because the patients were obese, but rather because those patients were younger and their stroke severity was less severe than those with a normal BMI. Thus, the ARM algorithm can be used to find novel and valuable linkages among risk factors and outcomes in the medical research field.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/jpm12010016/s1>, Figure S1: Flow chart of the participants; Table S1: Results of binary logistic regression analysis for predictors of poor functional outcome at 3 months in patients with acute ischemic stroke, considering the variable age and BMI as continuous values; Table S2: Results of binary logistic regression analysis for predictors, considering the interaction between the stratified age and BMI, of poor functional outcome at 3 months in patients with acute ischemic stroke Table S3: Results of the frequent pattern growth algorithms of four different datasets (BMI: Obese, Overweight, Normal, Underweight) with independent and dependent variables in patients having an acute ischemic stroke.

Author Contributions: Conceptualization, P.-J.K., C.K. and J.-J.L.; Data curation, S.-H.L., J.-H.S. and Y.K. (Youngmi Kim); Investigation, J.-H.S., D.-K.K., J.-P.J. and Y.K. (Youngsuk Kwon); Software, J.-H.K., H.Y. and H.-J.A.; Supervision, C.K. and J.-J.L.; Validation, Y.K. (Youngmi Kim), J.-H.K. and H.-J.A.; Visualization, J.-P.J.; Writing—original draft, P.-J.K. and C.K.; Writing—review & editing, P.-J.K. and Y.K. (Youngsuk Kwon). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Research Fund of Korea (NRF-2019R1G1A1097707, website: <https://www.nrf.re.kr/index/> 12 December 2021), by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by Ministry of Health & Welfare, Republic of Korea (grant number: HR21C0198), and carried out with the support of R&D Program for Forest Science Technology (Project No. 2021397C10-2123-0107, 2021397B10-2123-0107) provided by Korea Forest Service (Korea Forestry Promotion Institute). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Institutional Review Board Statement: The study design was approved by the Institutional Review Board/Ethics Committee of the Chuncheon Sacred Heart Hospital (No. 2020-02-014).

Informed Consent Statement: Participants or their guardians gave written informed consent for the enrollment in prospective registry and data acquisition for 3 months' outcome.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the policy of our IRB.

Conflicts of Interest: The authors declare that they have no competing interest.

References

1. GBD 2016 Lifetime Risk of Stroke Collaborators. Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. *N. Engl. J. Med.* **2018**, *379*, 2429–2437. [[CrossRef](#)]
2. Hubert, H.B.; Feinleib, M.; McNamara, P.M.; Castelli, W.P. Obesity as an independent risk factor for cardiovascular disease: A 26-year follow-up of participants in the Framingham Heart Study. *Circulation* **1983**, *67*, 968–977. [[CrossRef](#)]
3. Lavie, C.J.; Milani, R.V.; Ventura, H.O. Obesity and cardiovascular disease: Risk factor, paradox, and impact of weight loss. *J. Am. Coll. Cardiol.* **2009**, *53*, 1925–1932. [[CrossRef](#)]

4. Pagidipati, N.J.; Zheng, Y.; Green, J.B.; McGuire, D.K.; Mentz, R.J.; Shah, S.; Aschner, P.; Delibasi, T.; Rodbard, H.W.; Westerhout, C.M. Association of obesity with cardiovascular outcomes in patients with type 2 diabetes and cardiovascular disease: Insights from TECOS. *Am. Heart J.* **2020**, *219*, 47–57. [[CrossRef](#)]
5. Clark, A.L.; Fonarow, G.C.; Horwich, T.B. Obesity and the obesity paradox in heart failure. *Prog. Cardiovasc. Dis.* **2014**, *56*, 409–414. [[CrossRef](#)] [[PubMed](#)]
6. Kurth, T.; Gaziano, J.M.; Berger, K.; Kase, C.S.; Rexrode, K.M.; Cook, N.R.; Buring, J.E.; Manson, J.E. Body mass index and the risk of stroke in men. *Arch. Int. Med.* **2002**, *162*, 2557–2562. [[CrossRef](#)]
7. Suk, S.-H.; Sacco, R.L.; Boden-Albala, B.; Cheun, J.F.; Pittman, J.G.; Elkind, M.S.; Paik, M.C. Abdominal obesity and risk of ischemic stroke: The Northern Manhattan Stroke Study. *Stroke* **2003**, *34*, 1586–1592. [[CrossRef](#)]
8. Chen, H.-J.; Bai, C.-H.; Yeh, W.-T.; Chiu, H.-C.; Pan, W.-H. Influence of metabolic syndrome and general obesity on the risk of ischemic stroke. *Stroke* **2006**, *37*, 1060–1064. [[CrossRef](#)]
9. Shakiba, M.; Mansournia, M.A.; Kaufman, J.S. Estimating Effect of Obesity on Stroke Using G-Estimation: The ARIC study. *Obesity* **2019**, *27*, 304–308. [[CrossRef](#)]
10. Vemmos, K.; Ntaios, G.; Spengos, K.; Savvari, P.; Vemmou, A.; Pappa, T.; Manios, E.; Georgiopoulos, G.; Alevizaki, M. Association between obesity and mortality after acute first-ever stroke: The obesity–stroke paradox. *Stroke* **2011**, *42*, 30–36. [[CrossRef](#)]
11. Kim, B.J.; Lee, S.-H.; Jung, K.-H.; Yu, K.-H.; Lee, B.-C.; Roh, J.-K. Dynamics of obesity paradox after stroke, related to time from onset, age, and causes of death. *Neurology* **2012**, *79*, 856–863. [[CrossRef](#)]
12. Kim, C.K.; Ryu, W.-S.; Kim, B.J.; Lee, S.-H. Paradoxical effect of obesity on hemorrhagic transformation after acute ischemic stroke. *BMC Neurol.* **2013**, *13*, 1–6. [[CrossRef](#)] [[PubMed](#)]
13. Doehner, W.; Schenkel, J.; Anker, S.D.; Springer, J.; Audebert, H.J. Overweight and obesity are associated with improved survival, functional outcome, and stroke recurrence after acute stroke or transient ischaemic attack: Observations from the TEMPiS trial. *Eur. Heart J.* **2013**, *34*, 268–277. [[CrossRef](#)] [[PubMed](#)]
14. Andersen, K.K.; Olsen, T.S. The obesity paradox in stroke: Lower mortality and lower risk of readmission for recurrent stroke in obese stroke patients. *Int. J. Stroke* **2015**, *10*, 99–104. [[CrossRef](#)]
15. Park, H.; Lee, H.W.; Yoo, J.; Lee, H.S.; Nam, H.S.; Kim, Y.D.; Heo, J.H. Body mass index and prognosis in ischemic stroke patients with type 2 diabetes mellitus. *Front. Neurol.* **2019**, *10*, 563. [[CrossRef](#)] [[PubMed](#)]
16. Tang, H.; Mo, J.; Chen, Z.; Xu, J.; Wang, A.; Dai, L.; Cheng, A.; Wang, Y. Uric acid contributes to obesity-paradox of the outcome of ischemic stroke. *Front. Neurol.* **2019**, *10*, 1279. [[CrossRef](#)]
17. Khan, S.S.; Ning, H.; Wilkins, J.T.; Allen, N.; Carnethon, M.; Berry, J.D.; Sweis, R.N.; Lloyd-Jones, D.M. Association of body mass index with lifetime risk of cardiovascular disease and compression of morbidity. *JAMA Cardiol.* **2018**, *3*, 280–287. [[CrossRef](#)] [[PubMed](#)]
18. Agrawal, R.; Imieliński, T.; Swami, A. Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, Washington, DC, USA, 25–28 May 1993; pp. 207–216.
19. Altaf, W.; Shahbaz, M.; Guergachi, A. Applications of association rule mining in health informatics: A survey. *Artif. Intell. Rev.* **2017**, *47*, 313–340. [[CrossRef](#)]
20. Park, S.H.; Jang, S.Y.; Kim, H.; Lee, S.W. An association rule mining-based framework for understanding lifestyle risk behaviors. *PLoS ONE* **2014**, *9*, e88859. [[CrossRef](#)]
21. Burton, S.H.; Morris, R.G.; Giraud-Carrier, C.G.; West, J.H.; Thackeray, R. Mining useful association rules from questionnaire data. *Intell. Data Anal.* **2014**, *18*, 479–494. [[CrossRef](#)]
22. Creighton, C.; Hanash, S. Mining gene expression databases for association rules. *Bioinformatics* **2003**, *19*, 79–86. [[CrossRef](#)]
23. Nahar, J.; Tickle, K.S.; Ali, A.S.; Chen, Y.-P.P. Significant cancer prevention factor extraction: An association rule discovery approach. *J. Med. Syst.* **2011**, *35*, 353–367. [[CrossRef](#)]
24. Peng, S.; Shen, F.; Wen, A.; Wang, L.; Fan, Y.; Liu, X.; Liu, H. Detecting Lifestyle Risk Factors for Chronic Kidney Disease with Comorbidities: Association Rule Mining Analysis of Web-Based Survey Data. *J. Med. Internet Res.* **2019**, *21*, e14204. [[CrossRef](#)]
25. Kim, C.; Jang, M.U.; Oh, M.S.; Park, J.-H.; Jung, S.; Lee, J.-H.; Yu, K.-H.; Han, M.-K.; Kim, B.J.; Park, T.H. Off-hour effect on 3-month functional outcome after acute ischemic stroke: A prospective multicenter registry. *PLoS ONE* **2014**, *9*, e105799. [[CrossRef](#)]
26. Collaboration, A.P.C.S. The burden of overweight and obesity in the Asia–Pacific region. *Obes. Rev.* **2007**, *8*, 191–196.
27. Kim, J.Y.; Kang, K.; Kang, J.; Koo, J.; Kim, D.-H.; Kim, B.J.; Kim, W.J.; Kim, E.-G.; Kim, J.K. Executive Summary of Stroke Statistics in Korea 2018: A Report from the Epidemiology Research Council of the Korean Stroke Society. *J. Stroke* **2019**, *21*, 42–59. [[CrossRef](#)]
28. Lindley, R.I.; Wardlaw, J.M.; Whiteley, W.N.; Cohen, G.; Blackwell, L.; Murray, G.D.; Sandercock, P.A.; Committee, T.S.; Baigent, C.; Chadwick, D. Alteplase for acute ischemic stroke: Outcomes by clinically important subgroups in the Third International Stroke Trial. *Stroke* **2015**, *46*, 746–756. [[CrossRef](#)]
29. Han, J.; Pei, J.; Yin, Y.; Mao, R. Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data Min. Knowl. Discov.* **2004**, *8*, 53–87. [[CrossRef](#)]
30. Bayardo, R.J.; Agrawal, R.; Gunopulos, D. Constraint-based rule mining in large, dense databases. *Data Min. Knowl. Discov.* **2000**, *4*, 217–240. [[CrossRef](#)]

31. Whelton, S.P.; McAuley, P.A.; Dardari, Z.; Orimoloye, O.A.; Brawner, C.A.; Ehrman, J.K.; Keteyian, S.J.; Al-Mallah, M.; Blaha, M.J. Association of BMI, fitness, and mortality in patients with diabetes: Evaluating the obesity paradox in the Henry Ford exercise testing project (fit project) cohort. *Diab. Care* **2020**, *43*, 677–682. [[CrossRef](#)]
32. Bhaskaran, K.; dos-Santos-Silva, I.; Leon, D.A.; Douglas, I.J.; Smeeth, L. Association of BMI with overall and cause-specific mortality: A population-based cohort study of 3.6 million adults in the UK. *Lancet Diabetes Endocrinol.* **2018**, *6*, 944–953. [[CrossRef](#)]
33. Kim, Y.; Kim, C.K.; Jung, S.; Yoon, B.-W.; Lee, S.-H. Obesity-stroke paradox and initial neurological severity. *J. Neurol. Neurosurg. Psychiatry* **2015**, *86*, 743–747. [[CrossRef](#)]
34. Winter, Y.; Rohrmann, S.; Linseisen, J.; Lanczik, O.; Ringleb, P.A.; Hebebrand, J.; Back, T. Contribution of obesity and abdominal fat mass to risk of stroke and transient ischemic attacks. *Stroke* **2008**, *39*, 3145–3151. [[CrossRef](#)]
35. Ryu, W.-S.; Lee, S.-H.; Kim, C.K.; Kim, B.J.; Yoon, B.-W. Body mass index, initial neurological severity and long-term mortality in ischemic stroke. *Cerebrovasc. Dis.* **2011**, *32*, 170–176. [[CrossRef](#)]
36. Mitchell, A.B.; Cole, J.W.; McArdle, P.F.; Cheng, Y.-C.; Ryan, K.A.; Sparks, M.J.; Mitchell, B.D.; Kittner, S.J. Obesity increases risk of ischemic stroke in young adults. *Stroke* **2015**, *46*, 1690–1692. [[CrossRef](#)]
37. Sun, W.; Huang, Y.; Xian, Y.; Zhu, S.; Jia, Z.; Liu, R.; Li, F.; Wei, J.W.; Wang, J.-G.; Liu, M. Association of body mass index with mortality and functional outcome after acute ischemic stroke. *Sci. Rep.* **2017**, *7*, 1–8. [[CrossRef](#)]
38. Dehlendorff, C.; Andersen, K.K.; Olsen, T.S. Body mass index and death by stroke: No obesity paradox. *JAMA Neurol.* **2014**, *71*, 978–984. [[CrossRef](#)]
39. Szalkai, B.; Grolmusz, V.K.; Grolmusz, V.I.; Diseases, C.A.M. Identifying combinatorial biomarkers by association rule mining in the CAMD Alzheimer’s database. *Arch. Gerontol. Geriatr.* **2017**, *73*, 300–307. [[CrossRef](#)]
40. Schäfer, I.; Kaduszkiewicz, H.; Wagner, H.-O.; Schön, G.; Scherer, M.; van den Bussche, H. Reducing complexity: A visualisation of multimorbidity by combining disease clusters and triads. *BMC Public Health* **2014**, *14*, 1–14. [[CrossRef](#)]
41. Li, Q.; Zhang, Y.; Kang, H.; Xin, Y.; Shi, C. Mining association rules between stroke risk factors based on the Apriori algorithm. *Technol. Health Care* **2017**, *25*, 197–205. [[CrossRef](#)]
42. Lee, S.J.; Cartmell, K.B. An Association Rule Mining Analysis of Lifestyle Behavioral Risk Factors in Cancer Survivors with High Cardiovascular Disease Risk. *J. Pers. Med.* **2021**, *11*, 366. [[CrossRef](#)] [[PubMed](#)]
43. Lee, J. An insight on the use of multiple logistic regression analysis to estimate association between risk factor and disease occurrence. *Int. J. Epidemiol.* **1986**, *15*, 22–29. [[CrossRef](#)] [[PubMed](#)]
44. Antonie, M.-L.; Zaïane, O.R. Mining positive and negative association rules: An approach for confined rules. In Proceedings of the European Conference on Principles of Data Mining and Knowledge Discovery, Pisa, Italy, 20–24 September 2004; pp. 27–38.
45. Huang, Y.; Britton, J.; Hubbard, R.; Lewis, S. Who receives prescriptions for smoking cessation medications? An association rule mining analysis using a large primary care database. *Tob. Control* **2013**, *22*, 274–279. [[CrossRef](#)] [[PubMed](#)]