Identification of Metabolic Syndrome Based on Anthropometric, Blood and Spirometric Risk Factors Using Machine Learning

Sang Yeob Kim [1], Gyeong Hee Nam [1] and Byeong Mun Heo [2,*]

[1] National Center for Standard Reference Data, Korea Research Institute of Standards and Science, Daejeon 34113, Korea; sangyeob.kim@kriss.re.kr (S.Y.K.); ghnam@kriss.re.kr (G.H.N.)

[2] Division of Epidemiology and Health Index, Center for Genome Science, Korea National Institute of Health, Cheongju 28159, Korea

* Correspondence: munido@korea.kr; Tel.: +82-10-3177-4424

check for updates

Abstract: Metabolic syndrome (MS) is an aggregation of coexisting conditions that can indicate an individual's high risk of major diseases, including cardiovascular disease, stroke, cancer, and type 2 diabetes. We conducted a cross-sectional survey to evaluate potential risk factor indicators by identifying relationships between MS and anthropometric and spirometric factors along with blood parameters among Korean adults. A total of 13,978 subjects were enrolled from the Korea National Health and Nutrition Examination Survey. Statistical analysis was performed using a complex sampling design to represent the entire Korean population. We conducted binary logistic regression analysis to evaluate and compare potential associations of all included factors. We constructed prediction models based on Naïve Bayes and logistic regression algorithms. The performance evaluation of the prediction model improved the accuracy with area under the curve (AUC) and calibration curve. Among all factors, triglyceride exhibited a strong association with MS in both men (odds ratio (OR) = 2.711, 95% confidence interval (CI) [2.328–3.158]) and women (OR = 3.515 [3.042–4.062]). Regarding anthropometric factors, the waist-to-height ratio demonstrated a strong association in men (OR = 1.511 [1.311–1.742]), whereas waist circumference was the strongest indicator in women (OR = 2.847 [2.447–3.313]). Forced expiratory volume in 6s and forced expiratory flow 25–75% strongly associated with MS in both men (OR = 0.822 [0.749–0.903]) and women (OR = 1.150 [1.060–1.246]). Wrapper-based logistic regression prediction model showed the highest predictive power in both men and women (AUC = 0.868 and 0.932, respectively). Our findings revealed that several factors were associated with MS and suggested the potential of employing machine learning models to support the diagnosis of MS.

1. Introduction

Metabolic syndrome (MS) is a collection of at least three of the five risk factors that increase health problems (e.g., cardiovascular disease [CVD], stroke, cancer, and type 2 diabetes [T2D]) [1–4]. In the presence of MS, the risk of CVD is more than twofold, and the risk of T2D increases more than tenfold [5]. Thus, MS is a major cause of death and a high-risk disease in many people. Favorably, identifying the conditions and associated risk factors of comorbid severe illnesses is easy and inexpensive. Further, these routine checkups are more likely to keep the perturbed physiology in check as compared to the efforts necessary for overcoming a severe disease [6]. The incidence of MS is

increasing worldwide. In the United States, it increases by 1–2% every year [7]. The incidence is also steadily increasing in Asia-Pacific countries, such as China, Korea, and Taiwan [8]. MS is related to several factors, and lung disease, in particular, is related to various diseases that can cause MS [9–11].

Numerous studies have supported the association between anthropometric factors and MS [12–17]. For instance, forearm circumference and bioelectric-impedance-measured visceral fat are associated with MS. However, waist circumference (WC) is not. [12]. Some studies have shown several anthropometric factors to be predictors of components of MS. As a result, no single index was consistently the strongest predictor [13]. Waist-to-hip ratio (WHtR) is the most commonly associated risk factor for MS in Japan, whereas WC and body mass index (BMI) had the strongest associations with MS among other ethnicities [14,15]. Furthermore, WHtR was more strongly associated with MS than WC and BMI, in the identification of metabolic risk factors [16].

The associations of blood parameters and spirometric factors with MS have also been examined [18–24], wherein it was established that dietary patterns were associated with glucose (GLU) intolerance and MS [18]. Additionally, abnormal white blood cell (WBC) count was a vascular risk factor for MS [19]. Further, the incidence of MS was higher in subjects with chronic obstructive lung disease than in those with normal lung function, as determined by the Global Initiative for Chronic Obstructive Lung Disease (GOLD II–IV) guidelines. WC and blood pressure (BP) were also associated with the disease [20]. Low pulmonary function was associated with restrictive lung disease and MS risk factors [22]. Some studies have investigated the association between chronic obstructive pulmonary disease (COPD) and MS in men and women [23]. Lin et al. [24] studied the association between restrictive lung impairment and an increased risk of MS.

In medicine and genomics, numerous studies are conducted using machine learning models [25–37]. However, research that compares and analyzes anthropometric factors, blood parameters, urinary parameters, and spirometric factor anomalies for diagnosis of MS is lacking. Such evaluation is critical because in Asian countries, particularly, MS has become a concern due to rapid and constant changes in diet and lifestyle. Moreover, the study of MS has progressed through the use of machine learning [38].

We analyzed the association between various indicators and MS. Further, we aimed to compare and evaluate data sets for machine learning models to provide an algorithm for predicting MS. Measurements of variables are presented with their respective *p*-values and odds ratios (ORs), calculated using binary logistic regression analysis to identify the factors associated with MS. For evaluation of the machine learning models, Naïve Bayes (NB) and logistic regression algorithm predictive models were created for estimating predictive power. The results of this study provide basic knowledge on the association between MS and anthropometric and spirometric factors that can be used to predict and facilitate the prevention and management of MS. Moreover, the application of machine learning in medicine is expected to support the diagnosis of MS.

2. Materials and Methods

2.1. Subjects and Dataset

We obtained the data for this study from the Korea National Health and Nutrition Examination Survey (KNHANES V and VI) from 2010 to 2015. The KNHANES is a cross-sectional survey conducted initially by the Korea Centers for Disease Control and Prevention [39–41]. Datasets from the survey were approved by the Korea Ministry of Health and Welfare (2010-02CON-21-C, 2011-02CON-06-C, 2012-01EXP-01-2C, and 2013-07CON-03-4C). National Health and Nutrition Examination was conducted without deliberation by the Research Ethics Review Committee, as it corresponds to research conducted by the state for public welfare, according to Article 2 (1) of the Bioethics Law and Article 2 (2) 1 of the Enforcement Regulations of the same law. The research was conducted in accordance with principles of the Helsinki Declaration update of 2008. This study was approved by the Institutional Review Board of the Korea Research Institute of Standards and Science,

and included approval for the access and analysis of open-source data from the KNHANES V and VI with a waiver for the documentation of informed consent (IRB No. KRISS-IRB-2019-14).

MS was defined by the presence of three or more of the five conditions mentioned below [42]. The conditions are as follows (Table 1):

Table 1. Conditions of metabolic syndrome.

Subject	Conditions	Disease
Waist circumference	Men: ≥102 cm (≥40 inches) Women: ≥88 cm (≥35 inches)	Abdominal obesity
Blood pressure	SBP: ≥130 mmHg DBP: ≥85 mmHg	Hypertension
Triglycerides	≥150 mg/dL (1.7 mmol/L)	Hypertriglyceridemia
HDL-cholesterol	Men: <40 mg/dL (1.03 mmol/L) Women: <50 mg/dL (1.3 mmol/L)	Dyslipidemia
Glucose	≥100 mg/dL	Diabetes

SBP, systolic blood pressure; HDL-cholesterol, high-density lipid cholesterol; DBP, diastolic blood pressure.

We conducted an experimental study in three stages. First, we performed data integration based on the measured data (spirometric factors) from 2010 to 2015. Second, we extracted data according to the participants' demographic and clinical characteristics, including age, sex, alcohol consumption, smoking, income, anthropometric factors, blood parameters, and spirometric factors. MS was defined relative to five conditions, and we performed data translation and clarification (addressing missing, uncalculated, and unconverted values).

Finally, we performed data standardization for comparison and analysis. In the statistical analysis, we implemented three methods to analyze the associations of various indices with MS. A t-test was conducted to evaluate significant differences between men and women. We conducted a binary logistic regression analysis to identify significant associations with the measured variables (anthropometric and spirometric factors and blood parameters) and determined differences between normal subjects and those with MS. We used an area under the curve (AUC) from the receiver operating characteristic (ROC) curve to assess whether significant improvement in the identification of MS was achieved based on anthropometric factors, blood parameters, and spirometric factors. Since no method exists to calculate the AUC in a complex sampling analysis, we analyzed the recognized performance by the general AUC analysis. Finally, we constructed a prediction model using machine learning algorithms based on the feature selection method for predictive power by sensitivity, 1-specificity, precision, and AUC [43]. Figure 1 shows the design of this study for preprocessing and statistical analysis.
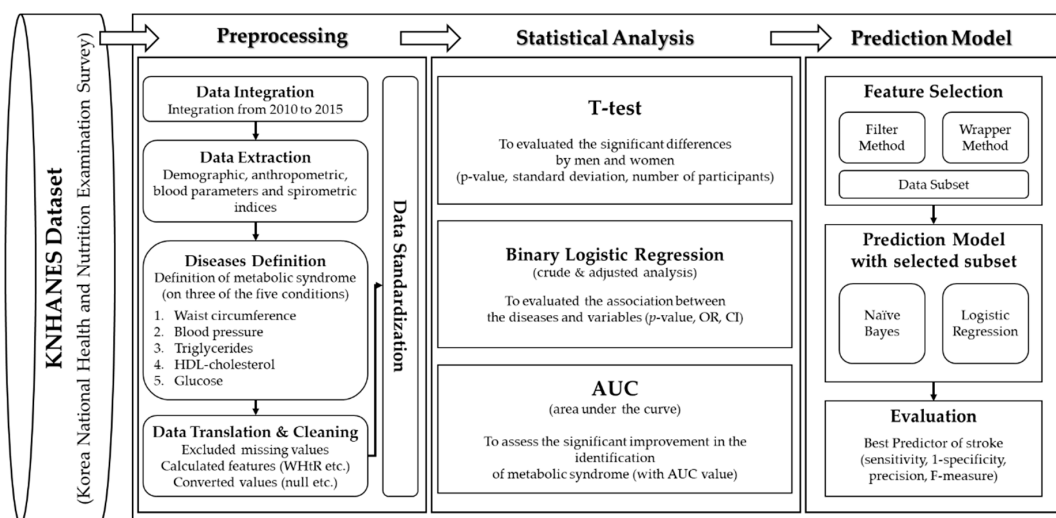
Figure 1. The metabolic syndrome prediction study was designed for preprocessing and statistical analysis. WHtR, waist-to-height ratio; OR, odds ratio; CI, confidence interval.

We focused on participants aged ≥ 40 years in this study because the prevalence of MS increases rapidly in this age group [44]. This yielded a sample size of 26,499 subjects. We excluded subjects with the following measurements: demographics (n = 3409), anthropometric factors (n = 85), blood parameters (n = 4930), uric test results (n = 394), and spirometric factors (n = 3703). Thus, the study comprised 13,978 subjects, and the final dataset consisted of 9893 normal subjects (4489 men and 5404 women) and 4085 subjects with MS (1749 men and 2340 women). Figure 2 presents a flow chart showing details of the sample selection procedure.

Table 2 shows the demographic and clinical characteristics of each group and detailed descriptions of all experimental factors in the t-test.

Table 2. Primary characteristics of all factors analyzed for metabolic syndrome.

Variable	Men		Women	
	Normal	Hypertension	Normal	Hypertension
Subjects, no. (%)	4489 (72.0%)	1745 (28.0%)	5404 (69.8%)	2340 (30.2%)
Age, mean (SD) [†]	57.01 (10.97)	58.95 (10.10)	55.32 (9.97)	62.36 (9.66)
Height (HT) (cm) [†]	168.48 (6.12)	168.08 (6.09)	155.9 (5.75)	154.54 (5.67)
Weight (WT) (kg) [†]	67.91 (9.24)	72.54 (10.49)	56.54 (7.38)	62.63 (9.24)
Waist circumference (WC), (cm) [†]	84.57 (7.58)	89.86 (7.83)	77.95 (7.53)	87.61 (8.44)
Waist-to-height ratio (WHtR) [†]	0.50 (0.05)	0.53 (0.04)	0.50 (0.05)	0.57 (0.06)
Body mass index (BMI) (kg/m^2) [†]	23.88 (2.64)	25.61 (2.89)	23.25 (2.69)	26.19 (3.35)
Systolic BP (mmHg) [†]	121.00 (15.08)	128.81 (15.23)	117.35 (16.57)	129.49 (16.36)
Diastolic BP (mmHg) [†]	77.80 (9.83)	81.2 (11.33)	74.25 (9.22)	77.46 (10.32)
Glucose (mg/dL) [†]	100.62 (21.35)	120.6 (32.51)	94.02 (14.06)	114.63 (30.41)
Hemoglobin A1c (HBA1C) (%) [†]	5.80 (0.77)	6.42 (1.16)	5.68 (0.54)	6.40 (1.07)
Total cholesterol (TC) (mg/dL) [*]	190.93 (34.24)	187 (39.07)	198.13 (34.45)	197.13 (39.73)
High-density lipid cholesterol (HDL) (mg/dL) [†]	48.65 (11.1)	40.82 (9.55)	54.90 (12.05)	45.26 (9.73)
Low-density lipid cholesterol (LDL) (mg/dL) [†]	118.38 (32.66)	107.48 (33.28)	122.33 (32.13)	118.32 (34.97)
Triglycerides (TG) (mg/dL) [†]	136.99 (95.58)	242.71 (176.66)	104.62 (57.33)	182.03 (105.83)
Aspartate aminotransferase (AST), (IU/L) [†]	24.58 (12.95)	26.89 (13.47)	21.26 (11.89)	24.12 (11.10)
Alanine aminotransferase (ALT), (IU/L) [†]	24.05 (19.12)	28.94 (17.03)	17.80 (12.43)	23.83 (16.51)
Hemoglobin (HB) (g/dL) [†]	15.09 (1.23)	15.24 (1.30)	13.12 (1.11)	13.38 (1.08)
Hematocrit (HCT) (%) [*]	44.51 (3.33)	44.73 (3.53)	39.54 (2.91)	40.07 (2.93)
Blood urea nitrogen (BUN) (mg/dL) [*]	15.69 (4.2)	16.12 (5.12)	14.46 (4.01)	15.42 (4.74)
Creatinine (CRT) (mg/dL) [*]	0.97 (0.18)	1.01 (0.31)	0.72 (0.15)	0.76 (0.29)
White blood cell (WBC) (Thous/uL) [†]	6.34 (1.72)	6.84 (1.80)	5.55 (1.52)	6.23 (1.69)
Red blood cell (RBC) (Mil/uL) [†]	4.79 (0.41)	4.84 (0.43)	4.31 (0.32)	4.39 (0.35)
Uric acid pH (UPH) [†]	8.51 (1.87)	8.26 (1.86)	6.13 (1.17)	5.71 (0.85)
Uric specific gravity (USG)	5.74 (0.82)	5.59 (0.76)	5.83 (0.88)	1.02 (0.01)
Urine creatinine (UCRT) (mg/dL)	1.02 (0.01)	1.02 (0.01)	1.02 (0.01)	110.86 (61.25)
Forced vital capacity (FVC) (L)	161.08 (76.55)	158.58 (79.37)	118.83 (65.57)	2.75 (0.46)
Forced expiratory volume in 1 s (FEV1)	4.19 (0.70)	4.01 (0.68)	3.00 (0.49)	2.16 (0.42)
Ratio of forced expiratory volume in 1 s to forced vital capacity (FEV1FVC)	3.14 (0.66)	2.98 (0.64)	2.39 (0.44)	0.79 (0.06)
Forced expiratory volume in 6 s (FEV6) (L)	0.75 (0.08)	0.74 (0.08)	0.79 (0.06)	2.67 (0.46)
Forced expiratory flow 25–75% (FEF25–75) (L/s) [*]	4.02 (0.71)	3.82 (0.68)	2.94 (0.50)	2.35 (0.99)
Peak expiratory flow (PEF) (L/s) [*]	2.81 (1.29)	2.8 (1.54)	2.49 (0.84)	5.78 (1.22)

[*] $p < 0.05$ and [†] $p < 0.0001$ indicate significant differences between men and women using the two-stage sample t-test. All data are presented as means ± standard deviations (±SD). However, the subjects are presented as the number of participants (%).

2.2. Measurements

All factors were measured according to established guidelines, as described previously [40,41]. Height measurements were quantified to the nearest 1 mm using a Seca 225 portable stadiometer (Seca, Hamburg, Germany). Weight measurements were quantified to the nearest 0.1 kg using an electronic scale (GL-6000-20; Caskorea, Seoul, Korea). WC measurements (between the sternum and the hips) were quantified to the nearest 1 mm using Seca 200. BMI was calculated as weight/height^2. The WHtR (a new indicator in this study) was calculated as WC/height. Blood parameters were measured using Hitachi Automatic Analyzer 7600-210 (Hitachi, Tokyo, Japan) with Pureauto SCHO-N (Sekisui, Tokyo, Japan) and S TG-N, S AST, S ALT, S GLU, and Cholestest N HDL, LDL (Sekisui, Tokyo, Japan). Glycated hemoglobin was measured using HLC-723G7 (Tosoh, Tokyo, Japan). White and red blood cell counts were measured using XE-2100D (Sysmex, Tokyo, Japan). Spirometric factors were measured using the Vmax series 2130 (SensorMedics, Yorba Linda, CA, USA). Spirometric factors

measured included forced vital capacity, forced expiratory volume in 1s, ratio of forced expiratory volume in 1 s to forced vital capacity, forced expiratory volume in 6s, forced expiratory flow 25–75% (FEF25-75), and peak expiratory flow by pulmonary function test.
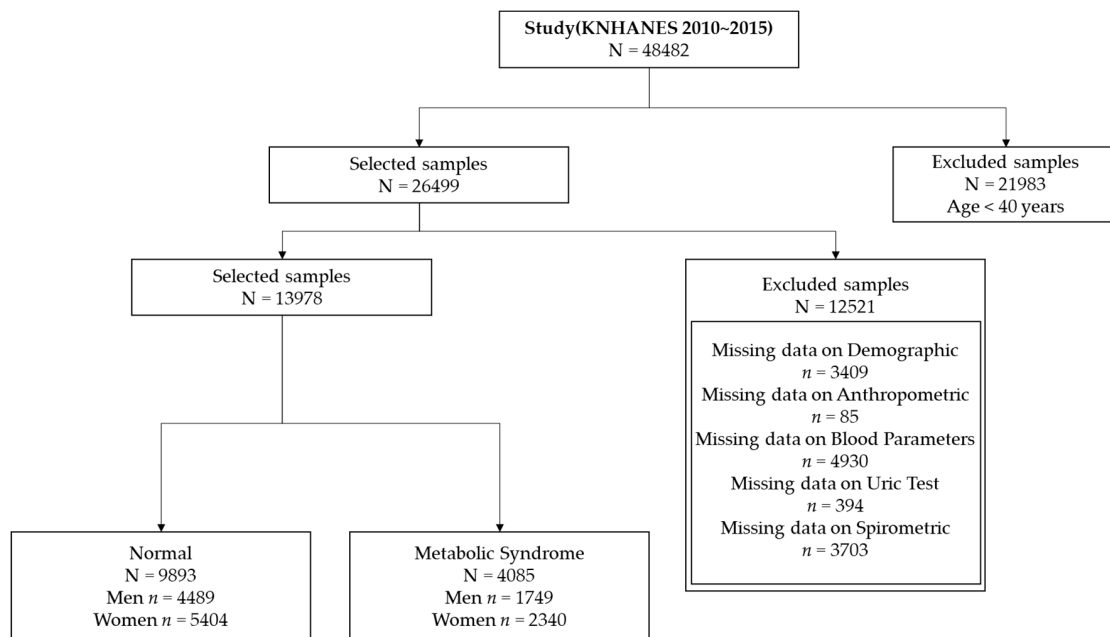
Figure 2. Sample selection procedure for association analysis.

2.3. Statistical Analysis

We used a complex sampling design to examine associations between MS and related factors. Stratified two-stage sampling was performed with the primary sampling unit (PSU), cluster, and weight values that are taken into consideration. The results could be biased due to the estimated factors, variance in the means, and prevalence rate from simple random sampling. Thus, we performed a complex sampling data analysis that utilized weighting, as described previously [40,41,45].

Statistical analyses were performed using SPSS 22 software for Windows (SPSS, Inc., Armonk, NY, USA). Binary logistic regression was used to predict or classify machine learning and data mining fields, but this algorithm is primarily used to analyze and examine the associations between diseases and various factors in medicine, public health, and epidemiology studies. It can provide *p*-values, OR values, and confidence intervals (CIs) for association analyses. Thus, it was suitable for this algorithm and both crude and adjusted analyses of our data and objectives. In the adjusted analyses, we used age, BMI, frequency of alcohol consumption, smoking, income, recognized stress rate, and education level as the value adjustment factors. We analyzed and evaluated the association between normal subjects and those with MS. We used the data mining tool of the Waikato Environment for Knowledge Analysis (WEKA), constructed the prediction models, and evaluated the prediction performance. We used 10-fold cross validation to efficiently distribute training, test data sets and evaluate prediction models [46].

3. Results

3.1. Associations of MS with Anthropometric Factors, Blood Parameters, Urinary Parameters, and Spirometric Factors

Shown in Tables 3 and 4 are the associations of MS with the anthropometric and spirometric factors, and blood and urinary parameters in Korean men and women. Overall, 29.2% (N = 4085) of the 13,978 participants exhibited MS. The prevalence of MS was higher in women than in men (30.2% and 28.0%, respectively).

Table 3. Associations of anthropometric factors, blood parameters, urinary parameters, and spirometric factors with metabolic syndrome in men.

Variable	Crude		Adjusted		AUC
	p-Value	OR [95% CI]	*p*-Value	OR [95% CI]	
Age	<0.0001	1.272 [1.196–1.352]			0.555 [0.540–0.570]
Anthropometric factors					
HT	<0.05	0.918 [0.858–0.982]	0.8655	0.993 [0.914–1.078]	0.480 [0.464–0.496]
WT	<0.0001	1.592 [1.485–1.706]	0.9151	1.008 [0.865–1.175]	0.623 [0.608–0.639]
WC	<0.0001	2.069 [1.922–2.228]	<0.0001	1.468 [1.287–1.674]	0.683 [0.668–0.697]
WHtR	<0.0001	2.164 [2.006–2.334]	<0.0001	1.511 [1.311–1.742]	0.694 [0.680–0.709]
BMI	<0.0001	1.902 [1.763–2.052]			0.667 [0.652–0.682]
Blood parameters					
GLU	<0.0001	2.348 [2.048–2.691]	<0.0001	2.165 [1.923–2.438]	0.776 [0.764–0.789]
HBA1C	<0.0001	1.906 [1.703–2.133]	<0.0001	1.797 [1.633–1.978]	0.712 [0.697–0.726]
TC	0.2372	0.957 [0.889–1.030]	0.0929	0.935 [0.865–1.011]	0.464 [0.448–0.481]
HDL	<0.0001	0.385 [0.347–0.427]	<0.0001	0.384 [0.342–0.431]	0.276 [0.261–0.290]
TG	<0.0001	2.719 [2.350–3.146]	<0.0001	2.711 [2.328–3.158]	0.787 [0.775–0.799]
AST	<0.05	1.144 [1.025–1.276]	<0.05	1.102 [1.018–1.193]	0.574 [0.558–0.590]
ALT	<0.05	1.250 [1.052–1.484]	0.1358	1.134 [0.961–1.338]	0.619 [0.603–0.634]
HB	<0.0001	1.155 [1.075–1.240]	<0.05	1.095 [1.009–1.188]	0.533 [0.517–0.550]
HCT	<0.05	1.088 [1.016–1.166]	0.5084	1.027 [0.950–1.109]	0.516 [0.500–0.532]
BUN	<0.05	1.117 [1.044–1.195]	0.2716	1.042 [0.968–1.123]	0.516 [0.500–0.532]
CRT	<0.05	1.199 [1.063–1.353]	<0.0001	1.261 [1.172–1.358]	0.547 [0.531–0.563]
WBC	<0.0001	1.316 [1.235–1.403]	<0.05	1.134 [1.046–1.230]	0.587 [0.571–0.603]
RBC	<0.05	1.143 [1.069–1.223]	<0.05	1.134 [1.046–1.230]	0.532 [0.516–0.548]
Urinary parameters					
UPH	<0.0001	0.823 [0.767–0.883]	<0.0001	0.852 [0.791–0.918]	0.451 [0.436–0.467]
USG	0.6728	1.014 [0.949–1.084]	0.6370	1.018 [0.945–1.096]	0.485 [0.469–0.501]
UCRT	0.6667	0.985 [0.919–1.056]	0.9480	1.003 [0.928–1.083]	0.484 [0.468–0.500]
Spirometric factors					
FVC	<0.0001	0.755 [0.706–0.807]	<0.0001	0.839 [0.769–0.915]	0.423 [0.407–0.438]
FEV1	<0.0001	0.772 [0.723–0.825]	<0.05	0.850 [0.770–0.940]	0.424 [0.408–0.439]
FEV1FVC	<0.05	0.933 [0.875–0.994]	0.6219	1.021 [0.941–1.106]	0.471 [0.455–0.487]
FEV6	<0.0001	0.733 [0.685–0.784]	<0.0001	0.822 [0.749–0.903]	0.413 [0.398–0.429]
FEF25–75	0.4195	0.971 [0.904–1.043]	<0.05	1.123 [1.035–1.219]	0.475 [0.459–0.491]
PEF	<0.05	0.882 [0.825–0.944]	0.4528	0.966 [0.884–1.057]	0.457 [0.441–0.473]

Adjusted for age, body mass index (BMI), alcohol consumption, smoking, income, recognized stress rate, and education level. The results are from crude and adjusted analyses using binary logistic regression.

In men, among all factors, triglycerides (TG) ($p < 0.0001$, OR = 2.719, 95% CI [2.350–3.146]) presented the strongest association with MS during crude analysis. Additionally, TG (adjusted $p < 0.0001$, adjusted OR = 2.711 [2.328–3.158]) exhibited the strongest association after adjusting for age, BMI, alcohol consumption, smoking, income, recognized stress rate, and education level. Among anthropometric factors, WHtR demonstrated the strongest association with MS in crude ($p < 0.0001$, OR = 2.164 [2.006–2.334]) and adjusted (adjusted $p < 0.0001$, adjusted OR = 2.069 [1.922–2.228]) analyses. In the second, WC had the next strongest association in the crude ($p < 0.0001$, OR = 2.164 [2.006–2.334]) and adjusted (adjusted $p < 0.0001$, adjusted OR = 1.468 [1.287–1.674]) analyses. Among blood parameters, TG (adjusted $p < 0.0001$, adjusted OR = 2.711 [2.328–3.158]) was the strongest negative indicator of MS in both crude and adjusted analyses. GLU ($p < 0.0001$, OR = 2.348 [2.048–2.691]) was associated with MS in crude analysis. However, HDL (adjusted $p < 0.0001$, adjusted OR = 0.384 [0.342–0.431]) showed a negative association with MS in the adjusted analysis. Among urinary parameters, uric acid pH (UPH) showed the highest negative association with MS in crude ($p < 0.0001$, OR = 0.823 [0.767–0.883]) and adjusted (adjusted $p < 0.0001$, adjusted OR = 0.852 [0.791–0.918]) analyses. Among spirometric factors, forced expiratory volume in 6s (FEV6) was negatively associated with MS in crude ($p < 0.0001$, OR = 0.733 [0.685–0.784]) and adjusted (adjusted $p < 0.0001$, adjusted OR = 0.822 [0.749–0.903]) analyses, followed by forced vital capacity (FVC) in both crude ($p < 0.0001$, OR = 0.755 [0.706–0.807]) and adjusted (adjusted $p < 0.0001$, adjusted OR = 0.839 [0.769–0.915]) analyses.

Table 4. Associations of anthropometric factors, blood parameters, urinary parameters, and spirometric factors with metabolic syndrome in women.

Variable	Crude		Adjusted		AUC
	p-Value	OR [95% CI]	*p*-Value	OR [95% CI]	
Age	<0.0001	2.070 [1.946–2.202]			0.696 [0.683–0.708]
Anthropometric factors					
HT	<0.0001	0.763 [0.719–0.808]	<0.0001	1.188 [1.097–1.287]	0.429 [0.415–0.443]
WT	<0.0001	2.197 [2.046–2.358]	<0.0001	1.411 [1.212–1.643]	0.700 [0.687–0.713]
WC	<0.0001	4.011 [3.647–4.411]	<0.0001	2.847 [2.447–3.313]	0.808 [0.797–0.819]
WHtR	<0.0001	4.166 [3.795–4.572]	<0.0001	2.564 [2.194–2.996]	0.813 [0.802–0.823]
BMI	<0.0001	2.922 [2.699–3.164]			0.758 [0.746–0.770]
Blood parameters					
GLU	<0.0001	4.593 [3.818–5.524]	<0.0001	3.377 [2.828–4.033]	0.810 [0.799–0.821]
HBA1C	<0.0001	3.504 [3.022–4.063]	<0.0001	2.466 [2.161–2.813]	0.776 [0.765–0.788]
TC	0.1381	1.050 [0.984–1.119]	0.4955	0.974 [0.903–1.051]	0.486 [0.472–0.501]
HDL	<0.0001	0.318 [0.290–0.348]	<0.0001	0.348 [0.313–0.386]	0.255 [0.244–0.267]
TG	<0.0001	3.723 [3.266–4.244]	<0.0001	3.515 [3.042–4.062]	0.796 [0.785–0.807]
AST	<0.0001	1.433 [1.274–1.612]	<0.05	1.174 [1.055–1.306]	0.601 [0.587–0.615]
ALT	<0.0001	1.588 [1.363–1.849]	<0.0001	1.301 [1.153–1.467]	0.671 [0.658–0.684]
HB	<0.0001	1.316 [1.229–1.409]	<0.0001	1.206 [1.119–1.300]	0.567 [0.553–0.581]
HCT	<0.0001	1.245 [1.165–1.330]	<0.05	1.122 [1.046–1.204]	0.551 [0.537–0.565]
BUN	<0.0001	1.247 [1.167–1.333]	0.6779	0.986 [0.920–1.056]	0.557 [0.543–0.571]
CRT	<0.05	1.265 [1.057–1.515]	<0.05	1.143 [1.006–1.299]	0.548 [0.533–0.562]
WBC	<0.0001	1.431 [1.346–1.522]	<0.0001	1.315 [1.225–1.413]	0.626 [0.613–0.640]
RBC	<0.0001	1.305 [1.222–1.395]	<0.0001	1.294 [1.206–1.389]	0.567 [0.553–0.581]
Urinary parameters					
UPH	<0.0001	0.862 [0.812–0.915]	<0.0001	0.834 [0.781–0.891]	0.413 [0.400–0.427]
USG	0.4582	0.979 [0.926–1.035]	0.1512	0.951 [0.887–1.019]	0.464 [0.451–0.478]
UCRT	<0.05	0.903 [0.852–0.957]	0.4165	0.971 [0.904–1.043]	0.485 [0.472–0.499]
Spirometric factors					
FVC	<0.0001	0.583 [0.544–0.624]	<0.05	0.882 [0.809–0.962]	0.465 [0.452–0.479]
FEV1	<0.0001	0.590 [0.551–0.631]	<0.05	0.898 [0.818–0.985]	0.351 [0.338–0.364]
FEV1FVC	<0.0001	0.855 [0.806–0.907]	0.5393	1.024 [0.949–1.106]	0.352 [0.339–0.365]
FEV6	<0.0001	0.566 [0.528–0.607]	<0.05	0.872 [0.798–0.954]	0.460 [0.447–0.474]
FEF25–75	<0.0001	0.843 [0.781–0.911]	<0.05	1.150 [1.060–1.246]	0.344 [0.331–0.357]
PEF	<0.0001	0.734 [0.689–0.781]	0.2660	1.045 [0.967–1.130]	0.433 [0.419–0.447]

Adjusted for age, body mass index (BMI), alcohol consumption, smoking, income, recognized stress rate, and education level. The results are from crude and adjusted analyses using binary logistic regression.

Among all the factors analyzed in women, GLU ($p < 0.0001$, OR = 4.593 [3.818–5.524]) showed the strongest association with MS. Moreover, TG (adjusted $p < 0.0001$, adjusted OR = 3.515 [3.042–4.062]) was associated with MS in adjusted analysis. Among anthropometric factors, WHtR ($p < 0.0001$, OR = 4.166 [3.795–4.572]) was strongly associated with MS in crude analysis. WC (adjusted $p < 0.0001$, adjusted OR = 2.847 [2.447–3.313]) was strongly associated in adjusted analysis. Among blood parameters, GLU exhibited the highest association with MS in crude analysis. And TG was strongly associated in the adjusted analysis. UPH ($p < 0.0001$, OR = 0.862 [0.812–0.915]) showed the strongest negative association with MS among the urinary parameters in crude analysis. The association with UPH (adjusted $p < 0.0001$, adjusted OR = 0.834 [0.781–0.891]) was maintained following adjusted analysis. Urine creatinine ($p < 0.05$, OR = 0.903 [0.852–0.957]) was strongly associated with crude analysis; however, this association disappeared following adjustment for confounders. FEV6 ($p < 0.0001$, OR = 0.583 [0.544–0.624]) showed highest negative association with MS among the spirometric factors in crude analysis. Furthermore, FEF25-75 (adjusted $p < 0.05$, adjusted OR = 1.150 [1.060–1.246]) showed MS association in adjusted analysis.

In the overall crude analysis, TG presented the strongest association with MS in men, whereas GLU exhibited the strongest association in women. WHtR was associated with both men and women, but WC was associated with adjusted analysis. FEV6 was also associated with both men and women. In the adjusted analysis, FEV6 and FEF25–75 associated with MS. These results suggested sex-based differences in blood parameters; however, other risk factors showed similar results in both sexes.

From the AUC analysis, Figures 3 and 4 show the predictive power of anthropometric and spirometric factors, blood and urinary parameters for MS in Korean men and women. Among all factors, TG (AUC = 0.787 [0.775–0.799]) showed strong predictive power in men, whereas WHtR (AUC = 0.813 [0.802–0.823]) demonstrated the highest AUC value in women. GUL and HBA1C exhibited strong predictive powers among the blood parameters in men (AUC = 0.776 [0.764–0.789] and AUC = 0.712 [0.697–0.726], respectively) and women (AUC = 0.810 [0.799–0.821] and AUC = 0.776 [0.765–0.788], respectively). Among the spirometric factors, FEV6 and FVC showed strongest negative predictive power for men (AUC = 0.413 [0.398–0.429] and AUC = 0.423 [0.407–0.438], respectively) and women (AUC = 0.460 [0.447–0.474] and AUC = 0.465 [0.452–0.479], respectively).
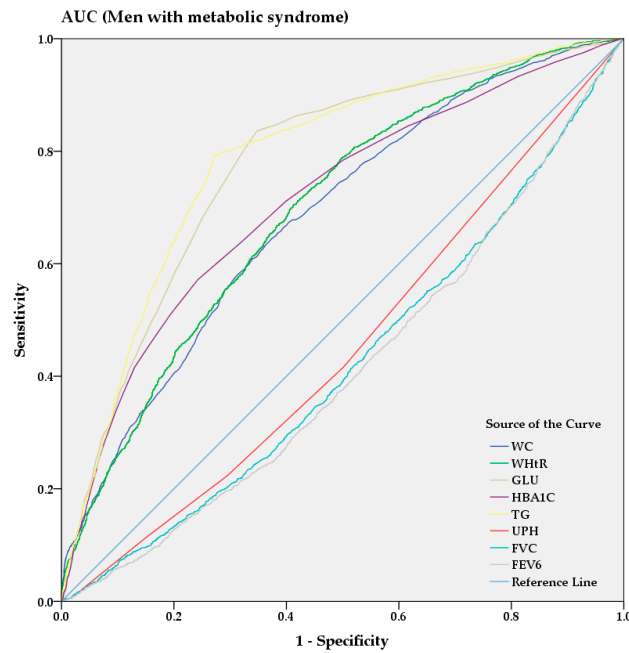
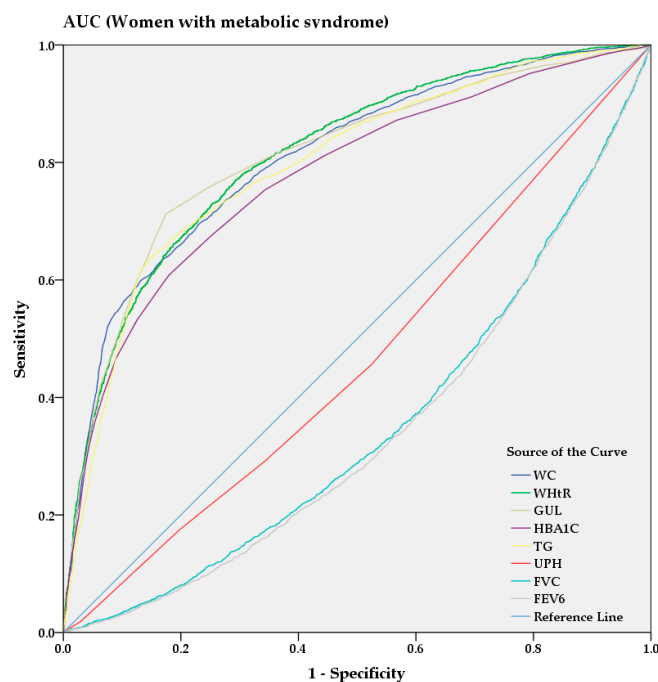**Figure 3.** AUC analysis of metabolic syndrome in men.

Figure 4. AUC analysis of metabolic syndrome in women.

3.2. Comparison of the Predictive Power of MS with Anthropometric Factors, Blood Parameters, Urinary Parameters, and Spirometric Factors

Each prediction model was constructed using NB and logistic regression algorithms by a subset of features using the wrapper and filter feature selection method for performance evaluation of the MS prediction model for both men and women. Table 5 shows detailed results of feature selection by the wrapper and filter method. The model built by Wrapper NB in men included 13 features: age, WT, WC, SBP, GLU, HDL, TG, ALT, WBC, USG, UCREA, FVC, and FEV6. Our results revealed a higher predictive power in women than in men for NB and logistic regression predictions.

Table 5. Feature selection using the wrapper and filter method for each model.

Sex	Method		Num. of Features	Features
Men	Wrapper	Naïve Bayes	13	Age, WT, WC, SBP, Glucose, HDL, TG, ALT, WBC, USG, UCREA, FVC, FEV6
		Logistic Regression	15	HT, WC, WHtR, SBP, DBP, Glucose, HBA1C, TC, HDL, TG, HB, HCT, CREA, WBC, RBC
	Filter	CFS	8	WHtR, SBP, Glucose, HDL, TG, AST, CREA, RBC
Women	Wrapper	Naïve Bayes	9	WC, WHtR, SBP, DBP, Glucose, HDL, TG, RBC, UPH
		Logistic Regression	12	Age, WT, WC, SBP, Glucose, TC, HDL, TG, AST, BUN, UPH, FEV1FVC
	Filter	CFS	11	WC, WHtR, SBP, Glucose, HBA1C, HDL, TG, ALT, CREA, WBC, FEF25-75

CFS: Correlation-based feature selection.

The performance of models was compared and evaluated based on sensitivity, 1-specificity, F-measure, AUC, area under the precision-recall curve (AUPRC), and root mean square error (RMSE) [47]. As shown in Table 6, for men, the sensitivity, 1-specificity, and F-measure of the prediction model (Wrapper LR [Logistic Regression]) were 0.926, 0.449, and 0.882 in the normal group and 0.551, 0.074, and 0.633 in the MS group, respectively. Among women, the sensitivity, 1-specificity, and F-measure were 0.930, 0.261, and 0.911 in the normal group and 0.739, 0.070, and 0.778 in the MS group, respectively.

Table 6. Predictive power analysis of the four models in men and women.

Sex	Prediction Model	Class	Sensitivity	1-Specificity	Precision	F-Measure
Men	Wrapper NB	Normal	0.904	0.488	0.826	0.863
		Metabolic	0.512	0.096	0.674	0.582
	Wrapper LR	Normal	0.926	0.449	0.841	0.882
		Metabolic	0.551	0.074	0.742	0.633
	Filter NB	Normal	0.919	0.598	0.798	0.854
		Metabolic	0.402	0.081	0.659	0.499
	Filter LR	Normal	0.924	0.469	0.835	0.877
		Metabolic	0.531	0.076	0.730	0.615
Women	Wrapper NB	Normal	0.905	0.262	0.888	0.896
		Metabolic	0.738	0.095	0.770	0.754
	Wrapper LR	Normal	0.930	0.261	0.892	0.911
		Metabolic	0.739	0.070	0.821	0.778
	Filter NB	Normal	0.923	0.334	0.864	0.893
		Metabolic	0.666	0.077	0.788	0.722
	Filter LR	Normal	0.927	0.275	0.886	0.906
		Metabolic	0.725	0.073	0.811	0.766

This table was created using data transformed by standardization. The results of detailed classification performance were grouped by class (normal and metabolic groups) using a confusion matrix. NB: Naïve Bayes, LR: logistic regression, metabolic: metabolic syndrome.
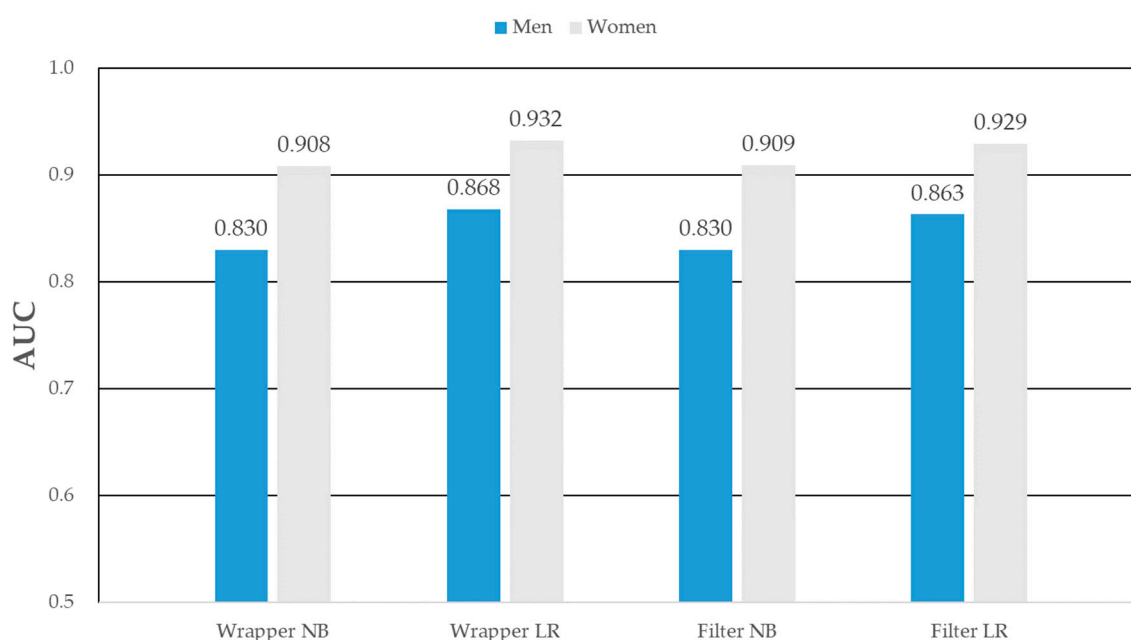
As shown in Table 7, among the men MS prediction models, Wrapper_LR's AUC [95% CI], AUPRC [95% CI], R^2 [95% CI], and RMSE [95% CI] were the predictive power with 0.868 [0.860–0.876], 0.712 [0.701–0.723], 0.960 [0.955–0.965], and 0.358 [0.346–0.370], respectively. Among the women, Wrapper_LR's AUC, AUPRC, R^2, and RMSE were the predictive power with 0.932 [0.926–0.938], 0.851 [0.843–0.859], 0.986 [0.983–0.988], and 0.308 [0.298–0.318], respectively. The wrapper_LR showed the best performance among other prediction models in both men and women. For men, the MS prediction model (the wrapper-based logistic regression) showed the strongest predictive power (AUC = 0.868), as detailed in Figure 5. In women, the wrapper based logistic regression model showed the highest predictive power (AUC = 0.932).

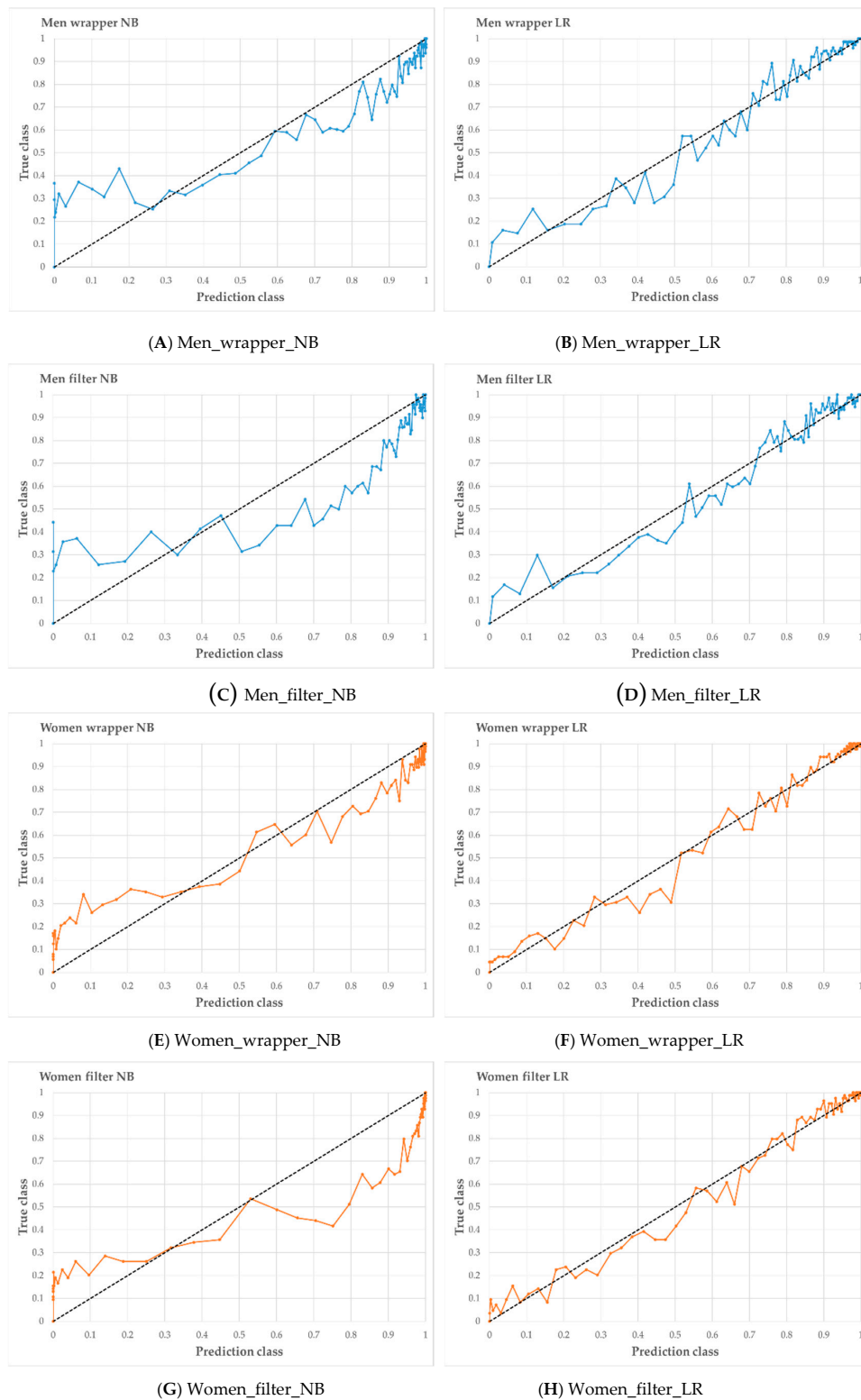**Table 7.** Performance analysis results of the four models in men and women.

Sex	Prediction Model	AUC	AUPRC	R^2	RMSE
Men	Wrapper NB	0.830 [0.821–0.839]	0.616 [0.604–0.628]	0.899 [0.892–0.907]	0.396 [0.384–0.408]
	Wrapper LR	0.868 [0.860–0.876]	0.712 [0.701–0.723]	0.960 [0.955–0.965]	0.358 [0.346–0.370]
	Filter NB	0.830 [0.821–0.839]	0.604 [0.592–0.616]	0.802 [0.792–0.812]	0.407 [0.395–0.419]
	Filter LR	0.863 [0.854–0.872]	0.703 [0.692–0.714]	0.952 [0.947–0.958]	0.361 [0.349–0.373]
Women	Wrapper NB	0.908 [0.902–0.914]	0.788 [0.779–0.797]	0.969 [0.965–0.973]	0.341 [0.330–0.352]
	Wrapper LR	0.932 [0.926–0.938]	0.851 [0.843–0.859]	0.986 [0.983–0.988]	0.308 [0.298–0.318]
	Filter NB	0.911 [0.905–0.917]	0.783 [0.774–0.792]	0.915 [0.908–0.921]	0.352 [0.341–0.363]
	Filter LR	0.929 [0.923–0.935]	0.843 [0.835–0.851]	0.986 [0.983–0.988]	0.311 [0.301–0.321]

The results are shown as each values with corresponding 95% CI. NB: Naïve Bayes, LR: logistic regression, R^2: coefficients of determination for calibration curve.

Figure 6 shows the calibration curve for each prediction model. The x-axis of the calibration plot is a predicted class, and the y-axis is plotted as the true class. In men, Wrapper_LR suggests a better calibration plot than Wrapper_NB. In women, Wrapper_LR suggests a better calibration plot than Wrapper_NB. In the feature selection method, the wrapper showed better results for prediction model development than CFS.

Figure 5. AUC of metabolic syndrome prediction models in men and women.

(**A**) Men_wrapper_NB

(**B**) Men_wrapper_LR

(**C**) Men_filter_NB

(**D**) Men_filter_LR

(**E**) Women_wrapper_NB

(**F**) Women_wrapper_LR

(**G**) Women_filter_NB

(**H**) Women_filter_LR

Figure 6. Calibration curve of four models in men (**A**–**D**) and women (**E**–**H**). When the line of the calibration plot is straight, the prediction model was presented the best performance.

4. Discussion

In this study, anthropometric and spirometric factors and blood parameters were examined as potential risk factors for MS. In a previous report on anthropometric factors and MS, Sagun et al. [12] conducted a study involving 387 subjects admitted to the obesity outpatient department of Istanbul

Medeniyet University Goztepe Training and Research Hospital in Turkey. The purpose of their study was to evaluate associations of WC, hip circumference, WHtR, waist-to-hip ratio (WHR), mid-upper arm circumference, forearm circumference, calf circumference, and body composition with MS. The OR of visceral fat, hip circumference, forearm circumference, and WHR were 2.19 [95% CI, 1.30–3.71], 1.89 [1.07–3.35], 2.47 [1.24–4.95], and 2.11 [1.26–3.53], respectively. WC was not related to MS. However, forearm circumference and bioelectric-impedance-measured visceral fat were associated with the disease.

Mooney et al. [13] determined that other anthropometric factors, such as WC, WHtR, percent body fat, fat mass index (FMI), and fat-free mass index (FFMI), were consistently better predictors as MS-associated factors than BMI. They obtained their data from 12,294 adults who took part in annual physical examinations provided by EHE International, Inc., New York, NY, USA. They showed that each anthropometric factor was related to metabolic risk factors using Pearson correlation analyses, linear regression analyses, and ROC curves, and no single index exhibited the strongest prediction consistently. BMI was identified as the strongest predictor of BP.

Hsieh et al. [14] investigated the association of anthropometric factors, such as BMI, WC, and WHtR, with MS in 8278 Japanese subjects (6141 men and 2137 women). The AUC values for BMI, WC, and WHtR were 0.677 [95% CI, 0.644–0.691], 0.687 [0.673–0.700], and 0.696 [0.682–0.710] respectively. WHtR was the most commonly associated risk factor in screening Japanese individuals for MS.

Shen et al. [15] evaluated whether WC correlated more strongly with MS components than percent fat and other related anthropometric factors, such as BMI, in 1010 healthy white and African-American men and women. Their results demonstrated that WC was most strongly associated with MS, followed by BMI.

Rodea-Montero et al. [16] identified the associations between anthropometric factors and MS in 110 Mexican obese adolescents. BMI, WC, and WHtR were evaluated as anthropometric factors. The investigators used ROC curves to identify significant predictors. The AUC values of BMI, WC, and WHtR were 0.651 [95% CI, 0.547–0.755], 0.704 [0.604–0.804], and 0.652 [0.544–0.759], respectively. They found that WHtR was associated more strongly than WC and BMI for identifying metabolic risk.

Williams et al. [18] demonstrated the association between GLU and MS. Their study enrolled 802 subjects aged 40–65 years who were randomly selected from a population-based sampling frame. Their study identified four dietary patterns using principal component analysis. These dietary patterns were related to other lifestyle factors, such as the socioeconomic group, smoking, alcohol intake, and physical activity. In component 1, there was a negative association with diabetes as one of the causes of MS. According to the results, dietary patterns were associated with GLU intolerance and MS.

Lao et al. [19] investigated the association between white blood cell (WBC) count and MS in old Chinese patients. The analyzed dataset (obtained from a medical checkup record) consisted of 3020 men and 7256 women aged 50–85 years. Vascular risk factors (e.g., WC, BMI, TG, TC, LDL, C-reactive protein, SBP, and DBP) were associated with WBCs both in men and women. The risk of MS increased significantly with higher WBC counts (OR = 1.86 [1.43–2.42]). There was a strong association between WBC count and vascular risk factors of MS.

Funakoshi et al. [20] investigated the association of airflow obstruction with MS in Japanese men. Their dataset consisted of 7189 subjects aged 45–88 years from spirometric lung function tests at medical checkups. The airflow obstruction was divided into two parts (GOLD I and GOLD II–IV) according to GOLD guidelines. The incidence rate of MS was higher in subjects with GOLD II–IV than in those with normal lung function (OR = 1.33 [1.01–1.76]). Additionally, the MS component WC was associated with MS (OR = 1.76 [1.24–2.50]) and BP (OR = 1.37 [1.08–1.74]).

Paek et al. [22] evaluated the association between impaired lung function and metabolic risk factors, enrolling 4001 subjects aged > 18 years in 2001 from the KNHANES dataset. Using multiple linear regression, they analyzed the association of low pulmonary function with MS. They also examined the associations of restrictive lung disease and obstructive lung disease with MS using multiple logistic regression adjusting for WHtR, sex, age, smoking, physical activity, alcohol consumption, and

socioeconomic status. WC, SBP, and TG were associated with FVC. They showed that the association of low pulmonary function with MS risk factors and restrictive lung disease (OR = 1.40 [1.01–1.98]) was also related to MS.

Park et al. [23] investigated the association between chronic obstructive pulmonary disease (COPD) and MS. Their dataset comprised 1215 subjects aged > 40 years from the KNHANES in 2001. The prevalence of MS was significantly high among COPD patients in both men and women (33.0% and 48.5% higher, respectively). In men, the risk of COPD (OR = 2.03 [1.08–3.80]) was associated with MS and abdominal obesity (OR = 1.95 [0.93–4.11]).

Lin et al. [24] studied the relationship between impaired lung function and MS in adults. Their study assessed 46,514 patients (21,669 men and 24,845 women) aged > 20 years, recruited from four nationwide MJ Health Screening Centers in Taiwan. The investigators examined associations between lung function test results and MS using multivariate logistic regression. They demonstrated the association between restrictive lung impairment and an increased risk of MS ($p < 0.01$, OR = 1.221), adjusting for age, sex, BMI, smoking, alcohol consumption, and physical activity.

This study showed results similar to those of previous studies [12–24] regarding the association of MS with anthropometric, blood, urinary, and spirometric factors. Specifically, TG exhibited the strongest association among all parameters in both men and women. With regard to anthropometric factors, WHtR showed the strongest association in both men and women. Concerning blood parameters, GLU showed a strong positive association. With respect to urinary parameters, UPH showed a negative association with MS. Regarding spirometric factors, FEV6 and FVC exhibited negative associations with MS.

Choe et al. [38] constructed and analyzed the five MS prediction models including the MLP (multilayer perceptron), NB, RF (random forest classification), CT (decision tree), and SVM (support vector machine). As for the AUC, the NB model showed the highest predictive power at 0.690. Kopitar et al. [47] evaluated and validated T2DM (type 2 diabetes mellitus) prediction model including lm (linear regression model), Glmnet (regularized generalized linear model), RF, XGBoot (extreme gradient boosting), and LightGBM (light gradient boosting machine) by RMSE, AUC, AUPRC, R^2, and calibration plot. RF showed the best performance according to RMSE. In our study, as for the results of performance analysis and calibration plot, the wrapper_LR model showed the best performance in both men (AUC = 0.868) and women AUC = 0.932), respectively.

Overall, our current findings show that anthropometric, blood, urinary, and spirometric factors might be involved in the induction of MS. Specifically, there was increased association of WC and WHtR, and decreased association of TG, UPH, FEV6, and FVC with MS. The wrapper based logistic regression model showed a high predictive power for both men and women.

Collectively, all the studies, including ours, identified different potent indicators of MS, demonstrating that no single factor could serve as an effective marker for the identifying MS. This confirms the concept that co-occurring conditions contribute to development of the disease. Thus, a subgroup of individuals with shared pathophysiology, and hence a common strong indicator, might be at high risk of a specific comorbid disease. Therefore, subgroup stratification of MS might be necessary for partitioning risk factor clusters involved, which should provide insight into personalizing pharmacological and lifestyle modifications as treatment approaches for managing the condition.

This study had several limitations: First, it was difficult to determine cause-and-effect associations due to the cross-sectional design of the study. Second, our results were limited to Korean adults because of the KNHANES dataset employed. Nonetheless, despite the identification of specific indicators of MS, which confirmed the findings of previous reports utilizing machine learning, our results indicated sex-based differences in MS risk factors.

5. Conclusions

MS is closely related to well-known major diseases, such as CVD, stroke, cancer, and T2D. Thus, it is increasingly attracting attention as an objective health indicator worldwide. In this

study, we investigated the associations between MS and anthropometric factors, blood parameters, and spirometric factors, and evaluated their predictive performance using prediction models. Binary logistic regression analysis was used for statistical analysis, and the wrapper and filter methods were used for feature selection. Based on the selected factors, prediction models were constructed using NB and logistic regression algorithms. The results showed that WHtR, WC, TG, GLU, UPH, FEV6, and FVC were associated with MS. In particular, TG in men and GLU in women were both highly significant risk factors in participants with MS. According to AUC analysis, TG in men and WHtR in women showed strongest predictive powers for the disease among all factors. FEV6 and FVC represented high predictive power among the spirometric factors. The wrapper-based logistic regression prediction model showed highest predictive performance in both men and women. In the calibration curve for verification of the predictive model, the wrapper_LR showed the highest performance in men and women. Thus, our results provide fundamental clinical information supporting the prevention and management of MS, and our model shows potential as a large-scale screening tool.

Author Contributions: Conceptualization, S.Y.K., B.M.H.; formal analysis, S.Y.K., G.H.N.; investigation, S.Y.K.; validation, S.Y.K.; writing—original draft preparation, S.Y.K., B.M.H.; writing—review and editing, S.Y.K., G.H.N., B.M.H. All authors have reviewed and agreed to the publication of the final version of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

MS, metabolic syndrome; AUC, area under the curve; AUPRC, area under the precision-recall curve; RMSE, root mean square error; HT, height; WT, weight; WC, waist circumference; WHtR, waist-to-height ratio; GLU, glucose; HBA1C, hemoglobin A1c; TC, total cholesterol; HDL, high-density lipid cholesterol, TG, triglyceride; AST, aspartate aminotransferase; ALT, alanine aminotransferase; HB, hemoglobin; HCT, hematocrit; BUN, blood urea nitrogen; CRT, creatinine; WBC, white blood cell; RBC, red blood cell; UPH, uric acid pH; USG, uric specific gravity; UCRT, uric creatinine; FVC, forced vital capacity; FEV1, forced expiratory volume in 1s; FEV1FVC, ratio of forced expiratory volume in 1s to forced vital capacity; FEV6, forced expiratory volume in 6s; FEF25-75, forced expiratory flow 25–75%; PEF, peak expiratory flow.

References

1. Cameron, A.J.; Shaw, J.E.; Zimmet, P.Z. The metabolic syndrome: Prevalence in worldwide populations. *Endocrinol. Metab. Clin. N. Am.* **2004**, *33*, 351–375. [CrossRef] [PubMed]
2. Wilson, P.W.F.; Meigs, J.B. Cardiometabolic risk: A Framingham perspective. *Int. J. Obes.* **2008**, *32*, S17–S20. [CrossRef] [PubMed]
3. Stocks, T.; Hemelrijck, M.V.; Manjer, J.; Bjørge, T.; Ulmer, H.; Hallmans, G.; Lindkvist, B.; Selmer, R.; Nagel, G.; Tretli, S.; et al. Blood pressure and risk of cancer incidence and mortality in the Metabolic Syndrome and Cancer Project. *Hypertension* **2012**, *59*, 802–810. [CrossRef] [PubMed]
4. Boden-Albala, B.; Sacco, R.L.; Lee, H.-S.; Grahame-Clarke, C.; Rundek, T.; Elkind, M.V.; Wright, C.; Giardina, E.-G.V.; DiTullio, M.R.; Homma, S.; et al. Metabolic syndrome and ischemic stroke risk. *Stroke* **2008**, *39*, 30–35. [CrossRef]
5. Korea Center for Disease Control and Prevention. Available online: http://www.cdc.go.kr/ (accessed on 21 August 2020).
6. Statistics Research Institute. Available online: http://kostat.go.kr/sri/srikor/index.action (accessed on 21 August 2020).
7. Beltrán-Sánchez, H.; Harhay, M.O.; Harhay, M.M.; McElligott, S. Prevalence and trends of metabolic syndrome in the adult U.S. population, 1999–2010. *J. Am. Coll. Cardiol.* **2013**, *62*, 697–703. [CrossRef]
8. Ranasinghe, P.; Mathangasinghe, Y.; Jayawardena, R.; Hills, A.P.; Misra, A. Prevalence and trends of metabolic syndrome among adults in the Asia-Pacific region: A systematic review. *BMC Public Health* **2017**, *17*, 101. [CrossRef]

9. Mannino, D.M.; Thorn, D.; Swensen, A.; Holguin, F. Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in COPD. *Eur. Respir. J.* **2008**, *32*, 962–969. [CrossRef]

10. Schroeder, E.B.; Welch, V.L.; Couper, D.; Nieto, F.J.; Liao, D.; Rosamond, W.D.; Heiss, G. Lung function and incident coronary heart disease: The Atherosclerosis Risk in Communities Study. *Am. J. Epidemiol.* **2003**, *158*, 1171–1181. [CrossRef]

11. Engström, G.; Hedblad, B.; Valind, S.; Janzon, L. Increased incidence of myocardial infarction and stroke in hypertensive men with reduced lung function. *J. Hypertens.* **2001**, *19*, 295–301. [CrossRef]

12. Sagun, G.; Oguz, A.; Karagoz, E.; Filizer, A.T.l.; Tamer, G.; Mesci, B. Application of alternative anthropometric measurements to predict metabolic syndrome. *Clinics* **2014**, *69*, 347–353. [CrossRef]

13. Mooney, S.J.; Baecker, A.; Rundle, A.G. Comparison of anthropometric and body composition measures as predictors of components of the metabolic syndrome in a clinical setting. *Obes. Res. Clin. Pract.* **2013**, *7*, e55–e66. [CrossRef] [PubMed]

14. Hsieh, S.D.; Muto, T. Metabolic syndrome in Japanese men and women with special reference to the anthropometric criteria for the assessment of obesity: Proposal to use the waist-to-height ratio. *Prev. Med.* **2006**, *42*, 135–139. [CrossRef]

15. Shen, W.; Punyanitya, M.; Chen, J.; Gallagher, D.; Albu, J.; Pi-Sunyer, X.; Lewis, C.E.; Grunfeld, C.; Heshka, S.; Heymsfield, S.B. Waist circumference correlates with metabolic syndrome indicators better than percentage fat. *Obesity* **2006**, *14*, 727–736. [CrossRef] [PubMed]

16. Rodea-Montero, E.R.; Evia-Viscarra, M.L.; Apolinar-Jiménez, E. Waist-to-height ratio is a better anthropometric index than waist circumference and BMI in predicting metabolic syndrome among obese Mexican adolescents. *Int. J. Endocrinol.* **2014**, *2014*, 1–9. [CrossRef]

17. Katzmarzyk, P.T.; Church, T.S.; Janssen, I.; Ross, R.; Blair, S.N. Metabolic syndrome, obesity, and mortality: Impact of cardiorespiratory fitness. *Diabetes Care* **2005**, *28*, 391–397. [CrossRef] [PubMed]

18. Williams, D.E.M.; Prevost, A.T.; Whichelow, M.J.; Cox, B.D.; Day, N.E.; Wareham, N.J. A cross-sectional study of dietary patterns with glucose intolerance and other features of the metabolic syndrome. *Br. J. Nutr.* **2000**, *83*, 257–266. [CrossRef]

19. Lao, X.Q.; Neil Thomas, G.; Jiang, C.; Zhang, W.; Adab, P.; Lam, T.H.; Cheng, K.K. White blood cell count and the metabolic syndrome in older Chinese: The Guangzhou Biobank Cohort Study. *Atherosclerosis* **2008**, *201*, 418–424. [CrossRef]

20. Funakoshi, Y.; Omori, H.; Mihara, S.; Marubayashi, T.; Katoh, T. Association between airflow obstruction and the metabolic syndrome or its components in Japanese men. *Intern. Med.* **2010**, *49*, 2093–2099. [CrossRef]

21. McDevitt, R.M.; Poppitt, S.D.; Murgatroyd, P.R.; Prentice, A.M. Macronutrient disposal during controlled overfeeding with glucose, fructose, sucrose, or fat in lean and obese women. *Am. J. Clin. Nutr.* **2000**, *72*, 369–377. [CrossRef]

22. Paek, Y.J.; Jung, K.S.; Hwang, Y.I.; Lee, K.S.; Lee, D.R.; Lee, J.U. Association between low pulmonary function and metabolic risk factors in Korean adults: The Korean National Health and Nutrition Survey. *Metabolism* **2010**, *59*, 1300–1306. [CrossRef]

23. Park, B.H.; Park, M.S.; Chang, J.; Kim, S.K.; Kang, Y.A.; Jung, J.Y.; Kim, Y.S.; Kim, C. Chronic obstructive pulmonary disease and metabolic syndrome: A nationwide survey in Korea. *Int. J. Tuberc. Lung Dis.* **2012**, *16*, 694–700. [CrossRef] [PubMed]

24. Lin, W.Y.; Yao, C.A.; Wang, H.C.; Huang, K.C. Impaired lung function is associated with obesity and metabolic syndrome in adults. *Obesity* **2006**, *14*, 1654–1661. [CrossRef] [PubMed]

25. Wulczyn, E.; Steiner, D.F.; Xu, Z.; Sadhwani, A.; Wang, H.; Flament-Auvigne, I.; Mermel, C.H.; Chen, P.-H.C.; Liu, Y.; Stumpe, M.C. Deep learning-based survival prediction for multiple cancer types using histopathology images. *PLoS ONE* **2020**, *15*, e0233678. [CrossRef]

26. Kopp, W.; Monti, R.; Tamburrini, A.; Ohler, U.; Akalin, A. Deep learning for genomics using Janggu. *Nat. Commun.* **2020**, *11*, 1–7. [CrossRef] [PubMed]

27. Sonogashira, M.; Shonai, M.; Iiyama, M. High-resolution bathymetry by deep-learning-based image superresolution. *PLoS ONE* **2020**, *15*, e0235487. [CrossRef]

28. Davagdorj, K.; Pham, V.H.; Theera-Umpon, N.; Ryu, K.H. XGBoost-based framework for smoking-induced noncommunicable disease prediction. *Int. J. Environ. Res. Public Health* **2020**, *17*, 6513. [CrossRef]

29. Munkhdalai, L.; Munkhdalai, T.; Ryu, K.H. GEV-NN: A deep neural network architecture for class imbalance problem in binary classification. *Knowl. Based Syst.* **2020**, 105534. [CrossRef]

30. Davagdorj, K.; Lee, J.S.; Pham, V.H.; Ryu, K.H. A comparative analysis of machine learning methods for class imbalance in a smoking cessation intervention. *Appl. Sci.* **2020**, *10*, 3307. [CrossRef]

31. Amarbayasgalan, T.; Park, K.H.; Lee, J.Y.; Ryu, K.H. Reconstruction error based deep neural networks for coronary heart disease risk prediction. *PLoS ONE* **2019**, *14*, e0225991. [CrossRef]

32. Batbaatar, E.; Ryu, K.H. Ontology-based healthcare named entity recognition from Twitter messages using a recurrent neural network approach. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3628. [CrossRef]

33. Ryu, K.S.; Bae, J.-W.; Jeong, M.H.; Cho, M.-C.; Ryu, K.-H.; Investigators, K.A.M.I.R. Risk scoring system for prognosis estimation of multivessel disease among patients with ST-segment elevation myocardial infarction. *Int. Hear. J.* **2019**, *60*, 708–714. [CrossRef] [PubMed]

34. Heo, B.M.; Ryu, K.H. Prediction of prehypertension and hypertension based on anthropometry, blood parameters, and spirometry. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2571. [CrossRef] [PubMed]

35. Piao, Y.; Piao, M.; Ryu, K.H. Multiclass cancer classification using a feature subset-based ensemble from microRNA expression profiles. *Comput. Biol. Med.* **2017**, *80*, 39–44. [CrossRef]

36. Kim, H.; Ishag, M.I.M.; Piao, M.; Kwon, T.; Ryu, K.H. A data mining approach for cardiovascular disease diagnosis using heart rate variability and images of carotid arteries. *Symmetry* **2016**, *8*, 47. [CrossRef]

37. Ryu, K.S.; Park, H.W.; Park, S.H.; Shon, H.S.; Ryu, K.H.; Lee, D.G.; Bashir, M.E.; Lee, J.H.; Kim, S.M.; Lee, S.Y. Comparison of clinical outcomes between culprit vessel only and multivessel percutaneous coronary intervention for ST-segment elevation myocardial infarction patients with multivessel coronary diseases. *J. Geriatr. Cardiol.* **2015**, *12*, 208.

38. Choe, E.K.; Rhee, H.; Lee, S.; Shin, E.; Oh, S.W.; Lee, J.E.; Choi, S.H. Metabolic Syndrome Prediction Using Machine Learning Models with Genetic and Clinical Information from a Nonobese Healthy Population. *Genom. Inform.* **2018**, *16*, e31. [CrossRef] [PubMed]

39. Korea Centers for Disease Control and Prevention: Korea National Health and Nutrition Examination Survey. Available online: https://knhanes.cdc.go.kr/knhanes/eng/index.do (accessed on 20 January 2020).

40. Ministry of Health and Welfare of Korea; Korea Centers for Disease Control and Prevention. *The Fifth Korea National Health and Nutrition Examination Survey Data User Guide (KNHANES V) 2010–2012*; Korea Centers for Disease Control and Prevention Press: Cheongju, Korea, 2012.

41. Ministry of Health and Welfare of Korea; Korea Centers for Disease Control and Prevention. *The Sixth Korea National Health and Nutrition Examination Survey Data User Guide (KNHANES VI) 2013–2015*; Korea Centers for Disease Control and Prevention Press: Cheongju, Korea, 2015.

42. Grundy, S.M.; Cleeman, J.I.; Daniels, S.R.; Donato, K.A.; Eckel, R.H.; Franklin, B.A.; Gordon, D.J.; Krauss, R.M.; Savage, P.J.; Smith, S.C.; et al. Diagnosis and Management of the Metabolic Syndrome. *Circulation* **2005**, *112*, 2735–2752. [CrossRef]

43. Saeys, Y.; Inza, I.; Larrañaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* **2007**, *23*, 2507–2517. [CrossRef]

44. Statistics Korea Web Sites. Available online: http://kostat.go.kr (accessed on 1 September 2020).

45. Kim, S.Y.; Nam, G.H. Assessment of Anthropometric and Body Composition Risk Factors in Patients with both Hypertension and Stroke in the Korean Population. *Appl. Sci.* **2020**, *10*, 3046. [CrossRef]

46. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI), Montreal, QC, Canada, 20–25 August 1995; Volume 2, pp. 1137–1145.

47. Kopitar, L.; Kocbek, P.; Cilar, L.; Sheikh, A.; Stiglic, G. Early detection of type 2 diabetes mellitus using machine learning-based prediction models. *Sci. Rep.* **2020**, *10*, 11981. [CrossRef]