Medicine®

OBSERVATIONAL STUDY

OPEN

A Model for Predicting the Future Risk of Incident Erosive Esophagitis in an Asymptomatic Population Undergoing Regular Check-ups

Soo Hoon Kang, MD, Yaeji Lim, PhD, Hyuk Lee, MD, Joungyoun Kim, PhD, Sangah Chi, BS, Yang Won Min, MD, Byung-Hoon Min, MD, Jun Haeng Lee, MD, Hee Jung Son, MD, Seungho Ryu, MD, Poong-Lyul Rhee, MD, and Jae J. Kim, MD

Abstract: Erosive esophagitis is a major risk factor for Barrett esophagus and esophageal adenocarcinoma. Information regarding the putative risk factors for developing erosive esophagitis is considerably heterogeneous; thus, a risk model is required to clinically predict the incidence of erosive esophagitis. This study was to derive and validate a predictive model for the incidence of developing erosive esophagitis after negative index endoscopy in a population subjected to routine health check-ups. This retrospective cohort study of health check-ups included 11,535 patients who underwent repeated screening endoscopy after >3 years from a negative index endoscopy. We used logistic regression analysis to predict the incidence of erosive esophagitis, and a Simple Prediction of Erosive Esophagitis Development score for risk assessment was developed and internally validated using the split-sample approach. The development and validation cohorts included 5765 patients (675 with erosive esophagitis [11.7%]) and 5770 patients (670 with erosive esophagitis [11.6%]), respectively. The final model included sex, smoking behavior, body mass index, hypertension, and the triglyceride level as variables. This model predicted 667 cases of erosive esophagitis, yielding an expected-to-observed ratio of 1.00 (95% confidence interval [CI], 0.92–1.07). A simplified 5-item risk scoring system based on coefficients was developed, with a risk of erosive esophagitis of 6.2% (95% CI, 5.2–7.1) for the low-risk group (score ≤2), 15.1% (95% CI, 13.5–16.6) for the intermediate-risk group (score ≤3, 4), and 18.2% (95% CI, 15.2–21.3) for the high-risk group (score ≥5). The discriminative performance of the risk-prediction score was consistent in the derivation cohort and validation cohort (c-statistics 0.68 and 0.64, respectively); the calibration was good (Brier score 0.099 and 0.1, respectively). In conclusion, a simple risk-scoring model using putative risk factors can predict the future incidence of developing erosive esophagitis in asymptomatic populations.

(*Medicine* 95(4):e2591)

Abbreviations: Apo = apolipoprotein, BMI = body mass index, DBP = diastolic blood pressure, GERD = gastroesophageal reflux disease, GI = gastrointestinal, HbA1c = glycated hemoglobin, HDL = high-density lipoprotein, LDL = low-density lipoprotein, SBP = systolic blood pressure.

INTRODUCTION

Gastroesophageal reflux disease (GERD) is a common gastrointestinal (GI) disorder that frequently occurs in the primary care setting, with a high direct and indirect economic burden on society.[1,2] Endoscopic erosive esophagitis is a major risk factor for Barrett esophagus, and the risk of esophageal adenocarcinoma substantially increases in patients with previously diagnosed erosive esophagitis.[3–5] From the viewpoint of long-term complications of erosive esophagitis, treatment and prevention of esophagitis is necessary, whether it is symptomatic or silent esophagitis, although the natural history for silent GERD remains unclear.[6]

Massive endoscopic screening of the entire population to detect erosive esophagitis would not be feasible because of physical risks and cost effectiveness.[7] Furthermore, a symptom-based approach has limits in terms of heterogeneity of symptoms, low correlation with the degree of endoscopic esophagitis, and a relatively high prevalence of silent esophagitis.[6,8,9] Therefore, the comprehensive elucidation of risk factors for developing erosive esophagitis in the future would be helpful for determining a high-risk group and for subsequent individualized decision for screening. A large number of observational studies showed various putative risk factors for erosive esophagitis, including demographic and metabolic parameters, even if they were somewhat heterogeneous.[10–13] As these analyses were limited by significant temporal bias, the interpretation of parameters, including lifestyle changes, may be suboptimal, considering the positive benefit of only long-term intervention. Currently, there is no available risk predictive model for the incidence of erosive esophagitis after negative index endoscopy, especially one that uses accessible clinical primary care parameters.

Editor: Tuomo Rantanen.
Received: July 28, 2015; revised: December 23, 2015; accepted: December 30, 2015.
From the Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea (SHK, HL, YWM, B-HM, JHL, HJS, P-LR, JJK); Department of Statistics, Pukyong National University, Pusan, Republic of Korea (YL); Biostatistics and Clinical Epidemiology Center, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea (JK, SC); Center for Health Promotion, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea (HJS); and Department of Occupational and Environmental Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University, School of Medicine, Seoul, Republic of Korea (SR).
Correspondence: Hyuk Lee, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 135-710, Republic of Korea (e-mail: leehyuk@skku.edu).
Co-correspondence: Hee Jung Son, Center for Health Promotion, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea (e-mail: hjls.son@samsung.com).
SHK and YL contributed equally to this study.
Conflicts of interest: No conflicts of interest exist for SHK, YL, HLK, JK, SC, YWM, B-HM, JHL, HJS, SR, P-LR, and JJK.

ISSN: 0025-7974
DOI: 10.1097/MD.0000000000002591

To determine the comprehensive risk factors for the incidence of developing erosive esophagitis after negative index endoscopy, we have analyzed a large-scale health check-up cohort. We performed data analysis to derive and validate a risk-prediction model for new-onset erosive esophagitis. The results of the model were used to develop a simple scoring system that estimates the likelihood of developing erosive esophagitis. In addition, this scoring system was internally validated using split-sample method.

MATERIALS AND METHODS

Design Overview

We performed a retrospective cohort analysis of database records for patients who entered the health check-up program for upper GI cancer at the Center for Health Promotion, Samsung Medical Center, in Korea. This comprehensive health-screening program included anthropometric measurements, annual or biennial endoscopy, various laboratory studies, and an epidemiological questionnaire on lifestyle factors, medication, and chronic disease. Patients paid voluntarily for their health check-ups, or were partly supported by an affiliated company. The study was approved by the institutional review board of Samsung Medical Center, and due to the retrospective nature of the study, the requirement for informed consent was waived.

Study Sample

In total, 19,217 patients who underwent upper endoscopic examination for screening from January 2006 through December 2008 were enrolled. All patients were asymptomatic at the time of index endoscopy. Patients were included if they underwent repeated screening endoscopy after at least 3 years until July 2014, considering previous data from the kinetic curve showing the progression to erosive esophagitis.[14] Patients were excluded when they were diagnosed with erosive esophagitis, Barrett esophagus, malignant disease of the upper GI tract, and active or healing peptic ulcer disease on index baseline endoscopy; when they had prior gastroesophageal surgery; when they had not fully completed the epidemiological questionnaire; or when their records were missing. The final cohort included 11,535 patients (Figure 1), and none of them needed additional medical treatment, as they did not have clinically noticeable GI symptoms at the time of index endoscopy.

Endpoint, Definitions, and Covariates

The primary endpoint was the development of erosive esophagitis after negative index endoscopy during screening. Erosive esophagitis was defined if definite erosions (mucosal breaks) were present, and were classified according to the Los Angeles classification system.[15] All participants had their body mass index (BMI), body fat, and waist circumference measured by previously described techniques.[16] Weight and height were measured in the morning with participants wearing light clothing, but no shoes, and BMI was calculated as weight in kilograms divided by the square of the height in meters. According to the World Health Organization's classification for BMI, participants were stratified into underweight (BMI $<18.5\,kg/m^2$), normal (BMI $18.5–23.0\,kg/m^2$), overweight (BMI $23.0–25.0\,kg/m^2$), and obesity (BMI $\geq 25.0\,kg/m^2$).[17] The waist circumference was measured midway between the lower border of the rib cage and iliac crest when participants were standing at the end of normal expiration. Percentage of body fat was measured by bioelectrical impedance analysis (In Body 3.0; Biospace, Seoul, Korea). We measured blood pressure and blood markers such as levels of fasting glucose, glycated hemoglobin (HbA1c), total cholesterol, triglyceride, low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, apolipoprotein (Apo)-A1, and ApoB. Blood samples were collected from the antecubital vein after overnight nothing per oral. Total cholesterol, LDL, HDL, triglyceride, and fasting glucose were measured using enzymatic or colorimetric methods. HbA1c levels were measured using a high-performance liquid chromatography method with a Tosoh Glycohemoglobin Analyzer (Tosoh Bioscience Inc, Tokyo, Japan). Serum glucose levels were measured using the hexokinase/glucose-6-phosphate dehydrogenase method with a Hitachi 7600 Modular Dp-110 autoanalyzer (Hitachi, Tokyo, Japan). The average interassay and intraassay coefficients of variation for quality control were 6.5% and 2.1% for fasting glucose, and 2.5% and 2.5% for HbA1c levels, respectively. Systolic blood pressure and diastolic blood pressure were measured after a patient had rested for 5 min in a sitting position. Structured questionnaires included self-reported comorbidities of diabetes, hypertension, or dyslipidemia. Regarding alcohol intake, patients were classified into current drinker, former drinker, or never drinker. Patients were also classified as current smoker, former smoker, or never smoker. Regular exercise was defined as doing physical exercise of at least moderate intensity >3 times per week, for at least 30 minutes each time. Medication history of antihypertensive agents, aspirin, and nonsteroidal anti-inflammatory drug use was also collected.

FIGURE 1. Study flow diagram.

Power Calculations and Statistical Analysis

Power estimates focused on identifying independent predictors in a multivariate logistic regression model for erosive esophagitis. Assuming an overall prevalence of erosive esophagitis of 8.0% for a Korean population,[18] a sample of 5765 patients would provide 98% power of detecting an adjusted odds ratio (OR) of 1.5 for a predictor with a prevalence of 25%.

The original dataset was randomly partitioned to generate derivation and validation cohorts. To split the data randomly, we generated the random numbers and used the index of the number to split the data. The model was built with the derivation cohort using logistic regression analysis. Univariate logistic regression analysis was conducted to identify risk factors associated with erosive esophagitis, and corresponding unadjusted relative risks were computed. In multivariate logistic regression analysis, we included risk factors with P values <0.05 from univariate analysis. Multivariate logistic regression analysis was used to verify the relationship between erosive esophagitis and variables shown as significant risk factors in univariate analysis. Additionally, stepwise selection analysis was performed to elucidate the most influential risk factors without multicollinearity between the independent variables for erosive esophagitis; a risk factor was entered into the model if the P value was <0.05, and it stayed in the model if the P value was <0.05. The model was internally validated using the validation cohort. The performance of the model was assessed considering discrimination and calibration.[19,20] The ability to discriminate the development of erosive esophagitis was evaluated by the area under the receiver-operating characteristic curve. Model calibration, in the sense of directly comparing a model's predicted probabilities to observed probabilities, was assessed graphically and was tested using the Hosmer–Lemeshow goodness-of-fit test of the model's predictions.[21] As the Hosmer–Lemeshow goodness-of-fit test is sensitive to sample size, we chose smaller subsets chosen at random (n = 500) to avoid escalation of power with large sample sizes.[22] Results of multivariate logistic regression were used to develop a simplified risk score for predicting the development of erosive esophagitis after negative index endoscopy in asymptomatic patients. Model-adjusted coefficients were rounded to the nearest one-half integer and were multiplied by 2 to avoid decimals. The performance of the risk score was assessed in the validation set using the concordance statistic. To measure the calibration of the risk scoring system, we calculated the Brier score.[23] Statistical analysis was performed using SAS 9.4 (SAS Corp.) and R (R Foundation for Statistical Computing). Differences with a P value <0.05 were considered statistically significant.

RESULTS

Patient Characteristics

The study population (11,535 participants) was divided into the derivation cohort (5765 participants) and validation cohort (5770 participants). Overall, there were 1345 cases of erosive esophagitis. General characteristics of the cohorts are presented in Table 1. Characteristics of each cohort were well balanced. The incidence of erosive esophagitis was not different between cohorts (11.7% in derivation cohort and 11.6% in validation cohort; $P = 0.871$, testing equality of rate).

TABLE 1. Population Characteristics of the Derivation and Validation Cohorts

Characteristics (Mean ± SD or n, %)	Derivation Cohort (n = 5765)	Validation Cohort (n = 5770)	*P*
Age	50.9 ± 8.9	50.9 ± 8.8	0.8958
Sex (male)	3523 (61.1)	3513 (60.9)	0.8033
Anthropometric profile			
Body mass index, kg/m^2	23.9 ± 2.8	23.9 ± 2.9	0.5036
Waist circumference, cm	83.1 ± 9.1	83.2 ± 9.3	0.7327
Body fat, kg	16.4 ± 5.0	16.4 ± 5.1	0.9958
Blood pressure profile			
SBP, mm Hg	115.5 ± 15.5	115.4 ± 15.6	0.7854
DBP, mm Hg	71.2 ± 10.7	71.4 ± 10.7	0.2731
Lipid profile			
Total cholesterol, mg/dL	192.2 ± 33.1	193.0 ± 33.2	0.1048
Triglyceride, mg/dL	128.1 ± 77.6	129.2 ± 79.9	0.4801
LDL, mg/dL	123.4 ± 29.2	124.2 ± 29.3	0.1476
HDL, mg/dL	56.6 ± 13.7	56.6 ± 14.0	0.6437
ApoA1, mg/dL	143.2 ± 24.3	142.7 ± 24.3	0.3282
ApoB, mg/dL	91.5 ± 21.2	91.5 ± 20.8	0.9685
Insulin resistance profile			
Fasting glucose, mg/dL	93.3 ± 18.6	93.0 ± 18.0	0.2866
HbA1c	5.5 ± 0.7	5.5 ± 0.7	0.6838
Medication			
Aspirin	661 (11.5)	680 (11.8)	0.5903
NSAID	138 (2.4)	153 (2.7)	0.3764
Antihypertensive agent	1229 (21.0)	1243 (21.2)	0.7638
Comorbidity			
Hypertension	1268 (22.0)	1284 (22.3)	0.7383
Dyslipidemia	1059 (18.4)	1029 (17.8)	0.4548
Diabetes mellitus	1236 (21.4)	1168 (20.2)	0.1135
Metabolic syndrome	1121 (19.4)	1096 (19.0)	0.5396
Lifestyle-related activities			
Alcohol drinking			0.9639
Never	1644 (28.8)	1657 (29.0)	
Former	395 (6.9)	391 (6.9)	
Current	3663 (64.2)	3658 (64.1)	
Smoking			0.1883
Never	3060 (53.1)	2974 (51.5)	
Former	1709 (29.6)	1738 (30.1)	
Current	996 (17.3)	1058 (18.3)	
Regular physical exercise	4739 (86.5)	4700 (85.6)	0.1596
Incidence of erosive esophagitis	675 (11.7)	670 (11.6)	0.8713

Apo = apolipoprotein, DBP = diastolic blood pressure, HbA1c = glycated hemoglobin, HDL = high-density lipoprotein, LDL = low-density lipoprotein, SBP = systolic blood pressure.

Development of a Prediction Model

Univariate associations between potential risk factors and the development of erosive esophagitis after negative index endoscopy are presented in Table 2. Sex; BMI; waist circumference; body fat; systolic blood pressure; diastolic blood pressure; levels of triglyceride, HDL, fasting glucose, and HbA1c; current history of hypertension, dyslipidemia, and

TABLE 2. Estimated Relative Risk of Developing Erosive Esophagitis by Univariate Analysis in the Derivation Cohort (n = 5765)

	Crude Model			Age-adjusted Model		
Parameters	**Coefficient**	**RR (95% CI)**	***P***	**Coefficient**	**RR (95% CI)**	***P***
Age	0.0075	1.01 (0.99–1.02)	0.0611			
Sex	1.2693	3.56 (2.91–4.35)	<0.0001	1.2647	3.54 (2.89–4.33)	<0.0001
Anthropometric profile						
Body mass index	0.0979	1.10 (1.08–1.13)	<0.0001	0.0972	1.10 (1.08–1.13)	<0.0001
Waist circumference	0.0425	1.04 (1.04–1.05)	<0.0001	0.0425	1.04 (1.04–1.05)	<0.0001
Body fat	0.0151	1.02 (1.00–1.03)	0.0282	0.0143	1.01 (1.00–1.03)	0.0395
Blood pressure profile						
SBP	0.0045	1.00 (1.00–1.01)	0.0490	0.0038	1.00 (1.00–1.01)	0.1032
DBP	0.0137	1.01 (1.01–1.02)	<0.0001	0.0133	1.01 (1.01–1.02)	<0.0001
Lipid profile						
Total cholesterol	−0.0007	0.99 (0.99–1.00)	0.5088	−0.0008	0.99 (0.99–1.00)	0.4546
LDL	−0.0008	0.99 (0.99–1.00)	0.5001	−0.0001	0.99 (0.99–1.00)	0.4126
Triglyceride	0.5413	1.71 (1.47–2.00)	<0.0001	0.5394	1.72 (1.47–2.00)	<0.0001
HDL	−0.0161	0.98 (0.98–0.99)	<0.0001	−0.0160	0.98 (0.98–0.99)	<0.0001
ApoA1	−0.0026	0.99 (0.99–1.00)	0.2651	−0.0026	0.99 (0.99–1.00)	0.2649
ApoB	0.0029	1.00 (0.99–1.01)	0.2662	0.0026	1.00 (0.99–1.01)	0.2726
Insulin resistance profile						
Fasting glucose	0.0065	1.01 (1.00–1.01)	<0.0001	0.0063	1.01 (1.00–1.01)	<0.0001
HbA1c	0.1623	1.18 (1.09–1.27)	<0.0001	0.1533	1.17 (1.08–1.26)	0.0001
Comorbidity						
Hypertension	0.3910	1.48 (1.27–1.72)	<0.0001	0.3785	1.46 (1.24–1.71)	<0.0001
Dyslipidemia	0.2222	1.25 (1.05–1.48)	0.0100	0.2006	1.22 (1.03–1.45)	0.0217
Diabetes mellitus	0.4166	1.52 (1.22 -1.88)	0.0002	0.3868	1.47 (1.18–1.84)	0.0006
Medication						
Antihypertensive agent	0.3654	1.44 (1.23–1.68)	<0.0001	0.3528	1.42 (1.21–1.68)	<0.0001
Aspirin	0.1497	1.16 (0.95–1.43)	0.1559	0.1029	1.11 (0.89–1.37)	0.3477
NSAID	−0.1082	0.90 (0.55–1.46)	0.6612	−0.1547	0.86 (0.53–1.39)	0.5329
Lifestyle-related activities						
Alcohol drinking						
Never		1.00 (reference)			1.00 (reference)	
Former	0.7808	2.18 (1.56–3.06)	<0.0001	0.7927	2.21 (1.58–3.09)	<0.0001
Current	0.9613	2.62 (2.10–3.25)	<0.0001	0.9794	2.66 (2.14–3.31)	<0.0001
Smoking						
Never		1.00 (reference)			1.00 (reference)	
Former	0.6864	1.99 (1.68–2.35)	<0.0001	0.6670	1.95 (1.64–2.31)	<0.0001
Current	0.9787	2.66 (2.23–3.18)	<0.0001	0.9949	2.70 (2.26–3.23)	<0.0001
Regular physical exercise						
Yes		1.00 (reference)			1.00 (reference)	
No	0.2556	1.29 (1.02–1.64)	0.0350	0.2158	1.24 (0.97–1.58)	0.0816

Apo = apolipoprotein, CI = confidence interval, DBP = diastolic blood pressure, HbA1c = glycated hemoglobin, HDL = high-density lipoprotein, LDL = low-density lipoprotein, RR = relative risk, SBP = systolic blood pressure.

diabetes mellitus; use of antihypertensive agents; alcohol intake; smoking behavior; and physical exercise had *P* values <0.05 and were included as covariates in the multivariate logistic regression models. Those with a prognostic significance in univariate analysis were entered into the multivariate model, and stepwise regression was performed to eliminate multicollinearity. Five variables maintained their prognostic significance after multivariate analysis (Table 3): sex, hypertension, smoking behavior, BMI, and triglyceride level.

Model Discrimination and Calibration

The model was applied to derivation cohort patients resulting in a concordance statistic of 0.68 (95% confidence interval [CI], 0.66–0.70) and a Hosmer–Lemeshow goodness-of-fit statistics of 4.362 ($P = 0.823$). The model was then applied to patients from the validation cohort, resulting in a concordance statistic of 0.64 (95% CI, 0.62–0.66), indicating moderate discrimination (Figure 2A). Based on the derivation model, the probability for developing erosive esophagitis in the validation cohort was used to divide patients into deciles. In each of the deciles, the number of expected cases of erosive esophagitis was compared to the observed number of erosive esophagitis cases (Figure 3). Although Figure 3B indicates that the probability of the incidence of erosive esophagitis was slightly underestimated for low and intermediate-risk patients in the validation cohort, the Hosmer–Lemeshow goodness-of-

TABLE 3. Risk Factors Obtained From Multivariate Stepwise Regression Analysis for the Predictive Score Model For Developing Erosive Esophagitis in the Derivation Cohort (n = 5765)

Parameters	β Coefficient	RR (95% CI)	*P*	Risk Score
Sex				
Female		Reference group		0
Male	1.1529	2.86 (2.28–3.58)	<0.0001	2
Hypertension				
No		Reference group		0
Yes	0.2962	1.28 (1.08–1.51)	0.0019	1
Smoking				
Never/former		Reference group		0
Current	0.3637	1.34 (1.13–1.60)	0.0002	1
BMI				
<18.5		Reference group		0
18.5 – <23.0	0.1792	1.18 (0.52–2.68)	0.6804	0
23.0 – <25.0	0.3586	1.39 (0.61–3.16)	0.4089	1
≥25.0	0.3259	1.34 (0.59–3.04)	0.4527	1
Triglyceride				
<150 mg/dL		Reference group		0
≥150 mg/dL	0.3248	1.31 (1.12–1.54)	0.0003	1

BMI = body mass index, CI = confidence interval, RR = relative risk.

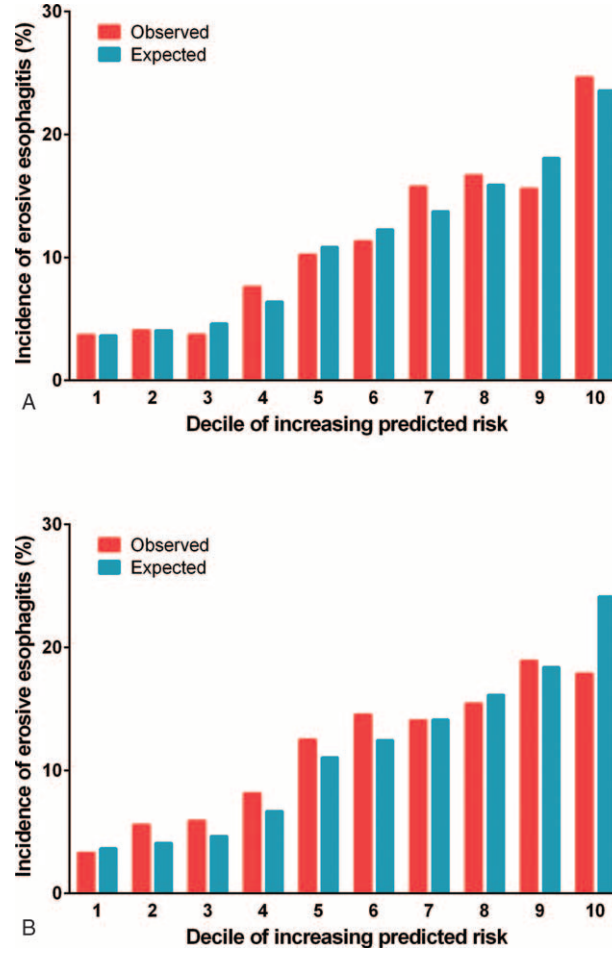

FIGURE 3. Expected versus observed incidence of developing erosive esophagitis by decile of the predicted risk. The derivation (A) and validation cohorts (B).

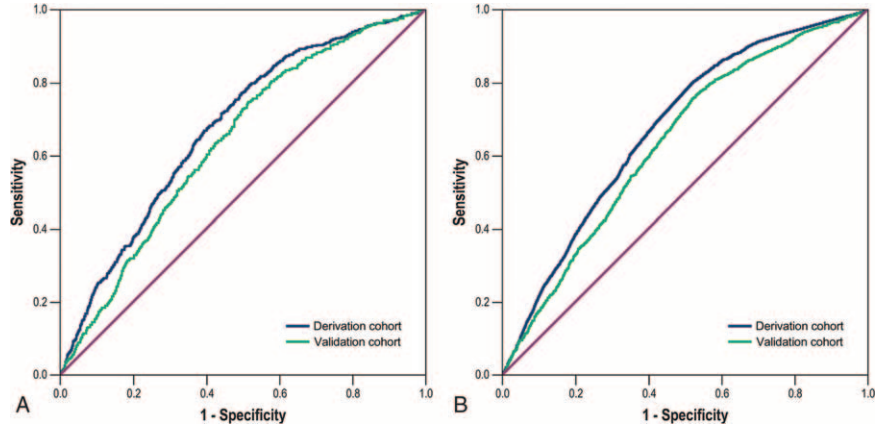

FIGURE 2. Receiver-operating characteristic curve analysis for the discriminative ability of the prediction model (A) and the Simple Prediction of Erosive Esophagitis Development scoring system (B) for the incidence of developing erosive esophagitis after negative index endoscopy in the derivation and validation cohorts.

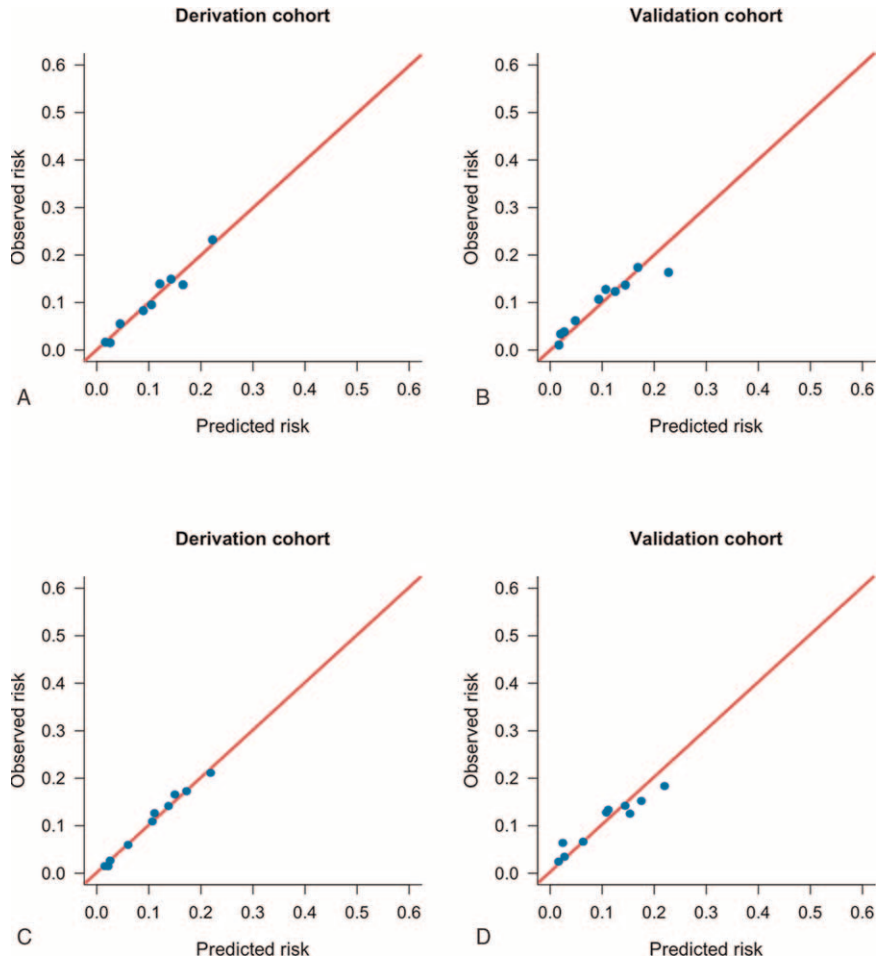

FIGURE 4. Calibration plot. A prediction model for the incidence of developing erosive esophagitis after negative index endoscopy in the derivation (A) and validation cohorts (B). The Simple Prediction of Erosive Esophagitis Development scoring system for the incidence of developing erosive esophagitis after negative index endoscopy in both the cohorts (C and D). The diagonal line indicates perfect calibration (ie, the predicted probabilities of erosive esophagitis are equal to the estimated probabilities of erosive esophagitis).

fit test showed no significant difference between the expected and observed number of erosive esophagitis ($\chi^2 = 12.949$, $P = 0.114$). The calibration curves (Figure 4A and B) show good agreement between the expected probability and observed incidence of erosive esophagitis risk across the observed range of risk.

Results of the model calibration performed in each category of risk factors of the validation dataset are shown in Table 4. Overall, this prediction model predicted that 667 participants would develop erosive esophagitis, compared with 670, who were observed for an expected-to-observed ratio of 1.00 (95% CI, 0.92–1.07). There was a significant overprediction in women, whereas an underestimation in men. Risk was also significantly underestimated in participant with current smoking, high BMI more than 25, and high level of triglyceride.

Simplified Risk Scoring Model

To calculate a risk score, we assigned each of the 6 prognostic variables a number of points proportional to its regression coefficient (Table 3). Summing the points yielded total risk scores ranging from 0 to 6, with higher scores indicating a greater predicted risk for the incidence of erosive esophagitis. This Simple Prediction of Erosive Esophagitis Development (SPEED) score predicted the incidence of erosive esophagitis well in the derivation and validation cohorts across all risk classes. As shown in Figure 2B, when applied to the derivation and validation cohorts, the risk-prediction score yielded a concordance statistic of 0.68 (95% CI, 0.66–0.70) and 0.64 (95% CI, 0.62–0.66), respectively, indicating moderate discrimination. For the cut-off point of 2.5, which maximizes the sum of sensitivity and specificity, the sensitivity, specificity, accuracy, negative predictive value, and positive predictive value were 0.81, 0.46, 0.51, 0.95, and 0.17 in the derivation cohort, and 0.77, 0.46, 0.50, 0.94, and 0.16 in the validation cohort, respectively. This scoring model also showed good calibration for predicting erosive esophagitis according to the Brier score (0.099 and 0.1 in the derivation and validation cohorts, respectively) (Figure 4C and D).

Figure 5 depicts the ratio of the expected-to-observed risk for developing erosive esophagitis in the validation cohort using the simplified score. The score calculated for each person from

TABLE 4. Comparison of the Number of Erosive Esophagitis Cases Predicted by the Model With the Number Observed in the Validation Cohort (n = 5770)

Parameters	Expected Number of Erosive Esophagitis	Observed Number of Erosive Esophagitis	E/O Ratio	95% CI
Overall	666.86	670	1.00	0.92–1.07
Sex				
Female	262.87	127	2.07	1.74–2.46
Male	403.98	543	0.74	0.68–0.81
Hypertension				
No	516.32	495	1.04	0.96–1.14
Yes	150.54	175	0.86	0.74–1.00
Smoking				
Never/former	543.95	473	1.15	1.05–1.26
Current	122.91	197	0.62	0.54–0.72
BMI				
<18.5	12.37	10	1.24	0.67–2.30
18.5 < ≤ 22.9	198.25	126	1.57	1.32–1.87
22.9 < ≤ 24.9	191.03	196	0.97	0.85–1.12
24.9 <	265.22	338	0.78	0.71–0.87
Triglyceride				
<150 mg/dL	488.42	448	1.09	0.99–1.20
≥150 mg/dL	178.43	222	0.80	0.70–0.92

BMI = body mass index, CI = confidence interval, E = expected, O = observed.

the validation set estimated the likelihood of detecting erosive esophagitis as 3.9% for patients with a score of 0, and 23.8% for patients with a score of 6 in the complete dataset. Patients were then divided into 3 categories based on the score distribution. Scores were collapsed into 3 categories: ≤2, 3 or 4, and ≥5 to identify potential low, intermediate, and high-risk individuals, respectively. Classification of the derivation cohort according to the risk score resulted in the assignment of 43.1% of patients to the low-risk group, 44.7% to the intermediate-risk group, and 12.2% to the high-risk group (Table 5). Results were similar for the validation cohort: 43.3% of patients were in the low-risk group, 42.9% in the intermediate-risk group, and 13.8% in the high-risk group. In the derivation cohort, the incidence rates of erosive esophagitis for the low, intermediate, and high-risk groups were 5.0%, 15.2%, and 22.7%, respectively. In the validation cohort, the incidence rates for the low, intermediate, and high-risk groups were 6.1%, 15.1%, and 18.2%, respectively.

DISCUSSION

In this study that involved a large-scale population that underwent regular health check-ups in the form of repeated screening endoscopy, we developed and validated a simple risk-prediction score (SPEED) to estimate the risk of developing the incidence of erosive esophagitis among asymptomatic patients. The final model includes male sex, current smoker, BMI, hypertension, and serum triglyceride levels as predictive variables. In internal validation, the model was well calibrated in that it predicted 667 cases of erosive esophagitis of 670 that occurred (E/O ratio 1.0). This simplified risk score is easy to use with readily available clinical, demographic, and laboratory values in primary healthcare services. To our knowledge, this is the first score designed to predict the incidence of developing erosive esophagitis in an asymptomatic population.

As erosive esophagitis is assumed an initial step in the development of esophageal adenocarcinoma, screening to identify patients with erosive esophagitis may be an effective strategy to prevent progression.[3,24] However, in terms of cost effectiveness of endoscopic screening, trying to identify patient groups in an entire population, as of now, is not recommended. With respect to Barrett esophagus as a long-term complication of erosive esophagitis, guidelines recommend selective screening for erosive esophagitis patients with multiple risk factors.[7,25–30] However, unlike symptom-based analysis of factors associated with progression to Barrett esophagus, analyses to predict the incidence of erosive esophagitis should be conducted in a population without symptoms or endoscopic esophagitis. Indeed, there is an inflammatory change in the cardiac mucosa in the asymptomatic, moderately overweight population, without evidence of typical symptom, suggesting the limitation of symptom-based risk stratification.[31] Therefore, comprehensive analysis for every possible factor is required for the development of a prediction model for developing erosive esophagitis.

Various epidemiologic studies regarding risk factors for erosive esophagitis pointed out that parameters including male sex, overweight, obesity, or smoking were consistently associated with erosive esophagitis.[10–13,18,32–35] Apart from this, clinical factors such as alcohol intake or dietary fat intake were suggested by some cross-sectional analyses.[36,37] However, these included factors have been derived from patients who were already diagnosed with erosive esophagitis. Therefore, these are not suitable for selecting high-risk patients beforehand and for making preemptive modifications of lifestyle-related factors. Recent prospective cohort studies, even if these were symptom-based studies rather than endoscopic studies, showed BMI gain and smoking to be risk factors; therefore, smoking cessation or weight reduction might be beneficial.[38–40] Providing preventive intervention or education at the time of negative endoscopy and using a prediction model with baseline clinical parameters are of vital importance, and may be worthwhile in the primary care setting.[41]

In our cohort study, the triglyceride level and diagnosis of hypertension as individual components of metabolic syndrome were independently associated with the incidence of erosive esophagitis.[42,43] Clinical parameters such as current smoker, male sex, and BMI were also definite risk factors for developing erosive esophagitis. In contrast to a previous cross-sectional study on a healthy population, the influence of waist circumference in our data was not evident.[44] Overall, this cohort analysis elucidated the temporal relationship between these putative risk factors and the eventual development of esophagitis. In particular, male sex, which has been suggested by many other studies, was identified as strong contributing factors invested with 2 points of risk score. The attributable risk of other factors was similar, given 1 point. Regarding performance measurement, even if this risk-prediction scoring system yielded moderate discrimination, this prediction model is of high value, considering there has never been a model like this, and it was derived from an asymptomatic population. Moreover, the present model was well calibrated overall, as verified by the validation set, which means that the observed risk of erosive esophagitis fit the predicted risk well. In the present model, there was a distinct difference in the incidence of erosive esophagitis among groups classified according to the risk score, giving the approximate 20% incidence in the high-risk group.

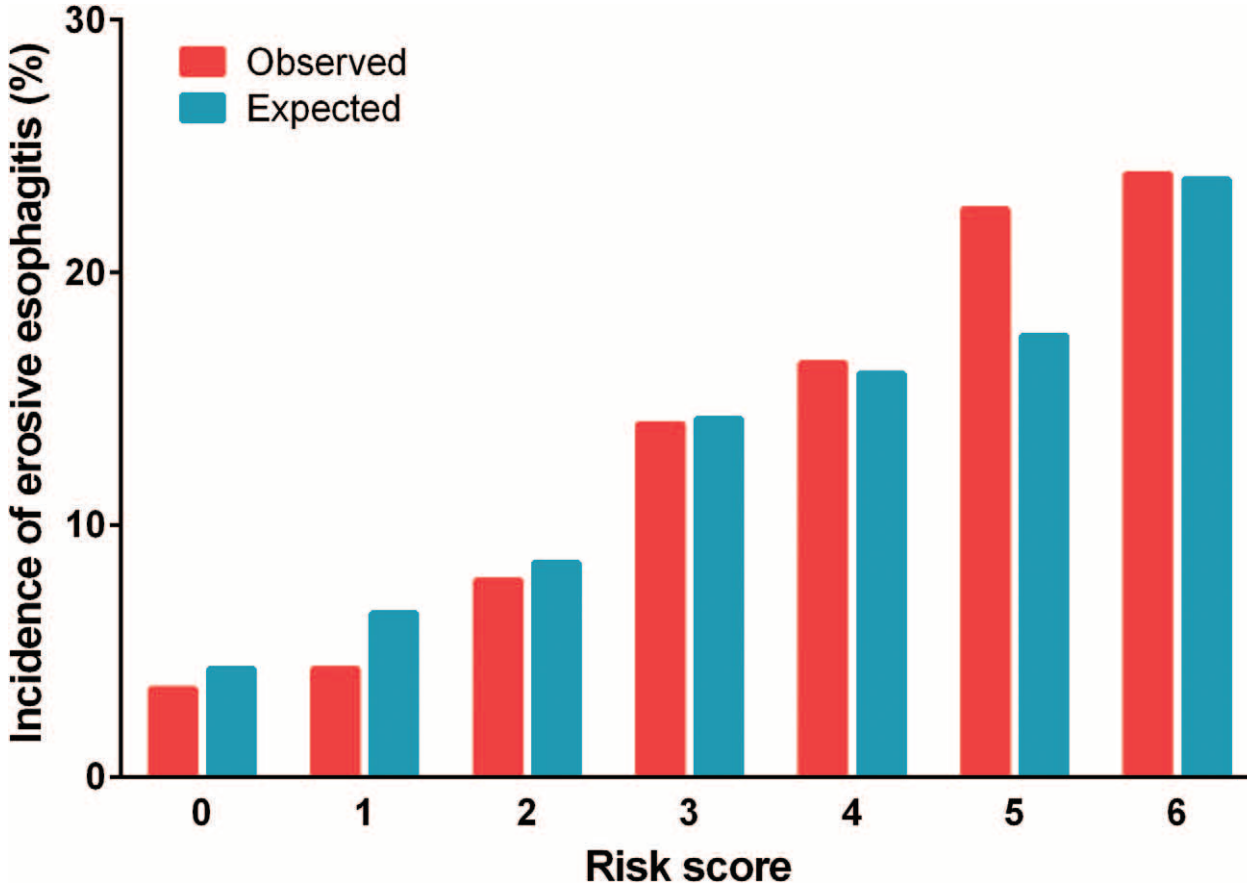

FIGURE 5. Ratio of expected-to-observed risk for developing erosive esophagitis. The simplified score is used in the validation cohort (n = 5770).

Clinical relevance of this risk-prediction model may lead to active participation in future endoscopic screening and correction of modifiable lifestyle factors in a selected high-risk group. Furthermore, for predicting progression from nonerosive reflux disease to erosive esophagitis, this risk score may be useful after combining the key symptom-based parameter to this scoring system in the future.

Regarding the clinical use of this prediction model, it is crucial to consider the epidemiologic aspect, as there is a significant difference in the incidence rate of erosive esophagitis between western and Asian populations. Whereas the reported prevalence of erosive esophagitis has varied from 3.4% to 16.3% in Asia, western studies revealed a rate of 12.1% to 28.5%. Barrett esophagus is regarded as a complication of chronic erosive esophagitis and has the potential to develop into esophageal adenocarcinoma. A population-based cohort study found the annual risk of esophageal adenocarcinoma to be 0.12% among patients with Barrett esophagus.[45] In the United States, however, the prevalence of Barrett esophagus was reported to be 6.8%.[46] Moreover, the incidence of esophageal adenocarcinoma in the United States increased by more than 460% in white men and 335% in white women between 1975 and 2004, outpacing squamous cell carcinoma to become the most predominant esophageal cancer in that country.[47] On

TABLE 5. Risk of the Incidence of Erosive Esophagitis in the Derivation and Validation Cohorts According to the Risk Group

Risk Groups	Derivation Cohort (n = 5765)			Validation Cohort (n = 5770)		
	n (%)	Incidence % (95% CI)	P^*	n (%)	Incidence % (95% CI)	P^*
Low	2487 (43.1)	5.0 (4.1–5.9)	Reference	2501 (43.3)	6.2 (5.2–7.1)	Reference
Intermediate	2576 (44.7)	15.2 (13.5–16.7)	<0.0001	2474 (42.9)	15.1 (13.5–16.6)	<0.0001
High	702 (12.2)	22.7 (19.1–26.2)	<0.0001	795 (13.8)	18.2 (15.2–21.3)	<0.0001

CI = confidence interval.
* By testing the difference incidence rate between low-risk group and other groups.

the contrary, the reported prevalence of Barrett esophagus in Asia shows that the disease is still rare across most of the continent. A Korean nationwide prospective multicenter study showed that Barrett esophagus was diagnosed in 1% of the population.[48] Although the incidence of esophageal adenocarcinoma has increased in Asia, it has done so only slightly. This seems to be because most Barrett esophagus cases in Asia are of the short segment type, and that this condition is associated with a limited risk of esophageal adenocarcinoma.[49] Because of this difference in the risk of developing premalignant or malignant esophageal disease from esophagitis, our risk-prediction model may be more useful for western populations. Moreover, anthropometric parameters that comprise of risk-scoring models should be defined differently in western populations. The current BMI cut-off points established by the WHO are $\geq 25\,kg/m^2$ for being overweight and $\geq 30\,kg/m^2$ for obesity. However, there is a high prevalence of various metabolic diseases in Asian populations even among those with BMIs lower than $25\,kg/m^2$.[50] Although Asian BMI criteria were applied to our risk scoring system, it is likely that they would require modification for western countries.

The major strength of the study was that it is the first cohort study involving participants with negative index screening endoscopy, which properly predicted the future risk of developing erosive esophagitis. In this analysis, a comprehensive approach to baseline anthropometric, demographic, and laboratory risk factors was conducted to minimize temporal bias. An additional value of the current scoring system is that it was created using a large and asymptomatic homogeneous population representative of primary care patients. Therefore, this validation of the risk-prediction score can lead to a useful clinical tool.

We acknowledge some potential limitations of this analysis. First, interobserver variations were not evaluated in the endoscopic diagnosis of esophagitis. However, all investigators in this study were highly experienced in endoscopic diagnosis. Second, our risk-prediction score was not validated on an external data set, so there are concerns about overfitting. Third, there seemed to be a recall error in self-reported risk factors, although this error is unlikely to be biased by outcomes because all baseline data were collected before endoscopy.

In conclusion, for the first time, a simple risk-prediction score (SPEED) has been developed based on sex, smoking habits, BMI, hypertension, and the triglyceride level for predicting the incidence of developing erosive esophagitis in asymptomatic patients. Application of this score may help to identify patients at an increased risk of erosive esophagitis in the future, and assist in stratifying asymptomatic populations into low, intermediate, and high-risk categories for erosive esophagitis, which will improve patient management through appropriate individualized education and screening. This risk score needs to be validated in a western population.

ACKNOWLEDGMENTS

Author contributions: Soo Hoon Kang—acquisition of data, and analysis and interpretation of data; Yaeji Lim—acquisition of data and statistical analysis; Hyuk Lee and Hee Jung Son—study concept and design, and drafting of the manuscript; Joungyoun Kim and Sangah Chi—statistical analysis; Yang Won Min, Byung-Hoon Min, Jun Haeng Lee, Seungho Ryu, Poong-Lyul Rhee, and Jae J. Kim—critical revision of the manuscript.

REFERENCES

1. Sandler RS, Everhart JE, Donowitz M, et al. The burden of selected digestive diseases in the United States. *Gastroenterology.* 2002;122:1500–1511.
2. Peery AF, Dellon ES, Lund J, et al. Burden of gastrointestinal disease in the United States: 2012 update. *Gastroenterology.* 2012;143:1179–1187e1171-1173.
3. Ronkainen J, Talley NJ, Storskrubb T, et al. Erosive esophagitis is a risk factor for Barrett's esophagus: a community-based endoscopic follow-up study. *Am J Gastroenterol.* 2011;106:1946–1952.
4. Spechler SJ. Barrett esophagus and risk of esophageal cancer: a clinical review. *JAMA.* 2013;310:627–636.
5. Chang JT, Katzka DA. Gastroesophageal reflux disease, Barrett esophagus, and esophageal adenocarcinoma. *Arch Intern Med.* 2004;164:1482–1488.
6. Fass R, Dickman R. Clinical consequences of silent gastroesophageal reflux disease. *Curr Gastroenterol Rep.* 2006;8:195–201.
7. Committee ASoPMuthusamy VR, Lightdale JR, et al. . The role of endoscopy in the management of GERD. *Gastrointest Endosc.* 2015;81:1305–1310.
8. Katz PO, Gerson LB, Vela MF. Guidelines for the diagnosis and management of gastroesophageal reflux disease. *Am J Gastroenterol.* 2013;108:308–328quiz 329.
9. Kahrilas PJ, Shaheen NJ, Vaezi MF, et al., American Gastroenterological A. American Gastroenterological Association Medical Position Statement on the management of gastroesophageal reflux disease. *Gastroenterology.* 2008;135:1383–13911381-1385, 1391 e.
10. Avidan B, Sonnenberg A, Schnell TG, et al. Risk factors for erosive reflux esophagitis: a case-control study. *Am J Gastroenterol.* 2001;96:41–46.
11. Labenz J, Jaspersen D, Kulig M, et al. Risk factors for erosive esophagitis: a multivariate analysis based on the ProGERD study initiative. *Am J Gastroenterol.* 2004;99:1652–1656.
12. El-Serag HB, Graham DY, Satia JA, et al. Obesity is an independent risk factor for GERD symptoms and erosive esophagitis. *Am J Gastroenterol.* 2005;100:1243–1250.
13. Nam SY, Choi IJ, Ryu KH, et al. Abdominal visceral adipose tissue volume is associated with increased risk of erosive esophagitis in men and women. *Gastroenterology.* 2010;139:1902–1911e1902.
14. Lee YC, Yen AM, Tai JJ, et al. The effect of metabolic risk factors on the natural course of gastro-oesophageal reflux disease. *Gut.* 2009;58:174–181.
15. Lundell LR, Dent J, Bennett JR, et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. *Gut.* 1999;45:172–180.
16. Willett WC, Dietz WH, Colditz GA. Guidelines for healthy weight. *N Engl J Med.* 1999;341:427–434.
17. World Health Organization. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser.* 2000;894:i–xii,1–253.
18. Kim N, Lee SW, Cho SI, et al., Gerd Study Group of Korean College of H., Upper Gastrointestinal R. The prevalence of and risk factors for erosive oesophagitis and non-erosive reflux disease: a nationwide multicentre prospective study in Korea. *Aliment Pharmacol Ther.* 2008;27:173–185.
19. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996;15:361–387.
20. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010;21:128–138.

21. Hosmer DW, Hosmer T, Le Cessie S, et al. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med.* 1997;16:965–980.
22. Paul P, Pennell ML, Lemeshow S. Standardizing the power of the Hosmer-Lemeshow goodness of fit test in large data sets. *Stat Med.* 2013;32:67–80.
23. Redelmeier DA, Bloch DA, Hickam DH. Assessing predictive accuracy: how to compare Brier scores. *J Clin Epidemiol.* 1991;44:1141–1146.
24. Shaheen N, Ransohoff DF. Gastroesophageal reflux, Barrett esophagus, and esophageal cancer: scientific review. *JAMA.* 2002;287:1972–1981.
25. Shaheen NJ, Weinberg DS, Denberg TD, et al., Clinical Guidelines Committee of the American College of P. Upper endoscopy for gastroesophageal reflux disease: best practice advice from the clinical guidelines committee of the American College of Physicians. *Ann Intern Med.* 2012;157:808–816.
26. Rubenstein JH, Morgenstern H, Appelman H, et al. Prediction of Barrett's esophagus among men. *Am J Gastroenterol.* 2013;108: 353–362.
27. Rubenstein JH. Risk factors for Barrett's esophagus. *Curr Opin Gastroenterol.* 2014;30:408–414.
28. Rubenstein JH, Thrift AP. Risk factors and populations at risk: selection of patients for screening for Barrett's oesophagus. *Best Pract Res Clin Gastroenterol.* 2015;29:41–50.
29. Thrift AP, Garcia JM, El-Serag HB. A multibiomarker risk score helps predict risk for Barrett's esophagus. *Clin Gastroenterol Hepatol.* 2014;12:1267–1271.
30. Thrift AP, Hilal J, El-Serag HB. Metabolic syndrome and the risk of Barrett's oesophagus in white males. *Aliment Pharmacol Ther.* 2015;41:1182–1189.
31. Robertson EV, Derakhshan MH, Wirz AA, et al. Central obesity in asymptomatic volunteers is associated with increased intrasphincteric acid reflux and lengthening of the cardiac mucosa. *Gastroenterology.* 2013;145:730–739.
32. Chung SJ, Kim D, Park MJ, et al. Metabolic syndrome and visceral obesity as risk factors for reflux oesophagitis: a cross-sectional case-control study of 7078 Koreans undergoing health check-ups. *Gut.* 2008;57:1360–1365.
33. Hsu CS, Wang PC, Chen JH, et al. Increasing insulin resistance is associated with increased severity and prevalence of gastro-oesophageal reflux disease. *Aliment Pharmacol Ther.* 2011;34:994–1004.
34. Singh S, Sharma AN, Murad MH, et al. Central adiposity is associated with increased risk of esophageal inflammation, metaplasia, and adenocarcinoma: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol.* 2013;11:1399–1412e1397.
35. Nam SY, Choi IJ, Nam BH, et al. Obesity and weight gain as risk factors for erosive oesophagitis in men. *Aliment Pharmacol Ther.* 2009;29:1042–1052.
36. Anderson LA, Cantwell MM, Watson RG, et al. The association between alcohol and reflux esophagitis, Barrett's esophagus, and esophageal adenocarcinoma. *Gastroenterology.* 2009;136: 799–805.
37. Mulholland HG, Cantwell MM, Anderson LA, et al. Glycemic index, carbohydrate and fiber intakes and risk of reflux esophagitis, Barrett's esophagus, and esophageal adenocarcinoma. *Cancer Causes Control.* 2009;20:279–288.
38. Hallan A, Bomme M, Hveem K, et al. Risk factors on the development of new-onset gastroesophageal reflux symptoms. A population-based prospective cohort study: the HUNT study. *Am J Gastroenterol.* 2015;110:393–400quiz 401.
39. Ness-Jensen E, Lindam A, Lagergren J, et al. Weight loss and reduction in gastroesophageal reflux. A prospective population-based cohort study: the HUNT study. *Am J Gastroenterol.* 2013;108: 376–382.
40. Ness-Jensen E, Lindam A, Lagergren J, et al. Tobacco smoking cessation and improved gastroesophageal reflux: a prospective population-based cohort study: the HUNT study. *Am J Gastroenterol.* 2014;109:171–177.
41. Kaltenbach T, Crockett S, Gerson LB. Are lifestyle measures effective in patients with gastroesophageal reflux disease? An evidence-based approach. *Arch Intern Med.* 2006;166:965–971.
42. Niigaki M, Adachi K, Hirakawa K, et al. Association between metabolic syndrome and prevalence of gastroesophageal reflux disease in a health screening facility in Japan. *J Gastroenterol.* 2013;48:463–472.
43. Moki F, Kusano M, Mizuide M, et al. Association between reflux oesophagitis and features of the metabolic syndrome in Japan. *Aliment Pharmacol Ther.* 2007;26:1069–1075.
44. Yasuhara H, Miyake Y, Toyokawa T, et al. Large waist circumference is a risk factor for reflux esophagitis in Japanese males. *Digestion.* 2010;81:181–187.
45. Hvid-Jensen F, Pedersen L, Drewes AM, et al. Incidence of adenocarcinoma among patients with Barrett's esophagus. *N Engl J Med.* 2011;365:1375–1383.
46. Sharma P, Wani S, Romero Y, et al. Racial and geographic issues in gastroesophageal reflux disease. *Am J Gastroenterol.* 2008;103:2669–2680.
47. Ho KY. From GERD to Barrett's esophagus: is the pattern in Asia mirroring that in the West? *J Gastroenterol Hepatol.* 2011;26: 816–824.
48. Lee IS, Choi SC, Shim KN, et al. Prevalence of Barrett's esophagus remains low in the Korean population: nationwide cross-sectional prospective multicenter study. *Dig Dis Sci.* 2010;55:1932–1939.
49. Lee HS, Jeon SW. Barrett esophagus in Asia: same disease with different pattern. *Clin Endosc.* 2014;47:15–22.
50. Low S, Chin MC, Ma S, et al. Rationale for redefining obesity in Asians. *Ann Acad Med Singapore.* 2009;38:66–69.