Open Access Full Text Article

ORIGINAL RESEARCH

Validation of a risk prediction model for Barrett's esophagus in an Australian population

Colin J Ireland[1]
Andrea L Gordon[2]
Sarah K Thompson[3]
David I Watson[4]
David C Whiteman[5]
Richard L Reed[6]
Adrian Esterman[1,7]

[1]School of Nursing and Midwifery, Division of Health Sciences, University of South Australia, Adelaide, SA, Australia; [2]School of Pharmacy and Medical Science, Division of Health Sciences, University of South Australia, Adelaide, SA, Australia; [3]Discipline of Surgery, University of Adelaide, Adelaide, SA, Australia; [4]Department of Surgery, Flinders University, Bedford Park, SA, Australia; [5]Population Health Department, QIMR Berghofer Medical Research Institute, Herston, QLD, Australia; [6]Discipline of General Practice, Flinders University, Bedford Park, SA, Australia; [7]Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, QLD, Australia

Correspondence: Colin J Ireland
School of Nursing and Midwifery, University of South Australia, GPO Box 2471, Adelaide, South Australia 5001, Australia
Tel +61 4 2969 0502
Fax +61 8 8302 2168
Email Colin.Ireland@mymail.unisa.edu.au

Background: Esophageal adenocarcinoma is a disease that has a high mortality rate, the only known precursor being Barrett's esophagus (BE). While screening for BE is not cost-effective at the population level, targeted screening might be beneficial. We have developed a risk prediction model to identify people with BE, and here we present the external validation of this model.

Materials and methods: A cohort study was undertaken to validate a risk prediction model for BE. Individuals with endoscopy and histopathology proven BE completed a questionnaire containing variables previously identified as risk factors for this condition. Their responses were combined with data from a population sample for analysis. Risk scores were derived for each participant. Overall performance of the risk prediction model in terms of calibration and discrimination was assessed.

Results: Scores from 95 individuals with BE and 636 individuals from the general population were analyzed. The Brier score was 0.118, suggesting reasonable overall performance. The area under the receiver operating characteristic was 0.83 (95% CI 0.78–0.87). The Hosmer–Lemeshow statistic was $p=0.14$. Minimizing false positives and false negatives, the model achieved a sensitivity of 74% and a specificity of 73%.

Conclusion: This study has validated a risk prediction model for BE that has a higher sensitivity than previous models.

Keywords: Barrett's esophagus, risk prediction model, screening, validation

Introduction

Esophageal adenocarcinoma has a 5-year survival rate of <20%,[1] making it a lethal disease. The proportion of patients with metastatic spread at the time of diagnosis has been increasing,[2] potentially contributing to low survival rates. Barrett's esophagus (BE), a condition where some of the lining of the distal esophagus undergoes a metaplastic change to resemble tissue similar to that normally found in the intestine,[3] is the only known precursor to esophageal adenocarcinoma.[1,4] BE progresses stepwise via low- and high-grade dysplasia to cancer.[5]

Screening for BE presents an opportunity to detect individuals with a greater risk of adenocarcinoma, facilitating targeted endoscopic surveillance potentially resulting in earlier diagnosis of high-grade dysplasia or early cancer, thereby allowing more effective interventions. Population-level screening for BE has not been shown to be cost-effective.[4,6] Current recommendations are that screening for BE could be considered for men with chronic gastroesophageal reflux disease and two or more risk factors including >50 years, central obesity, first-degree family history of BE

or esophageal adenocarcinoma, a smoking history, and Caucasian race.[7,8] This recommendation would still require substantial resources if screening were to be widespread.[9] Methods to further facilitate targeted screening have not been evaluated and might be of benefit.

We have previously developed and published a risk prediction model with internal validation.[10] Briefly, a case–control study was undertaken in 2015 to identify potential variables that might better predict BE. The resulting risk prediction model included the variables such as age, gender, individual history of hypertension, individual history of acid regurgitation, first-degree family history of reflux, number of alcoholic drinks per week, and body mass index (BMI). The model showed good discrimination (area under the receiver operating characteristic [AUC] 0.82; 95% CI 0.78–0.87) and calibration (Hosmer–Lemeshow test p=0.67) during development.[10]

However, for a model to be clinically applicable, it needs to be externally validated, as model performance generally decreases when applied to different populations.[11] The purpose of this current study was to externally validate the accuracy of the previously developed risk prediction tool.[10]

Materials and methods

This study is reported according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) checklist for prediction model development and validation.[12] A cohort study was undertaken to validate the previously developed model. In November 2016, potential new participants with BE were mailed study information, a self-reporting questionnaire, and a reply-paid envelope to return completed questionnaires. After 3 weeks, those who had not returned the questionnaire received the information again as a reminder; no further contact was made following this. Retrospective population data, people without BE, were obtained from a study undertaken in Queensland and reported in 2009.[13] Participants were recruited from different institutions than the initial development group for the risk prediction model.

Ethics statement

Ethical approval was granted by the Southern Adelaide Clinical Human Research Ethics Committee (#433.14), the University of South Australia Human Research Ethics Committee (#0000033915), and QIMR Berghofer Medical Research Institute Human Research Ethics Committee (#P514). For the participants with BE, consent to use data was implied if the questionnaire was returned.

Study participants

Potential participants with BE were identified using the endoscopy database from the Royal Adelaide Hospital, a major tertiary hospital within South Australia, Australia. BE definition was the same used in the development of the model – that is, BE ≥2 cm with intestinal metaplasia (IM). Histological reports were used to confirm IM.

The population data were provided by QIMR Berghofer Medical Research Institute in Queensland, Australia. The data had been collected for a previous study of BE and the participants had provided written consent for the data to be used in subsequent studies. These participants lived in Queensland and had been recruited at random from the Australian Electoral Roll (voter registration is compulsory in Australia).[13] They did not undergo an endoscopy to exclude BE, but respondents who answered "yes" to the question regarding previous BE diagnosis were omitted from this study (1.24%). This is comparable with the prevalence of BE within the previous model development group (1.3%)[10] and aligns with the suggested prevalence of BE (1.2%–1.6%) within the general population.[3]

Questionnaire

The self-reported questionnaire for BE patients consisted of 10 questions, 8 relating to the 7 predictors within the model (height and weight were collected to calculate BMI) and 2 additional questions to capture participant demographic characteristics. These questions used in this validation study were identical to those from the derivation study. Importantly, the questionnaire items used for this validation study and the derivation study were based on those used in the Queensland study,[13] which serves as an independent validation here.

Statistical analysis

Returned questionnaire information was entered into the Stata 14 software package (StataCorp, 2015, Stata Statistical Software: Release 14; StataCorp LP, College Station, TX, USA). BMI was then calculated (weight [kg]/Height [m]^2). For BE participants, variables recording an individual's history of reflux and hypertension were modified to calculate pre-BE status. For example, those diagnosed with hypertension following the diagnosis of BE were considered not to have hypertension for analysis, consistent with the process followed during the development phase.[10] This information was then added to the population sample data.

The Queensland control population dataset contained all but one of the items required for the prediction equation, namely, family history of reflux. As such, this variable

was imputed by developing a predictive model for family history of reflux from the original series of controls from the development dataset, and then applying this model to the new Queensland control population. To do this, a bootstrapped aggregated stepwise logistic regression process was undertaken to reduce the list of 12 potential predictors of family history of reflux to a smaller set of predictors.[14] This involved taking 1000 bootstrapped samples and selecting those variables included in the model 500 or more times. Five variables remained at this stage as the best joint predictors of a family history of reflux, namely individual acid reflux, smoking, BMI, height, and history of hernia. Logistic regression, with a family history of reflux as the dependent variable and the other variables as independent variables, was then undertaken. A receiver operating characteristic analysis was performed, and this identified that a cutoff of ≥0.237 in the predicted value was indicative of a family history of reflux.

The resulting equation from the development population to obtain the predicted value for a family history of reflux was then applied to the Queensland population dataset, using the following formula: predicted = exp(x)/(1 + exp(x)), where x is the output from the equation for each individual. The 0.237 cutoff was used to convert family history of reflux into a 0 (no)/1 (yes) variable resulting in 80.24% without and 19.76% with a family history of reflux, which is comparable to the development population (75.64% without and 24.36% with).

Validation

The risk prediction model was run to generate participants' probability of having BE. The formula used is as follows:

Probability of BE = exp(f)/(1 + exp(f))

where "f" is $-7.003635 + (0.0645714 \times age) + (0.8431104 \times gender) + (-0.9861185 \times history of hypertension) + (1.051748 \times history of acid regurgitation) + (1.186615 \times family history of reflux) + (0.2971902 \times alcoholic drinks per week) + (0.6125564 \times BMI)$.[10] Table 1 provides the different predictor scorings.

Example of model in use

The following scenario of a BE patient demonstrates the model in practice: a 44-year-old male reports that he has been experiencing a sour taste in his mouth at times and he has been using antacid tablets (Quick-eze®; Nestlé Australia Ltd, Rhodes, NSW, Australia) for symptom relief. Lately, they have not been as effective in relieving the symptoms. His father had suffered from reflux; he drank an average of

Table 1 Scoring of predictors for BE risk prediction model

Risk factor	Measure	Scoring
Age (years)	Continuous	Continuous
Gender	Female	0
	Male	1
History of hypertension (treated with medication)	No	0
	Yes	1
History of acid regurgitation (a sour taste from acid rising up into your mouth or throat)	No	0
	Yes	1
First-degree family history of reflux	No	0
	Yes	1
Alcoholic drinks per week	Do not drink	0
	≤4	1
	5–13	2
	14–27	3
	≥28	4
BMI (kg/m²)	<25.00	0
	25.00–29.99	1
	≥30.00	2

Abbreviations: BE, Barrett's esophagus; BMI, body mass index.

30 beers per week and his BMI was 28. Other observations were unremarkable and no other issues were found.

Using the coding in Table 1 and the model formula:

f = −7.003635 + (0.0645714×**44**) + (0.8431104×**1**) + (−0.9861185×**0**) + (1.051748×**1**) + (1.186615×**1**) + (0.2971902×**4**) + (0.6125564×**1**).

The probability of BE=0.67; therefore, this person has a 67% chance of having BE. The bold fonts indicate the score that has been given for each variable.

Assessment

The overall risk prediction model performance was initially assessed using the Brier score, a test of how well a probability prediction performs, by measuring the mean square error of the predicted probability of the patient having BE compared to their actual disease status, and is used for binomial outcomes. A score of zero (0) represents perfect performance, while a score of 0.25 represents a 50% chance of having BE.[15] The model was then assessed for discrimination and calibration. Discrimination, the ability of the model to predict someone with BE or not, was assessed using the AUC.[16] An AUC of 0.5 indicates random chance, the closer to 1.0, the more accurate the test is at discriminating a case from a non-case in any pair of individuals,[16] in this case BE. Interpretation of AUC results can be further refined: 0.51–0.69 least accurate, 0.7–0.9 moderately accurate, and >0.9 highly accurate.[17] Calibration, how well the predicted

probabilities align with the observed probability of being an actual case, was assessed using the calibration curve, where a slope of 1 equals good alignment and a slope <1 indicates model overfitting,[15] and the Hosmer–Lemeshow goodness-of-fit test, where a larger *p*-value indicates a better fit[1] and the Lowess smoother calibration plot.

Results

Two hundred seventy-one questionnaires were mailed to BE patients and 95 were returned (35% response rate). Data were available from 644 general population sample participants (58% response rate);[13] of these, 8 had indicated a diagnosis of BE and hence were not included in this analysis, leaving 636 population samples.

Demographics

Mean age for BE patients was 66.8 years (SD 10.5), with the population sample being 57.9 years (SD 11.3), *p*<0.001. No statistical difference was seen in gender of both groups (70.53% in the BE group and 64.47% in the population group were male [*p*=0.30]). Table 2 provides a summary of the demographic and clinical variables between patients with BE and the population sample.

Testing of the imputation formula for family history of reflux

The imputation equation was assessed. The AUC was 0.74 (95% CI 0.67–0.82), Brier score was 0.154, and the calibration curve slope was 0.99 (*p*<0.001) with an intercept of 4.14e-07 (*p*=1.0). This indicated that the imputation was moderately accurate at discriminating family history of reflux. Overall calibration was also reasonably good; however, there was a slight overprediction at the upper end of the calibration curve.

Validation

To assess overall risk prediction model performance, the Brier score was obtained (0.118), which indicated reasonable overall performance.[15] The calibration curve slope was 1.14 (*p*<0.001) with an intercept of −1.08 (*p*<0.001), indicating the model is not overfitted. The Hosmer–Lemeshow statistic was *p*=0.14, identifying that the predicted risks are not significantly different from observed risks. The Lowess smoother calibration plot was derived (Figure 1). This shows the tool to be well calibrated in predicting those who are at ultrahigh risk (needing to be investigated) and those who are predicted to be at low risk (not needing investigation). False positives will be identified at the upper middle of the prediction range.

Table 2 Demographic and clinical information

Variables	Barrett's esophagus patients, n=95	Population sample, n=636	*p*-value[a]
Age, years, mean (SD)	66.8 (10.5)	57.9 (11.3)	<0.001
Gender, n (%)			
Male	67 (70.53)	410 (64.47)	0.30
Female	28 (29.47)	226 (35.53)	
History of hypertension, n (%)			
Yes	33 (34.74)	91 (14.31)	<0.001
No	62 (65.26)	539 (84.75)	
Missing	0	6 (0.94)	
History of acid regurgitation, n (%)			
Yes	73 (76.84)	25 (3.93)	<0.001
No	22 (23.16)	555 (87.26)	
Missing	0	56 (8.81)	
First-degree family history of reflux, n (%)			
Yes	54 (56.84)	113 (17.77)	<0.001
No	39 (41.05)	459 (72.17)	
Missing	2 (2.11)	64 (10.06)	
Alcoholic drinks per week, n (%)			
Do not drink	23 (24.21)	67 (10.55)	<0.001
≤4	28 (29.47)	160 (25.16)	
5–13	35 (36.84)	191 (30.03)	
14–27	6 (6.32)	98 (15.41)	
≥28	3 (3.16)	76 (11.94)	
Missing	0	44 (6.92)	
BMI, kg/m², n (%)			
<25.00	21 (22.11)	230 (36.17)	0.001
25.00–29.99	36 (37.89)	263 (41.35)	
≥30.00	37 (38.95)	136 (21.38)	
Missing	1 (1.05)	7 (1.10)	

Note: [a]*p*-values were calculated using *t*-test for continuous variables and Fisher's exact test for categorical variables.
Abbreviation: BMI, body mass index.

The AUC was 0.83 (95% CI 0.78–0.87; Figure 2), with 74% sensitivity and 73% specificity. To maximize sensitivity and specificity, the Youden Index was calculated,[18] providing the empirical cut off point (0.317955) to minimize false positives and false negatives.

Clinical applicability of the model needs to be considered with any prediction tool. The current study has calculated the number of patients who would be sent for endoscopy if various probability thresholds were applied. Positive predictive values (PPV) and negative predictive values (NPV) were also calculated for each cut off point (Table 3). PPV identifies the probability of a person whose test results are positive will actually have the disease, while NPV shows the probability of the person whose test results are negative will be disease free.[19] These were based on the estimated prevalence of undiagnosed BE within the general population (1.6%).[3] This
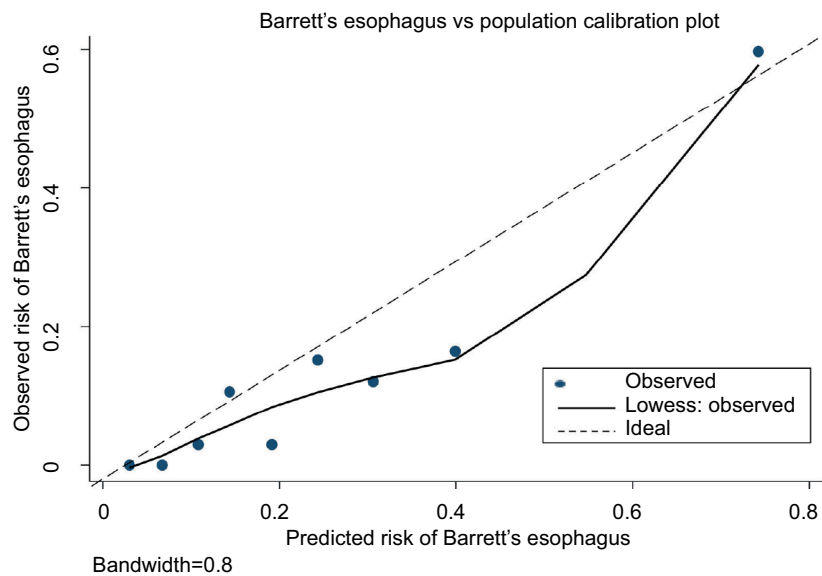
Figure 1 Logistic calibration for predicted versus observed risk.
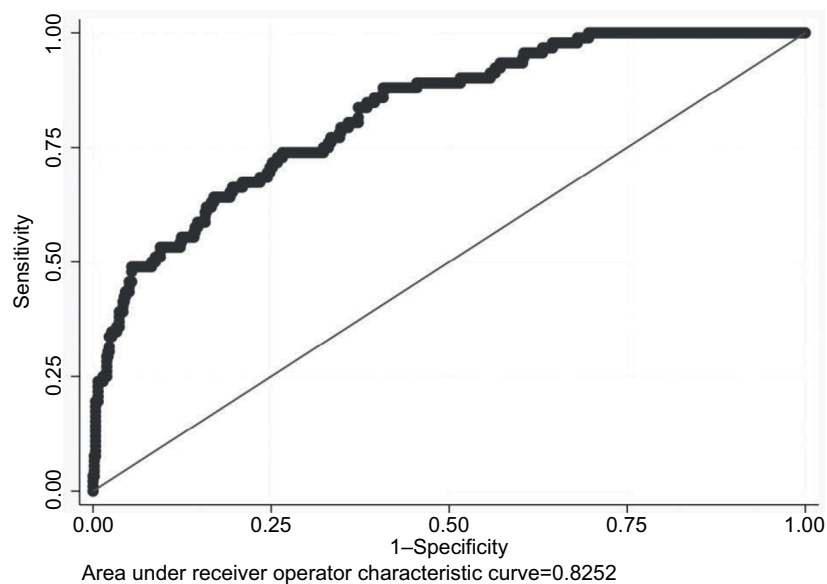

Figure 2 Area under the receiver operator characteristic BE patients versus population.
Abbreviation: BE, Barrett's esophagus.

indicates that the number of false positives will be high, while the number of false negatives will be negligible.

Discussion

We have validated a risk prediction model for BE, developed for use within primary health care to identify individuals who are at potentially higher risk and may benefit from further investigation.

Five BE risk prediction tools have been developed over the last 15 years, with varying results; however, none have been sufficiently accurate to implement in clinical practice.[1,20–23] Three of these models have now been validated (Table 4).[1,22,24] These results are on the low end of moderate accuracy at predicting BE.[17] It has also been reported that tests with an AUC ≤0.75 have minimal clinical usefulness.[25] Therefore, none of the above mentioned models meet this clinical usefulness

Table 3 Probability threshold performance in a general population

Probability threshold for having BE	Sensitivity (%)	Specificity (%)	Patients sent for endoscopy (%)	Positive predictive value (%)	Negative predictive value (%)
≥0.1	100	26	76	0.7	100
≥0.2	89	51	51.2	2.9	99.7
0.27[a]	77	67	35.1	3.7	99.4
≥0.3	74	71	31	4.0	99.4
0.317955[b]	74	73	29	4.3	99.4
≥0.4	64	82	19.9	5.5	99.3
≥0.5	53	88	13.6	6.7	99.1
≥0.6	49	94	7.5	11.7	99.1
≥0.7	30	98	3.0	19.6	98.9
≥0.8	15	100	0.5	100	98.6
≥0.9	3	100	0.1	100	98.4

Notes: [a]Empirical cutoff from model development to minimize false positives and false negatives. [b]Empirical cutoff from validation data to minimize false positives and false negatives.
Abbreviation: BE, Barrett's esophagus.

Table 4 Comparison of previous risk prediction models for BE

Author (year)	Country	Sample size development	Development AUC	Sample size validation	Validation AUC
Gerson et al,[21] (2001)	USA	517	0.72	NA	Not performed
Locke et al,[20] (2003)	USA	1009	0.76	NA	Not performed
Thrift et al,[1] (2012)	Australia/USA	706	0.70	593	0.61
Rubenstein et al,[23] (2013)[a]	USA	822	0.72	1: 716	1: 0.71
				2: 302	2: 0.70
				3: 256	3: 0.72
				4: 118	4: 0.70
Liu et al,[22] (2014)	UK	1603	0.81	478	0.64

Note: [a]Validation undertaken by an independent author on four datasets.[24]
Abbreviations: AUC, area under the receiver operating characteristic; BE, Barrett's esophagus; NA, not applicable.

criterion. These models used statistical significance to determine the final variables, potentially negating some variables that could have improved the overall AUC.[26,27]

Using statistical significance alone (*p*<0.05) to select variables could potentially remove important predictors that have an association with the outcome, especially in small datasets.[26,27] Variables associated with but not necessarily having a causal link to the outcome should also be considered in the selection process, as these may not always be statistically significant during the analysis.[27]

To obtain the best performing predictors within our model, bootstrapping with stepwise logistic regression followed by aggregation was used rather than statistical significance. This resulted in the inclusion of three predictors that have not been included in previous models: history of hypertension, family history of reflux, and the number of alcoholic drinks per week. The remaining variables within our model have been used in various combinations in other

prediction models.[1,20–23] This achieved an AUC of 0.83 on validation, indicating a moderately accurate test that might be clinically useful.[17,25]

Selecting a referral cut off point requires a balance between disease severity and cost. For a disease with a high mortality rate that could be reduced with early treatment, a higher sensitivity compared with specificity may be accepted for screening tests. However, the consequence is more people being sent for investigation, and a high false-positive rate. A higher specificity could be acceptable for diseases with low consequence as not identifying the disease may have minimal impact.[28] While BE is not fatal, it can progress to adenocarcinoma (which carries a high mortality rate) at a rate between 0.3% and 0.6% per year.[29] As the prevalence of BE is low within the general population with low adenocarcinoma progression rates,[29] a higher specificity and PPV should be considered to reduce the number of costly and invasive endoscopies undertaken unnecessarily.

If there were an intermediate step between the risk prediction tool and endoscopy, identifying the 30 out of 100 people who warrant further investigation may be more acceptable. An example is the cytosponge, currently under investigation.[30] This risk prediction model coupled with the use of nonendoscopic screening methods for BE (e.g., the cytosponge) has the potential of being clinically useful.

Current BE screening recommendations suggest that men with chronic gastroesophageal reflux disease and two or more risk factors, including >50 years, central obesity, first-degree family history BE or esophageal adenocarcinoma, a smoking history, and Caucasian race, should be considered.[7,8] Others have suggested that this be reviewed and the age be lowered to <50 years of age for males,[31] particularly if they have had gastroesophageal reflux disease symptoms from an early age.[32] If this recommendation was adopted, it would place further pressure on already limited resources. We contend that using the proposed screening strategy described earlier could reduce this.

Potential limitations to this study include that the population sample did not undergo an endoscopy to exclude BE; however, the prevalence identified is comparable to the suggested prevalence within the general population. As the estimated prevalence of BE within the general population is low,[3] the magnitude of this effect would be minimal, with 2 of the 636 participants used as controls potentially having BE. The population sample had been collected previously without asking about a family history of reflux, and therefore, this variable had to be imputed. This may produce different results compared with having this variable collected at the time. However, the model without family history of reflux had a pseudo-R-squared of 0.543, and the addition of this variable only added another 0.003 to the R-squared, indicating that a family history of reflux plays only a minor contribution to the predictive model. The development and validation of the model occurred within Australia, although from different geographical locations; thus, the results may vary if undertaken in different countries due to differing sociodemographic factors. Finally, the low response rate within the BE group (35%) could potentially introduce a nonresponse bias. Ideally, a prospective trial should be undertaken to further validate these results.

In summary, this study externally validated a BE risk prediction model that shows higher sensitivity than previous models. These results are promising; however, a prospective trial should be undertaken to investigate the application of the model within clinical practice. If the model continues to perform well in prospective trials, it has the potential to improve the identification of patients at greater risk of having BE.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Thrift AP, Kendall BJ, Pandeya N, Vaughan TL, Whiteman DC; Study of Digestive Health. A clinical risk prediction model for Barrett esophagus. *Cancer Prev Res (Phila).* 2012;5(9):1115–1123.
2. Stavrou EP, McElroy HJ, Baker DF, Smith G, Bishop JF. Adenocarcinoma of the oesophagus: incidence and survival rates in New South Wales, 1972–2005. *Med J Aust.* 2009;191(6):310–314.
3. Zimmerman TG. Common questions about Barrett esophagus. *Am Fam Physician.* 2014;89(2):92–98.
4. Whiteman DC, Appleyard M, Bahin FF, et al. Australian clinical practice guidelines for the diagnosis and management of Barrett's esophagus and early esophageal adenocarcinoma. *J Gastroenterol Hepatol.* 2015; 30(5):804–820.
5. Gopal DV, Reichelderfer M, Gaumnitz EA, Harter J, Jobe BA. Barrett's esophagus: is screening and surveillance justified? *Dis Manag Health Out.* 2004;12(6):353–361.
6. Kadri SR, Lao-Sirieix P, O'Donovan M, et al. Acceptability and accuracy of a non-endoscopic screening test for Barrett's oesophagus in primary care: cohort study. *BMJ.* 2010;341:c4372.
7. Shaheen NJ, Falk GW, Iyer PG, Gerson LB; American College of Gastroenterology. ACG clinical guideline: diagnosis and management of Barrett's esophagus. *Am J Gastroenterol.* 2016;111(1):30–50.
8. Fitzgerald RC, di Pietro M, Ragunath K, et al. British society of gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut.* 2014;63(1):7–42.

9. Vaughan TL, Fitzgerald RC. Precision prevention of oesophageal adenocarcinoma. *Nat Rev Gastroenterol Hepatol.* 2015;12(4):243–248.

10. Ireland CJ, Fielder AL, Thompson SK, Laws TA, Watson DI, Esterman A. Development of a risk prediction model for Barrett's esophagus in an Australian population. *Dis Esophagus.* 2017;30(11):1–8.

11. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med.* 2016;35(2):214–226.

12. Collins GS, Reitsma JB, Altman DG, Moons KM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): the tripod statement. *Ann Intern Med.* 2015;162(1):55–63.

13. Smith KJ, O'Brien SM, Green AC, Webb PM, Whiteman DC; Study of Digestive Health. Current and past smoking significantly increase risk for Barrett's esophagus. *Clin Gastroenterol Hepatol.* 2009;7(8): 840–848.

14. Brieman L. Bagging predictors. *Mach Learn.* 1996;24(2):123–140.

15. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.*2010;21(1):128–138.

16. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem.* 2008;54(1):17–23.

17. Greiner M, Pfeiffer D, Smith RD. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev Vet Med.* 2000;45(1–2):23–41.

18. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *BiomJ.*2005;47(4):458–472.

19. Buttner P, Muller R. *Epidemiology.* 1st ed. South Melbourne, Australia: Oxford University Press; 2011.

20. Locke GR, Zinsmeister AR, Talley NJ. Can symptoms predict endoscopic findings in GERD? *Gastrointest Endosc.* 2003;58(5): 661–670.

21. Gerson LB, Edson R, Lavori PW, Triadafilopoulos G. Use of a simple symptom questionnaire to predict Barrett's esophagus in patients with symptoms of gastroesophageal reflux. *Am J Gastroenterol.* 2001;96(7):2005–2012.

22. Liu X, Wong A, Kadri SR, et al. Gastro-esophageal reflux disease symptoms and demographic factors as a pre-screening tool for Barrett's esophagus. *PLoS One.* 2014;9(4):e94163.

23. Rubenstein JH, Morgenstern H, Appelman H, et al. Prediction of Barrett's esophagus among men. *Am J Gastroenterol.* 2013;108(3):353–362.

24. Thrift AP, Vaughan TL, Anderson LA, Whiteman DC, El-Serag HB. External validation of the Michigan Barrett Esophagus Prediction Tool. *Clin Gastroenterol Hepatol.* 2017;15(7):1124–1126.

25. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. *CJEM.* 2006;8(1):19–20.

26. Hendriksen JM, Geersing GJ, Moons KG, de Groot JA. Diagnostic and prognostic prediction models. *J Thromb Haemost.* 2013;11(Suppl 1):129–141.

27. Moons KG, Kengne AP, Woodward M, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio) marker. *Heart.* 2012;98(9):683–690.

28. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatr.*2007;96(5):644–647.

29. Schoofs N, Bisschops R, Prenen H. Progression of Barrett's esophagus toward esophageal adenocarcinoma: an overview. *Ann Gastroenterol.* 2017;30(1):1–6.

30. Tan WK, di Pietro M, Fitzgerald RC. Past, present and future of Barrett's oesophagus. *Eur J Surg Oncol.* 2017;43(7):1148–1160.

31. Desilets DJ, Nathanson BH, Navab F. Barrett's esophagus in practice: gender and screening issues. *J Men's Health.* 2014;11(4):177–182.

32. Thrift AP, Kramer JR, Qureshi Z, Richardson PA, El-Serag HB. Age at onset of GERD symptoms predicts risk of Barrett's esophagus. *Am J Gastroenterol.* 2013;108(6):915–922.