Open

Assessing the Feasibility of Targeted Screening for Esophageal Adenocarcinoma Based on Individual Risk Assessment in a Population-Based Cohort Study in Norway (The HUNT Study)

Shao-Hua Xie, BMed, PhD[1], Eivind Ness-Jensen, MD, PhD[1,2,3], Nils Medefelt[1] and Jesper Lagergren, MD, PhD[1,4]

OBJECTIVES: **Unselected screening for oesophageal adenocarcinoma (OAC) is not justified due to the low absolute risk in the general population. This study aimed to evaluate a risk prediction model in identifying high-risk individuals who might be considered for targeted screening.**

METHODS: **A population-based cohort of 62,576 participants was recruited in 1995–1997 in Nord-Trøndelag County, Norway (HUNT) and followed up until 31 December 2015. A model for predicting individuals' absolute risk of OAC was developed using competing-risk regression. The Lorenz curve was used to assess the concentration of OAC patients in high-risk individuals and the feasibility of targeted screening based on individual risk assessment.**

RESULTS: **During 1,085,137 person-years of follow-up, 29 incident cases of OAC occurred. The model included risk factors for OAC, in which male sex, older age, gastro-oesophageal reflux symptoms, obesity, and tobacco smoking predicted higher risk of OAC. The area under the receiver operating characteristic curve for 10-year risk of OAC was 0.71 (95% confidence interval 0.57–0.85) and for 15-year risk was 0.84 (95% confidence interval 0.76–0.91) after 10-fold cross-validation, with good agreements between observed and predicted risks. The Lorenz curve indicated that 33% of all OAC cases would have occurred in the 5% of the population with the highest risks within 15 years, and 61% of all cases in the top 10% of the population.**

CONCLUSIONS: **Individual risk assessment based on known risk factors for OAC has the potential to identify a selected high-risk group of individuals who may benefit from screening for early detection.**

SUPPLEMENTARY MATERIAL accompanies this paper at https://doi.org/10.1038/s41395-018-0069-9

Am J Gastroenterol (2018) 113:829–835. https://doi.org/10.1038/s41395-018-0069-9

INTRODUCTION

The past four decades have witnessed a rapidly increasing incidence of oesophageal adenocarcinoma (OAC) in many Western countries, e.g., in North America, Europe, and Australia [1–5]. OAC is usually diagnosed at advanced stages and thus carries a poor prognosis with an overall 5-year survival lower than 20% in most developed countries [2, 5–7]. Detection at an early tumour stage would substantially reduce the mortality [2, 6]. Endoscopy provides an opportunity for early detection of OAC or its pre-malignant condition Barrett's oesophagus with dysplasia [8, 9]. However, universal endoscopic screening is not feasible given the low incidence in the general population, considerable costs, and complications and inconveniences associated with endoscopy [9, 10]. Instead, it would be preferable to target a limited group of individuals with high absolute risk who may benefit from endoscopic screening.

Risk prediction modelling based on information on risk factors has shown promising usefulness in selecting individuals at high absolute risk of cardiovascular disease and breast, colorectal, and lung cancers [11–14]. Yet, only two prediction models have been

[1]Upper Gastrointestinal Surgery, Department of Molecular Medicine and Surgery, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden. [2]Department of Public Health and Nursing, HUNT Research Centre, NTNU, Norwegian University of Science and Technology, Levanger, Norway. [3]Medical Department, Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway. [4]School of Cancer and Pharmaceutical Sciences, King's College London, London, UK. **Correspondence:** S.-H.X. (email: shaohua.xie@ki.se)

Received 27 November 2017; accepted 20 February 2018; published online 11 May 2018

© The Author(s) 2018

developed for OAC, one in an Australian population and the other in a Swedish population [15–17]. The models provided by these studies were promising, but remain to be validated in external populations, and because they were based on case–control studies, they also require validation in terms of risk calibration in a prospective design.

Therefore, we aimed to evaluate a prediction model for identifying individuals at high absolute risk of OAC based on a prospective population-based cohort study in Norway.

METHODS

Study design and data source

The Nord-Trøndelag Health Study (HUNT) is a prospective cohort study including the majority of adult residents in Nord-Trøndelag County in Norway. HUNT consists of a series of population-based health surveys initiated in the mid-1980s. These surveys include data from questionnaires and clinical measurements on basic demographics, health-related behaviours, symptoms, diseases, and body composition of the participants [18]. The present study was based on data from the second survey (HUNT2) conducted in 1995–1997, which was the first survey that included the assessment of gastro-oesophageal reflux symptoms. Among 93,898 adult residents aged over 20 years in the county, 64,975 (69%) participated by answering the questionnaires and taking part in the clinical measurements. For the purpose of the present study, 2399 participants with any prior cancer diagnosis were excluded, and thus the final study cohort consisted of 62,576 participants. All participants were followed up from the date of taking part in the HUNT2 survey until the date of diagnosis of OAC, death, emigration, or the end of the study (31 December 2015), whichever occurred first. Information on incident OAC cases, death, and emigration was retrieved through linkage to the nationwide Cancer Registry of Norway, and Statistics Norway, respectively. The linkages of individuals' data between registries were enabled by the unique personal identity number, which was available for each study participant.

Predictors and model development

The pre-defined predictors evaluated in the model were five well-established risk factors for OAC, i.e., age, sex, gastro-oesophageal reflux symptoms, obesity, and tobacco smoking status [2, 5]. Data on age and sex were collected at the time of participation in HUNT. Gastro-oesophageal reflux symptoms were assessed using questionnaires, where participants were asked: 'To what degree have you had heartburn or acid regurgitation during the previous 12 months?' and replied with one of three response alternatives: 'no complaints', 'minor complaints', or 'severe complaints' [19, 20]. Use of this questionnaire has been well validated as representing gastro-oesophageal reflux disease in a previous validation study in HUNT. The severe complaints answer responded well to reflux disease with a high specificity of 99.5%. Among the participants who reported minor symptoms in the validation study, 15% had daily symptoms or had used anti-reflux medication on a daily basis, 10% had at least weekly symptoms, and 75% had symptoms less frequently than weekly [19]. Weight and height were measured under standardised conditions by trained personnel at screening stations, providing an objective assessment of body mass index (BMI), which was calculated by dividing weight in kilograms by the squared height in metres (kg/m^2). Obesity was defined according to the definition stated by the World Health Organisation as a BMI of 30 or above. Participants were asked using the questionnaires about their smoking status and were classified as current daily smokers, previous daily smokers, or never daily smokers [21].

Competing-risk regression was used to assess associations between the candidate predictors and the risk of OAC and to estimate the absolute risk of OAC in individual participants. The competing-risk regression is based on the method of Fine and Gray and estimates the sub-hazard ratio (SHR) for OAC in the presence of competing risk from death [22–24]. The absolute risks of OAC within 10 years and within 15 years from participation were projected for individuals with different combinations of risk factors by the cumulative incidence functions derived from the competing-risk regression. The cumulative incidence function estimates the probability of developing OAC in the presence of competing risks from death within a specific period of time. We dichotomised sex, reflux symptoms (yes, either minor or severe, or no), obesity (BMI <30 or ≥ 30), and daily tobacco smoking (ever or never) and grouped age into four categories (<50, 50–59, 60–69, or ≥ 70 years). Interactions were not considered due to the limited number of OAC cases.

Analysis of performance and validation

We assessed the model performance in terms of discriminative accuracy using two measurements, i.e., the area under the receiver operating characteristic curve (AUC) and Somers' *D* statistic. The AUC measures the model's ability to discriminate participants developing OAC from those who do not, after a specific time period of follow-up based on predicted risk at baseline, whereas Somers' *D* estimates the association between the predicted risks and observed outcomes [25]. To overcome the problem of overfitting when the model performance is evaluated with the same dataset as used in the model building, the statistics were re-calculated using the 10-fold cross-validation strategy. We randomly divided all participants into 10 approximately equal groups, and predicted the risks for the participants in each group from a model excluding this group. The unbiased AUC and Somers' *D* statistics were calculated with the predicted risks based on the 10-fold cross-validation. The calibration was examined by plotting the agreement between the cumulative predicted risks and cumulative observed proportions of OAC cases across tenths of predicted risk. Those with predicted risks below the 60th percentile were combined into one group because no OAC cases occurred. Due to the limited number of OAC patients in each risk category, statistical calibration test or calibration slope were not calculated [26].

Lorenz curve

The Lorenz curve is an established graphical representation of the distribution of income or of wealth in econometrics, which has also been proposed for evaluating the 'concentration' of disease risks in the population [17, 27]. In this study, we used the Lorenz curve to assess the concentration of OAC patients occurring in individuals at high risk based on the prediction model. We ranked

all participants based on their predicted risks from the model. For each level of predicted risk, we calculated the cumulative proportion of participants with this level of risk or below to the entire cohort and, correspondingly, the cumulative proportion of OAC patients occurring among these participants within 10 years or 15 years of follow-up. We plotted the cumulative proportion of the population (x-axis) against the cumulative proportion of cases in each risk level (y-axis). The greater the degree to which cases are concentrated in a small number of high-risk individuals, the greater the deflection of the Lorenz curve downward from the line of equality, and the higher yield of screening programme targeting high-risk subsets of the population.

External validation of the Swedish prediction model

Using data from HUNT2, we externally assessed a previous prediction model derived from a nationwide population-based case–control study in Sweden which included 189 histologically confirmed incident cases of OAC and 820 age- and sex-matched control subjects [16]. We modified the Swedish model so that the included predictors, i.e. age, sex, reflux symptoms, BMI, and tobacco smoking, were in the same categories as in the current study (Supplementary Table **1**). Methods of model development have been described in detail previously [16]. This Swedish model predicts the 5-year risk of OAC, which was used as a proxy for long-term risk hereby and validated against the occurrence of OAC in 15 years among participants of the HUNT2 cohort. The performance of this Swedish model was assessed in terms of both discriminative accuracy and risk calibration.

All statistical analyses were performed using the statistical software SAS version 9.4 (SAS Institute, Cary, NC). All P values are two-sided, and a P value of less than 0.05 was considered statistically significant.

Ethical approval

The study was approved by the Regional Committee for Medical and Health Research Ethics in Norway (reference number 2016/2040/REK sør-øst A). All participants in HUNT gave written informed consent when participating stating that their data could be used in future medical research, including linkage to other registries.

RESULTS

Study participants

The final study cohort of 62,576 participants included 29,509 (47.2%) males and 33,067 (52.8%) females, with a mean age at baseline of 49.6 years. Some baseline characteristics of the study participants are shown in Table 1. Among all participants, 16.4% were obese (BMI ≥ 30), 28.2% had reflux symptoms, and 55.2% were daily tobacco smokers. During a total of 1,085,137 person-years of follow-up, 29 incident cases of OAC occurred.

Table 1 Basic characteristics of study participants

Characteristic at baseline	Number of participants (%)	Number of person-years
Total	62,576 (100)	1,085,137
Sex		
Male	29,509 (47.2)	502,955
Female	33,067 (52.8)	582,182
Age, years		
<50	33,718 (53.9)	647,632
50–59	10,605 (17.0)	195,264
60–69	8473 (13.5)	137,076
≥70	9780 (15.6)	105,164
Mean ± standard deviation	49.6±17.1	
Body mass index		
<30	51,855 (82.9)	908,454
≥30	10,236 (16.4)	171,537
Missing	485 (0.8)	5146
Mean ± standard deviation	26.3±4.1	
Reflux symptoms		
No	38,931 (62.2)	689,832
Minor	14,637 (23.4)	256,689
Severe	3024 (4.8)	52,477
Missing	5984 (9.6)	86,139
Daily tobacco smoking		
Never	26,643 (42.6)	469,403
Former	16,675 (26.6)	281,828
Current	17,911 (28.6)	313,882
Missing	1347 (2.2)	20,024

Competing-risk regression

The competing-risk regressions showed that male sex, older age, presence of gastro-oesophageal reflux symptoms, obesity, and daily tobacco smoking predicted higher risk of OAC (Table 2). The predicted absolute risk of OAC varied greatly according to individuals' risk factor profiles; the 15-year risk ranged from 3.6/100,000 to 292.5/100,000 (Table 3). The highest 15-year absolute risk of OAC (292.5/100,000) was predicted for males aged 60–69 years who had reflux symptoms, were obese, and smoked daily. In this group, 342 individuals are needed to be surveyed to detect one OAC case within 15 years. An interactive web tool to estimate an individual's absolute risk of OAC in 15 years based on this model, the Oesophageal Adenocarcinoma Risk Assessment Tool, can be accessed at https://sites.google.com/view/oacrisk. A screen shot of this online risk assessment tool is shown in Supplementary Figure **1**.

Model performance

Table 4 shows the discriminative accuracy of the model with and without 10-fold cross-validation. The AUC for the risk of OAC within 10 years was 0.81 (95% confidence interval (CI) 0.70–0.91) and for the risk within 15 years was 0.88 (95% CI 0.83–0.93). The AUC statistics declined after cross-validation for the 10-year risk (AUC 0.71, 95% CI 0.57–0.85) and slightly also for the 15-year

© The Author(s) 2018

Table 2 Associations between risk factors and oesophageal adenocarcinoma from competing-risk regression, expressed as sub-hazard ratios (SHR) with 95% confidence interval (CI) (*N* = 62,576)

Variables at baseline	Number of cases	Crude SHR (95% CI)	Adjusted SHR (95% CI)[a]
Sex			
Female	9	1.0 (Reference)	1.0 (Reference)
Male	20	2.5 (1.1–5.5)	1.9 (0.8–4.2)
Age, years			
<50	8	1.0 (Reference)	1.0 (Reference)
50–59	6	2.4 (0.8–6.9)	2.1 (0.7–6.0)
60–69	8	4.0 (1.5–10.6)	3.2 (1.1–8.9)
≥70	7	3.0 (1.1–8.3)	3.1 (1.1–8.9)
Body mass index			
<30	21	1.0 (Reference)	1.0 (Reference)
≥30	8	1.9 (0.9–4.4)	1.8 (0.7–4.1)
Reflux symptoms			
No	9	1.0 (Reference)	1.0 (Reference)
Yes	18	4.4 (2.0–9.8)	3.7 (1.6–8.4)
Daily tobacco smoking			
Never	6	1.0 (Reference)	1.0 (Reference)
Ever	23	3.0 (1.2–7.3)	2.1 (0.8–5.5)

[a]Participants with missing values of predictors ($n = 7210$) were excluded from the multivariate competing-risk regression

risk (AUC 0.84, 95% CI 0.76–0.91). The calibration plot showed good agreement between observed and predicted cumulative risks of OAC within 15 years (Fig. 1).

Lorenz curve

Figure 2 shows the Lorenz curve constructed with individuals' estimated risks of OAC within 10 years and within 15 years based on the prediction model. The analysis indicated substantial concentration of OAC cases in individuals at high risk as estimated by the prediction model. Among all OAC cases occurring within 15 years, 33% would have occurred in the 5% of the population of the highest risks, and 61% in the top 10% of the population. The thresholds of the 15-year risk to include the top 5, 10, and 20% of the population at the highest risks and the corresponding numbers of individuals to be surveyed to detect one OAC case are also illustrated in Fig. 2.

External validation of the Swedish prediction model

Applying the Swedish risk prediction model to the HUNT2 population gave an AUC of 0.89 (95% CI 0.84–0.94) for the 15-year risk of OAC, and showed good calibration between observed and predicted cumulative risks (Supplementary Figure **2**). However, most data points in the calibration plot are above the line of perfect agreement (Supplementary Figure **2**), indicating that the predicted cumulative risks are generally lower than the observed cumulative risks in individuals with relatively lower risks.

DISCUSSION

This study provides a model for predicting individuals' absolute risk of OAC based on a panel of established risk factors derived from a prospective cohort study. The model had good discriminative accuracy, particularly for predicting 15-year risk of OAC. The agreement between the observed and predicted risks showed good calibration. The Lorenz curve suggested a clear concentration of OAC cases in high-risk individuals based on the developed prediction model, and thus indicated a high-yield target group for potential endoscopy screening.

To our knowledge, this is the first risk prediction model of OAC developed from a prospective study. The cohort design counteracted information misclassification as compared with standard case–control studies and enabled evaluation of the model performance in terms of risk calibration. Other strengths of this study include the population-based design, objectively assessed anthropometric measures, well-validated data on reflux symptoms [19, 20], and complete and long follow-up. A weakness is the small number of OAC cases occurring during follow-up, which is due to the limited number of person-years of follow-up in this low-incidence population of OAC. This reduced the precision of the predicted risk estimates and the ability to categorise the predictors into several groups. In addition, lack of information on use of anti-reflux medications in this study might have led to reduced model precision. A larger sample size with longer follow-up and more detailed information on risk factors may enable greater granularity in terms of further classifying reflux symptoms based on severity or frequency, more precise categorisation of BMI, and quantitative assessment of tobacco smoking and, consequently, improve the performance of the model in identifying individuals with extremely high risk of OAC. Although the model performance was assessed with the cross-validation strategy, there remains a risk of model over-fitting because the model performance was assessed with data on individuals in the same cohort as used to build the model. Therefore, an external validation of the prediction model in an independent population would be ideal. Moreover, as the Lorenz curve is dependent on the distribution of risk factors in the population and may differ across populations, the curve shown in this study may not be generalisable to other populations.

To the best of our knowledge, there are only two previous prediction models for estimating the absolute risk of OAC, and these were based on data from case–control studies in Australia and Sweden [15, 16]. The predictors in these models included the five predictors used in the present study together with other factors. These earlier models also showed good discriminative accuracy with the AUC statistics ranging from 0.75 to 0.85 after cross-validation, depending on which predictors were included. The two previous case–control studies were unable to assess the calibration of the predicted and observed risks. External validation of the previous Swedish model using data from the HUNT cohort provided an AUC even higher than that for the original logistic models, which could be explained by contribution from the two additional important predictors, i.e., age and sex, in the final model. When applying the Swedish model to the HUNT population, the calibration plot indicated over-fitting of the model, i.e.,

Table 3 Estimated 15-year cumulative risk of oesophageal adenocarcinoma with selected profiles of risk factors

Profile	Sex	Age (years)	Reflux symptoms	Obesity	Daily tobacco smoking	15-Year risk per 100,000	Number to survey to detect one case within 15 years
1	Female	<50	No	No	Never	3.6	27,908
2	Female	50–59	Yes	No	Ever	58.8	1701
3	Female	50–59	Yes	Yes	Ever	102.9	915
4	Female	60–69	Yes	Yes	Ever	157.4	635
5	Male	<50	No	No	Never	6.7	15,008
6	Male	<50	Yes	No	Never	24.6	4061
7	Male	50–59	Yes	No	Never	50.9	1962
8	Male	<50	No	Yes	Ever	25.0	3994
9	Male	≥70	Yes	No	Ever	162.9	614
10	Male	60–69	Yes	Yes	Ever	292.5	342

Table 4 Discriminative ability of the model for predicting 10-year and 15-year risks of oesophageal adenocarcinoma, expressed as area under the receiver operating characteristic curve (AUC) with 95% confidence interval (CI) and Somers' *D*

	10-Year risk		15-Year risk	
	AUC (95% CI)	Somers' *D*	AUC (95% CI)	Somers' *D*
Derivation cohort	0.81 (0.70–0.91)	0.61	0.88 (0.83–0.93)	0.76
10-Fold cross-validation	0.71 (0.57–0.85)	0.41	0.84 (0.76–0.91)	0.67

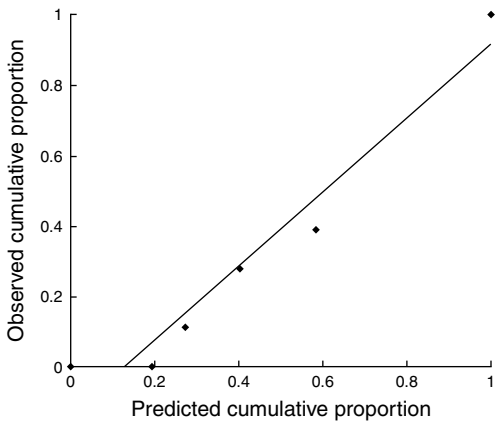
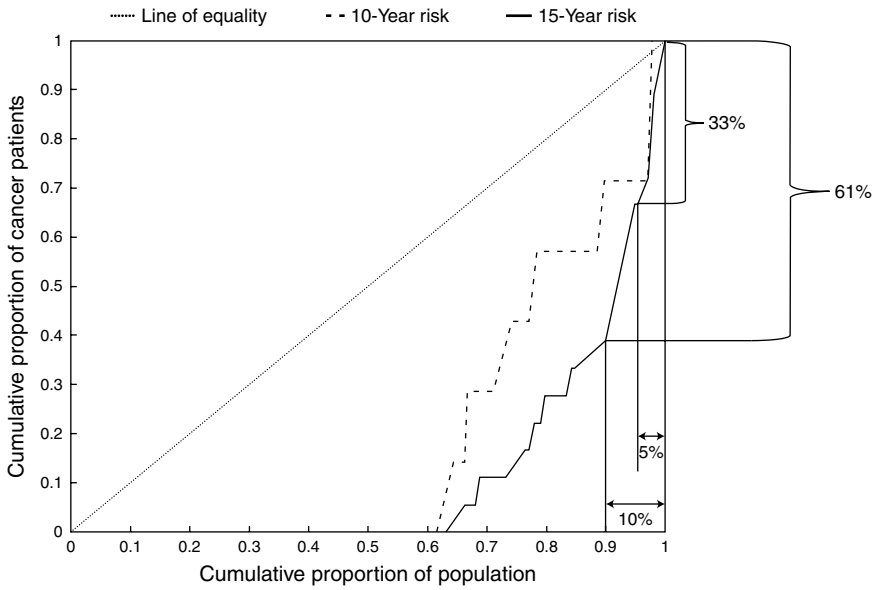


Fig. 1 Calibration of observed cumulative proportion of oesophageal adenocarcinoma patients and cumulative predicted risk in 15 years

underestimated risks in low-risk individuals and overestimated risks in high-risk individuals. However, the calibration plots in this study should be interpreted with caution considering the small number of observed cancer patients in each risk category. Based on information on the five major risk factors for OAC only, the prediction model provided in this study is the simplest, and yet the discriminative ability was promising with a high AUC (0.84) for 15-year risk after cross-validation. Compared with risk prediction models for other cancers, the models used for OAC, including the present one, seem to have higher discriminative accuracy as measured by the AUC statistics. The Lorenz curve in the present study indicated higher concentration compared with the curve based on the previous Swedish study in which OAC cases and control subjects were matched by age and sex, confirming the importance of considering age and sex for targeted prevention and screening.

Despite decades of efforts to develop the treatment, the prognosis of OAC has improved to only a limited extent [5–7]. However, unselective endoscopic screening or surveillance to detect early and thus curable OAC is not justified given the low absolute risk of OAC even in reflux patients (~20/100,000 per year) [10]. In this study, the Lorenz curve indicated that targeting the top 10% individuals based on individual risk assessment would have identified 61% of all OAC cases occurring within 15 years and probably more cases with even longer follow-up, while in as many as 28% of the total population who had reflux, only 62% of all OAC cases occurred. Therefore, an individual risk assessment based on prediction modelling of some key factors could help better tailor screening programmes compared with relying on reflux symptoms only (surveying 10% versus 28% of the population to have the same yield). In contrast, none of the OAC cases in this study were from the bottom 60% of the population with the lowest estimated risks. These individuals generally corresponded to non-smoking men and women younger than 50 years and without reflux symptoms, and thus should not be considered for screening. The Clinical Guidelines Committee of the American College of Physicians recommends endoscopic investigation in men aged over 50 years with long-term reflux symptoms and additional risk factors, but not in the general population [28]. The conceptual basis for this recommendation is similar to an individual risk assessment based on multiple risk factors in seeking for a more precise high-risk group, but a risk prediction model would provide a more objective tool to assist clinical decision making and public health recommendations. Particularly, despite the male predominance in the

© The Author(s) 2018



Population at high risk	% Of cancer patients	Threshold 15-year risk	Number to survey to detect 1 case
Top 5%	33%	191/100,000	523
Top 10%	61%	109/100,000	915
Top 20%	72%	53/100,000	1894

Fig. 2 Lorenz curve for prediction of oesophageal adenocarcinoma based on individual risk assessment. The horizontal axis denotes cumulative proportion of all cohort members according to their predicted risks from the lowest to the highest; the vertical axis denotes the cumulative proportion of oesophageal adenocarcinoma cases which occurred from the corresponding proportion of population

incidence of OAC, female patients still account for a substantial proportion (1/7 in European populations) of all OAC cases [2]. The current guidelines do not recommend screening in women, regardless of reflux symptoms, considering the low average risk in women, while a risk prediction model may provide more individualised risk assessment and identify women with considerably high risk who may still benefit from screening. Because information on the risk factors included in the present model can be easily obtained, the model can be straightforwardly applied in clinical practice, particularly with the help of a user-friendly online risk assessment tool, and is readily amenable for external validation. Further validation and evaluation of the risk prediction model is recommended particularly in populations with particularly high incidence of OAC, e.g., in the United Kingdom [29].

Thresholds of the predicted risk are necessary for clinical and public health practices, but these need to be carefully determined by taking into account the incidence of OAC in the population, perspectives of the healthcare providers and practitioners, and potential benefits and harms for the individuals. However, choosing a risk threshold is beyond the scope of this study.

In summary, this population-based cohort study in Norway provides a model for predicting individuals' risk of developing OAC by combining information on five readily identifiable risk factors. The model had good performance in terms of discriminative ability and calibration, but further external validation in larger populations is needed, preferably in countries with high incidence of OAC. Individual risk assessment based on known risk factors for OAC has the potential of identifying individuals at high risk of OAC who may benefit from tailored screening and surveillance for detection of dysplastic Barrett's oesophagus or early and curable OAC.

ACKNOWLEDGEMENTS

The Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between HUNT Research Centre (Faculty of Medicine, NTNU, Norwegian University of Science and Technology), Nord-Trøndelag County Council, Central Norway Health Authority, and the Norwegian Institute of Public Health. The study has used data from the Cancer Registry of Norway and the Norwegian Cause of Death Registry. The interpretation and reporting of these data are the sole responsibility of the authors, and no endorsement by the Cancer Registry of Norway or the Norwegian Cause of Death Registry is intended nor should be inferred. We thank Dr. Giola Santoni in our research group for her technical help in developing the interactive online risk assessment tool.

CONFLICT OF INTEREST

Guarantor of the article: Shao-Hua Xie, BMed, PhD.

Specific author contributions: S-HX and JL jointly conceived the study. S-HX performed data analysis and prepared the first draft of the paper. EN-J was responsible for the data collection and NM contributed to data analysis and the writing of the manuscript. EN-J and JL contributed to the interpretation of the results, and reviewed

the manuscript. All authors had full access to all of the data (including statistical reports and tables) in the study and S-HX can take responsibility for the integrity of the data and the accuracy of the data analysis.

Financial support: This study was supported by the Swedish Research Council (grant numbers 521-2014-2536; 2015-06275) to JL; the Swedish Cancer Society (grant number CAN 2015/460) to JL; the Bengt Ihres Foundation (grant numbers SLS-595591; SLS-687761) to S-HX; and Ruth and Richard Julin Foundation to S-HX. The funding bodies had no role in the study design, the collection, analysis, and interpretation of data, or the writing of the article and the decision to submit it for publication.

Potential competing interests: None.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ The incidence of oesophageal adenocarcinoma has been increasing rapidly in many Western populations.
- ✓ Oesophageal adenocarcinoma is usually diagnosed at advanced stages with poor prognosis, stressing the need for early detection.
- ✓ Unselected endoscopic screening is not feasible given the low absolute risk in the general population.
- ✓ Risk prediction is a promising approach but has rarely been developed for oesophageal adenocarcinoma.

WHAT IS NEW HERE

- ✓ We developed a prediction model for individuals' absolute risk of oesophageal adenocarcinoma using data from a population-based cohort study.
- ✓ The model was based on information on a limited panel of established and readily identifiable risk factors.
- ✓ The prediction model had good discriminative accuracy and risk calibration.
- ✓ The prediction model could identify a limited group of high-risk individuals who may be considered for endoscopic screening.

REFERENCES

1. Edgren G, Adami HO, Weiderpass E, Nyren O. A global assessment of the oesophageal adenocarcinoma epidemic. Gut. 2013;62:1406–14.
2. Coleman HG, Xie SH, Lagergren J. The epidemiology of esophageal adenocarcinoma. Gastroenterology. 2018;154:390–405.
3. Xie SH, Mattsson F, Lagergren J. Incidence trends in oesophageal cancer by histological type: an updated analysis in Sweden. Cancer Epidemiol. 2017;47:114–7.
4. Xie SH, Rabbani S, Petrick JL, Cook MB, Lagergren J. Racial and ethnic disparities in the incidence of esophageal cancer in the United States, 1992-2013. Am J Epidemiol. 2017;186:1341–51.
5. Lagergren J, Lagergren P. Recent developments in esophageal adenocarcinoma. CA Cancer J Clin. 2013;63:232–48.
6. Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal cancer. Lancet. 2017;390:2383–96.
7. De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D, et al. Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE–5-a population-based study. Lancet Oncol. 2014;15:23–34.
8. Lao-Sirieix P, Fitzgerald RC. Screening for oesophageal cancer. Nat Rev Clin Oncol. 2012;9:278–87.
9. Sami SS, Ragunath K, Iyer PG. Screening for Barrett's esophagus and esophageal adenocarcinoma: rationale, recent progress, challenges, and future directions. Clin Gastroenterol Hepatol. 2015;13:623–34.
10. Vaughan TL, Fitzgerald RC. Precision prevention of oesophageal adenocarcinoma. Nat Rev Gastroenterol Hepatol. 2015;12:243–8.
11. Matsuno RK, Costantino JP, Ziegler RG, Anderson GL, Li H, Pee D, et al. Projecting individualized absolute invasive breast cancer risk in Asian and Pacific Islander American women. J Natl Cancer Inst. 2011;103:951–61.
12. Win AK, Macinnis RJ, Hopper JL, Jenkins MA. Risk prediction models for colorectal cancer: a review. Cancer Epidemiol Biomark Prev. 2012;21:398–410.
13. Katki HA, Kovalchik SA, Berg CD, Cheung LC, Chaturvedi AK. Development and validation of risk models to select ever-smokers for CT lung cancer screening. JAMA. 2016;315:2300–11.
14. Damen JA, Hooft L, Schuit E, Debray TP, Collins GS, Tzoulaki I, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. BMJ. 2016;353:i2416.
15. Thrift AP, Kendall BJ, Pandeya N, Whiteman DC. A model to determine absolute risk for esophageal adenocarcinoma. Clin Gastroenterol Hepatol. 2013;11:138–44 e2.
16. Xie SH, Lagergren J. A model for predicting individuals' absolute risk of esophageal adenocarcinoma: moving toward tailored screening and prevention. Int J Cancer. 2016;138:2813–9.
17. Xie SH, Lagergren J, Lagergren J. Using the Lorenz curve to assess the feasibility of targeted screening for esophageal adenocarcinoma. Epidemiology. 2017;28:e11–e2.
18. Krokstad S, Langhammer A, Hveem K, Holmen TL, Midthjell K, Stene TR, et al. Cohort profile: the HUNT Study, Norway. Int J Epidemiol. 2013;42:968–77.
19. Nilsson M, Johnsen R, Ye W, Hveem K, Lagergren J. Obesity and estrogen as risk factors for gastroesophageal reflux symptoms. JAMA. 2003;290:66–72.
20. Ness-Jensen E, Lindam A, Lagergren J, Hveem K. Changes in prevalence, incidence and spontaneous loss of gastro-oesophageal reflux symptoms: a prospective population-based cohort study, the HUNT study. Gut. 2012;61:1390–7.
21. Ness-Jensen E, Lindam A, Lagergren J, Hveem K. Tobacco smoking cessation and improved gastroesophageal reflux: a prospective population-based cohort study: the HUNT study. Am J Gastroenterol. 2014;109:171–7.
22. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999;94:496–509.
23. Dignam JJ, Zhang Q, Kocherginsky M. The use and interpretation of competing risks regression models. Clin Cancer Res. 2012;18:2301–8.
24. Lin G, So Y, Johnston G. Analyzing survival data with competing risks using SAS® software Available from: https://support.sas.com/resources/papers/proceedings12/344-2012.pdf.
25. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. 2010;21:128–38.
26. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. Stat Med. 2015;34:1659–80.
27. Mauguen A, Begg CB. Using the Lorenz curve to characterize risk predictiveness and etiologic heterogeneity. Epidemiology. 2016;27:531–7.
28. Shaheen NJ, Weinberg DS, Denberg TD, Chou R, Qaseem A, Shekelle P, et al. Upper endoscopy for gastroesophageal reflux disease: best practice advice from the clinical guidelines committee of the American College of Physicians. Ann Intern Med. 2012;157:808–16.
29. Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. Gut. 2015;64:381–7.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, and provide a link to the Creative Commons license. You do not have permission under this license to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.