Title: Development and validation of a risk prediction model for esophageal squamous cell carcinoma using cohort studies

Running head: Risk prediction model for esophageal squamous cell carcinoma

Authors:

Qiao-Li Wang (MD, PhD)[1], Eivind Ness-Jensen (MD, PhD)[1,2,3], Giola Santoni (PhD)[1], Shao-Hua Xie (MD, PhD)[1], Jesper Lagergren (MD, PhD)[1,4]

Affiliations:

[1] Upper Gastrointestinal Surgery, Department of Molecular medicine and Surgery, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden.

[2] HUNT Research Centre, Department of Public Health and Nursing, 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, United Kingdom.

Corresponding author:

Dr Qiao-Li Wang, Upper Gastrointestinal Surgery, Department of Molecular medicine and Surgery, Karolinska Institutet, Retzius väg 13 a, Floor 4, 17171, Stockholm, Sweden, Tel.: +46 8-524 841 41, E-mail: qiaoli.wang@ki.se

Word count: 3243

Guarantor of the article: First author Qiao-Li Wang and senior author Jesper Lagergren take full responsibility for the conduct of the study and have had access to the data and have control of the decision to publish.

Specific author contributions: The respective roles of each author are the following. Study concept and design: all authors; data management: Qiao-Li Wang, Giola Santoni; data analysis: Qiao-Li Wang; interpretation of results: all authors; drafting the manuscript: Qiao-Li Wang; critical revision of the manuscript for valuable intellectual content: all authors. All authors have contributed to and approved the final version of the manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Financial support: This work was supported by the Swedish Research Council (521-2014-2536; 2015-06275, both Jesper Lagergren), the Swedish Cancer Society (CAN 2015/460, Jesper Lagergren), and the Robert Lundberg Memorial Foundation (2018-01044, Qiao-Li Wang). The study sponsors had no role in the study design, data collection, analysis, or interpretation, the writing of the report, or the decision to submit the manuscript for publication.

Potential competing interests: All authors declare no conflicts of interest.

Abbreviations: AUC, area under the receiver operating characteristic curve; BMI, body mass index; CI, confidence interval; CIF, cumulative incidence function; ESCC, esophageal squamous cell carcinoma; HR, hazard ratio; HUNT, Nord-Trøndelag Health Study; SD, standard deviation.

STUDY HIGHLIGHTS

What is known

- Esophageal squamous cell carcinoma (ESCC) has a poor prognosis and early detection improves the survival
- Universal endoscopic screening is unfeasible, but prediction model could identify high-risk individuals for tailed surveillance
- There are only a few risk prediction models for ESCC, and none has been externally validated

What is new here

- An ESCC risk prediction model has been developed and externally validated by two national populations
- The model included five readily available predictors, showing good discrimination and calibration accuracy, and clinical usefulness
- A web-based risk assessment tool and a scoring system were derived for individual risk estimation

ABSTRACT

Objectives Esophageal squamous cell carcinoma (ESCC) carries a poor prognosis, but earlier tumour detection would improve survival. We aimed to develop and externally validate a risk prediction model based on exposure to readily available risk factors to identify high-risk individuals of ESCC.

Methods Competing risk regression modelling was used to develop a risk prediction model. Individuals' absolute risk of ESCC during follow-up was computed with the cumulative incidence function. We used prospectively collected data from the Nord-Trøndelag Health Study (HUNT) for model derivation and the UK Biobank cohort for validation. Candidate predictors were age, sex, tobacco smoking, alcohol consumption, body-mass index, education, cohabitation, physical exercise, and employment. Model performance was validated internally and externally by evaluating model discrimination using the area under the receiver operating characteristic curve (AUC) and model calibration.

Results The developed risk prediction model included age, sex, smoking, alcohol, and body-mass index. The AUC for 5-year risk of ESCC was 0.76 (95% confidence interval 0.58-0.93) in the derivation cohort, and 0.70 (95% confidence interval 0.64-0.75) in the validation cohort. The calibration showed close agreement between the predicted cumulative risk and observed probabilities of developing ESCC. Higher net benefit was observed when applying the risk prediction model than considering all participants as being at high risk, indicating good clinical usefulness. A web-tool for risk calculation was developed: https://sites.google.com/view/escc-ugis-ki.

Conclusions This ESCC risk prediction model showed good discrimination and calibration and validated well in an independent cohort. This readily available model can help select high-risk individuals for preventive interventions.

Keywords

Esophageal neoplasm; risk assessment; screening; surveillance; early diagnosis

INTRODUCTION

Esophageal cancer is the seventh most common cancer (572,000 new cases in 2018) and the sixth leading cause of cancer-related death (508,000 deaths) globally (1). Esophageal squamous cell carcinoma (ESCC) accounts for 87% of all histological types of esophageal cancer (2, 3), and 14% of all ESCC cases occur in Western countries (2). Although the overall incidence of ESCC has decreased in Western populations during the last few decades, the absolute number of cases is increasing because of the aging of many populations (4). ESCC is characterised by poor prognosis, with a population-based 5-year survival rate below 10-20%. The survival is closely related to tumour stage, with 5-year survival rates ranging from >95% for cancer in situ to <5% for stage IV, stressing that earlier tumour detection would greatly improve the survival (5-9). General endoscopy screening would enable early detection of ESCC, but is not cost-effective or clinically feasible given the low incidence of ESCC in Western populations and thus is not advocated by clinical guidelines (10). Yet, to detect ESCC at a curable stage, identifying a limited high-risk group of individuals by means of a valid prediction model based on readily accessible variables and implementing tailored endoscopic screening and surveillance programs could be a more effective and feasible approach (3, 11).

The few available risk prediction models for ESCC are derived from hospital-based studies or case-control studies which are prone to selection bias, or are based on predictor variables that are difficult to measure (12-16). Prediction models in non-Western populations do not apply to Western populations because of the different ESCC risk factor profiles (15). The only risk prediction model of ESCC in a Western population was derived from our Swedish case-control study (16), but the results were not externally validated and the predictor variables were retrospectively collected.

The aim of the present study was to develop models to predict the absolute risk of developing ESCC based on a panel of readily available risk factors using data from two prospective and population-based cohort studies in Western populations, one from Norway and the other from the United Kingdom.

METHODS

Data sources

The cohort Nord-Trøndelag Health Study (HUNT) in Norway was used to derive an ESCC risk model and the cohort UK Biobank database in the United Kingdom was used for external validation. The incidence of ESCC is similar in the populations of Norway and the United Kingdom (4). The HUNT is a large open population-based cohort initiated in the year 1984 with more than 30 years of follow-up. And the UK Biobank cohort is a closed cohort which included 502,628 participants in 2006-2010. Detailed description of two cohorts were in supplementary documents.

Study design

Both the derivation cohort (HUNT) and validation cohort (UK Biobank) are ongoing prospective cohorts. Follow-up time started from the date of study entry until the date of ESCC, death, loss to follow-up, or end of the study, whichever came first. Data were available until December 31, 2016 in HUNT and May 31, 2015 in the UK Biobank. The inclusion criteria for the present study were age 40 years or over when entering the study and no cancer diagnosed within nine years before study entry. The study was approved by the Regional Committee for Medical and Health Research Ethics in Norway and the North West Multi-Centre Research Ethics Committee in the United Kingdom.

Candidate predictors

Information about the candidate predictor variables was collected at the time of study entry. The selection of candidate predictors was based on subject knowledge and literature review of risk factors for ESCC. Predefined predictors were four well-established risk factors for ESCC, i.e. age, sex, tobacco smoking, and alcohol consumption. These variables were included in the models without further evaluation. Other available variables of potential interest were years of formal education, body-mass index (BMI), cohabitation status, physical exercise, and employment status. Hot drinking and poor oral hygiene have

been proposed as risk factors of ESCC in some studies (17, 18), but they were not considered in this study because studies from Western populations have not found any such associations (19, 20). The questionnaires used for assessing all nine predictor variables in both cohorts are shown in Supplementary Table 1. Adequate statistical power was achieved given the large sample size and long follow-up of the cohorts.

Model derivation

Competing risk regression models were used to calculate sub-hazards ratios for associations between the candidate predictor variables and risk of ESCC while taking the competing event of all-cause mortality into account (21). To facilitate interpretation and application, each of the predictor variables was introduced into the model as a categorical variable. The cutoff pointes for the variables were determined prior to analysis and by consideration of clinical usefulness and by optimizations from the two cohorts. A basic model included only the four well-established risk factors age, sex, tobacco smoking, alcohol consumption. A full model was developed based on a stepwise selection strategy (22). Detailed information can be found in supplementary document. From the full competing risk model, the individual cumulative incidence function (CIF) was estimated for all combinations of predictors. The CIF estimated the probability of developing ESCC over time while taking the competing risk of mortality into account (21). The risk of ESCC was computed from the CIF at three time points: 5, 10, and 15 years after inclusion. A 5-year risk of ESCC scoring system (ESCaScore) was developed by using the parameter estimates from the full model. Each estimated β-coefficient parameter (nature logarithm of sub-hazards ratios) was re-weighted by dividing the total β estimates and then multiplying by 100, rounding to the nearest integer. The final scoring system had a minimum value of 0 as the lowest ESCC risk and a maximum value of 100 as the highest ESCC risk. Complete case analyses were conducted for all analyses.

Assessment of model performance

Model performance was examined in terms of discrimination and calibration. The discriminative ability was assessed by the area under the receiver operating characteristic curve (AUC) statistics and the Somers' D statistics (23). The AUC summarizes the model's predictive accuracy of discriminating individuals who developed ESCC from those who did not. The AUC was evaluated at years 5, 10, and 15 after study entry in the derivation cohort (HUNT) and at year 5 in the validation cohort (UK Biobank), depending on risk factor distribution. Somers' D statistics assessed the strength and direction of associations between predicted probabilities and observed outcomes. Over-fitting of the prediction model could occur if the model had low bias but high variance. To adjust for possible over-fitting of the model, we assessed the model with leave-one-out internal cross-validation, which produced AUC and Somers' D with the predicted probability of each participant (24). Calibration was evaluated by plotting the observed frequencies of ESCC versus the cumulative predicted risk by deciles of the predicted risks (25, 26). The obtained curve was compared to that of a model with an ideal calibration, characterized by calibration in-the-large *a* (intercept) of 0 and calibration slope *b* of 1. The same assessment criteria were applied to test the performance of the full model, the basic model, and the ESCaScore model.

External validation

We applied the three developed risk prediction models above to the external validation cohort (UK Biobank) based on the individuals' risk pattern. Model performance was tested by examining the discriminative ability using AUC statistics and the Somers' D statistics, and the calibration accuracy using the calibration plot. Sensitivity, specificity, and positive predictive value of the full model were computed using the external validation cohort across the ESCC scores. The optimal score cut-off point for classifying individuals at high risk and low risk of ESCC was selected based on three indexes: the Youden Index defined as (sensitivity + specificity -1); distance to the ideal point (0, 1) in the AUC defined as square root of $((1- \text{sensitivity})^2+(1- \text{specificity})^2)$; and sensitivity-specificity equality defined as the absolute value of

(sensitivity - specificity). Higher Youden index, shorter distance to the ideal point, and lower sensitivity-specificity equality indicated better cut-off points.

Decision curve analysis

To assess the clinical usefulness of the risk prediction models, decision curve analysis was conducted using data from the external validation cohort (the UK Biobank) (27). The decision curve analysis compared the net benefit of using the proposed prediction model versus the strategy of assuming the entire population at high risk or at low risk of ESCC. We computed the net benefit by weighting the true positive rate against the false positive rate across different absolute risk thresholds, where the relative weights were based on the absolute risk threshold (27). Higher net benefit corresponded to higher clinical value of the prediction model. The net benefit was defined as:

$$Net\ benefit = \frac{true\ positive\ value - (\frac{\%\ risk\ threshold}{100 - \%\ risk\ threshold} \times false\ positive\ value)}{total\ number\ of\ participants}$$

RESULTS

Participants and incidence of ESCC

In HUNT, 77,476 participants met the inclusion criteria and were enrolled in the derivation cohort. These contributed 1.42 million person-years at risk of ESCC. The mean age at study entry was 55 years (standard deviation [SD] 12 years). From the UK Biobank validation cohort, 477,535 participants and 3.03 million person-years at risk were included. The mean age was 56 years (SD 8 years). Characteristics of the participants in the two cohorts are presented in Table 1. With 53 observed ESCC cases during the follow-up in HUNT, an overall incidence rate of 3.73 per 100,000 person-years was estimated (Table 2). In the UK Biobank, 105 participants developed ESCC, resulting in an incidence rate of 3.46 per 100,000 person-years. In both cohorts, the ESCC incidence increased with age.

Competing risk regression

After predictor selection, five predictors were selected for the final model: age, sex, tobacco smoking, alcohol consumption, and BMI. Table 3 shows the sub-hazard ratios of ESCC in both cohorts. Older age, female sex, current daily tobacco smoking, consumption of alcohol ≥3 times per week, and BMI ≤25 was associated with an increased risk of ESCC. The predicted cumulative risk of ESCC within 15 years of follow-up varied across patterns of predictor combinations (Supplementary Table 2). The highest incidence (301.3 per 100,000 person-years) was found in women older than 60 years with BMI ≤25 who were current smokers and drank alcohol ≥3 times per week. In this relative high-risk group, 332 individuals need to be surveyed to detect one ESCC case within 15 years of follow-up.

Model performance and external validation

Within 5 years of follow-up, the full risk prediction model had an AUC of 0.76 (95% CI 0.58-0.93) and Somers' D statistics of 0.51 in the derivation cohort (Table 4). After internal cross-validation, the AUC was 0.67 (95% CI 0.45-0.89). When evaluating the 15-year risk of ESCC, the AUC was 0.73 (95% CI 0.62-

0.85) after internal cross-validation. The calibration plots indicated close agreements between the predicted cumulative risks and observed probabilities of incidence of ESCC within 5, 10, and 15 years of follow-up in the derivation cohort (Figure 1, panel a, c, and d). The calibration in-large *a* and calibration slope *b* was -0.005 and 0.987 for the 5-year risk model and -0.024 and 1.035 for the 15-year risk model, respectively. When applying the full risk prediction model for 5-year risk in the validation cohort, the AUC was 0.70 (95% CI 0.64-0.85) and Somers' D statistics was 0.39. The validation showed good model calibration with calibration in-large *a* of -0.121 and calibration slope *b* of 1.110 (Figure 1, panel b). The basic model, excluding BMI, had a similarly internal and external model performance (Supplementary Table 3, Supplementary Figure 1).

ESCaScore and interactive web tool

The ESCaScore scale based on the full risk prediction model is shown in Table 5. The ESCaScore model exhibited a good discriminative accuracy with AUC of 0.71 (95% CI 0.52-0.90) in the derivation cohort after cross-validation and 0.70 (95% CI 0.64-0.75) in the external validation cohort (Table 6). The calibration of the ESCaScore was excellent, with calibration in-large *a* of -0.004 and slope *b* of -1.009 (Supplementary Figure 2) using data from the validation cohort.

To best discriminate high-risk persons from low-risk persons, the optimal cut-off analysis identified a maximum Youden index of 0.31 and the minimum distance to (0,1) of 0.50 (Supplementary Table 4). This cut-off corresponded to a risk threshold of 20.9 cases of ESCC per 100,000 persons and to an ESCaScore of 49. At this cut-off, the model had a sensitivity of 56.8%, a specificity of 74.0%, and a positive predictive value of 39 cases of ESCC per 100,000.

We constructed an interactive web tool for estimating individual 5-year risk of ESCC based on the full model. This tool can be accessed at https://sites.google.com/view/escc-ugis-ki. A screen shot of the web tool is presented in Supplementary Figure 3.

Decision curve analysis

Figure 2 shows the net benefit curve for the full model within 5 years of follow-up in the validation cohort. The horizontal axis is the risk threshold used to define high risk and the vertical axis is the net benefit at a given threshold. The net benefit curve implies that for every 100,000 participants where we apply the prediction model at a threshold of 18.9/100,000, for instance, there will be a net 5 more true positives than false positives identified (Figure 2, Supplementary Table 4). The curve of applying the prediction model crosses the curve of treating all as low risk (i.e., no screening) at a 5-year absolute risk threshold of 49/100,000 and crosses the curve of treating all as high risk (i.e., screening all individuals) at 9/100,000. It indicates that the model had higher net benefits than the strategy assuming all patients either at high risk or at low risk of ESCC for risk thresholds between 9 and 49 cases of ESCC per 100,000 persons.

DISCUSSION

This study developed and externally validated a model for assessing the absolute cumulative risk of ESCC, based on the five variables age, sex, tobacco smoking, alcohol consumption, and BMI, showing good discriminative and calibration accuracy, and clinical usefulness. A web-based calculation tool and an ESCaScore scale were developed for individual risk assessment.

Among strengths of this study was the use of two large prospective and independent cohorts from Western populations. The information was collected at least five years before the ESCC diagnosis and all ESCC cases were identified through high-quality cancer registries, which counteracted bias from selection, information, and detection. The prediction models were developed from a Norwegian population with long and complete follow-up and validated in a larger sample from an English population with similar incidence of ESCC. The models were extensively calibrated (calibration plot, calibration in-large, and calibration slope) with good agreements between observed and predicted risk in both cohorts. These calibration methods were preferred over the Hosmer-Lemeshow tests because the latter are sensitive to sample size and powerless in detecting overfitting of predictor effects (28-30). Given the long follow-up in the cohort studies, another strength was the use of competing risk regression, because competing risk from mortality could otherwise bias the results. The decision curve analysis indicated that the prediction model had good clinical usefulness in classifying individuals at high risk. The prediction model is also easy-to-use because it is based on only five readily assessed predictors. The web-based calculator tool and the ESCaScore calculator could be used for quick risk classification. The risk prediction model should have good generalizability in Western populations, because they share the risk factors for ESCC.

There are also several limitations of the study. We were limited to the data collected in both cohorts and other possible predictors than the nine assessed could have been of interest. However, well-known risk factors were included. Two of the prediction variables were self-reported (tobacco smoking and alcohol consumption), which might introduce misclassification (31). However, given the prospective cohort design,

any misclassification due to self-reporting should occur at random and thus only underestimate associations. Although we harmonized the exposure variables in the derivation and validation cohorts as closely as possible, some discrepancies between these cohorts were possible. Yet, the derived models performed well in the external cohort analyses. Because the UK Biobank had shorter follow-up than the HUNT cohort, external validation was only possible for the prediction model of 5-year ESCC risk. Another limitation was missing data for some variables, but the missing rate of the predictors was limited (less than 14% in the derivation cohort and less than 1% in validation cohort), and the model developed from the derivation cohort validated well in the validation cohort. Finally, the relative low incidence of ESCC in the study populations reduced the statistical power, which was particularly evident in some subgroup analyses.

Few earlier prediction models have been developed for ESCC, with AUCs ranging from 0.58 to 0.81 (12-16). Direct comparison with these models is difficult because of differences in study design, predictor definitions, and risk factor patterns in the study populations. A cross-sectional study identified individuals with increased risk of esophageal squamous dysplasia based on environmental exposures combined with physical examinations, but the AUC was low (0.58) (12). By combining environmental and genetic risk factors of ESCC, a prediction model based on a Chinese case-control study found an AUC of 0.71 (13). A case-control study from Iran included ten known risk factors in that region and built a prediction model with an AUC of 0.77 (14). Limitations of these three studies above include the hospital-based design and the use of predictors difficult to measure. A population-based cross-sectional study from China based on lifestyle risk factors and preclinical symptoms of ESCC found an AUC of 0.80 in individuals aged below 60 years, and AUC of 0.68 in those above 60 years (15). Yet, that study was limited by not having model presentation and the model difficult to use. The only prediction model developed in a Western population came from our Swedish population-based case-control study with a high AUC (0.81), which was based on the variables of age, sex, tobacco smoking, alcohol overconsumption, education, duration of living with a partner, and place of residence during childhood (16). But the retrospective data collection in that case-control study introduces a risk of misclassification of predictors. Moreover, none of the five previous

prediction models evaluated the discriminative accuracy in an external independent population or examined the model calibration.

During the last decades, much effort has been devoted to improving the treatment of ESCC. Yet, the overall 5-year survival rate remains lower than 10-20% (32), which is mainly due to late tumor detection. Earlier detection at a curable stage would improve the survival (33), but universal endoscopy screening is not cost-effective and unfeasible in Western populations given the low incidence. Instead, risk prediction models could be used for risk stratification targeting the absolute high-risk individuals of ESCC. With the developed prediction model, we suggest that endoscopic screening could be performed above a risk threshold of 18.9 per 100,000 or with an ESCaScore above 46. The present prediction model may help policy makers and healthcare to determine public health recommendations and tailor screening strategies for high-risk individuals of ESCC and may assist clinical decision-making regarding the need for endoscopy.

In conclusion, this study developed and externally validated a readily retrieved risk prediction model for estimation of the absolute risk of ESCC in two Western cohorts. The model derived from one cohort, based on the five variables age, sex, tobacco smoking, alcohol consumption, and BMI, showed good discrimination and calibration, and validated well in an independent external cohort. The model can be used to inform healthcare about individuals' risk of developing ESCC and to target primary or secondary interventions.

Acknowledgment

The Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology), Nord-Trøndelag County Council, Central Norway Health Authority, and the Norwegian Institute of Public Health.

REFERENCE

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Arnold M, Soerjomataram I, Ferlay J, et al. Global incidence of oesophageal cancer by histological subtype in 2012. Gut. 2015;64(3):381-7.
3. Lagergren J, Smyth E, Cunningham D, et al. Oesophageal cancer. Lancet. 2017.
4. Wang QL, Xie SH, Wahlin K, et al. Global time trends in the incidence of esophageal squamous cell carcinoma. Clin Epidemiol. 2018;10:717-28.
5. Lagergren J, Mattsson F. Diverging trends in recent population-based survival rates in oesophageal and gastric cancer. PLoS One. 2012;7(7):e41352.
6. Klint A, Engholm G, Storm HH, et al. Trends in survival of patients diagnosed with cancer of the digestive organs in the Nordic countries 1964-2003 followed up to the end of 2006. Acta Oncol. 2010;49(5):578-607.
7. Baastrup R, Sorensen M, Hansen J, et al. Social inequality and incidence of and survival from cancers of the oesophagus, stomach and pancreas in a population-based study in Denmark, 1994-2003. Eur J Cancer. 2008;44(14):1962-77.
8. Enzinger PC, Mayer RJ. Esophageal cancer. N Engl J Med. 2003;349(23):2241-52.
9. O'Connell J B, Maggard MA, Liu JH, et al. A report card on outcomes for surgically treated gastrointestinal cancers: are we improving? J Surg Res. 2004;121(2):214-21.
10. Wang KK, Wongkeesong M, Buttar NS. American Gastroenterological Association technical review on the role of the gastroenterologist in the management of esophageal carcinoma. Gastroenterology. 2005;128(5):1471-505.
11. 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(12):2813-9.
12. Wei WQ, Abnet CC, Lu N, et al. Risk factors for oesophageal squamous dysplasia in adult inhabitants of a high risk region of China. Gut. 2005;54(6):759-63.
13. Chang J, Huang Y, Wei L, et al. Risk prediction of esophageal squamous-cell carcinoma with common genetic variants and lifestyle factors in Chinese population. Carcinogenesis. 2013;34(8):1782-6.
14. Etemadi A, Abnet CC, Golozar A, et al. Modeling the risk of esophageal squamous cell carcinoma and squamous dysplasia in a high risk area in Iran. Arch Iran Med. 2012;15(1):18-21.
15. Liu MF, Liu Z, Cai H, et al. A Model To Identify Individuals at High Risk for Esophageal Squamous Cell Carcinoma and Precancerous Lesions in Regions of High Prevalence in China. Clin Gastroenterol Hepatol. 2017;15(10):1538-46.
16. Wang QL, Lagergren J, Xie SH. Prediction of individuals at high absolute risk of esophageal squamous cell carcinoma. Gastrointest Endosc. 2019;89(4):726-32 e2.
17. Islami F, Boffetta P, Ren JS, et al. High-temperature beverages and foods and esophageal cancer risk--a systematic review. Int J Cancer. 2009;125(3):491-524.
18. Ahrens W, Pohlabeln H, Foraita R, et al. Oral health, dental care and mouthwash associated with upper aerodigestive tract cancer risk in Europe: the ARCAGE study. Oral Oncol. 2014;50(6):616-25.
19. Terry P, Lagergren J, Wolk A, et al. Drinking hot beverages is not associated with risk of oesophageal cancers in a Western population. Br J Cancer. 2001;84(1):120-1.
20. Ljung R, Martin L, Lagergren J. Oral disease and risk of oesophageal and gastric cancer in a nationwide nested case-control study in Sweden. Eur J Cancer. 2011;47(14):2128-32.
21. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999;94(446):496-509.
22. 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):683-90.

23. 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):128-38.
24. Thrift AP, Kendall BJ, Pandeya N, et al. A model to determine absolute risk for esophageal adenocarcinoma. Clin Gastroenterol Hepatol. 2013;11(2):138-44 e2.
25. F.E. H. Regression modeling strategies. New York: Springer-Verlag; 2001.
26. Steyerberg EW. Clinical Prediction Models - A Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.
27. Vuong K, Armstrong BK, Weiderpass E, et al. Development and External Validation of a Melanoma Risk Prediction Model Based on Self-assessed Risk Factors. JAMA Dermatol. 2016;152(8):889-96.
28. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. Crit Care Med. 2007;35(9):2052-6.
29. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating New York: Springer; 2009.
30. Collins GS, de Groot JA, Dutton S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. BMC Med Res Methodol. 2014;14:40.
31. Li K, Husing A, Sookthai D, et al. Selecting High-Risk Individuals for Lung Cancer Screening: A Prospective Evaluation of Existing Risk Models and Eligibility Criteria in the German EPIC Cohort. Cancer Prev Res (Phila). 2015;8(9):777-85.
32. Codipilly DC, Qin Y, Dawsey SM, et al. Screening for esophageal squamous cell carcinoma: recent advances. Gastrointest Endosc. 2018;88(3):413-26.
33. Lagergren J, Smyth E, Cunningham D, et al. Oesophageal cancer. Lancet. 2017;390(10110):2383-96.

Table 1. Distribution of basic characteristics of study participants in the derivation cohort (HUNT) and the validation cohort (UK Biobank)

	Derivation cohort		Validation cohort	
Characteristics at baseline	**Number of participants (%)**	**Person-years**	**Number of participants (%)**	**Person-years**
Total	77,476 (100.0)	1,420,142	477,535 (100.0)	3,033,715
Age (in years)				
40-49	35,372 (45.7)	713,678	115,523 (24.2)	737,034
50-59	16,474 (21.3)	355,762	160,358 (33.6)	1,023,607
≥ 60	25,630 (33.1)	350,701	201,654 (42.2)	1,273,074
Sex				
Male	37,260 (48.1)	662,109	218,156 (45.7)	1,386,724
Female	40,216 (51.9)	758,033	259,379 (54.3)	1,646,991
Daily tobacco smoking				
Never	27,628 (35.7)	515,794	260,673 (54.6)	1,654,253
Former	18,473 (23.8)	339,304	163,228 (34.2)	1,036,193
Current	21,128 (27.3)	388,592	50,828 (10.6)	325,851
Missing	10,247 (13.2)	176,452	2,806 (0.6)	17,418
Alcohol consumption				
<3 times/week	62,083 (80.1)	1,148,488	269,666 (56.5)	1,711,480
≥3 times/week	4,938 (6.4)	76,674	206,419 (43.2)	1,313,368
Missing	10455 (13.5)	194,980	1,450 (0.3)	8,868
Body mass index				
>25	44,207 (57.1)	788,459	317,659 (66.5)	2,018,957
≤25	31,445 (40.6)	614,556	156,906 (32.9)	996,300
Missing	1,824 (2.4)	17,127	2,970 (0.6)	18,458
Formal education years				
>10	32,480 (41.9)	631,287	225,645 (47.3)	1,426,555
≤10	35,397 (45.7)	676,014	245,462 (51.4)	1,563,322
Missing	9,599 (12.4)	112,841	6,428 (1.4)	43,838
Cohabitation status				
Living with a partner	50,188 (64.8)	978,741	344,812 (72.2)	2,192,614
Living with relatives	5,098 (6.6)	87,377	40,485 (8.5)	255,573
Alone	7,357 (9.5)	104,097	88,054 (18.4)	559,562
Missing	14,833 (19.2)	249,927	4,184 (0.9)	25,966
Physical exercise				
≥60 minutes/week	13,936 (18.0)	239,230	181,524 (38.0)	1,148,983
<60 minutes/week	53,512 (69.1)	995,892	270,368 (56.6)	1,722,065
Missing	10,028 (12.9)	185,019	25,643 (5.4)	162,667
Employment status				
Employed	47,705 (61.6)	966,613	276,946 (58.0)	1,765,536
Unemployed	29,384 (37.9)	449,102	195,167 (40.9)	1,234,490
Missing	38,7 (0.5)	4,426	5,422 (1.1)	33,690

Abbreviation: HUNT, Nord-Trøndelag Health Study

Table 2. Number of esophageal squamous cell carcinoma cases, person-years of follow up, and incidence rates in the derivation cohort (HUNT) and the validation cohort (UK Biobank), overall and by age and sex

	Derivation cohort			Validation cohort		
Variables	**Number of cases**	**Person-years**	**Incidence rate per 100,000 person-years (95% CI)**	**Number of cases**	**Person-years**	**Incidence rate per 100,000 person-years (95% CI)**
Overall	53	1,420,142	3.73 (2.85-4.89)	105	3,033,715	3.46 (2.86-4.19)
Age (in years)						
40-49	14	713,678	1.96 (1.16-3.31)	6	737,034	0.81 (0.37-1.81)
50-59	14	355,762	3.94 (2.33-6.64)	27	1,023,607	2.64 (1.81-3.85)
≥60	25	350,701	7.13 (4.82-10.55)	72	1,273,074	5.66 (4.49-7.13)
Sex						
Male	27	662,109	4.08 (2.80-5.95)	46	1,386,724	3.32 (2.48-4.43)
Female	26	758,033	3.43 (2.34-5.04)	59	1,646,991	3.58 (2.78-4.62)

Abbreviation: HUNT, Nord-Trøndelag Health Study; CI, confidence interval

Table 3. Risk factors for esophageal squamous cell carcinoma from competing risk modelling in the derivation cohort (HUNT) and the validation cohort (UK Biobank) [a]

	Derivation cohort		Validation cohort	
Variable at baseline	**Crude sub-hazard ratio (95% CI)**	**Adjusted sub-hazard ratio (95% CI)**	**Crude sub-hazard ratio (95% CI)**	**Adjusted sub-hazard ratio (95% CI)**
Age (in years)				
40-49	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
50-59	1.88 (0.89-3.98)	1.88 (0.84-4.22)	3.24 (1.34-7.84)	3.03 (1.24-7.41)
≥60	2.06 (1.04-4.08)	2.68 (1.24-5.77)	6.93 (3.01-15.95)	7.06 (3.06-16.34)
Sex				
Male	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Female	0.89 (0.52-1.52)	1.04 (0.55-1.95)	1.08 (0.74-1.59)	1.23 (0.83-1.81)
Daily tobacco smoking				
Never	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Former	1.79 (0.60-5.34)	1.84 (0.63-5.34)	1.88 (1.23-2.88)	1.60 (1.04-2.47)
Current	6.57 (2.73-15.77)	6.42 (2.70-15.25)	2.59 (1.51-4.45)	2.92 (1.70-5.02)
Alcohol consumption				
<3 times/week	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
≥3 times/week	1.86 (0.72-4.76)	1.58 (0.61-4.09)	1.38 (0.94-2.02)	1.32 (0.90-1.93)
Body mass index				
>25	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
≤25	1.74 (1.01-2.99)	1.66 (0.93-2.97)	1.42 (0.96-2.10)	1.50 (1.00-2.24)

Abbreviation: HUNT, Nord-Trøndelag Health Study; CI, confidence interval

[a] *Participants with missing values of predictor variables were excluded from the multivariate competing risk regression (n=13,468 in HUNT and n=5,779 in UK Biobank).*

Table 4. Performance of a full prediction model for esophageal squamous cell carcinoma within 5, 10, and 15 years of follow-up in the derivation cohort (HUNT) and within 5 years in the validation cohort (UK Biobank)

Outcomes	Area under the receiver-operating characteristic curve (95% CI)	Somers' D
Derivation cohort		
5 years	0.76 (0.58-0.93)	0.51
10 years	0.74 (0.61-0.88)	0.49
15 years	0.77 (0.66-0.87)	0.53
Internal cross-validation		
5 years	0.67 (0.45-0.89)	0.35
10 years	0.68 (0.51-0.85)	0.36
15 years	0.73 (0.62-0.85)	0.46
External validation cohort		
5 years	0.70 (0.64-0.75)	0.39

Abbreviation: HUNT, Nord-Trøndelag Health Study; CI, confidence interval

Table 5. A scoring system for estimating the 5-year risk of esophageal squamous cell carcinoma (ESCaScore)

Predictors	β estimate	Score
Age (in years)		
40-49	Reference	0
50-59	0.63	16
≥ 60	0.98	26
Sex		
Male	Reference	0
Female	0.04	1
Daily tobacco smoking		
Never	Reference	0
Former	0.61	16
Current	1.86	48
Alcohol consumption		
<3 times/week	Reference	0
≥3 times/week	0.45	12
Body mass index		
>25	Reference	0
≤25	0.51	13
Total score ranges		0-100

Table 6. Model performance of the scoring system (ESCaScore) for predicting 5-year risk of esophageal squamous cell carcinoma in the derivation cohort (HUNT) and the validation cohort (UK Biobank)

Outcomes	Area under the receiver-operating characteristic curve (95% CI)	Somers' D
Derivation cohort	0.76 (0.58-0.93)	0.51
Internal cross-validation	0.71 (0.52-0.90)	0.43
External validation	0.70 (0.64-0.75)	0.39

Abbreviation: HUNT, Nord-Trøndelag Health Study; CI, confidence interval

FIGURE LEGEND

Figure 1. Full model calibration of observed cumulative proportion of esophageal squamous cell carcinoma (ESCC) and predicted cumulative risk of ESCC within 5, 10, and 15 years of follow-up in the derivation cohort (HUNT) and within 5 years in the validation cohort (UK Biobank)

Note: a denotes the intercept; b denotes the slope.

Figure 2. Decision curve obtained from plotting the net benefit of detecting esophageal squamous cell carcinoma (ESCC) at different 5-year absolute risk thresholds in the validation cohort (UK Biobank)

LEGENDS OF SUPPLEMENTARY MATERIALS

Supplementary Document. Detailed description of cohort profiles, potential predictors, and predictor selection methods

Supplementary Table 1. Evaluation of candidate predictor variables in a derivation cohort (HUNT) and a validation cohort (UK Biobank)

Supplementary Table 2. Estimated 15-year cumulative risk of esophageal squamous cell carcinoma (ESCC) per 100,000 person-years and number need to survey to detect one case, depending on combination of risk factors

Supplementary Table 3. Performance of a basic model for predicting risks of esophageal squamous cell carcinoma within 5, 10, and 15 years of follow-up in the derivation cohort (HUNT) and within 5 years in the validation cohort (UK Biobank)

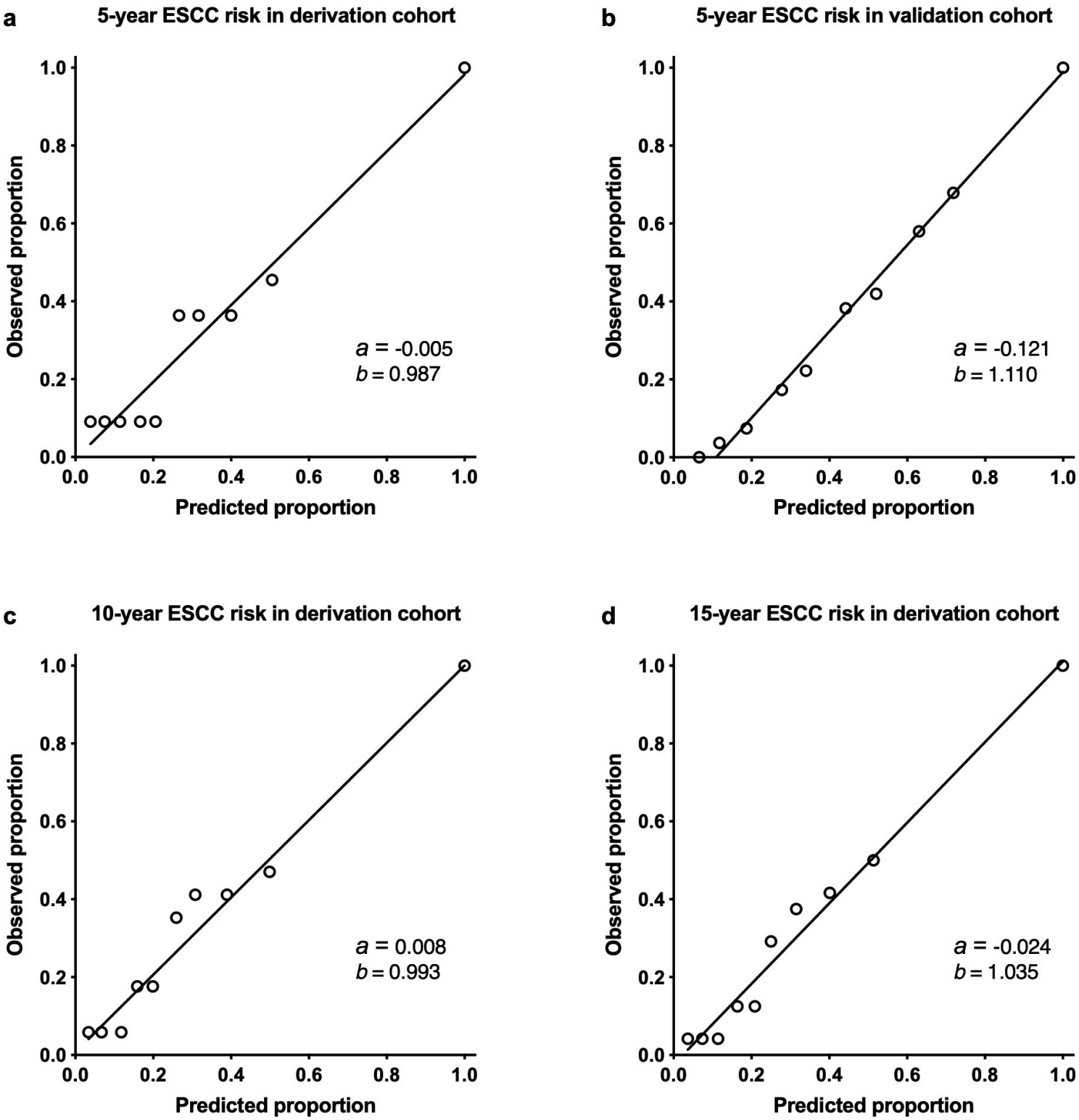
Supplementary Table 4. Sensitivity, specificity, positive predictive value, and cut-off evaluation criteria for predicting the risk of esophageal squamous cell carcinoma (ESCC) within 5 years of follow-up in the validation cohort (UK Biobank)

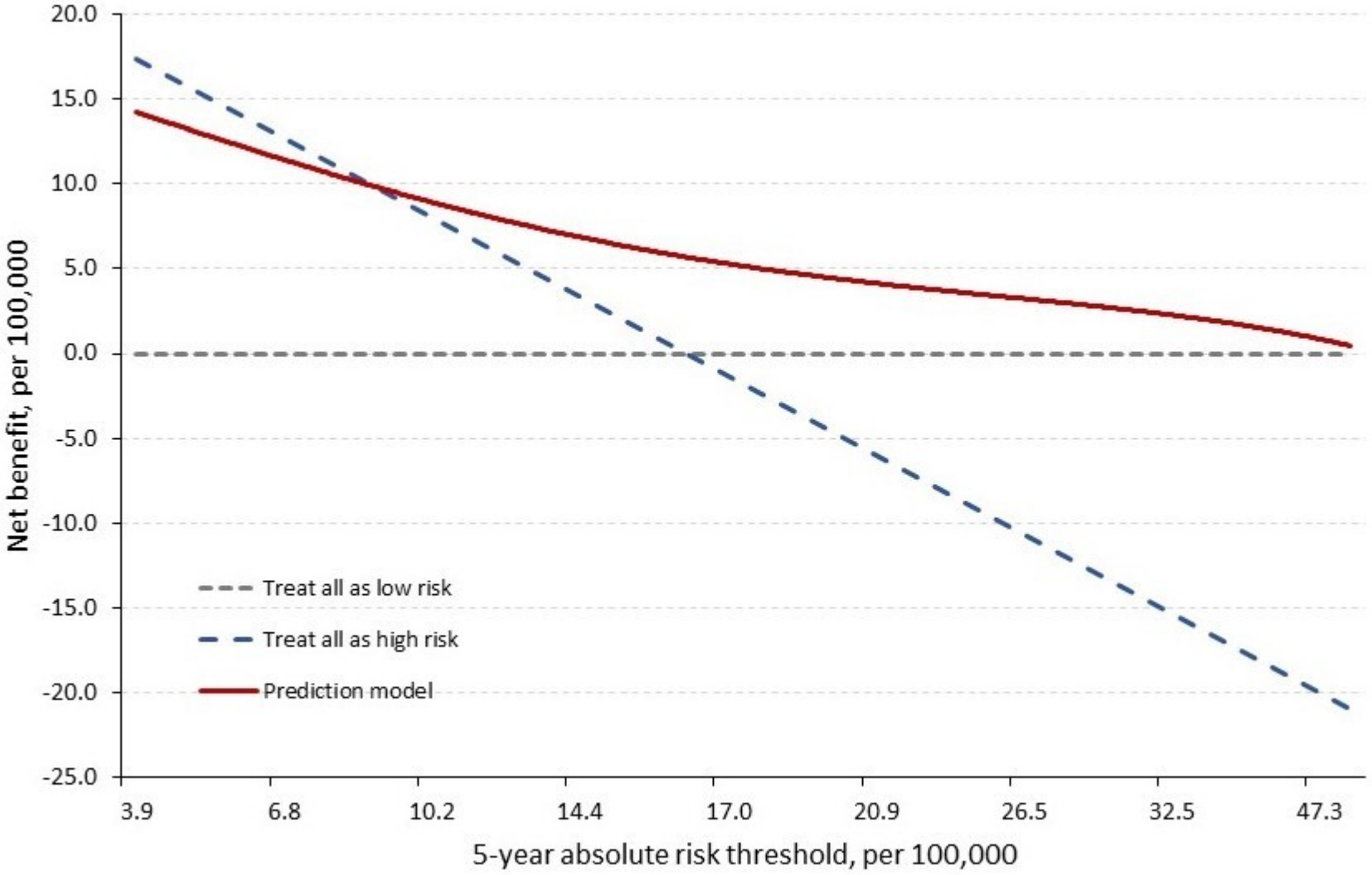
Supplementary Figure 1. Basic model calibration of the observed cumulative proportion of esophageal squamous cell carcinoma (ESCC) and predicted cumulative risk of ESCC within 5, 10, and 15 years of follow-up in the derivation cohort (HUNT) and within 5 years in the validation cohort (UK Biobank)

Supplementary Figure 2. ESCaScore model calibration of observed cumulative proportion of esophageal squamous cell carcinoma (ESCC) and predicted cumulative risk of ESCC within 5 years of follow-up in the derivation cohort (HUNT) and the validation cohort (UK Biobank)

Supplementary Figure 3. Screenshot of web-based 5-year esophageal squamous cell carcinoma risk assessment tool

Note: original website can be reached here: https://sites.google.com/view/escc-ugis-ki


a
5-year ESCC risk in derivation cohort
Observed proportion
Predicted proportion
a = -0.005
b = 0.987
b
5-year ESCC risk in validation cohort
Observed proportion
Predicted proportion
a = -0.121
b = 1.110
c
10-year ESCC risk in derivation cohort
Observed proportion
Predicted proportion
a = 0.008
b = 0.993
d
15-year ESCC risk in derivation cohort
Observed proportion
Predicted proportion
a = -0.024
b = 1.035


20.0
15.0
10.0
5.0
0.0
-5.0
-10.0
-15.0
-20.0
-25.0
Net benefit, per 100,000
Treat all as low risk
Treat all as high risk
Prediction model
3.9
6.8
10.2
14.4
17.0
20.9
26.5
32.5
47.3
5-year absolute risk threshold, per 100,000

SUPPLEMENTARY MATERIAL

Development and validation of a risk prediction model for esophageal squamous cell carcinoma using cohort studies

Authors:

Qiao-Li Wang (MD, PhD)[1], Eivind Ness-Jensen (MD, PhD)[1,2,3], Giola Santoni (PhD)[1], Shao-Hua Xie (MD, PhD)[1], Jesper Lagergren (MD, PhD)[1,4]

Affiliations:

[1] Upper Gastrointestinal Surgery, Department of Molecular medicine and Surgery, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden.

[2] HUNT Research Centre, Department of Public Health and Nursing, 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, United Kingdom.**LEGENDS OF SUPPLEMENTARY MATERIALS**

Supplementary Document. Detailed description of cohort profiles, potential predictors, and predictor selection methods

Supplementary Table 1. Evaluation of candidate predictor variables in a derivation cohort (HUNT) and a validation cohort (UK Biobank)

Supplementary Table 2. Estimated 15-year cumulative risk of esophageal squamous cell carcinoma (ESCC) per 100,000 person-years and number need to survey to detect one case, depending on combination of risk factors

Supplementary Table 3. Performance of a basic model for predicting risks of esophageal squamous cell carcinoma within 5, 10, and 15 years of follow-up in the derivation cohort (HUNT) and within 5 years in the validation cohort (UK Biobank)

Supplementary Table 4. Sensitivity, specificity, positive predictive value, and cut-off evaluation criteria for predicting the risk of esophageal squamous cell carcinoma (ESCC) within 5 years of follow-up in the validation cohort (UK Biobank)

Supplementary Figure 1. Basic model calibration of the observed cumulative proportion of esophageal squamous cell carcinoma (ESCC) and predicted cumulative risk of ESCC within 5, 10, and 15 years of follow-up in the derivation cohort (HUNT) and within 5 years in the validation cohort (UK Biobank)

Supplementary Figure 2. ESCaScore model calibration of observed cumulative proportion of esophageal squamous cell carcinoma (ESCC) and predicted cumulative risk of ESCC within 5 years of follow-up in the derivation cohort (HUNT) and the validation cohort (UK Biobank)

Supplementary Figure 3. Screenshot of web-based 5-year esophageal squamous cell carcinoma risk assessment tool

Supplementary Document. Detailed description of cohort profiles, potential predictors, and predictor selection methods

Cohort profiles

Derivation cohort: The HUNT is a large population-based cohort initiated in the year 1984 with more than 30 years of follow-up. This longitudinal cohort is based on a series of health surveys of the adult population of Nord-Trøndelag County, which is representative of the Norwegian population at large (1). Detailed

information about HUNT can be found elsewhere (2, 3). Briefly, all residents in the county over 20 years of age were invited to participate in this open cohort, and the number of participants was 77,212 in the data collection in 1984-1986 (HUNT1, 89% participation rate), 65,237 in 1995-1997 (HUNT2, 70% participation rate), and 50,807 in 2006-2008 (HUNT3, 54% participation rate). At each assessment period, basic information and a wide range of health-related data were collected through written questionnaires and in-person clinical and laboratory examinations. Information about cancer and mortality were collected by linkages to the Norwegian Cancer Registry and the Norwegian Cause of Death Registry, respectively, using the unique national identity number assigned to each Norwegian resident. Participants aged 40 years or more were eligible for the current study.

Validation cohort: The UK Biobank includes participants recruited from 22 assessment centers in 2006-2010. Details of the cohort are presented in previous publications (4, 5). In brief, 9.2 million British residents aged between 40 and 69 years who lived within 25 English miles of any of the 22 assessment centers and registered in the National Health Service were invited to participate. Among these, 502,628 (5.5%) participated and provided information regarding their lifestyle, medical history, and physical measures, and donated biological samples. The participants in the UK Biobank is representative of the general population in the United Kingdom regarding the distribution of age, sex and ethnicity, but not regarding lifestyle variables, physical measures, or health-related factors, possibly due to "healthy volunteer" selection bias (6). Information about cancer incidence and mortality is retrieved by linkages to the national cancer registries and national death registries in the United Kingdom.

Potential predictors

The study included variables of age (40-49, 50-59, or ≥60 years), sex (male or female), tobacco smoking status (never, former, or current smoker), alcohol consumption (<3 or ≥3 times/week), years of formal education (>10 or ≤10 years), BMI (>25 or ≤25), cohabitation status (living with partner, other relatives, or alone), physical exercise (≥60 or <60 minutes per week), and employment status (employed or unemployed).

Predictor selection

Except for the predictors in the basic model, the other five candidate predictors were evaluated by a stepwise selection method, where a variable was considered for inclusion only if its p-value was below 0.1. Starting with the predictor with the lowest p-value, new models were evaluated by adding one predictor at a time. This predictor was removed if it exceeded the threshold p-value of 0.1 in the new model. Potential interactions between the predictors were also assessed (7). Using the likelihood ratio test, eliminated predictors were re-entered into the final multivariable model one by one to ensure that none of the preliminary excluded predictors statistically significantly improved the goodness of fit (7). The final criterion for predictor inclusion in the model was based on the Akaike Information Criteria (8, 9).

References

1. Holmen J, Midthjell K, Krüger Ø, et al. The Nord-Trøndelag Health Study 1995-97 (HUNT 2): Objectives, contents, methods and participation. Norsk Epidemiologi. 2003;13(1):13.
2. 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(2):171-7.
3. Ness-Jensen E, Gottlieb-Vedi E, Wahlin K, et al. All-cause and cancer-specific mortality in GORD in a population-based cohort study (the HUNT study). Gut. 2018;67(2):209-15.
4. Ganna A, Ingelsson E. 5 year mortality predictors in 498,103 UK Biobank participants: a prospective population-based study. Lancet. 2015;386(9993):533-40.
5. Lyall DM, Celis-Morales C, Ward J, et al. Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank: A Mendelian Randomization Study. JAMA Cardiol. 2017;2(8):882-9.
6. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. Am J Epidemiol. 2017;186(9):1026-34.
7. 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):683-90.
8. Thrift AP, Kendall BJ, Pandeya N, et al. A model to determine absolute risk for esophageal adenocarcinoma. Clin Gastroenterol Hepatol. 2013;11(2):138-44 e2.
9. Vrieze SI. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol Methods. 2012;17(2):228-43.

Supplementary Table 1. Evaluation of candidate predictor variables in a derivation cohort (HUNT) and a validation cohort (UK Biobank)

	Questions	
Variables	**HUNT**	**UK Biobank**
Age	Age when participating in HUNT	Age at study entry
Sex	Participant's gender	Participant's gender
Tobacco smoking	Re-constructed from reported answers about current or previous daily smoking	Reconstructed from questionnaires of “Do you smoke tobacco now?" and "In the past, how often have you smoked tobacco?"
Alcohol consumption	How often did you drink alcohol (beer, wine or spirits) during the LAST 14 DAYS? / How many times a month do you normally drink alcohol? / About how often in the last 12 months did you drink alcohol?	About how often do you drink alcohol?
Formal education	What is your highest level of education?	Which of the following qualifications do you have?
Body mass index	Calculated by the participant's weight in kg/(height in meter*height in meter)	Calculated by the participant's weight in kg/(height in meter*height in meter)
Cohabitation status	Who do you live with?	How are people in the household related to the participant?
Physical exercise	Reconstructed from questions of "How often do you exercise?" and "How long do you exercise each time?"	Reconstructed from questions of "Number of days/week of each physical activity 10+ minutes?" and "Duration of each activity?"
Employment status	Are you currently employed?	Which of the following describes your current situation?

Abbreviation: HUNT, Nord-Trøndelag Health Study

Supplementary Table 2. Estimated 15-year cumulative risk of esophageal squamous cell carcinoma (ESCC) per 100,000 person-years and number need to survey to detect one case, depending on combination of risk factors

Sex	Age	Daily tobacco smoking	Alcohol consumption	Body mass index	Cumulative risk of ESCC	Number need to survey to detect one case
Male	40-49	Never	No	>25	6.5	15,411
Male	50-59	Former	Yes	>25	35.3	2,829
Male	50-59	Current	Yes	≤25	204.8	488
Male	≥60	Never	No	≤25	28.8	3,469
Male	≥60	Current	Yes	≤25	291.0	344
Female	40-49	Never	No	>25	6.7	14,880
Female	50-59	Former	Yes	>25	36.6	2,731
Female	50-59	Current	Yes	≤25	212.1	471
Female	≥60	Never	No	≤25	29.9	3,350
Female	≥60	Current	Yes	≤25	301.3	332

Supplementary Table 3. Performance of a basic model for predicting risks of esophageal squamous cell carcinoma within 5, 10, and 15 years of follow-up in the derivation cohort (HUNT) and within 5 years in the validation cohort (UK Biobank)

Outcomes	Area under the receiver-operating characteristic curve (95% CI)	Somers' D
Derivation cohort		
5 years	0.76 (0.59-0.93)	0.52
10 years	0.76 (0.63-0.88)	0.51
15 years	0.77 (0.68-0.87)	0.54
Internal cross-validation		
5 years	0.69 (0.48-0.90)	0.38
10 years	0.69 (0.53-0.84)	0.37
15 years	0.74 (0.63-0.85)	0.48
External validation cohort		
5 years	0.68 (0.62-0.74)	0.36

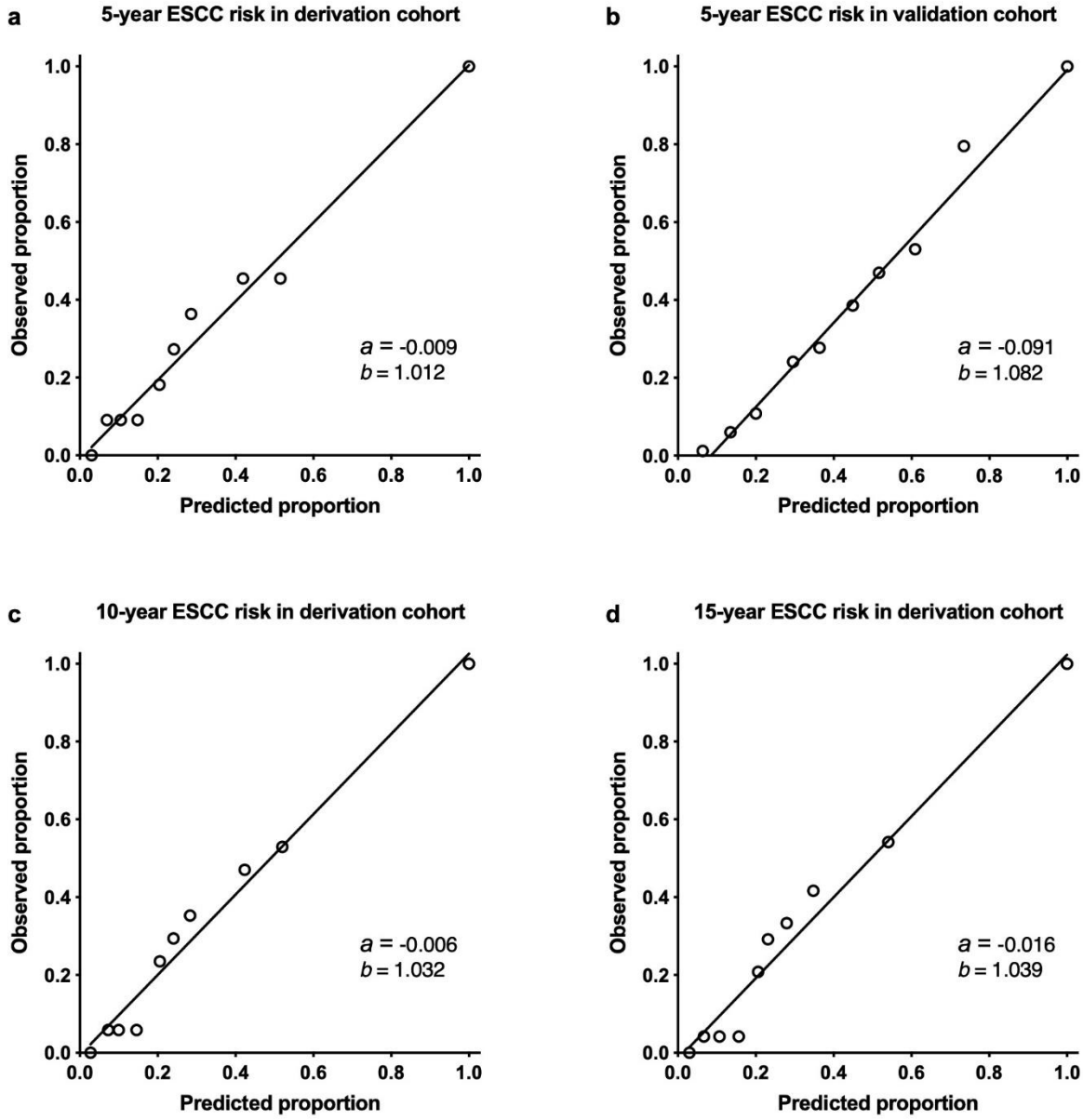
Abbreviation: HUNT, Nord-Trøndelag Health Study; CI, confidence interval

Supplementary Table 4. Sensitivity, specificity, positive predictive value, and cut-off evaluation criteria for predicting the risk of esophageal squamous cell carcinoma (ESCC) within 5 years of follow-up in the validation cohort (UK Biobank)

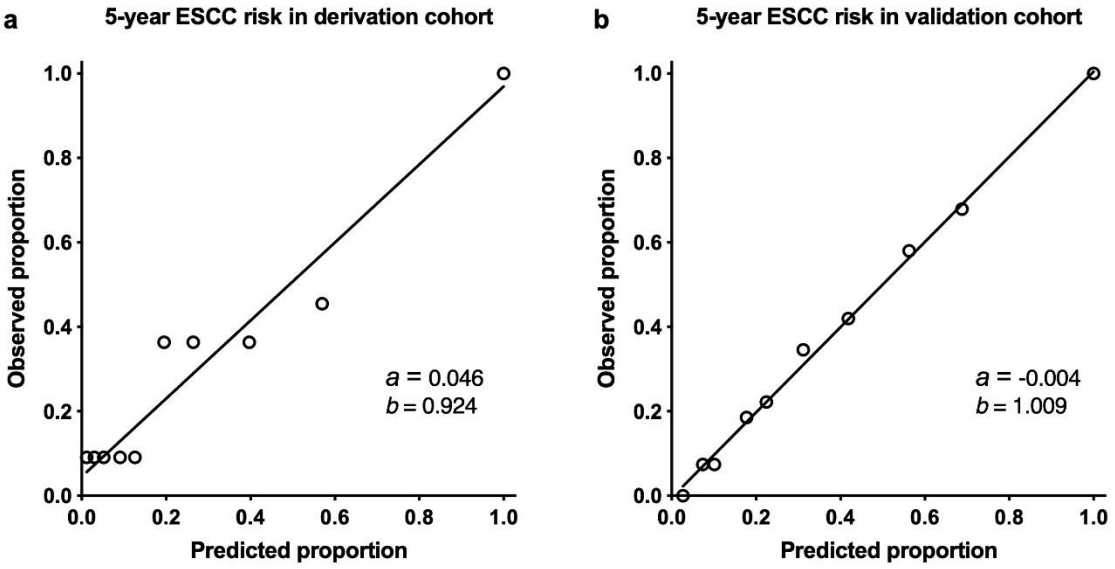
ESCa Score	Risk threshold per 100,000	True positive (count)	True negative (count)	False positive (count)	False negative (count)	Sensitivity (%)	Specificity (%)	Positive predictive value, per 100,000	Youden Index	Distance to (0,1)	Sensitivity-specificity equality
100	118.6	2	448,227	1,183	79	2.5	99.7	169	0.02	0.98	0.97
99	114.7	4	446,696	2,714	77	4.9	99.4	147	0.04	0.95	0.94
90	84.4	4	445,264	4,146	77	4.9	99.1	96	0.04	0.95	0.94
89	81.6	6	443,719	5,691	75	7.4	98.7	105	0.06	0.93	0.91
88	78.9	8	442,232	7,178	73	9.9	98.4	111	0.08	0.90	0.89
87	76.2	9	439,777	9,633	72	11.1	97.9	93	0.09	0.89	0.87
86	73.7	12	436,448	12,962	69	14.8	97.1	92	0.12	0.85	0.82
78	56.1	12	434,568	14,842	69	14.8	96.7	81	0.12	0.85	0.82
77	54.2	13	431,843	17,567	68	16.0	96.1	74	0.12	0.84	0.80
76	52.4	15	428,711	20,699	66	18.5	95.4	72	0.14	0.82	0.77
75	50.6	15	425,841	23,569	66	18.5	94.8	64	0.13	0.82	0.76
74	49.0	15	421,453	27,957	66	18.5	93.8	54	0.12	0.82	0.75
73	47.3	15	420,113	29,297	66	18.5	93.5	51	0.12	0.82	0.75
68	39.9	21	413,673	35,737	60	25.9	92.0	59	0.18	0.74	0.66
67	38.6	25	408,535	40,875	56	30.9	90.9	61	0.22	0.70	0.60
65	36.0	26	405,288	44,122	55	32.1	90.2	59	0.22	0.69	0.58
64	34.8	27	402,237	47,173	54	33.3	89.5	57	0.23	0.67	0.56
62	32.5	27	400,395	49,015	54	33.3	89.1	55	0.22	0.68	0.56
61	31.4	27	397,858	51,552	54	33.3	88.5	52	0.22	0.68	0.55
60	30.4	27	395,315	54,095	54	33.3	88.0	50	0.21	0.68	0.55
58	28.4	27	389,818	59,592	54	33.3	86.7	45	0.20	0.68	0.53
57	27.4	27	386,952	62,458	54	33.3	86.1	43	0.19	0.68	0.53
56	26.5	30	381,834	67,576	51	37.0	85.0	44	0.22	0.65	0.48
55	25.6	34	369,251	80,159	47	42.0	82.2	42	0.24	0.61	0.40
54	24.8	42	348,622	100,788	39	51.9	77.6	42	0.29	0.53	0.26
52	23.1	44	340,541	108,869	37	54.3	75.8	40	0.30	0.52	0.21

51	22.4	45	335,166	114,244	36	55.6	74.6	39	0.30	0.51	0.19
49	**20.9**	**46**	**332,560**	**116,850**	**35**	**56.8**	**74.0**	**39**	**0.31**	**0.50**	**0.17**
48	20.2	46	329,565	119,845	35	56.8	73.3	38	0.30	0.51	0.17
46	18.9	47	325,230	124,180	34	58.0	72.4	38	0.30	0.50	0.14
45	18.2	47	316,941	132,469	34	58.0	70.5	35	0.29	0.51	0.12
44	17.6	47	306,156	143,254	34	58.0	68.1	33	0.26	0.53	0.10
43	17.0	49	291,718	157,692	32	60.5	64.9	31	0.25	0.53	0.04
42	16.5	53	265,539	183,871	28	65.4	59.1	29	0.25	0.54	0.06
41	15.9	53	259,088	190,322	28	65.4	57.7	28	0.23	0.55	0.08
40	15.4	59	246,336	203,074	22	72.8	54.8	29	0.28	0.53	0.18
39	14.9	60	231,189	218,221	21	74.1	51.4	27	0.26	0.55	0.23
38	14.4	63	218,273	231,137	18	77.8	48.6	27	0.26	0.56	0.29
33	12.1	64	208,899	240,511	17	79.0	46.5	27	0.25	0.57	0.33
32	11.7	65	200,988	248,422	16	80.2	44.7	26	0.25	0.59	0.36
30	10.9	66	185,400	264,010	15	81.5	41.3	25	0.23	0.62	0.40
29	10.6	68	167,947	281,463	13	84.0	37.4	24	0.21	0.65	0.47
28	10.2	70	151,199	298,211	11	86.4	33.6	23	0.20	0.68	0.53
27	9.9	75	124,587	324,823	6	92.6	27.7	23	0.20	0.73	0.65
26	9.5	75	104,682	344,728	6	92.6	23.3	22	0.16	0.77	0.69
25	9.2	75	101,243	348,167	6	92.6	22.5	22	0.15	0.78	0.70
17	7.0	78	74,890	374,520	3	96.3	16.7	21	0.13	0.83	0.80
16	6.8	80	57,062	392,348	1	98.8	12.7	20	0.11	0.87	0.86
14	6.3	81	45,422	403,988	0	100.0	10.1	20	0.10	0.90	0.90
13	6.1	81	35,411	413,999	0	100.0	7.9	20	0.08	0.92	0.92
12	5.9	81	27,425	421,985	0	100.0	6.1	19	0.06	0.94	0.94
1	4.1	81	12,780	436,630	0	100.0	2.8	19	0.03	0.97	0.97
0	3.9	81	0	449,410	0	100.0	0.0	18	0.00	1.00	1.00

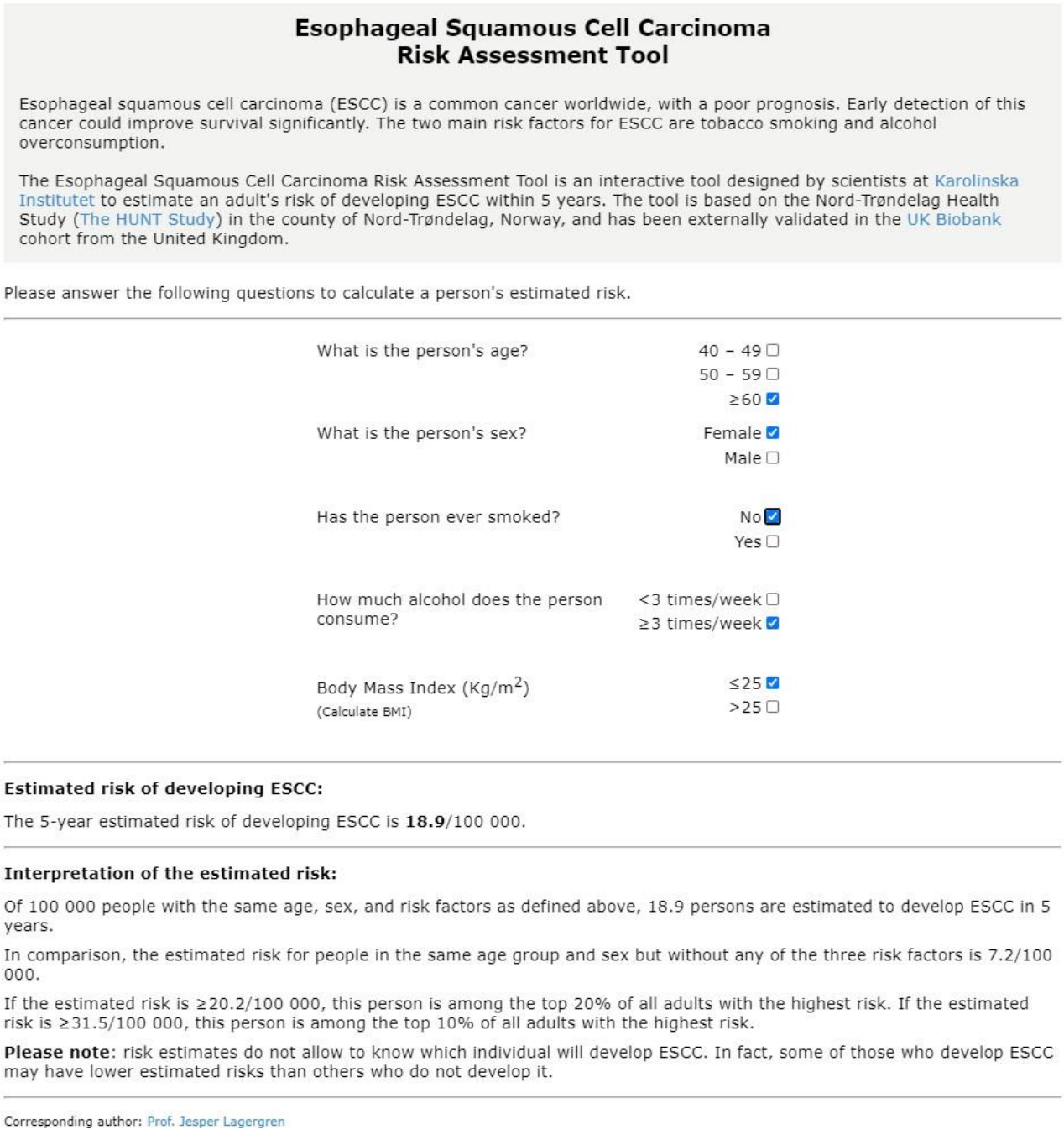
Supplementary Figure 1. Basic model calibration of the observed cumulative proportion of esophageal squamous cell carcinoma (ESCC) and predicted cumulative risk of ESCC within 5, 10, and 15 years of follow-up in the derivation cohort (HUNT) and within 5 years in the validation cohort (UK Biobank)



Supplementary Figure 2. ESCaScore model calibration of observed cumulative proportion of esophageal squamous cell carcinoma (ESCC) and predicted cumulative risk of ESCC within 5 years of follow-up in the derivation cohort (HUNT) and the validation cohort (UK Biobank)



Supplementary Figure 3. Screenshot of web-based 5-year esophageal squamous cell carcinoma risk assessment tool



Esophageal Squamous Cell Carcinoma Risk Assessment Tool

Esophageal squamous cell carcinoma (ESCC) is a common cancer worldwide, with a poor prognosis. Early detection of this cancer could improve survival significantly. The two main risk factors for ESCC are tobacco smoking and alcohol overconsumption.

The Esophageal Squamous Cell Carcinoma Risk Assessment Tool is an interactive tool designed by scientists at Karolinska Institutet to estimate an adult's risk of developing ESCC within 5 years. The tool is based on the Nord-Trøndelag Health Study (The HUNT Study) in the county of Nord-Trøndelag, Norway, and has been externally validated in the UK Biobank cohort from the United Kingdom.

Please answer the following questions to calculate a person's estimated risk.

What is the person's age?	40 – 49 ☐ 50 – 59 ☐ ≥60 ☑
What is the person's sex?	Female ☑ Male ☐
Has the person ever smoked?	No ☑ Yes ☐
How much alcohol does the person consume?	<3 times/week ☐ ≥3 times/week ☑
Body Mass Index (Kg/m^2) (Calculate BMI)	≤25 ☑ >25 ☐

Estimated risk of developing ESCC:

The 5-year estimated risk of developing ESCC is **18.9**/100 000.

Interpretation of the estimated risk:

Of 100 000 people with the same age, sex, and risk factors as defined above, 18.9 persons are estimated to develop ESCC in 5 years.

In comparison, the estimated risk for people in the same age group and sex but without any of the three risk factors is 7.2/100 000.

If the estimated risk is ≥20.2/100 000, this person is among the top 20% of all adults with the highest risk. If the estimated risk is ≥31.5/100 000, this person is among the top 10% of all adults with the highest risk.

Please note: risk estimates do not allow to know which individual will develop ESCC. In fact, some of those who develop ESCC may have lower estimated risks than others who do not develop it.

Corresponding author: Prof. Jesper Lagergren

Note: original website can be reached here: https://sites.google.com/view/escc-ugis-ki