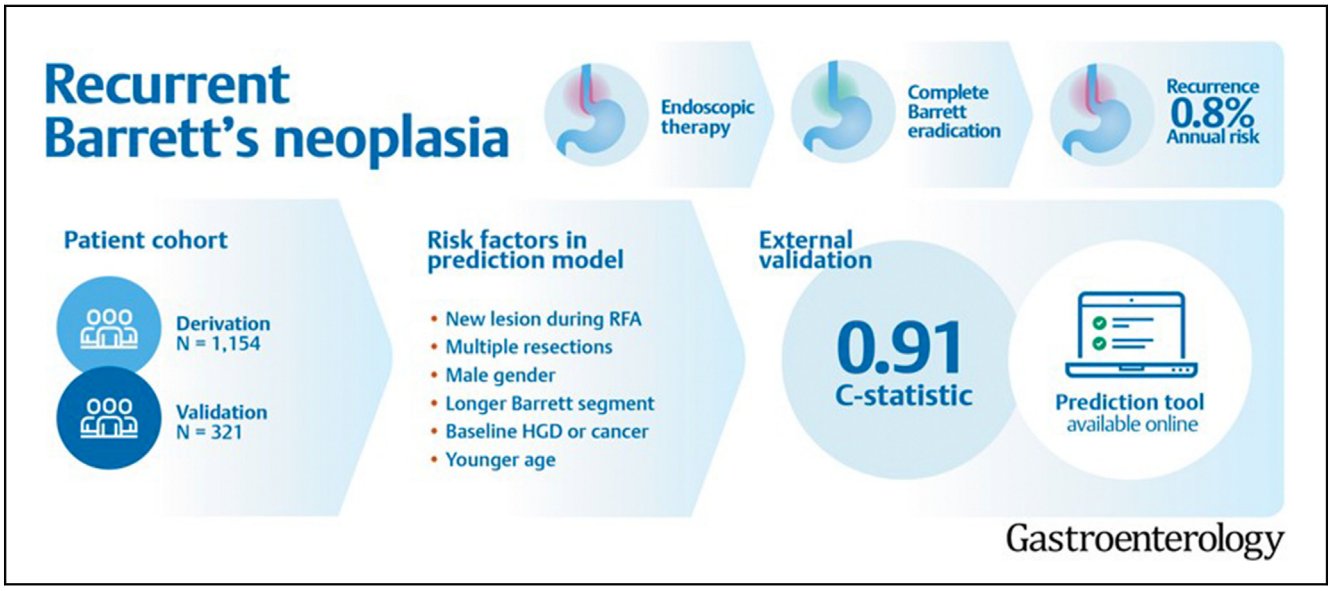
Gastroenterology 2022;163:285–294

Dysplastic Recurrence After Successful Treatment for Early Barrett's Neoplasia: Development and Validation of a Prediction Model

Sanne N. van Munster,[1,2] Esther Nieuwenhuis,[2] Raf Bisschops,[3] Hilde Willekens,[3] Bas L. A. M. Weusten,[2,4] Lorenza Alvarez Herrero,[2] Auke Bogte,[4] Alaa Alkhalaf,[5] Ed B. E. Schenk,[5] Erik J. Schoon,[6,7] Wouter Curvers,[6] Arjun D. Koch,[8] Pieter Jan F. de Jonge,[8] Tjon J. Tang,[9] Wouter B. Nagengast,[10] Jessie Westerhof,[10] Martin H. M. G. Houben,[11] Stefan Seewald,[12] Martinus J. C. Eijkemans,[13] Jacques J. G. H. M. Bergman,[1] and Roos E. Pouw[1]

[1]Department of Gastroenterology and Hepatology, Amsterdam Gastroenterology Endocrinology and Metabolism, Amsterdam University Medical Centers, Amsterdam, The Netherlands; [2]Department of Gastroenterology and Hepatology, Sint Antonius Hospital, Nieuwegein, The Netherlands; [3]Department of Gastroenterology and Hepatology, University Hospitals Leuven, Katholieke Universiteit, Leuven, Belgium; [4]Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; [5]Department of Gastroenterology and Hepatology, Isala Clinics, Zwolle, The Netherlands; [6]Department of Gastroenterology and Hepatology, Catharina Hospital, Eindhoven, The Netherlands; [7]School for Oncology and Developmental Biology, Faculty of Health, Maastricht University, Maastricht, The Netherlands; [8]Department of Gastroenterology and Hepatology, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands; Department of Gastroenterology and Hepatology, IJsselland Hospital, Cappelle aan den Ijssel, The Netherlands; [10]Department of Gastroenterology and Hepatology, University Medical Center Groningen, Groningen University, Groningen, The Netherlands; [11]Department of Gastroenterology and Hepatology, Haga Teaching Hospital, Den Haag, The Netherlands; [12]Centre of Gastroenterology, Klinik Hirslanden, Zürich, Switzerland; and [13]Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht University, Utrecht, The Netherlands



See editorial on page 39.

BACKGROUND & AIMS: The combination of endoscopic resection and radiofrequency ablation is the treatment of choice for eradication of Barrett's esophagus (BE) with dysplasia and/or early cancer. Currently, there are no evidence-based recommendations on how to survey patients after successful treatment, and most patients undergo frequent follow-up endoscopies. We aimed to develop and externally validate a prediction model for visible dysplastic recurrence, which can be used to personalize surveillance after treatment. **METHODS:** We collected data from the Dutch Barrett Expert Center Registry, a nationwide registry that captures outcomes from all patients with BE undergoing endoscopic treatment in the Netherlands in a centralized care setting. We used predictors related to demographics, severity of reflux, histologic status at baseline, and treatment characteristics. We built a Fine and Gray survival model with least absolute shrinkage and selection operator penalization to predict the incidence of visible dysplastic recurrence after initial successful treatment. The model was validated externally in patients with BE treated in

Switzerland and Belgium. **RESULTS:** A total of 1154 patients with complete BE eradication were included for model building. During a mean endoscopic follow-up of 4 years, 38 patients developed recurrent disease (1.0%/person-year). The following characteristics were independently associated with recurrence (strongest to weakest predictor): a new visible lesion during treatment phase, higher number of endoscopic resection treatments, male sex, increasing BE length, high-grade dysplasia or cancer at baseline, and younger age. External validation showed a C-statistic of 0.91 (95% confidence interval, 0.86–0.94) with good calibration. **CONCLUSIONS:** This is the first externally validated model to predict visible dysplastic recurrence after successful endoscopic eradication treatment of BE with dysplasia or early cancer. On external validation, our model has good discrimination and calibration. This model can help clinicians and patients to determine a personalized follow-up strategy.

Keywords: Barrett's Esophagus; Esophageal Adenocarcinoma; Endoscopic Eradication Therapy; Radiofrequency Ablation.

A combination of endoscopic resection (ER) for any visible abnormalities followed by endoscopic radiofrequency ablation (RFA) for the remaining flat Barrett's esophagus (BE) is the treatment of choice for BE with dysplasia and/or early cancer.[1] This combination has been found to be safe and effective for eradicating dysplasia and/or early BE cancer and allows for complete eradication of BE. It has been reported that 90%–95% of patients achieve complete eradication of all visible Barrett's epithelium (complete eradication of BE [CE-BE]).[1–4]

Because of the risk of recurrence after CE-BE, endoscopic follow-up is performed to identify and treat recurrences at early stages to prevent progression to advanced cancer. Reported recurrence risks vary widely, from 1% to 20% per person-year.[5–8] These differences can be partially explained by heterogeneous definitions for CE-BE and for recurrence. Centralization of BE treatment may play a role as well, with lower recurrence rates reported for patients treated in expert centers.[9] Most studies published to date are limited by small sample size and short duration of follow-up.[6,7]

With the lack of reliable data on the risk of recurrence, recommendations for follow-up after CE-BE are based on expert opinion. The strategies derive from the pre-ablation era, when, after ER for visible abnormalities, endoscopic follow-up was initiated of the remaining BE. With RFA, eradication of the residual BE can be accomplished in the vast majority of patients. Hence, it is reasonable to assume that this should reduce recurrence rates and that surveillance intervals can be widened.

Understanding the clinical and treatment determinants of recurrent disease may have important implications for development of follow-up regimens. The objective of the current study was to develop and externally validate a prognostic model to predict visible recurrent dysplasia after CE-BE to further develop personalized post-RFA surveillance strategies.

WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT

Endoscopic therapy is the treatment of choice for eradication of early neoplasia in Barrett's esophagus. Follow-up after successful treatment is performed frequently with strategies based on expert opinion because reliable long-term data are lacking.

NEW FINDINGS

Dysplasia recurs at 1.0% per patient-year. We developed an externally validated prognostic model on the basis of easily available characteristics, to predict dysplastic recurrence, with good performance.

LIMITATIONS

The low recurrence risk has a risk for overfitting, we attempted to minimize this by external validation.

IMPACT

The prediction model may allow personalized surveillance after treatment. It is conceivable that in the near future we will advise more stringent surveillance for some, although we will not start surveillance for others.

Methods

The model was built using data from the Barrett Expert Center (BEC) registry (Netherlands Trial Register, NL7039).[10] This registry captures outcomes for all patients with early Barrett's neoplasia in the Netherlands since 2008 who underwent endoscopic treatment. Treatment for early BE neoplasia in the Netherlands is centralized in 9 BECs, which means that every patient in the Netherlands is treated in one of these tertiary referral centers. All BECs follow a joint treatment and follow-up protocol; the endoscopists and pathologists (1–2 per center, depending on the volume) participated in a joint training program and the minimum caseload is 10 new patients with high-grade dysplasia (HGD)/esophageal adenocarcinoma (EAC) per center per year.

The model was validated on 2 external, separate databases (Supplementary Material). The Zurich database is a prospective database including all patients with BE treated in Hirslanden Klinik, Zurich, Switzerland. The Leuven database is a prospective database that included all patients treated at the University Hospital Leuven, Belgium.[11] Both centers have a tertiary referral function for treatment of BE neoplasia. Endoscopists working in these centers were jointly trained with the endoscopists from the Dutch expert centers and participated in jointly organized European training programs since 2010.

Abbreviations used in this paper: **BE, Barrett's esophagus; BEC, Barrett Expert Center; CE-BE, complete eradication of Barrett's esophagus; CI, confidence interval; EAC, esophageal adenocarcinoma; ER, endoscopic resection; HGD, high-grade dysplasia; HR, hazard ratio; IM, intestinal metaplasia; LASSO, least absolute shrinkage and selection operator; LGD, low-grade dysplasia; RFA, radiofrequency ablation.**

Most current article

0016-5085
https://doi.org/10.1053/j.gastro.2022.03.020

Treatment and Follow-Up Protocol

All patients underwent endoscopic workup and staging at baseline using high-definition white light endoscopy and optical chromoscopy, with careful inspection and documentation of the Prague C&M criteria[12]; presence of visible lesions; esophagitis; esophageal stenosis; or other abnormalities.

Visible abnormalities, that is, nonflat lesions and/or lesions with irregular mucosal patterns with a suspicion for neoplasia, were removed with ER for histologic staging, followed by 3-monthly RFA until all BE was eradicated endoscopically. Random 4-quadrant biopsies were then performed <1 cm below the neo-squamocolumnar junction. Successful treatment was defined as CE-BE. A failure for CE-BE had either persisting visible BE endoscopically or persisting dysplasia in biopsies just below the cardia. In line with prior studies, patients with complete endoscopic eradication of BE yet with persisting focal intestinal metaplasia (IM) in the biopsies distal to the neo-squamocolumnar junction were defined as CE-BE.[2]

All CE-BE patients entered follow-up. All follow-up endoscopies were performed with high-definition white light endoscopy and optical chromoscopy. The Dutch regimen changed over time in terms of surveillance intervals and histologic sampling (Supplementary Table 1). In 2008, surveillance was performed 3-monthly in year 1, annually from year 2 to year 5, and every 2–3 years afterwards. In 2015, we abandoned the extra 3-month endoscopies in year 1 due to low clinical relevance.

In 2008, we started with random biopsies from the neosquamous epithelium along the length of the initial BE, and from the cardia <1 cm below the neosquamocolumnar junction. In 2013, we abandoned the random neosquamous epithelium biopsies and in 2016, we abandoned the random cardia biopsies during follow-up endoscopies. From 2016 onward, we only performed histologic sampling from endoscopic abnormalities.

Data Collection and Data Management

Information regarding baseline characteristics, the treatment phase and long-term follow-up was collected in a joint database. Follow-up data were collected until January 1, 2020. Database quality control was performed by checking data against source documents for all patients who reached a primary end point and for 50% of the remaining patients. Data and/or images for all patients who reached a primary end point were discussed in interactive meetings with the research study group. All fields were examined for missing data, nonlogical values, and outliers, which were completed or corrected.

The BEC registry was merged with the nonpublic microdata from Statistics Netherlands for survival outcomes, including date and cause of death.

Study Population

For the current study, we included all patients from the BEC, Leuven, and Zurich registries who underwent at least 1 RFA treatment and achieved CE-BE before December 31, 2018, to ensure sufficient duration of follow-up. For the BEC registry, this is the same cohort of patients as published recently with the aim to report long-term outcomes.[10]

Study End Points

The primary end point was recurrent disease, defined as a histologic finding of low-grade dysplasia (LGD), HGD, or EAC in the esophagus or cardia during follow-up. This diagnosis could be established either on biopsy samples or on endoscopic resection specimens. Progression to advanced EAC (>T1 EAC and/or lymph node and/or distant metastasis) was included in this definition.

To assess the robustness of our outcomes, sensitivity analysis was performed, with recurrence of HGD or EAC as outcome and recurrence of LGD considered as sustained eradication.

Definition and Description of Potential Predictors

We included patient and treatment characteristics that would be known to the physician at the time of CE-BE and with clinically or biologically plausible effects on the risk for recurrent disease. These included demographics (age at the time of first treatment and sex); characteristics defining the severity of reflux disease (eg, baseline BE length, poor healing and/or poor squamous regeneration during treatment, and persisting reflux esophagitis at the end of treatment), characteristics defining histologic abnormalities (worst pathology at baseline, presence of a new visible lesion, ie, "incident lesion" during ablation), and characteristics of the treatment course (eg, number of treatment sessions and persisting IM in the cardia after treatment). Poor healing was defined as incomplete healing (active ulcers) at least 3 months after treatment, resulting in postponement of treatment and/or incomplete squamous regeneration (<50%) after treatment. Persisting reflux esophagitis at the end of treatment was defined as endoscopically visible evidence of reflux esophagitis Los Angeles classification grade B or higher.[13]

For all patients, information on all of these variables were available except for 35 patients (3%) in the BEC cohort, in whom no cardia biopsies were obtained at the end of treatment. We therefore included this variable in 2 ways in our analysis: first as a categorical variable with 3 levels ("no IM in cardia biopsies," "IM in cardia biopsies," or "no cardia biopsies performed") and by adding a new variable with single imputation for those patients without biopsies. The Leuven and Zurich registry had no missing values.

Statistics

Baseline characteristics were analyzed using standard descriptive statistics. Continuous variables were presented as mean (SD) and as median with interquartile range for normally distributed and skewed data, respectively. Categorical variables were presented as numbers with percentages and 95% confidence intervals (CIs) were obtained using internal bootstrapping.

The prognostic model was developed using a Fine and Gray survival model. The time-to-event analysis was time between last treatment endoscopy and occurrence of the event of interest (recurrent dysplasia), the competing risk (unrelated death), or censoring (the last follow-up endoscopy). Because recurrences are generally asymptomatic and therefore only detected at regularly scheduled surveillance endoscopies, the true timing of recurrent disease is unknown. To correct for this interval censoring, timing of recurrence was defined as the

moment in the middle of the interval between the last endoscopy without recurrence, and the first endoscopy with recurrence.

To select potential predictors, we used the least absolute shrinkage and selection operator (LASSO) algorithm and hazard ratios (HRs) were estimated by means of this method.[14] The functional form (linear vs nonlinear relations with the outcome) was checked for all continuous variables. The proportional hazard assumption was checked using the Schoenfeld residuals.

Model building consisted of leave-one-out cross-validation for choosing the LASSO penalty. In addition, we performed leave-one-out cross-validation for internal validation to quantify statistical optimism in performance. The final model was assessed for overall model performance (Brier score), discrimination (Harrell's C-statistic), and calibration in both internal and external validation. Bootstrapping was performed to obtain a 95% CI for the C-statistic. Details about all steps performed in the development and validation of the model can be found in the Supplementary Material.

The Fine and Gray model was considered the best model for our dataset because this model can take competing risks into account and LASSO variable selection was preferred, given the low number of events.[15] For sensitivity analysis, we also fitted a Fine and Gray model with backward selection with variable selection based on Akaike Information Criterion. In addition, 2 Cox proportional hazard models were built: one with variable selection based on LASSO and one based on backward selection using Akaike Information Criterion.

No formal sample size was calculated for our primary analysis using LASSO penalization, and the number of predictors in our model was much smaller than the number of outcomes. Data collection was carried out using R, version 3.6.3 with the following packages: cmprsk, crrp, survival, glmnet, shiny, ROCR, survminer, prodlim, ggplot2.

Ethics

The Institutional Review Board of the Amsterdam University Medical Centers (former AMC) declared that the BEC registry was not subject to the Medical Research Involving Human Subjects Act (*wet op medisch-wetenschappelijk onderzoek met mensen* in Dutch) and waived the need for formal ethical review and patient-informed consent. Patients were approached through an opt-out card with the possibility to object against participation. For the prospective part of the registry, all patients gave written informed consent. Written informed consent for prospective registration was also obtained in Leuven after approval by the Ethical Committee of the University Hospitals Leuven (S52432). In Zurich, written informed consent was deemed unnecessary for prospective registration by the ethical board. All authors had access to the study data and had reviewed and approved the final manuscript.

Results

Definition and Baseline Characteristics of the Barrett Expert Center Cohort

A total of 1154 patients reached complete endoscopic and histologic remission of BE after RFA ± ER and were included for the current follow-up study (Figure 1). The mean follow-up was 4 (±2) years with 4 (±2) endoscopy per patient. We had a substantial number of patients with long-term follow-up in our cohort: 370 patients had follow-up over 5 years and 112 patients over 8 years. Overall, this contributed to 4690 person-years of follow-up. Only 17 patients (2%) were lost to follow-up. Baseline characteristics are reported in Table 1.

Recurrent Disease

Among the 1154 patients in our study, visible recurrent LGD, HGD, or EAC occurred in 38 patients. The worst histologic grade of recurrence was LGD (n = 14), HGD (n = 7), or EAC (n = 17). The annual recurrence risk was 1.0% (95% CI, 0.8–1.4) for recurrent LGD or worse and 0.7% (95% CI, 0.4–1.0) for recurrent HGD or worse. All recurrences were detected as visible BE and/or nonflat abnormalities during endoscopy. Recurrence occurred at a median of 30 months (interquartile range, 22–40 months) after CE-BE was established. Recurrences have been described in detail previously.[10]

Figure 2 shows the regular Kaplan–Meier estimate for recurrent disease (ie, considering unrelated death as uninformative censoring) and the cumulative incidence curve (ie, considering unrelated death as competing event).

Unadjusted Associations Between Potential Predictors and Recurrence

In univariable analysis, patients with longer pretreatment BE segments were more likely to develop recurrence during follow-up (Table 2). Also, a higher number of ER treatments, a higher number of RFA treatments, and development of an incident lesion during the treatment phase were associated with a higher risk for recurrence. Although not statistically significant, patients with HGD or EAC at baseline had a 2.5 times higher chance of developing recurrence compared with patients with LGD at baseline.

Effect estimates for the Fine and Gray analysis were comparable with those resulting from regular Cox analysis.

Multivariate Model Building and Predictive Performance

Selected variables for the multivariable LASSO model were age, sex, baseline pathology, BE length, number of ER treatments, and incident lesions (Table 2). Younger age was associated with a higher risk for recurrence (HR, 1.01); as was male sex (HR, 1.37); HGD or EAC at baseline compared with LGD (HR, 1.02); increasing length of BE (HR, 1.16); higher number of ER treatments (HR, 1.18); and an incident lesion (HR, 2.88). Model assumptions were met (Supplementary Figure 1).

For example, a 50-year-old man with a 10-cm-long BE with EAC and 2 ER sessions, including 1 for an incident lesion, had a predicted risk for recurrence of 16% during the first 2 years, which increased to 48% during 7 years. This is an extreme example and we only had 3 such patients (3 of 1154 [0.3%]) in our cohort.

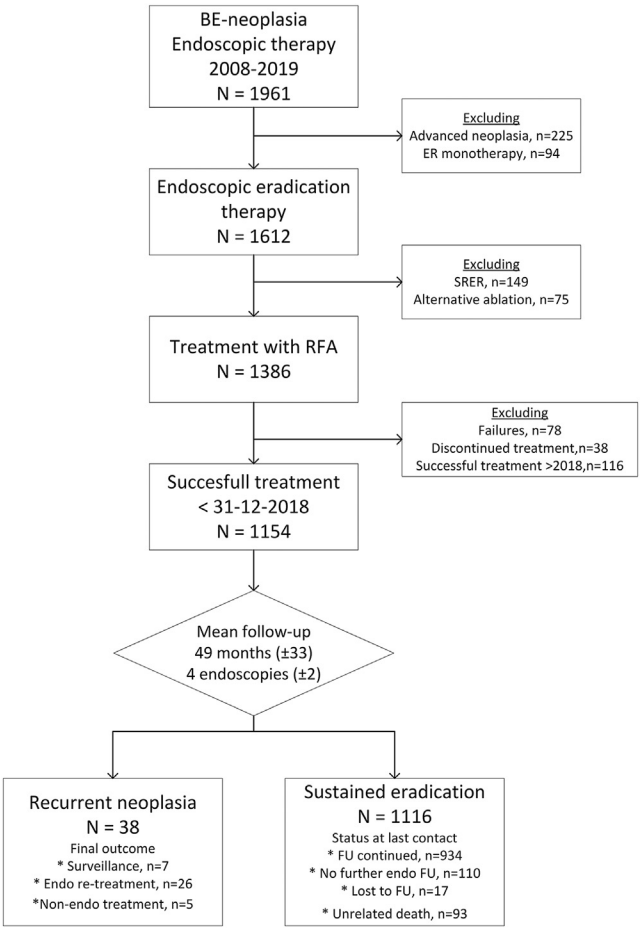


Figure 1. Patient flow and definition of our study cohort.

In contrast, a 65-year-old man with C2M5 BE who underwent a single ER for HGD followed by RFA had a 2% risk of developing recurrence during the first 2 years, and this increased to a cumulative 8% risk during 7 years. This is a much more representative case for our population and approximately 50% of our cohort had a comparable or lower risk.

The optimism-corrected concordance index for the prediction model was 0.76 (95% CI, 0.73–0.79). The lambda plot, coefficient plot, and calibration plots can be found in Supplementary Figures 2, 3, and 4.

Sensitivity Analysis

To test the robustness of our findings, sensitivity analyses were performed using different statistical models and a different definition of the outcome (Supplementary Table 2). Using backward regression techniques, fewer variables were included in the models. The most important variables that were selected in all 4 models were BE length and incident lesion during the treatment phase.

External Validation

Our model was externally validated on the Leuven and Zurich RFA registries, including 204 and 117 patients with successful RFA ± prior ER, respectively. Baseline characteristics for the 3 cohorts were comparable, with the exception of 2 variables (Table 1). The proportion of patients with a visible lesion at baseline appeared higher in Zurich (81%) compared with BEC (62%) and Leuven (60%). Furthermore, the proportion of baseline LGD diagnosis was lower in Leuven (8%) compared with the BEC (26%) and Zurich (27%) registry.

In Leuven, 14 of 204 patients (7%) developed recurrence during a median of 40 months (interquartile range, 19–78 months); worst histology was LGD (n = 2 [14%]), HGD/low-risk EAC (n = 11 [79%]), or advanced EAC (n = 1 [7%]). The annual risk was 1.6 (95% CI, 0.9–3.0) for recurrence of LGD+ and 1.3 (95% CI, 0.8–2.5) for HGD+. The risk for unrelated death was 21 of 204 (10%).

In Zurich, 5 of 117 patients (4%) developed recurrence during median 42 months (18–70 months), consisting of LGD (n = 2), HGD/low-risk EAC (n = 2), or advanced EAC (n = 1). The annual risk was 1.0 (95% CI, 0.4–2.0) for LGD, HGD, or EAC combined and 0.6 (95% CI, 0.2–1.8) for HGD or EAC combined. The risk for unrelated death was 4 of 117 (4%).

We combined the Leuven and Zurich datasets and assessed overall performance, discrimination, and validation of the created model on this external dataset. The Brier score was 0.38 (95% CI, 0.10–0.74), with lower scores indicating better overall performance (range, 0–1). The C-index was 0.92 (95% CI, 0.86–0.94), with higher scores indicating better discrimination (range, 0–1). The calibration plot at 5 years is shown in Figure 3 and indicates that for a predicted risk for recurrence within 5 years of <10%, the predicted and observed risks were comparable, but for predicted risks >10%, the model tended to underestimate the actual risk. Model performance for the 2 external datasets separately is shown in Supplementary Table 3 and Supplementary Figure 5.

Discussion

This is the first study to develop and externally validate a prediction model for visible dysplastic recurrence after successful endoscopic treatment of early Barrett's neoplasia on a large dataset with significant long-term follow-up data, and is an important step toward personalized post-treatment surveillance. We included 1154 Dutch patients with a mean follow-up of 4 years per patient for model building, and validated the model on data from 321 patients with a comparable length of follow-up treated in Belgium and Switzerland. We fitted a model for the incidence of recurrent LGD, HGD, or EAC, taking into account the risk for unrelated death, and we found 6 factors that independently predicted recurrence. The created model could discriminate well between patients with and without recurrence in an external dataset with excellent discrimination (C-statistic of 0.92) and good calibration, especially for low predicted risks, as in the majority of patients. Our model is easy to use (https://barrett-recurrence.shinyapps.io/Barrett/) (Figure 4) and may guide individualized post-treatment surveillance for patients with BE.

Recurrence rates after treatment vary widely between different studies. This can be partially explained by

Table 1. Baseline Characteristics for 1154 Patients Included in Our Cohort, Stratified for the Primary Outcome

Variable	BEC registry			Leuven validation registry	Zurich validation registry
	All patients (n = 1154)	No recurrence (n = 1116)	Recurrence (n = 38)	(n = 204)	(n = 117)
Demographic characteristics					
Sex, *male*, n (%)	947 (82)	914 (82)	33 (87)	173 (85)	100 (86)
Age, *y*, mean ± SD	64 ± 9	64 ± 9	62 ± 9	63 ± 11	64 ± 10
Baseline Barrett					
BE length, *cm*, mean ± SD	3 ± 3	3 ± 3	5 ± 4	3 ± 3	4 ± 3
Circumferential maximum, *cm*, mean ± SD	5 ± 3	5 ± 3	7 ± 3	5 ± 3	6 ± 4
Reflux esophagitis, n (%)	34 (3)	34 (3)	0 (0)	—	—
Visible lesion, n (%)	717 (62)	691 (62)	26 (68)	117 (57)	98 (84)
Worst pathology, n (%)					
LGD	306 (26)	302 (27)	4 (11)	16 (8)	32 (27)
HGD	363 (32)	350 (31)	13 (34)	123 (60)	48 (41)
EAC	485 (42)	464 (42)	21 (55)	65 (32)	37 (32)
m-EAC	455 (39)	434 (39)	21 (55)	60 (29)	34 (29)
sm1-EAC	30 (3)	30 (3)	0 (0)	5 (3)	3 (3)
Treatment					
ER, n (%)	719 (62)	691 (62)	28 (74)	122 (60)	95 (81)
Poor regression after ER, n (%)	34 (3)	32 (3)	2 (5)		4 (3)
No. of RFA sessions, median (IQR)	2 (1–3)	2 (1–3)	3 (3–4)	2 (1–3)	1 (1–2)
Poor healing, n (%)	80 (7)	76 (7)	4 (11)	13 (7)	2 (2)
Incident lesion, n (%)	72 (6)	62 (6)	9 (24)	15 (7)	4 (3)
Esophagitis after treatment, n (%)	109 (9)	103 (9)	6 (16)	—	—
Cardia biopsies, at the end of treatment, n (%)					
No IM	1045 (91)	1010 (91)	35 (92)	—	102 (87)
IM	74 (6)	72 (6)	2 (5)	—	15 (13)
No biopsies	35 (3)	34 (3)	1 (3)	—	0 (0)
Duration of treatment, *mo*, median (IQR)	9 (5–13)	5 (8–13)	11 (9–16)	6 (3–11)	11 (8–15)
No. of endoscopies, median (IQR)	3 (2–4)	3 (2–4)	4 (3–5)	3 (2–4)	2 (2–3)

IQR, interquartile range.

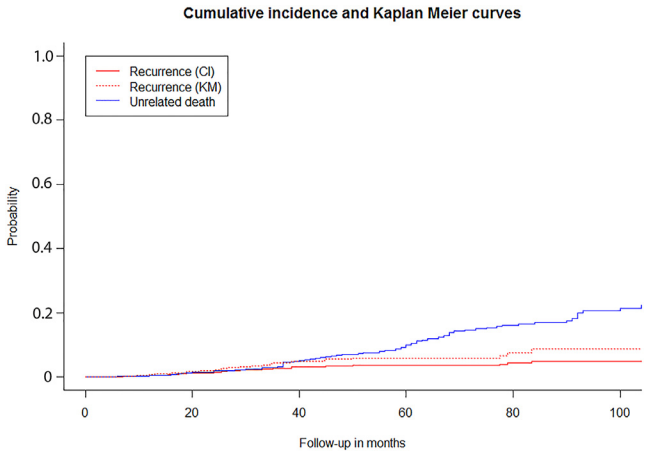


Figure 2. Kaplan–Meier (KM) and cumulative incidence curve. KM (*dashed red line*) and cumulative incidence curve (CI) (*red line*) for the risk of recurrent dysplasia, plotted against the risk of unrelated death (*blue line*).

heterogeneous definitions for successful treatment, as well as for recurrence; by different indications of treatment; differences in follow-up duration; differences in treatment protocols; varying expertise; and, potentially, due to actual differences in recurrence risks. We included only patients with dysplastic BE at baseline as indication for treatment, in line with current guidelines.[1] We defined *recurrence* as recurrent visible LGD, HGD, or EAC, but not nondysplastic BE.[2,11] Most studies have reported the incidence of recurrent nondysplastic BE, but have not generally reported the rate of recurrence with dysplasia.[6,8,16] Recurrent LGD or worse appears to be a more suitable end point because recurrence of nondysplastic BE is usually limited to a small surface area; can be treated easily and has minimal risk for progression. The clinical relevance of this end point appears low.

For our prediction model, we hypothesized that the following 3 overarching themes are associated with recurrence: the severity of reflux disease; the severity of

Table 2. Univariable and Multivariable Fine and Gray Model

Covariate	Univariable analysis, HR (95% CI)	Multivariable LASSO model, HR
Age	0.99 (0.95–1.01)	0.99
Sex, *male*	0.66 (0.26–1.68)	0.88
Worst pathology	2.52 (0.89–7.09)	1.02
BE length	1.18 (1.12–1.26)	1.16
Incident lesion	4.34 (2.05–9.31)	2.88
Poor healing	1.46 (0.52–4.10)	—
Persisting esophagitis	1.57 (0.67–3.70)	—
No. of ERs	1.63 (1.17–2.26)	1.18
No. of RFAs	1.33 (1.04–1.70)	—
Persisting IM in cardia	1.34 (0.48–3.87)	—
Baseline hazard for 2 y	NA	0.985
Baseline hazard for 5 y	NA	0.962

NOTE. Age was modeled in years. Sex was coded as 1 for female. Worst pathology was coded as 1 for HGD or worse. BE length was the maximum extent of BE at baseline in cm. Incident lesion was defined as a dysplastic visible lesion requiring resection that was noted during the ablation phase. Poor healing was defined as incomplete healing (active ulcers) or incomplete squamous regeneration (<50%) resulting in postponement of treatment. Persisting esophagitis was defined as active reflux esophagitis grade B or higher at the moment of complete eradication. The number of endoscopic resections, radiofrequency ablation, and total treatment endoscopies were modeled continuously. The model is available at: https://barrett-recurrence.shinyapps.io/Barrett/
NA, not applicable

histologic abnormalities; and abnormalities during the treatment course. We defined these overarching themes in several baseline and treatment characteristics. Other studies have not assessed the full range of potential predictors.[6–9]

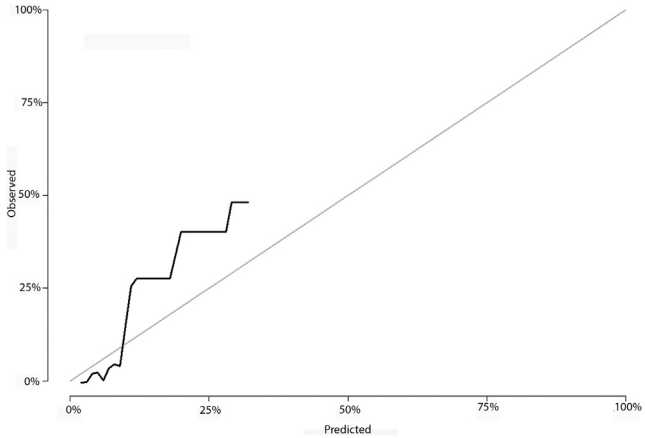


Figure 3. Calibration plot at 5 years for external validation in the Zurich (n = 117) and Leuven (n = 204) RFA registries. The *horizontal axis* represents the predicted recurrence risk and the *vertical axis* the observed recurrence risk. The *gray line* represents perfect calibration, with the predicted risk equal to the observed risk. The *plot* indicates that for a predicted risk for recurrence within 5 years, the predictions are accurate. For higher predicted risks, the model tends to underestimate the risk for recurrence.

In line with prior studies, we found an increasing risk for recurrence along with increasing BE length and higher baseline histologic grades. We also found additional predictive variables, with occurrence of incident lesions during the treatment phase as most important predictor. The occurrence of an incident lesion during the treatment phase (ie, "a pop-up lesion") might indicate multifocal dysplasia, which should have required endoscopic resection at baseline, and/or disease progression during treatment.

Most follow-up studies after RFA have used Cox regression for presentation of results. We used a Fine and Gray analysis, which takes into account competing risks. A significant drawback of Cox regression is that it censors patients who die from unrelated causes, which is, in fact, a violation of the prerequisite of Cox regression that censoring is uninformative. A patient with continued endoscopic surveillance who is censored after the last contact (because the next endoscopy is scheduled in the future, that is, after the moment of data collection), is considered the same type of censoring as a patient who died of an unrelated cause. The first patient may indeed develop recurrence during continued follow-up with a risk comparable with that of the other patients in the dataset (uninformative censoring), whereas a deceased patient has zero risk of developing recurrent disease in the future (no uninformative censoring). A Fine and Gray approach considers this difference and models the risk for the outcome, taking into account patients who died of an unrelated cause. This might explain the inverse association between age and

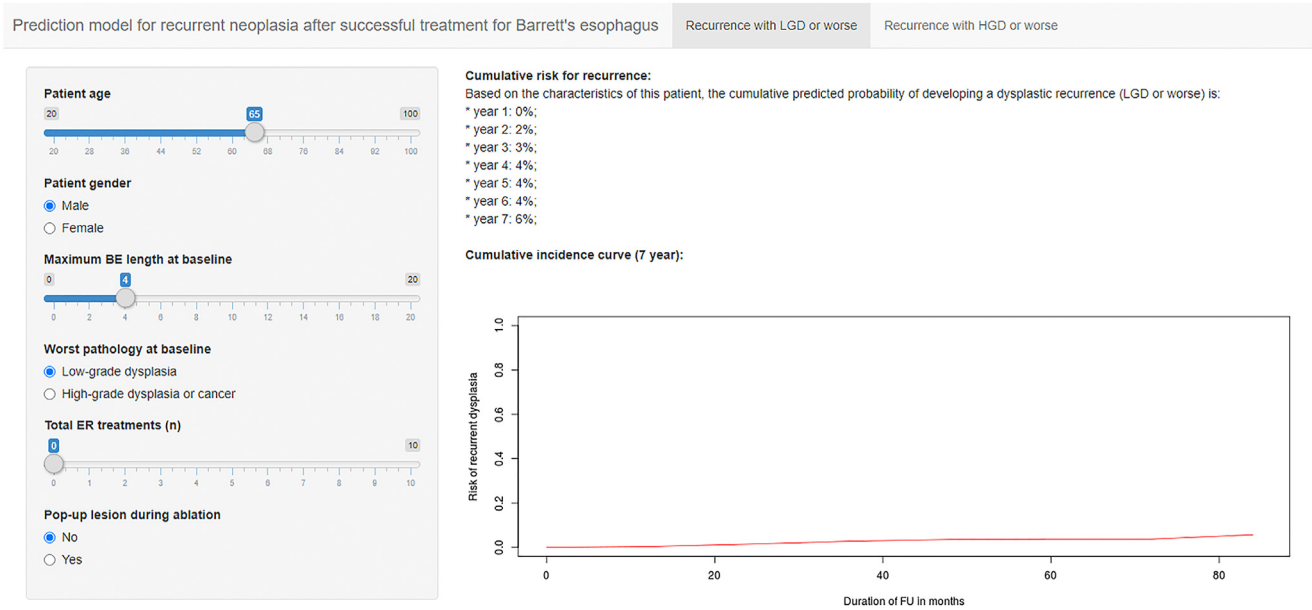


Figure 4. Easy-to-use online prediction model. The online risk prediction tool is available at https://barrett-recurrence.shinyapps.io/Barrett. This example shows recurrence risk for a 65-year-old man with a flat Barrett segment at baseline of C2M4 with LGD and underwent successful eradication therapy with RFA.

recurrence in the model. With increasing age, the risk of dying (from unrelated causes) will also increase and, as a result, the risk for developing recurrence will go down.

We selected predictive variables based on LASSO penalization, whereas other studies used backward or forward selection in their multivariable analyses. In short, LASSO penalization is a regression analysis that performs both variable selection and regularization to prevent overfitting. This technique is especially beneficial for model building with a large number of parameters in relation to the number of events, as in the current study. Our prediction model was externally validated on 2 separate datasets from expert centers in Leuven and Zurich. Baseline characteristics were comparable, with the exception of a higher proportion of patients with a visible lesion in Zurich; and a lower proportion of patients with LGD at baseline in Leuven. Recurrence risks were comparable among the datasets.

Our final model included 6 predictors and had good discrimination in internal and external validation (Harrell's C-statistic 0.76 and 0.92, respectively). We performed several sensitivity analyses, varying the model (ie, Cox regression vs Fine and Gray), the method for variable selection (ie, LASSO vs backward regression), and the outcome (ie, combining LGD, HGD, and EAC vs excluding LGD as an end point). Overall, our findings appeared robust in sensitivity analysis, but some differences are worth further elaboration. Consistently through all models, increasing BE length and an incident lesion during the treatment phase significantly predicted recurrence. In the Fine and Gray models, but not the Cox models, younger age was associated with recurrence. This difference might be explained by the fact that the Fine and Gray model take into account competing risks (ie, unrelated death). A younger patient is less likely to die from other causes but instead will enter a long follow-up period with a higher risk for recurrence compared with an older patient with a significant risk of unrelated death but not recurrent disease. Using LASSO penalization instead of backward regression, baseline histologic grade and number of ER sessions were also included in the model. The other way around, all variables selected with backward regression were also selected using LASSO. It is known that backward regression is uncertain for a model with a limited number of outcomes, such as our model. Based on Akaike Information Criterion and C-statistics, LASSO outperformed stepwise backward regression and the Fine and Gray model had improved performance compared with the Cox model.

This work has some limitations. We found 38 recurrences in a dataset of 1154 patients and this low number might limit the performance of our model with a risk for overfitting. This is reflected in the difference in the area under the curve for the 2 external validation datasets. Still, overall external validation showed good model performance. Other limitations include treatment of patients who developed nondysplastic BE during follow-up, which might underestimate the true dysplastic recurrence risk in our dataset. However, this occurred in only 6% of patients, either with short segment BE (0.4%) or tiny BE islands (5.6%).[10] The model did not correct for interval censoring, but this was corrected by defining recurrence in the middle of the interval between the last follow-up endoscopy and the endoscopy with recurrence. Our model used data from expert centers only, and this may limit the generalizability, although guidelines recommend treatment in expert centers only. Follow-up protocols changed over time, resulting in fewer endoscopies and less sampling, this may potentially affect the moment recurrences were found, but this appears unlikely to influence the incidence for recurrent dysplasia,

the end point of our study. We had no data on p53 staining, which makes it impossible to distinguish true recurrent disease from potential treatment failure that was initially missed. We built an easy-to-use prediction model with readily available parameters, but this may have led to impaired predictive value compared with a model with more detailed parameters that are not routinely performed in our country, such as extent of neoplasia and p53 staining. Although pathologists were highly experienced and extensively trained, the Dutch pathologists were jointly trained but the pathologists from Leuven and Zurich participated in other programs, which may have decreased interobserver agreement. Finally, concrete clinical recommendations for personalized follow-up cannot yet be provided on the basis of this study only.

This study also has important strengths. We built our model using a nationwide cohort that included all patients with BE with endoscopic treatment in the Netherlands. BE care in the Netherlands is centralized and performed in BECs only, with specifically trained endoscopists and pathologists, a common treatment and follow-up protocol, and a required annual case load. This resulted in homogeneous care and collection of high-quality data, with no missing data and only 2% of patients lost to follow-up. Our model was the first model to take into account competing risks in model development, and we systematically assessed a wide range of predictors. Finally, our model showed excellent discrimination in external validation in 2 high-quality, independent datasets of 322 patients treated in BECs in Europe. The centralized setting of our study reflects current guidelines in the Netherlands[17] and Europe,[1] which recommend restricting treatment of BE neoplasia to expert centers.

Pending the following steps, the current model can already be used by the endoscopist to assess a patient's individualized risk and to discuss surveillance intervals for patient-centered care (https://barrett-recurrence.shinyapps.io/Barrett).

In conclusion, we developed and externally validated a model to predict visible dysplastic recurrence after initial successful endoscopic treatment of BE-related neoplasia in a setting of centralized care. Based on 6 clinical features, our model showed excellent model performance in external validation. This model may help to determine personalized surveillance intervals.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://doi.org/10.1053/j.gastro.2022.03.020.

References

1. Weusten B, Bisschops R, Coron E, et al. Endoscopic management of Barrett's esophagus: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy 2017;49:191–198.
2. Phoa KN, Pouw RE, Bisschops R, et al. Multimodality endoscopic eradication for neoplastic Barrett oesophagus: results of an European multicentre study (EURO-II). Gut 2016;65:555–562.
3. Phoa KN, van Vilsteren FG, Weusten BL, et al. Radiofrequency ablation vs endoscopic surveillance for patients with Barrett esophagus and low-grade dysplasia: a randomized clinical trial. JAMA 2014;311:1209–1217.
4. Shaheen NJ, Sharma P, Overholt BF, et al. Radiofrequency ablation in Barrett's esophagus with dysplasia. N Engl J Med 2009;360:2277–2288.
5. Wani S, Puli SR, Shaheen NJ, et al. Esophageal adenocarcinoma in Barrett's esophagus after endoscopic ablative therapy: a meta-analysis and systematic review. Am J Gastroenterol 2009;104:502–513.
6. Pasricha S, Bulsiewicz WJ, Hathorn KE, et al. Durability and predictors of successful radiofrequency ablation for Barrett's esophagus. Clin Gastroenterol Hepatol 2014;12:1840–1847.e1.
7. Cotton CC, Haidry R, Thrift AP, et al. Development of evidence-based surveillance intervals after radiofrequency ablation of Barrett's esophagus. Gastroenterology 2018;155:316–326.e6.
8. Sami SS, Ravindran A, Kahn A, et al. Timeline and location of recurrence following successful ablation in Barrett's oesophagus: an international multicentre study. Gut 2019;68:1379–1385.
9. Tan MC, Kanthasamy KA, Yeh AG, et al. Factors associated with recurrence of Barrett's esophagus after radiofrequency ablation. Clin Gastroenterol Hepatol 2019;17:65–72.e5.
10. van Munster S, Nieuwenhuis E, Weusten B, et al. Long-term outcomes after endoscopic treatment for Barrett's neoplasia with radiofrequency ablation +/- endoscopic resection: results from the national Dutch database in a 10-year period. Gut 2022;71:265–276.
11. Vliebergh JH, Deprez PH, de Looze D, et al. Efficacy and safety of radiofrequency ablation of Barrett's esophagus in the absence of reimbursement: a multicenter prospective Belgian registry. Endoscopy 2019;51:317–325.
12. Sharma P, Dent J, Armstrong D, et al. The development and validation of an endoscopic grading system for Barrett's esophagus: the Prague C & M criteria. Gastroenterology 2006;131:1392–1399.
13. Lundell LR, Dent J, Bennett JR, et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. Gut 1999;45:172–180.
14. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Series B Stat Methodol 1996;58:267–288.
15. Fu Z, Parikh CR, Zhou B. Penalized variable selection in competing risks regression. Lifetime Data Anal 2017;23:353–376.
16. Phoa KN, Pouw RE, van Vilsteren FG, et al. Remission of Barrett's esophagus with early neoplasia 5 years after radiofrequency ablation with endoscopic resection: a Netherlands cohort study. Gastroenterology 2013;145:96–104.

17. Barrett-Oesofagus Richtlijn. 2017. Nederlandse Vereniging van Maag-Darm-Leverartse. Available at: https://www.mdl.nl/sites/www.mdl.nl/files/richlijnen/Richtlijnen%20Barrett%20oesofagus%20-%20jan%202018%20-%20tbv%20website.pdf. Accessed March 28, 2022.

Received November 14, 2021. Accepted March 10, 2022.

Correspondence
Address correspondence to: Jacques J. G. H. M. Bergman, MD, PhD, Amsterdam University Medical Centers, Vrije Universiteit Medical Center, Boelelaan 1105, 1081 HV Amsterdam, the Netherlands. e-mail: j.bergman@amsterdamumc.nl.

CRediT Authorship Contributions
Sanne van Munster, MD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Lead; Methodology: Equal; Project administration: Equal; Validation: Lead; Visualization: Equal; Writing – original draft: Lead).

Esther Nieuwenhuis, MD (Data curation: Equal; Investigation: Equal; Writing – review & editing: Equal).

Raf Bisschops, Professor (Data curation: Equal; Writing – review & editing: Equal).

Hilde Willekens, MSc (Data curation: Supporting; Writing – review & editing: Equal).

Bas Weusten, Professor (Conceptualization: Equal; Data curation: Equal; Writing – review & editing: Equal).

Lorenza Alvarez Herrero, MD, PhD (Data curation: Equal; Writing – review & editing: Equal).

Auke Bogte, MD, PhD (Data curation: Equal; Writing – review & editing: Equal).

Alaa Akhalaf, MD, PhD (Data curation: Equal; Writing – review & editing: Equal).

Ed Schenk, MD, PhD (Data curation: Equal; Writing – review & editing: Equal).

Erik Schoon, Professor (Data curation: Equal; Writing – review & editing: Equal).

Wouter Curvers, MD, PhD (Data curation: Equal; Writing – review & editing: Equal).

Arjun Koch, MD, PhD (Data curation: Equal; Writing – original draft: Equal).

Pieter-Jan de Jonge, MD, PhD (Data curation: Equal; Writing – review & editing: Equal).

Thjon Tang, MD, PhD (Data curation: Equal; Writing – review & editing: Equal).

Wouter Nagengast, Md, PhD (Data curation: Equal; Methodology: Equal; Writing – original draft: Equal).

Jessie Westerhof, MD, PhD (Data curation: Equal; Writing – review & editing: Equal).

Martin Houben, MD, PhD (Data curation: Equal; Writing – review & editing: Equal).

Conflicts of interest
These authors disclose the following: Bas L. A. M. Weusten received financial support for Institutional Review Board (IRB)-approved research from C2Therapeutics/Pentax Medical. Jacques J. G. H. M. Bergman received financial support for IRB-approved research from C2Therapeutics/Pentax Medical, Medtronic, and Aqua Medical. R. Bisschops received financial support for IRB-approved research from Pentax Medical, Fujifilm, Medtronic, and CD diagnostics and speaker's fees from Pentax Medical, Fujifilm, and Medtronic. The remaining authors disclose no conflicts.