An Accurate Cancer Incidence in Barrett’s Esophagus: A Best Estimate Using Published Data and Modeling

Sonja Kroep, Iris Lansdorp-Vogelaar, Joel H. Rubenstein, Harry J. de Koning, Reinier Meester, John M. Inadomi, and Marjolein van Ballegooijen

Department of Public Health, Erasmus Medical Center University, Rotterdam, The Netherlands; Division of Gastroenterology, University of Michigan, Ann Arbor, Michigan; and Division of Gastroenterology, Department of Medicine, University of Washington, Seattle, Washington

BACKGROUND & AIMS: Published estimates for the rate of progression from Barrett’s esophagus (BE) to esophageal adenocarcinoma (EAC) vary. We used simulation modeling to reconcile published data and more accurately estimate the incidence of EAC among people with BE. METHODS: We calibrated the ERASMUS/UW model (a collaboration between Erasmus Medical Center, Rotterdam, the Netherlands and the University of Washington, Seattle) for EAC to match the 0.18% annual rate of progression from population-based studies. This model was then used to simulate the design of prospective studies, introducing more endoscopic surveillance. We used the model to predict rates of progression for both types of studies and for different periods of follow-up, and compared the predicted rates with published data.

RESULTS: For the first 5 years of follow-up, the model reproduced the 0.19% mean annual rate of progression observed in population-based studies; the same disease model predicted a 0.36% annual rate of progression in studies with a prospective design (0.41% reported in published articles). After 20 years, these rates each increased to 0.63% to 0.65% annually, corresponding with a 9.1% to 9.5% cumulative cancer incidence. Between these periods, the difference between the progression rates of both study designs decreased from 91% to 5%.

CONCLUSIONS: In the first 5 years after diagnosis, the rate of progression from BE to EAC is likely to more closely approximate the lower estimates reported from population-based studies than the higher estimates reported from prospective studies in which EAC is detected by surveillance. Clinicians should use this information to explain to patients their short-term and long-term risks if no action is taken, and then discuss the risks and benefits of surveillance.

Keywords: Esophageal Cancer; Population-Based Modeling; Microsimulation; Early Detection.

In the last 4 decades, esophageal adenocarcinoma (EAC) has become an important cancer in terms of incidence and mortality due to a spectacular rise in incidence in the Western world. The discrepancy between EAC incidence rates by sex is large; the incidence of EAC in men is at least 7-fold that in women, and this difference is seen consistently in all Western countries. Barrett’s esophagus (BE) is the precursor of EAC and its estimated prevalence in the total population ranges between 1.6% and 6.8%, depending on study.1

Given its relevance for clinical management of BE, many have studied the progression rate from BE to EAC. Usually, one calculated the annual progression rate by dividing the number of EAC cases observed in a specified BE cohort by the number of person-years of follow-up in the specified cohort. The resulting estimates for the annual progression rate from BE to EAC vary widely in literature, within a range of 0.07%–3.6%, and the annual progression rates from BE to HGD/EAC are 1.5-2.5 fold higher.2–4 Apart from selection bias, publication bias, study size, and cohort characteristics, which all contribute to the difficulty of comparing these estimates, there are important differences in study design. Earlier, a plausible annual progression was estimated at between 0.41% and 0.5% based on meta-analyses that were restricted to prospective studies.2,4,5 However, several large recently published population-based studies suggest that the progression rate is actually much lower (approximately 0.18%).6–8 The length of follow-up in both designs is regularly between 5 and 7 years.

The uncertainty in the progression rate from BE to EAC is a major driver of the controversy in the management of BE because the (cost-) effectiveness of surveillance and treatment depends on the incidence and progression rate of cancer in BE.7 Identifying the true risk of progression to cancer would greatly improve the decision-making process. In this study, we hypothesize that the difference in

Abbreviations used in this paper: BE, Barrett’s esophagus; EAC, esophageal adenocarcinoma; HGD, high-grade dysplasia; LGD, low-grade dysplasia; ND, no dysplasia.
estimates of progression rates between prospective and population-based studies could be explained by the practice of endoscopic surveillance performed in prospective studies, resulting in earlier detection of preclinical cancers. The aim of this study was to provide a more accurate estimate for the clinical cancer incidence in BE by reconciling published data.

**Materials and Methods**

This section includes the description of the methods used in our study. We will first describe the features of the microsimulation model, and then discuss which characteristics are applied to the model in order to reproduce the population-based study design. Next, we discuss the assumptions for model calibration using the characteristics of the population-based design, which is followed by a description of the model simulations including, the characteristics of the simulated population and the differences between the 2 designs. We conclude with an overview of the evaluation outcomes and sensitivity analysis.

**The ERASMUS/UW Esophageal Adenocarcinoma Model**

The ERASMUS/UW-EAC model was developed as part of a collaboration between the Erasmus Medical Center (Rotterdam, The Netherlands) and the University of Washington (Seattle, WA). The model was previously used in a comparative modeling study to explore the future trends of EAC incidence and in a study to evaluate the influence of the uncertainty in progression rates on hypothetical screening and treatment interventions.

A more detailed description of the model with the quantification of inputs and calibration targets is provided in the Supplementary Material. In brief, the model simulates the life histories of a large population of individuals from birth to death. Part of the population has symptomatic gastroesophageal reflux disease. People can develop nondysplastic (ND) BE (disease onset) depending on age and the presence of gastroesophageal reflux disease symptoms (a 6× increased risk compared with no gastroesophageal reflux disease symptoms). Once BE has developed, low-grade dysplasia (LGD) can develop in BE, which in time can progress to high-grade dysplasia (HGD). In addition, there is also a probability that HGD will regress toward LGD, and LGD toward ND BE. In HGD, malignant cells can arise, transforming to localized EAC that can progress sequentially into regional and distant EAC. In every preclinical cancer stage, there is a probability of the cancer being diagnosed due to the development of symptoms vs remaining asymptomatic and progressing undiagnosed into the next cancer stage (Figure 1). Once a cancer is diagnosed, individuals can die from it, depending on its stage and age. As a competing risk, persons can die of other causes at any moment during their lifetime.

**Reproducing Population-Based Design**

We adjusted the model to reproduce the characteristics of the population-based study design by implementing realistic surveillance and diagnostic inaccuracy, as observed in population-based studies. BE with or without LGD was detected at index endoscopy. The occurrence of subsequent surveillance was based on 2 population-based studies reporting percent of the study population having surveillance and the surveillance intervals as observed in de Jonge et al. In this study, 38% of the patients with BE only and 52% of those with LGD also, received surveillance, with an initial surveillance interval of 2.0 years and 1.4 years for ND and LGD patients, respectively. These surveillance intervals are different than the current recommended endoscopic surveillance intervals (3.0 years and 1.0 years for ND and LGD patients, respectively), but are more consistent with the interval recommendations before 2011. Because there were no adherence data on repeat surveillance endoscopy from population-based studies, we used El-Serag et al, reporting that 54% of the cohort followed recommended surveillance practice, and 46% of the cohort adhered to the first surveillance endoscopy only. In line with this study, we assumed that after each endoscopy the adherence to the next endoscopy will diminish with similar rates. For all surveillance intervals after the initial interval, we used the recommended intervals for all following endoscopies (every 3 years, annually, and every 3 months for ND, LGD, and HGD patients, respectively).

The assumed diagnostic inaccuracy for BE specimens was based on data gathered from literature (Figure 2). In practice, some degree of misdiagnosis of the dysplastic grade is very common because interpretation and grading of biopsy specimens is subjective. In addition, there is sampling error, defined as dysplastic or malignant tissue that is missed when obtaining the biopsy.

**Calibration of the Model**

The model was calibrated to the 2000–2009 incidence data from the Surveillance, Epidemiology, and End Results program; the proportion of LGD and HGD in the BE population; and the mean sojourn time of preclinical (asymptomatic) EAC before becoming clinical (symptomatic) EAC. The model was calibrated to fit the BE to EAC weighted annual
progression rate for males in recently published population-based studies (approximately 0.18% annually).6–8

Population Simulated

Because the mean age at BE diagnosis is 60–65 years in most studies, we have chosen to consider individuals within our simulated cohort with BE at age 65 years and follow them until EAC incidence, death by other causes, or end of follow-up. In addition, because most studies excluded HGD at baseline in their reports of surveillance and cancer incidence, we have included only diagnosed ND and LGD BE. However, this includes HGD and EAC patients underdiagnosed as ND and LGD.

Simulation of Cohort and Design

We subsequently used the calibrated model to simulate the population according to the following 3 scenarios: without surveillance, population-based study design, and prospective cohort design. The only difference between the 3 scenarios is the intensity of surveillance applied after the initial endoscopy. In the scenario without surveillance, none of the included BE patients received any surveillance. For the population-based study design, as for the calibration, 38% of the BE ND patients and 52% of LGD patients received initial surveillance, and in the prospective design, 100% of BE ND and LGD patients received initial surveillance. Repeat surveillance and diagnostic inaccuracy are assumed to be similar in the 2 designs and as stated here (see Reproducing Population-Based Design).

Outcomes

We evaluated 2 primary outcomes: annual progression rate to EAC diagnosis and annual progression rate to HGD detection or EAC diagnosis combined. In line with published studies, we removed the cancers that were detected within the first 6 months of follow-up. We distinguished clinical EAC and surveillance-detected EAC. We define clinical EAC as a cancer being diagnosed by symptoms, and surveillance-detected EAC is a cancer diagnosed during surveillance endoscopy.

In the first outcome (EAC cases only), when HGD is detected by surveillance, this is included as a case and considered as an end point (individual no longer considered in the follow-up), corresponding with the situation where HGD is treated. The annual progression rate was determined by the number of diagnosed cases divided by the number of person-years in follow-up. We included only the true-positive diagnosed EAC in the final outcomes, assuming that overdiagnosed or false-positive EAC cases would have been recognized as such during the further staging process and, therefore, were also not counted in the studies. We calculated the annual progression rate without any influence of surveillance (clinical progression rate). Additionally, for both designs and varying lengths of follow-up, the annual progression rate and the cumulative cancer incidence were calculated. The key outcome of our study is the difference in percentages between the predicted rates of both study designs for varying follow-up lengths. This difference was calculated by dividing the difference in predicted rates between the 2 designs by the predicted rate of the population-based design.

Sensitivity Analysis

Next, in addition to the progression rate throughout the disease to EAC, which has been addressed to by calibration as the central issue to our analysis, in the sensitivity analyses we addressed the following uncertain assumptions that would influence EAC incidence during follow-up: intensity of surveillance during follow-up, overdiagnosis of BE at baseline in the population-based design, and duration of surveillance over time.

Perfect Surveillance Scenario

Because the frequency of surveillance incurs large uncertainty, we explored the impact of the most extreme difference in surveillance between the 2 study types. In this scenario, we assumed no surveillance (0%) for the population-based study design and 100% surveillance (every 3 year, annually, and every 3 months for ND, LGD, and HGD patients, respectively) for all individuals during the total follow-up period for the prospective design.

False-Positive Barrett’s Esophagus Scenario

We have looked at the validity of BE diagnosis within population-based studies. For population-based studies with large registries, one might raise the question of how often a coded BE registration is correctly defined. We have
alternatively assumed that within the population-based design, only 61.9% of the registered BE cases has true BE in accordance with the findings of Corley et al. and that the other 38.1% are false-positive diagnosed individuals, which should not contribute to the denominator when calculating the progression rate.

**Halting Surveillance Scenario**

We looked at the impact of diminishing surveillance over time. Because of comorbidity, there is a higher probability that patients might not be fit for endoscopy at older ages. In practice, it is therefore not uncommon that surveillance is halted for older patients. In this sensitivity analysis, we stopped offering surveillance after a 10-year period, at age 75 years.

**Results**

The calibration to the population-based study, including realistic surveillance, resulted in an annual progression rate of 0.19% for BE (ND+LGD) to EAC, with a 5-year follow-up. Without surveillance, the annual progression rate to clinical EAC (clinical progression rate) would only have been 0.07%.

The same disease model predicted an annual progression rate to EAC for the prospective design of 0.36% after a 5-year follow-up, 91% higher than in the population-based design (Figure 3). These estimates were close to the published estimates of 0.41% annual progression rate, with a 128% difference between designs. Figure 4 shows that the number of EACs diagnosed by surveillance declined over time in both designs, due to the assumed diminishing number of surveillance endoscopies in both designs. While in the first 5 years of follow-up the proportion of surveillance detected cancers was 82% and 98% of the total cancer diagnosis for the population-based and prospective design, respectively, this proportion declined to 26% and 62% for years 15–20 of follow-up. The earlier diagnosis of cancer with surveillance is mainly relevant in the short term. After 20 years, the model predicted that 9.1% of the BE population would have developed EAC, with a difference between the designs of only 5% (Table 1).

A similar pattern was seen when evaluating the progression rate to HGD and EAC combined. Simulated progression rates were approximately 1.5 times higher than for EAC alone, and the difference between the population-based and prospective design decreased from 110% (0.27% vs 0.56%) after a 5-year follow-up to 13% (0.66% vs 0.74%) after 20 years. After 20 years, the model predicted a cumulative incidence of 9.6%—10.6% to HGD and EAC (Figure 5).

**Sensitivity Analysis**

**Perfect Surveillance Scenario.** Assuming no surveillance in the population-based study design and 100% surveillance in the prospective cohort design resulted in considerably larger differences in 5-year average annual progression rates (0.07% and 0.50% in the population-based and prospective design, respectively, see Figure 6A). This shows how sensitive the simulated progression rates are to intensity of surveillance. Without surveillance, 8.9% of the BE patients will develop EAC within 20 years. An additional 1.1% of the patients will be detected with EAC because of surveillance (Figure 6B).

**False-Positive Barrett’s Esophagus Scenario**

If we assumed that only 61.9% of the registered BE cases had true BE in the population-based design, and that the other 38.1% are false-positive diagnosed individuals, the progression rates in the population-based design increased 30%. The observed progression rate of 0.18% actually occurs only in the 61.9% of the population with BE. In other words, the progression rate in BE patients is, in reality, 0.24% after 5 years of follow-up (Figure 6C). Calibrating the model to this rate would result in an estimated 0.10% clinical progression rate without surveillance. Our results only show the effect of the false-positive BE on the population-based studies, which means that the results for the prospective cohort design are identical to the base case scenario. The inclusion of false-positive BE patients in the population-based studies has a large influence on the estimated progression rate.

**Halting Surveillance Scenario**

Halting surveillance after a 10-year period of follow-up only affects long-term outcomes because surveillance assumptions are identical to the base case for the first 10 years.

---

**Figure 3.** Cancer incidence rates for different lengths of follow-up in the realistic scenario. (A) Mean annual progression rates toward EAC and (B) cumulative cancer incidence toward EAC.

---
After 20 years of follow-up, the cancer incidence was 8.7% in the population-based study design and 8.9% in the prospective study design. Compared with our main analysis (with a cancer incidence of 9.1%–9.5% after 20 years), fewer cancers were detected in the long term. Halting surveillance could therefore be a favorable option because the cancers that are not diagnosed by screening because of the halted surveillance are also not diagnosed because of symptoms. Overdiagnosis because of early surveillance-detected EAC is minimized when halting surveillance (Figure 6).

Discussion

In a 60-year-old with BE (ND or LGD) during a 5-year follow-up period, we calculated a clinical progression rate to EAC of 0.07 cases/year. This annual rate increased to 0.37% when evaluated a 10-year-follow up. The 0.19% and 0.36% annual progression rates after 5 years estimated in population-based and prospective designs, respectively, converged to 9.1% and 9.5% cumulative cancer incidence after 20 years. The relative differences between progression rates of the 2 designs decreased from 91% and 21% after 5 and 10 years, to 9% and eventually 5% after 15 and 20 years of follow-up. Results including HGD show that, in the long run, at least 0.5% of the population-based cohort and 1.1% of the prospective cohort will be overdiagnosed with HGD that would not have progressed to cancer even if not detected and treated. The differences between progression rates of the 2 designs diminish at a slower rate when evaluating both HGD and EAC instead of only EAC.

We believe this is the first study in which the difference between observed progression rates is explained using population-based modeling. Our results suggest that the gap between the published progression rates from population-based studies and prospective studies can be largely explained by differences in study design, more specifically by the differences in the surveillance intensity. Performing endoscopic surveillance in individuals with BE leads to earlier detection of cancers and to a higher observed cumulative EAC incidence, especially if follow-up time is not very long (eg, 20 years). In the situation without surveillance, a proportion of these cancers (those not overdiagnosed) would have developed symptoms that led to
clinical cancer diagnosis. However, given closure of the study (eg, after 5 years), some of these cancers would not yet have been observed without surveillance. In our simulations, significant differences between the study designs mainly impacted the progression rates when the follow-up length is <10 years, which is explained by our assumption of diminishing surveillance over time. The true mean annual incidence of clinically apparent EAC in BE over the long term is approximately 0.64% per year, but much lower when observed over 5 years only (approximately 0.2% per year). Much of the later cancers might represent over-diagnosis of pseudodisease, as the risk of competing causes of death increase as the patient ages, making the diagnosis of cancer at those later time points less relevant to all-cause mortality.

Our study is limited by uncertainties. We have assumed the same surveillance intervals and diagnostic inaccuracy for EAC in those undergoing surveillance for both designs. However, the diagnostic accuracy might be higher in prospective studies because these trials are performed by more expert endoscopists using better-established biopsy and pathologist-grading protocols than in a typical population-wide practice. Next, the more exact frequency and intensity of surveillance were uncertain in both types of studies. We supported our base case assumptions using the few studies that quantitatively described surveillance. To evaluate the impact of our surveillance assumptions, we simulated varying scenarios in the sensitivity analysis in which surveillance is altered. These analyses showed that when assuming perfect adherence to the protocol for all study participants in the prospective design, the 5-year progression rate was 0.50% annually instead of 0.36%, with the base case less perfect surveillance compliance assumptions. This reflects the influence of the surveillance protocol on the predicted progression rate, indicating that the published progression rates for prospective studies are

<table>
<thead>
<tr>
<th>Follow-up length</th>
<th>Population based</th>
<th>Prospective</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 y</td>
<td>0.19</td>
<td>0.36</td>
<td>0.17</td>
</tr>
<tr>
<td>10 y</td>
<td>0.43</td>
<td>0.53</td>
<td>0.10</td>
</tr>
<tr>
<td>15 y</td>
<td>0.57</td>
<td>0.62</td>
<td>0.05</td>
</tr>
<tr>
<td>20 y</td>
<td>0.63</td>
<td>0.65</td>
<td>0.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Follow-up length</th>
<th>Population based</th>
<th>Prospective</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGD and EAC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 y</td>
<td>0.27</td>
<td>0.56</td>
<td>0.29</td>
</tr>
<tr>
<td>10 y</td>
<td>0.48</td>
<td>0.64</td>
<td>0.16</td>
</tr>
<tr>
<td>15 y</td>
<td>0.61</td>
<td>0.72</td>
<td>0.11</td>
</tr>
<tr>
<td>20 y</td>
<td>0.66</td>
<td>0.74</td>
<td>0.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Follow-up length</th>
<th>Population based</th>
<th>Prospective</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 y</td>
<td>0.9</td>
<td>1.8</td>
<td>0.91</td>
</tr>
<tr>
<td>10 y</td>
<td>4.0</td>
<td>4.8</td>
<td>0.81</td>
</tr>
<tr>
<td>15 y</td>
<td>7.0</td>
<td>7.6</td>
<td>0.61</td>
</tr>
<tr>
<td>20 y</td>
<td>9.1</td>
<td>9.5</td>
<td>0.41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Follow-up length</th>
<th>Population based</th>
<th>Prospective</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGD and EAC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 y</td>
<td>1.3</td>
<td>2.8</td>
<td>1.51</td>
</tr>
<tr>
<td>10 y</td>
<td>4.4</td>
<td>5.8</td>
<td>1.4</td>
</tr>
<tr>
<td>15 y</td>
<td>7.5</td>
<td>8.7</td>
<td>1.2</td>
</tr>
<tr>
<td>20 y</td>
<td>9.6</td>
<td>10.6</td>
<td>1.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Follow-up length</th>
<th>Population based</th>
<th>Prospective</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean number of person-years per BE patient</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 y</td>
<td>5.0</td>
<td>5.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10 y</td>
<td>9.2</td>
<td>9.1</td>
<td>0.1</td>
</tr>
<tr>
<td>15 y</td>
<td>12.4</td>
<td>12.3</td>
<td>0.1</td>
</tr>
<tr>
<td>20 y</td>
<td>14.6</td>
<td>14.5</td>
<td>0.1</td>
</tr>
</tbody>
</table>

**Figure 5.** HGD and cancer incidence rates for different lengths of follow-up in the realistic scenario. (A) Mean annual progression rates toward HGD and EAC and (B) cumulative cancer incidence toward HGD and EAC.
highly influenced by the intensity of the surveillance protocol on short term. We did not account for HGD treatment when calibrating the model on the BE to EAC progression rate. The reason is that the studies we used for the calibration were performed before the uptake of endoscopic therapy for dysplastic BE. If published progression estimates in prospective cohort studies show persistently lower progression rates than previously assumed (approximately 0.50%), we aim to research the effect of this treatment intervention in future. Finally, there were only a limited number of population-based studies available to provide us with information concerning the progression rate, and each study is subject to its own limitations. These limitations magnify the uncertainty about the size of the difference between the study designs.

The published estimates for annual progression rates for 5 years of follow-up differed 128% between population-based and prospective studies (0.18% and 0.41%, respectively). The model-predicted progression rates compared well with regard to the absolute level and resulted in a difference of 91%. Therefore, more than two-thirds of the difference between published estimates for the short-term progression rate to EAC can be explained by differences in study design. The remaining unexplained difference could easily be caused by uncertainties in surveillance intensity and diagnostic accuracy assumptions. In particular, the impact of uncertainties concerning surveillance is large, as shown in our sensitivity analysis; the most extreme estimated gap between designs by varying the surveillance intensity (0.07% vs 0.50% for no surveillance vs complete surveillance for all) was considerably larger than the observed gap between designs (0.18% and 0.41%).

Previous research examined whether publication bias, the selective reporting of studies featuring positive or extreme results, might have caused the wide range of progression rate published and, in particular, resulted in overestimation of cancer risk in literature. Previous research examined whether publication bias, the selective reporting of studies featuring positive or extreme results, might have caused the wide range of progression rate published and, in particular, resulted in overestimation of cancer risk in literature. The researchers found that reported cancer risk was strongly negatively associated with the size of the study, which could be explained by publication bias. Our study shows that this association might not be

Figure 6. Cancer incidence rates for different lengths of follow-up in the sensitivity analysis. Perfect surveillance scenario: (A) mean annual progression rates toward EAC and (B) cumulative cancer incidence toward EAC. False-positive BE scenario: (C) mean annual progression rates toward EAC and (D) cumulative cancer incidence toward EAC. Halting surveillance scenario: (E) mean annual progression rates toward EAC and (F) cumulative cancer incidence toward EAC.
caused so much by study size, but rather by study design, as large studies are more likely to be population based and characterized by low endoscopic surveillance.

Our findings have several implications for practice. First of all, when the progression rate is inferred from a study, the surveillance practiced in the study should be carefully accounted for, as well as the duration of follow-up. For example, calibrating a model to the rates observed in the cohort studies without including surveillance will overestimate progression rates and lead to cost-effectiveness analyses that support too intensive surveillance recommendations. Modelers must make the distinction between the progression rate to clinical cancer in the situation without surveillance, and progression rate with surveillance, including surveillance detected preclinical cancers. For calibration to the former, it is recommended that models use the lower end of the published estimates, approximating the clinical progression rate to prevent overestimation of cancer incidence without surveillance and thereby provide more valid cost-effectiveness calculations of surveillance and treatment. Next, especially toward patients, but also between professionals, the so-called clinical progression rate is the only definition that is unambiguous and relevant for risk interpretation: it is not influenced by surveillance. Even the direction of this influence can vary, depending on the intensiveness of the surveillance, combined with the aggressiveness of associated precancer treatment (eradication of dysplasia could decrease the observed EAC progression rate) and the frequency of overdiagnosis (detection of clinically insignificant disease would increase observed progression rates). The clinical progression rate is also simple; it only includes the clinically diagnosed cancers that will develop in a situation without interventions before clinical symptoms develop. The clinical progression rate is best approximated by the published population-based studies.

Although population-based studies best approximate the true clinical progression rate, this does not mean that there is no use for the rates from prospective cohort studies. These studies can provide us with valuable information on the impact of surveillance practice on cancer incidence that clinicians can use to inform their patients on the risks (cancer diagnosis at earlier age with the possibility of overdiagnosis) and benefits (less invasive treatment and improvement of prognosis) of surveillance. Prospective cohort studies offer the possibility of estimating duration and significance of screen detectable disease and the sensitivity of surveillance for these precursors. This typically could be done by modeling with simultaneous calibration to studies incorporating infrequent and intensive surveillance. To this end, a good description of the interventions and of surveillance- vs symptom-detected cancers by length of follow-up in empirical studies is mandatory.

In conclusion, the published differences between the lower progression rates reported from population-based studies and the higher progression rates observed in prospective studies can be explained by detection bias from endoscopic surveillance in the prospective studies. For a short time frame, the clinical progression rate from BE to EAC, reflecting the individual risk, is likely to be closer to the estimations in the population-based studies than prospective studies because less surveillance was performed in the former studies. Clinicians informing their BE patients about their cancer risk can best use this clinical progression rate, which is not influenced by surveillance-detected cancers.

Supplementary Material
Note: To access the supplementary material accompanying this article, visit the online version of Gastroenterology at www.gastrojournal.org, and at http://dx.doi.org/10.1053/j.gastro.2015.04.045.

References

Received July 15, 2014. Accepted April 27, 2015.

Reprint requests
Address requests for reprints to: S. Kroep, PhD, Department of Public Health, Erasmus Medical Center University, Room Na 22-23, PO Box 2040, 3000 CA, Rotterdam, The Netherlands. e-mail: s.kroep@erasmusmc.nl; fax: +31-10-7038475.

Acknowledgments
The authors would like to thank the Cancer Intervention and Surveillance Modeling Network. (CISNET).

Conflicts of interest
The authors disclose no conflicts.

Funding
This work supported by grant U01 CA152926 by the National Cancer Institute (NCI).
The Erasmus/UW Model

Model Overview
The Erasmus/UW model (previously the UW/MISCAN-EAC model) is a semi-Markov microsimulation model for EAC. The population is simulated individual by individual, and each person can evolve through discrete disease states. However, instead of modeling yearly transitions with associated transition probabilities, the Erasmus/UW model generates durations in states. With the assumption of exponential distribution of the duration in each state, this way of simulating leads to results similar to a Markov model with yearly transition probabilities. The advantage of the MISCAN (microsimulation screening analysis) approach is that durations in a certain state need not necessarily be a discrete value but can be continuous. MISCAN uses the Monte Carlo method to simulate all events in the program. Possible events are birth and death of a person, Barrett’s incidence, and transitions from one state of disease to another.

The basic structure of the Erasmus/UW model is separated into 3 main parts:

- demography part
- natural history part
- screening part

These parts are not physically separated in the program, but it is useful to consider them separately.

Demography Part
The individual life histories are simulated in the demography part of the model. For each person, a birth date and death date is simulated for other causes than EAC. The distribution of births and deaths can be adjusted to represent the simulated population.

Natural History Part
The natural history part of Erasmus/UW simulates the development of EAC in the population. The current models used in this article do not include an increasing secular trend for the EAC increasing incidence over time. Because the focus of this article is on effectiveness and efficiency, we decided that it could be best compared with the simulation of a cohort excluding secular trend effects before and after 2000—2009.

We assume that EAC develops through precursor BE. For each individual in the simulated population a personal risk index is generated. A majority of the population has symptomatic gastroesophageal reflux disease (GERD), putting them at higher risk for BE developing during their lifetime. The development of BE is generated according to this personal risk index and an age-specific incidence of onset. The sequence from the onset of BE to EAC diagnosis is governed by sojourn times between the different states. BE starts in a state with ND, thereafter dysplasia can develop. Two states of dysplasia are defined: LGD and HGD. There is a possibility that regression from HGD to LGD and from LGD to ND occurs. The probability to regress or progress is dependent on a transition rate matrix and is therefore also influenced by the sojourn time. The probability of regression, progression, and the according sojourn times can be calculated as follow:

\[
\text{Probability of regression in state } i = \frac{R_i}{R_i + R_p}, \quad \text{where } i: \text{current state LGD or HGD, } r: \text{regress, } p: \text{progress, } R: \text{rate}
\]

\[
\text{Probability of progression in state } i = \frac{R_p}{R_p + R_r}, \quad \text{where } i: \text{current state LGD or HGD, } r: \text{regress, } p: \text{progress, } R: \text{rate}
\]

\[
\text{Sojourn time in state } i = \frac{1}{R_p + R_r}, \quad \text{where } i: \text{current state LGD or HGD, } r: \text{regress, } p: \text{progress, } R: \text{rate}
\]

From HGD, malignant cells can arise that can transform from this stage to preclinical localized EAC, which can sequentially progress into regional and distant preclinical EAC. In each of these 3 states, there is a probability of the cancer being diagnosed. The sojourn times between these described states are exponentially distributed and, in some states (BE ND, BE LGD, and BE HGD), age dependent. Because most sojourn times extend beyond the demography-generated age of death from other causes, only a small proportion of the population develop EAC from BE. The survival after clinical diagnosis depends on the cancer stage, and the year of diagnosis (period effect reflecting survival improvement over time). A graphical representation is shown in Supplementary Figure 1.

Screening Part
The development of EAC can be interrupted by screening. Screening can detect BE, the dysplasia states, and preclinical cancers. BE and dysplasia can be removed using treatment. Usually the cancers will be found in an earlier stage than with clinical diagnosis. In this way, screening reduces EAC incidence or EAC death.

Integration of the 3 Model Components
For each individual, the demography part of the model simulates a time of birth and a time of death of other causes than EAC, creating a life history without EAC. Subsequently, the onset of BE is simulated for that individual. For most individuals, no dysplasia is generated. In the case of progressive BE, dysplasia can develop and HGD transforms into a malignant carcinoma, causing symptoms and eventually resulting in death from EAC. When a person dies from EAC before he or she dies from other causes, death age is adjusted accordingly.

After the life history of a person is adjusted for EAC, the history can also be adjusted for the effects of screening. During screening, BE with or without dysplasia is removed by a hypothetical treatment. This results in a combined life history for EAC in the presence of screening. BE is removed at the time of screening and this individual does not develop cancer because the precursor has been removed. Therefore, the person dies from other causes and the effect of screening is the difference in life-years in between the simulation without screening and the simulation with screening.
**Model Quantification**

For this analysis, we developed a model with realistic diagnostic accuracy (realistic model).

**Demography Parameters.** There are two types of demography parameters: birth tables and life tables. The life tables were derived from the life tables published by the National Center for Health Statistics. The parameters for the natural history are directly estimated from data or fit to reference data, based on expert opinion, or calibrated to fit the model. The mean prevalence rate of symptomatic GERD is around 20%2; therefore, we have a fixed input parameter for which 20% of the total population suffers from symptomatic GERD.

The onset of BE was fitted per age group, we call this an optimized parameter. The parameter is relaxed and its final value results from the optimal calibrated model. Asymptomatic BE (no GERD symptoms) is a calibration target of the model, calibrated to be 40% of the total prevalence of BE in the model.

The exponential scale parameter for the time from a preclinical state to clinical detection is restricted within the range of 2–9 years for each individual. When evaluating the whole simulated population, the mean time from onset of preclinical cancer to the diagnosis of clinical EAC is calibrated to be within the range of 4 to 5 years, which is a calibration target of the model.

Using published studies, we estimated the proportion of LGD and HGD in a BE population (Supplementary Table 1), being 2.2% HGD, 9.4% LGD, and 88.4% ND BE, which were used as calibration targets in the models. In addition, the EAC incidence is calibrated to the total Surveillance, Epidemiology, and End Results (SEER) program esophageal cancer incidence rates from 2000–2009.

The assumption that differs between the models is the natural history parameter of the yearly progression rate from BE (ND+LGD) to EAC. The realistic model assumed 0.18% progression rate with partial surveillance and adherence with realistic diagnostic inaccuracy endoscopies. The realistic model incorporated clinical EAC and detected EAC (EAC diagnosed at surveillance endoscopy).

In the calibration process, all input assumptions are taken into account. Supplementary Table 1 contains a summary of the model input, calibration values and its data sources, and the calibration results for the natural history. The BE prevalence is not specified and therefore allowed to optimize for each model, which results in varying estimations. Other optimized parameters are the exponential sojourn times and the regression probabilities.

**Surveillance Parameters**

**Realistic Model Calibration.** We have used a one-time screening examination at age 65 years in the calibration, in which every person is categorized as having BE ND or BE LGD, incorporating false-positive ND and LGD patients. Thirty-eight percent of the ND (52% of the LGD) diagnosed patients will go to surveillance, getting their next endoscopy after 2 years (1.4 years for LGD). At each endoscopy, the patient will be categorized into a dysplastic state or cancer, and can be detected. After the first surveillance endoscopy, 46% of the patients will not adhere to the next endoscopy, and will not be offered any subsequent surveillance in our model. When they develop symptomatic EAC, the patient will be diagnosed with clinical EAC. The remaining patients in the surveillance cohort will be offered surveillance every 3 years (ND), 1 year (LGD), and 3 months (HGD) until cancer detection or death.

The stage-specific survival of patients with surveillance-detected cancer is assumed to be the same as the survival of patients with cancers clinically diagnosed in the same stage.

**Calibration Process**

The Erasmus/UW model is calibrated to fit several calibration targets in order to optimize the unknown natural history parameters. For our model, the incidence of BE per 10-year age group, the sojourn times in BE ND, the transition rates of regression and progression in LGD and HGD, and the sojourn times in the preclinical states have to be optimized.

During the optimization, the Pearson $\chi^2$ goodness of fit function was minimized on the basis of the number of observed and expected rates. The deviation of each of the 4 main calibration targets (ie, SEER-EAC incidence rates per age group, annual progression rate from BE to EAC, proportions of dysplasia, and mean preclinical sojourn time) were summed to calculate the overall goodness of fit of the model given a certain set of parameters. The search for new parameters was performed following the Nelder-Mead simplex method.

The regression and progression transition probabilities are highly correlated with the progression rate from BE toward EAC, which hampers the ability to identify these parameters. Although there are a large number of feasible parameters solutions, there is one optimal parameter solution resulting in the best fit that can be found in the calibration process.

**Validation of Mortality Rates**

The model-projected EAC mortality rates are estimated using incidence-based mortality rates by cancer stage reported in SEER. A direct comparison of model-projected EAC mortality and EAC mortality in SEER is not possible, because SEER does not distinguish mortality from EAC vs mortality from esophageal squamous cell carcinoma. Even if this distinction were possible, the appropriateness of stage migration effects of screening could not be validated with SEER data: SEER does not distinguish method of diagnosis and the estimated percentage of EAC diagnosed by screening or surveillance (7.6%) is too small test a hypothesis of stage migration.
References
Supplementary Figure 1. Graphical representation of the natural history of EAC in the Erasmus/UW model.

Supplementary Table 1. Main Natural History Assumptions and Results of the Erasmus/UW Realistic Model

<table>
<thead>
<tr>
<th>Model parameter/value</th>
<th>Value in realistic model</th>
<th>Source</th>
<th>Parameter characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptomatic GERD prevalence</td>
<td>20% of the total population</td>
<td>Prevalence studies</td>
<td>Fixed input*</td>
</tr>
<tr>
<td>BE from symptomatic GERD population</td>
<td>60% of total BE is from symptomatic GERD population</td>
<td>Published estimates</td>
<td>Fixed input*</td>
</tr>
<tr>
<td>BE prevalence age 60–64 y</td>
<td>1.4%</td>
<td>BE onset per 10-year age group</td>
<td>Optimized parameter</td>
</tr>
<tr>
<td>Percent of LGD in total BE at age 60–65 y</td>
<td>8.2%</td>
<td>Derived from published studies</td>
<td>Calibration target: 9.4%</td>
</tr>
<tr>
<td>Percent of HGD in total BE at age 60–65 y</td>
<td>1.2%</td>
<td>Derived from published studies</td>
<td>Calibration target: 2.2%</td>
</tr>
<tr>
<td>Annual progression rate from diagnosed BE (ND+LGD) to clinical EAC</td>
<td>0.07%</td>
<td>Optimized parameter</td>
<td></td>
</tr>
<tr>
<td>Annual progression rate from diagnosed BE (ND+LGD) to clinical and detected EAC</td>
<td>0.18%</td>
<td>Published estimates</td>
<td>Calibration target: 0.18% in realistic model</td>
</tr>
<tr>
<td>Mean sojourn time from preclinical cancer to clinical cancer, given transition</td>
<td>5.0 y</td>
<td>Published estimates</td>
<td>Calibration target: 4–5 y</td>
</tr>
<tr>
<td>Mean time in BE to next transition</td>
<td>6.7 y</td>
<td>Exponential sojourn times:</td>
<td>Optimized parameter</td>
</tr>
<tr>
<td>Mean time in LGD to next transition</td>
<td>1.0 y</td>
<td>LGD to BE ND or HGD</td>
<td></td>
</tr>
<tr>
<td>Mean time in HGD to next transition</td>
<td>1.1 y</td>
<td>HGD to LGD or preclinical localized</td>
<td></td>
</tr>
<tr>
<td>Regression transition probability</td>
<td></td>
<td>Regression transition probability</td>
<td>Optimized parameter</td>
</tr>
<tr>
<td>$P_{(LGD \to ND\ BE)}$</td>
<td>88%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P_{(HGD \to LGD)}$</td>
<td>15%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a* Fixed input: parameter is defined as a fixed input of the model.

*b* Optimized parameter: parameter is relaxed and is optimized during calibration of the model, the value is a result of the model.

*c* Calibration target: model is calibrated to fit the fixed calibration targets as good as possible. Model is also calibrated on the SEER-9 EAC incidence data from 2000–2009 for all males.