Gastroenterology 2024;■:1–3

Targeted Screening for Barrett's Esophagus and Esophageal Cancer: Post Hoc Analysis From the Randomized BEST3 Trial

Esophageal adenocarcinoma (EAC) has a 5-year survival of <20%, but a previous diagnosis of Barrett's esophagus (BE) is associated with earlier-stage cancer and improved survival.[1,2] The capsule sponge trefoil factor 3 (TFF3) test is a minimally invasive test that can be used to detect BE and early EAC.

The Barrett's Esophagus Screening Trial 3 (BEST3) was a large, pragmatic, multicenter randomized controlled trial comparing an offer of the capsule sponge TFF3 vs usual care among the primary care population aged ≥50 years with gastroesophageal reflux disease (GERD).[3] The results showed that the intervention group had 10× more BE/EAC diagnosed over 12 months, including some patients with early cancer who had curative endoscopic treatment. BE was defined as a coded diagnosis of BE ascertained through anonymous record linkage. This study aimed to use the BEST3 trial data to identify the optimal target population to screen for BE and stage 1 EAC (BE/EAC) to maximize diagnostic yield and minimize harm from overdiagnosis (Supplementary Methods).

Ethical approval for the BEST3 trial was obtained from the East of England–Cambridge East Research Ethics Committee (trial registration ISRCTN68382401). All participants gave informed consent before any individual-level patient data were collected and any clinical procedure was done. In addition, interview participants provided written consent before participating in interviews.

The trial was registered on the ISRCTN registry, with registration number ISRCTN68382401, on January 16, 2017. Registration was prospective before enrolling the first patient.

First, we estimated the number of missed BE/EACs in usual care (Supplementary Figure 1). We calculated the number of BE/EACs that would have been found among TFF3-positive patients if all patients agreed to a confirmatory endoscopy because 10 of 231 did not attend (step 1). Next, we used the data from participants who were TFF3-negative but who were invited at random for a poststudy research endoscopy and who had BE/EAC diagnosed (step 2). We used the expected BE/EAC proportion in the intervention arm to compute the expected number of BE/EACs among usual care (step 3). We deduce that although 13 of 6388 (0.3%) BE/EACs were detected in usual care, we would have expected 680 of 6388 (10.6%) individuals to have BE/EAC, suggesting that 98% of BE/EACs are undiagnosed (step 4). Assuming all 6388 participants in usual care were to accept an offer of capsule sponge TFF3 with the same rate of BE/EAC detected in the intervention arm (131 of 1654, intention-to-treat analysis),[3] we would detect 506 cases of BE/EAC, and hence, up to 74% (506 of 680) of cases of BE/EAC undiagnosed by the current management pathways could be screen-detected.

Next, we evaluated how best to enrich the population for screening. Among those who successfully swallowed the capsule sponge (n = 1654) age (odds ratio [OR], 1.05; 95% confidence interval [CI], 1.03–1.07), male sex (OR, 2.46; 95% CI, 1.67–3.64), and family history (OR, 1.81; 95% CI, 1.04–3.15) were positive predictors of BE/EAC. There was no interaction between age and sex ($P = .929$) or between sex and family history ($P = .804$). Using age and sex, we modeled the probability of a BE/EAC diagnosis among participants who successfully swallowed the capsule sponge (Supplementary Figure 2*A*). To target individuals with a 3% probability of harboring BE/EAC, women aged ≥60 years and men aged ≥50 years could be invited for screening. For ≥5% probability, women aged ≥65 years and men aged ≥55 years could be invited, and for a 10% threshold, only men aged ≥65 years should be invited because women never reach 10% probability. Although family history was predictive, it was only reported in <10% of participants and had more impact on men (Supplementary Figure 2*B* and *C*). Because of the low prevalence, we did not include family history in the modeling.

According to the Office of National Statistics, there are 67 million individuals in the United Kingdom with 9.68 million and 10.19 million male and female individuals, respectively, between ages 50 and 74 years.[4] Assuming 20% of adults have a history of GERD,[5] we estimate that the size of the screening population (men and women aged 50–74 years) is 3.98 million. Using the same prevalence of BE/EAC from BEST3 within the same age group (9.8% male, 4.8% female), we estimate that 288,040 cases could be detected through screening. Selecting a ≥3% probability of harboring BE/EAC requires screening 3.04 million of prevalence BE/EAC compared with 2.18 million for ≥5% and 644,000 for a ≥10% probability strategy (Figure 1). Overall, increasing from 3% to 5% substantially reduces the proportion of individuals to screen by 45.2% (1.8 million of 3.98 million) while still detecting 71.0% (204,480 of 288,040) of cases. Although BE/EAC is a male-predominant disease, increasing the threshold to 10% (men only) reduces the proportion to screen by 84% but misses around 72.3% of prevalent cases. Ultimately, the value of screening will also be reliant on uptake. In the BEST3 trial, 39% of invitees expressed interest, but only 25% attended their appointment once exclusion criteria and availability were considered. A major exclusion was anticoagulant medication, but this is no longer required

§Authors share co-senior authorship.

Abbreviations used in this paper: BE, Barrett's esophagus; BEST3, Barrett's Esophagus Screening Trial 3; CI, confidence interval; EAC, esophageal adenocarcinoma; GERD, gastroesophageal reflux disease; OR, odds ratio; TFF3, trefoil factor 3.

0016-5085
https://doi.org/10.1053/j.gastro.2024.04.030

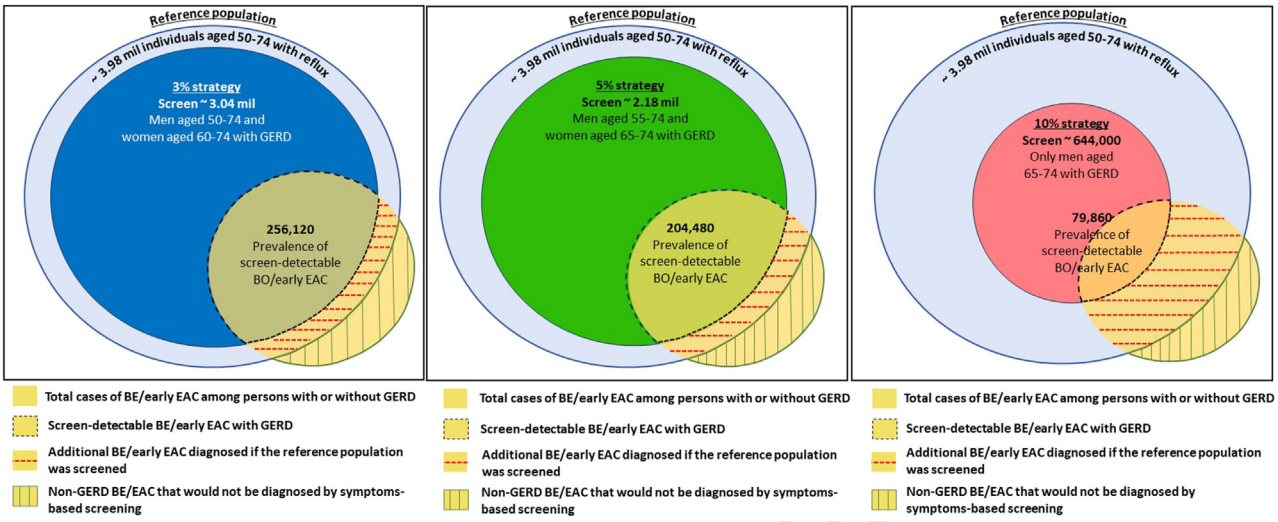


Figure 1. Proposed screening strategy. The reference population (3.98 million, *light blue circle*) was 20% of adults (men and women) aged 50–74 years with GERD, in whom we estimate around 288,040 cases of BE/EAC (using a similar prevalence of BE/EAC in individuals aged 50–74 years as seen in the BEST3 study). The outer black box represents the entire population aged 50–74 years with or without GERD. Screening adults with GERD using a 3% (3.04 million) probability of harboring BE/EAC and the estimated prevalence of BE/EAC among this population (*left*) is compared with increasing the stringency of the eligible population using a 5% (2.18 million) and 10% (644,000) probability strategy for detecting disease (*middle and right*). Horizontal red shading shows the cases of BE/EAC that would not be picked up as you move from one strategy to another, although the population to screen (*inner colored circles*) will decrease. The 5% strategy (*middle box*), would detect 71% (204,480 of 288,040) of BE/EACs but would substantially reduce the population to screen by 45.2% (1.8 million/3.98 million). Vertical green lines represent approximately 40% asymptomatic patients with BE/EAC that would not be picked up through symptom-based screening. Please note that the circle size is not exactly to scale. BO, Barrett's esophagus.

due to new safety data. Other well-established programs have benefitted from a well-advertised endorsement from public health agencies to increase uptake.

The major strength of this study is that the study results were not based on assumptions but extrapolations from the BEST3 randomized data.[3] There are several limitations. First, the BEST3 study was conducted on UK primary care population and may not be generalizable to other populations. The number of cases of BE/EAC diagnosed in usual care was low, reflecting the low rate of referrals by primary care physicians, which may not be true for other countries, where endoscopy referral rates may be higher. Second, the number of dysplastic BE/EAC in the intervention arm was small, which may lead to imprecision in our modeling. Third, the BEST3 trial only included patients on medication for GERD, and studies suggest that approximately 40% of BE/EACs are asymptomatic.[6,7] These cases are represented by the shaded area outside the Venn diagram (Figure 1) and will not be picked up through screening using GERD as a criterion. However, without enriching the population with GERD in addition to age and sex, it is expected that the 3% and 5% thresholds for BE/EAC would not be reached. Most guidelines mandate GERD as a criterion for case finding because screening all individuals with and without GERD would substantially increase the burden of screening.[8] If resources were available, then screening the entire population based on age criteria would diagnose a greater proportion of BE/EAC cases. Finally, we used the Office of National Statistics database for our population estimates, which does not consider the migration of populations within the United Kingdom.

Overall, this study shows that the current referral strategies identify only a fraction (13 of 680; 2%) of the projected BE and stage 1 EACs expected among a reflux population aged ≥50 years. Further, targeted screening of individuals with a 5% probability of having BE/EAC could substantially decrease the population to screen while still detecting >70% of BE/EACs. A large, randomized screening trial (BEST4) in the identified target population will determine whether a nonendoscopic screening approach can reduce morbidity and mortality from EAC.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at https://doi.org/10.1053/j.gastro.2024.04.030.

W. KEITH TAN
Early Cancer Institute
Department of Oncology
University of Cambridge
Cambridge, United Kingdom

ROBERTA MARONI
King's College London Cancer Prevention Trials Unit
Cancer Prevention Group
School of Cancer & Pharmaceutical Sciences
King's College London
Innovation Hub
Guys Cancer Centre

Guy's Hospital
Great Maze Pond, United Kingdom

JUDITH OFFMAN
Centre for Cancer Screening, Prevention, and Early Diagnosis
Wolfson Institute of Population Health
Queen Mary University of London
London, United Kingdom, *and*
School of Cancer & Pharmaceutical Sciences
King's College London
Innovation Hub
Guys Cancer Centre
Guy's Hospital
Great Maze Pond
London, United Kingdom

SHAHRIAR A. ZAMANI
Early Cancer Institute
Department of Oncology
University of Cambridge
Cambridge, United Kingdom

BEST3 TRIAL TEAM

PETER D. SASIENI[§]
King's College London Cancer Prevention Trials Unit
Cancer Prevention Group
School of Cancer & Pharmaceutical Sciences
King's College London
Innovation Hub
Guys Cancer Centre
Guy's Hospital
Great Maze Pond, United Kingdom, *and*
Centre for Cancer Screening, Prevention, and Early Diagnosis
Wolfson Institute of Population Health
Queen Mary University of London
London, United Kingdom

REBECCA C. FITZGERALD[§]
Early Cancer Institute
Department of Oncology
University of Cambridge
Cambridge, United Kingdom

References

1. Cummings LC, et al. Dig Dis Sci 2022;67:1036–1044.
2. Offman J, et al. Cancer Med 2022;11:1160–1171.
3. Fitzgerald RC, et al. Lancet 2020;396:333–344.
4. Office for National Statistics (ONS). Available at: https://www.ons.gov.uk/.
5. El-Serag HB, et al. Gut 2014;63:871–880.
6. Chak A, et al. Cancer 2006;107:2160–2166.
7. Gerson LB, et al. Gastroenterology 2002;123:461–467.
8. Vaughan TL, Fitzgerald RC. Nat Rev Gastroenterol Hepatol 2015;12:243–248.

Received September 13, 2023. Accepted April 23, 2024.

Correspondence
Address correspondence to: W. Keith Tan, XXXX. e-mail: XXXX; Peter D. Sasieni, XXXX. e-mail: XXXX; or Rebecca C. Fitzgerald, XXXX. e-mail: XXXX. Q5

Acknowledgments
The BEST3 Trial Team includes Massimiliano di Pietro, Maria O'Donovan, Beth Muldrew, Irene Debiram-Beecham, Marcel Gehrung, Monika Tripathi, Samuel G. Smith, Benoit Aigret, Fiona M. Walter, and Greg Rubin; from the Early Cancer Institute, Department of Oncology, University of Cambridge, Cambridge, United Kingdom. Q6

CRediT Authorship Contributions
W. Keith Tan, MBChB (Conceptualization: Lead; Data curation: Lead; Formal analysis: Lead; Writing – original draft: Lead; Writing – review & editing: Lead).
Roberta Maroni, MSc (Conceptualization: Supporting; Data curation: Supporting; Formal analysis: Supporting).
Judith Offman, PhD (Funding acquisition: Supporting; Investigation: Supporting; Supervision: Supporting).
Shahriar Zamani, PhD (Data curation: Supporting; Formal analysis: Supporting).
Peter D. Sasieni, PhD (Funding acquisition: Supporting; Investigation: Supporting; Methodology: Supporting; Project administration: Supporting; Supervision: Supporting; Writing – review & editing: Supporting).
Rebecca C. Fitzgerald, MD (Conceptualization: Supporting; Funding acquisition: Supporting; Investigation: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting).

Data Availability Statement
The trial protocol, statistical analysis plan, and statistical report has been previously made available in the original publication of the BEST3 trial in *The Lancet*. Datasets of the deidentified patients who swallowed the Cytosponge will be made available via the University of Cambridge data repository (https://www.data.cam.ac.uk/repository).

Conflict of interest
These authors disclose the following: Peter D. Sasieni reports fees paid to his organization from GRAIL, outside of the submitted work. Rebecca C. Fitzgerald is named on patents related to the Cytosponge-TFF3 test that have been licensed to Covidien (now Medtronic) and is a shareholder in Cyted Ltd. The remaining authors disclose no conflicts. Q7 Q8

Funding
The laboratory of Rebecca C. Fitzgerald is funded by a Core Programme Grant from the Medical Research Council (RG84369). The BEST3 study was primarily funded by Cancer Research UK (C14478/A21047). The National Institute for Health Research (NIHR) covered service support costs, National Health Service commissioners funded excess treatment costs, and Medtronic funded Cytosponge devices and TFF3 antibodies. This research was supported by the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). Cancer Research UK provides funding to the Cancer Prevention Trials Unit (C8162/A25356) and Cancer Prevention Group at King's College London (C8162/A16892). The funders had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit this article for publication. The researchers were independent from their funders. All authors had full access to the study data (including statistical reports and tables) and can take responsibility for the integrity of the data and the accuracy of the data analysis. Q9 Q16

Supplementary Methods

Study Design and Participants

The BEST3 study compared an offer of a capsule sponge TFF3 test vs usual care in the detection of BE/EAC among the primary care population.[e1,e2] Usual care was defined as the standard management pathway in primary care for GERD, whereby primary care physicians offer lifestyle advice, *Helicobacter pylori* testing, and acid-suppressant medications, with gastroscopy reserved only for those with refractory symptoms at the discretion of their primary care provider.[e3,e4] Inclusion criteria were age of ≥50 years and having primary care records of prescriptions of acid-suppressant therapy (proton-pump inhibitor or histamine 2 receptor antagonist) for ≥6 months in the year before inclusion in the study. Patients who were on acid suppression for concomitant use of nonsteroidal anti-inflammatory drugs were excluded.

A total of 13,222 participants enrolled in BEST3: 6388 in usual care and 6834 in the intervention arm. Of these, 1654 successfully swallowed the capsule sponge and produced a sample. Participants with a TFF3-positive result (n = 231) were invited for an endoscopy, of whom 221 (96%) attended. Patient demographics and clinical characteristics were collected during the capsule sponge appointment via a questionnaire. Follow-up was a weighted average of 12 months, and at the end of the study, a random subset of participants in both the usual care and those with a negative TFF3 test were invited for a research endoscopy. The outcome of BE was defined as a coded diagnosis of BE ascertained through anonymous record linkage of the BEST3 participants to the primary care database.

Outcomes

There were several outcomes for this study. The first was to estimate the proportion of undiagnosed BE or stage 1 EAC in the population with GERD using data extrapolated from the BEST3 study. Here, we restricted our analysis to stage 1 EAC given that the biggest impact of cancer screening is to identify early-stage cancer where endoscopic treatment could be applied. The second was to identify the clinical predictors for a diagnosis of BE/EAC and, from this, to determine the optimal age to screen under different screening strategies based on a ≥3%, ≥5%, or ≥10% probability of being diagnosed with the condition. The 3%, 5%, and 10% thresholds were selected through extrapolation from other cancer screening programs, although we acknowledge that the selection of these thresholds may not be fully transferrable between different cancer screening programs. In a recent modeling study, screening for colorectal cancer using a 3% probability threshold could maximize the benefits of screening and minimize the harms.[e5] On the other hand, a 5% probability threshold is consistent with the current screening strategy for lung cancer using annual low-dose computed tomography scans, as recommended by the American Association for Thoracic Surgery.[e6] We therefore modeled the relationship between a diagnosis of BE or EAC and age by sex and analyzed the results according to both a 3% and 5% probability threshold. The 10% threshold was arbitrarily selected for comparison purposes. The third outcome was to arithmetically simulate the number of expected screen-detectable BE/EAC cases using our screening criteria and applied to the UK population using data from the Office for National Statistics (ONS).[e7]

Statistical Analysis

We calculated the estimated proportion of missed BE/EAC in the usual care group using the following outputs from the BEST3 study: (1) the number of BE/EACs diagnosed in patients who successfully swallowed the capsule sponge, (2) the size of the usual care arm, and 3) the number of BE/EACs that were diagnosed in usual care. We then estimated the proportion of missed cases that could be detected by screening if all participants in usual care successfully swallowed the capsule sponge. Multivariable logistic regression was performed to identify clinical predictors of BE/EAC. As defined in the BEST3 study family history questionnaire, a positive family history was defined as the presence of a first or second degree relative with BE or esophageal cancer. Interaction tests were performed where appropriate. All covariates with $P < .1$ on univariate analysis were included in the multivariable regression. All tests were 2-sided, and P values of $<.05$ were considered statistically significant.

To avoid forcing linear variations of the log ORs with age, we modeled the relationship between the probability of a diagnosis of BE/EAC with age and sex by symmetric nearest neighbor smoothing. We plotted the resulting predicted probabilities of a BE/EAC diagnosis against ages 50–80 years only due to the small number of BE/EACs diagnosed beyond 80 years. We identified sex-specific age thresholds for targeted screening strategies based on a ≥3%, ≥5%, or ≥10% probability of harboring BE/EAC.

Finally, we performed arithmetic simulation using UK population-based data from the Office for National Statistics[e7] to estimate the number of people eligible for screening relative to the number and proportion of BE/EACs that could be detected according to these different probability thresholds. We calculated the numbers needed to screen by capsule sponge to diagnose 1 case of BE/EAC as the number of expected BE/EACs divided by the size of the population to screen. All analyses were performed with Stata version 16 (StataCorp LLC) and R version 4.2 (R Foundation for Statistical Computing), and the *mrunning* command was used for the symmetric nearest neighbor smoothing.[e8]

References

e1. Fitzgerald RC, di Pietro M, O'Donovan M, et al. Lancet 2020;396:333–344.

e2. Offman J, Muldrew B, O'Donovan M, et al. BMC Cancer 2018;18:784.

e3. Fitzgerald RC, di Pietro M, Ragunath K, et al. Gut 2014; 63:7–42.

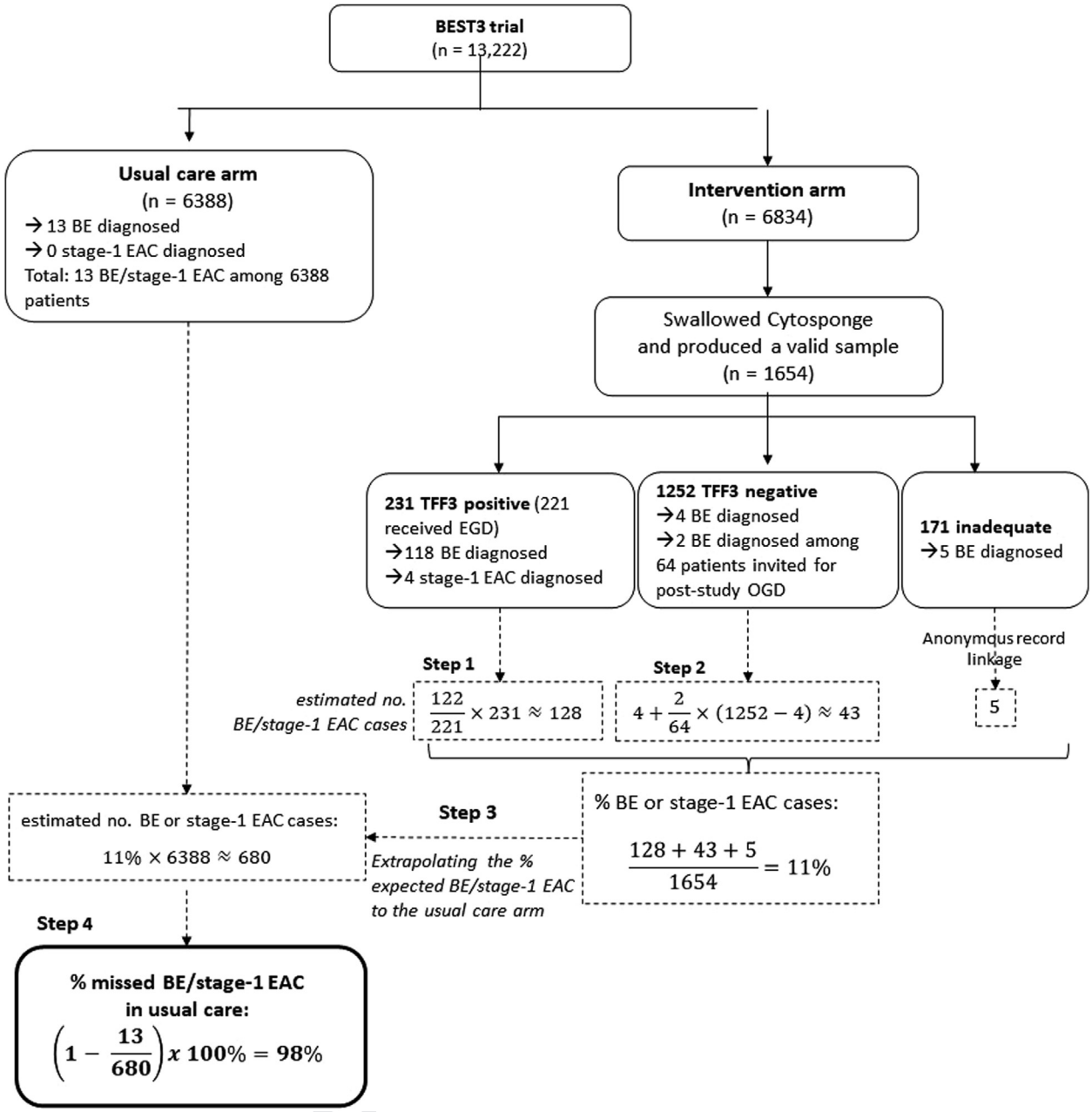
e4. National Institute of Health and Care Excellence. Available at: https://cks.nice.org.uk/topics/dyspepsia-unidentified-cause/.

e5. Buskermolen M, Cenin DR, Helsingen LM, et al. BMJ 2019;367:l5383.

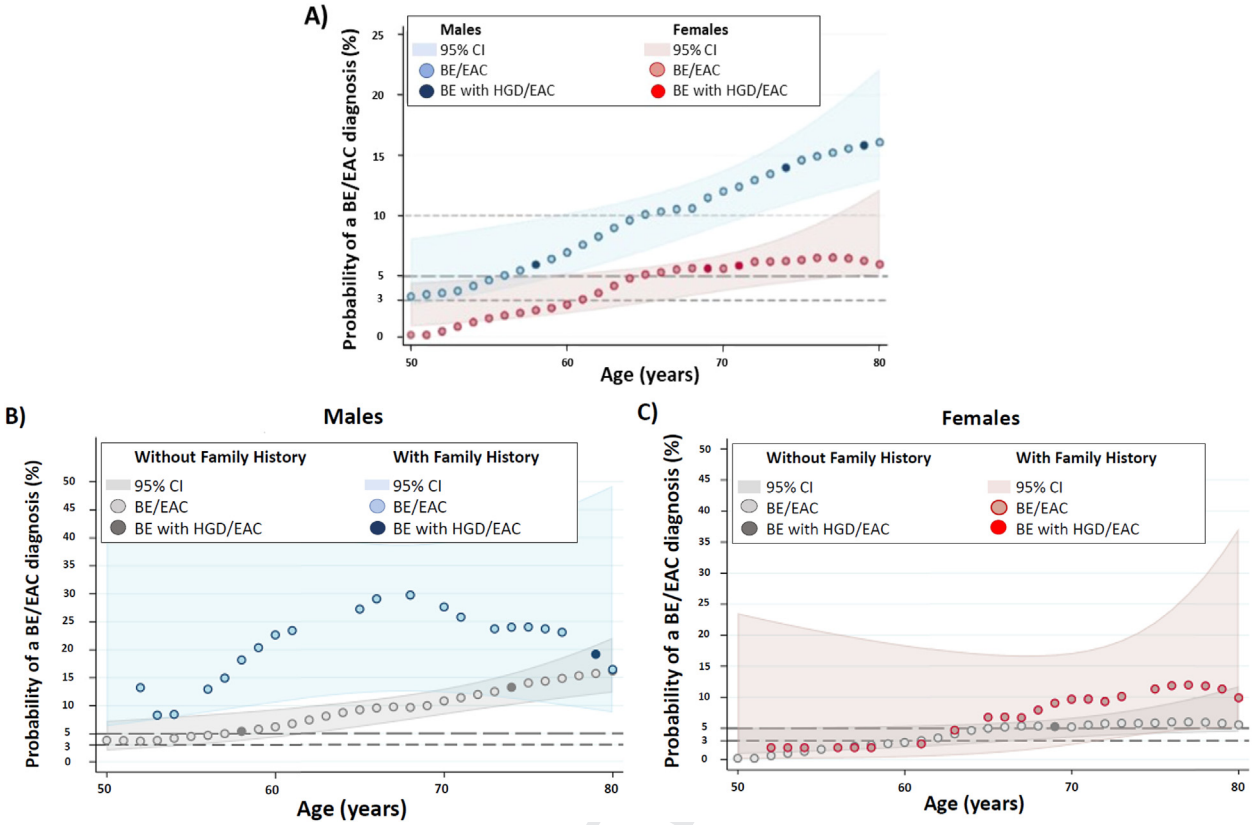
e6. Jaklitsch MT, Jacobson FL, Austin JH, et al. J Thorac Cardiovascular. Surg 2012;144:33–38.

e7. Office for National Statistics. Available at: https://www.ons.gov.uk/.

e8. Royston P, Cox NJ. A multivariable scatterplot smoother. Stata J 2005;5:405–412.



Supplementary Figure 1. Visualization of the arithmetic calculations to estimate the number of missed diagnoses of BE/EAC in the usual care arm. Note: there were 3 diagnoses of advanced-stage esophageal adenocarcinoma in the usual care arm that did not contribute to these calculations. EGD, esophagogastroduodenoscopy.



Supplementary Figure 2. Probability of a diagnosis of BE or EAC, estimated using symmetric nearest neighbor smoothing, with 95% CI. (*A*) When stratified by age and sex, taking a 5% probability, the starting age to screen men was around 55 years, whereas for women, it was around 65 years. Women do not reach 10% probability regardless of age. (*B*) When stratified by family history, this had a larger influence on men, such that the men with positive family history had a probability of BE/EAC of >5% even at age 50 years, but the probability of having BE/EAC reaches almost 30% between ages 65 and 70 years. (*C*) For women, the influence of family history was less, and those with positive family history seemed to only have an increased probability of BE/EAC above age 65 years, such that it peaks at around 12% at ages 75–80 years. Positive family history, however, was only reported in <10% of participants, hence resulting in a wide CI. The age and sex of patients with HGD and EAC are marked in dark blue and superimposed to BE cases. Note: due to low numbers, we did not include patients aged 80 years or older in the plots. Horizontal dashed lines represent the 3%, 5%, and 10% probabilities, respectively, of being diagnosed with BE or EAC. HGD, high-grade dysplasia.