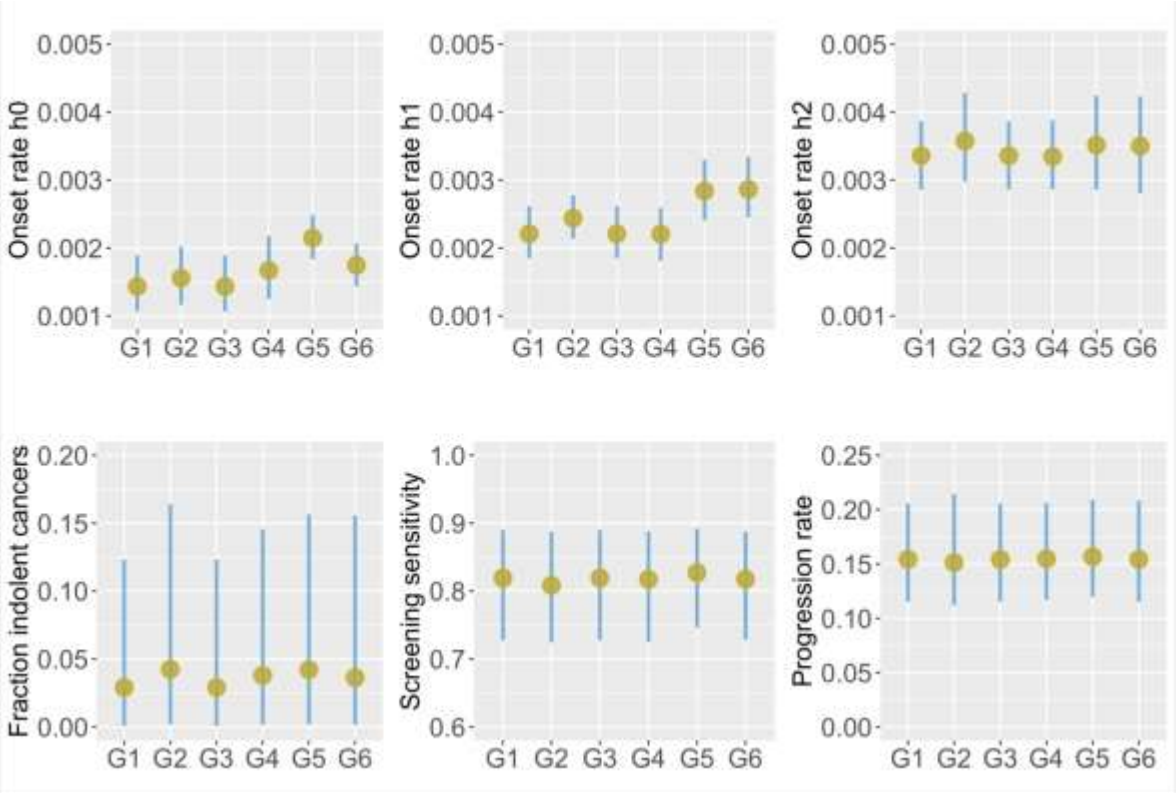
Supplement 2: Supplementary Material*

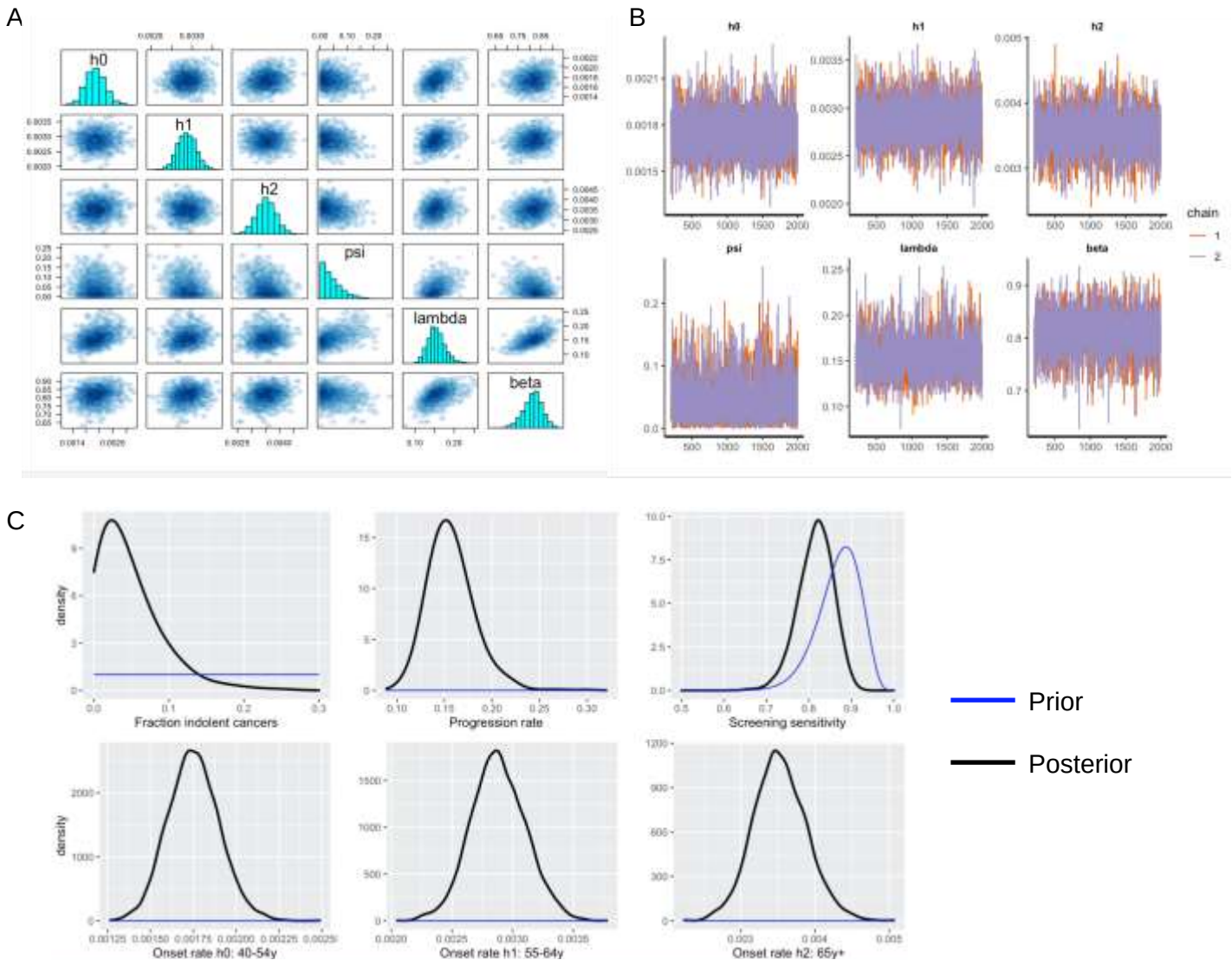
Ryser MD, Lange J, Inoue L, et al. Estimation of breast cancer overdiagnosis in a U.S. breast screening cohort. Ann Intern Med. 1 March 2022. [Epub ahead of print]. doi:10.7326/M21-3577

Content

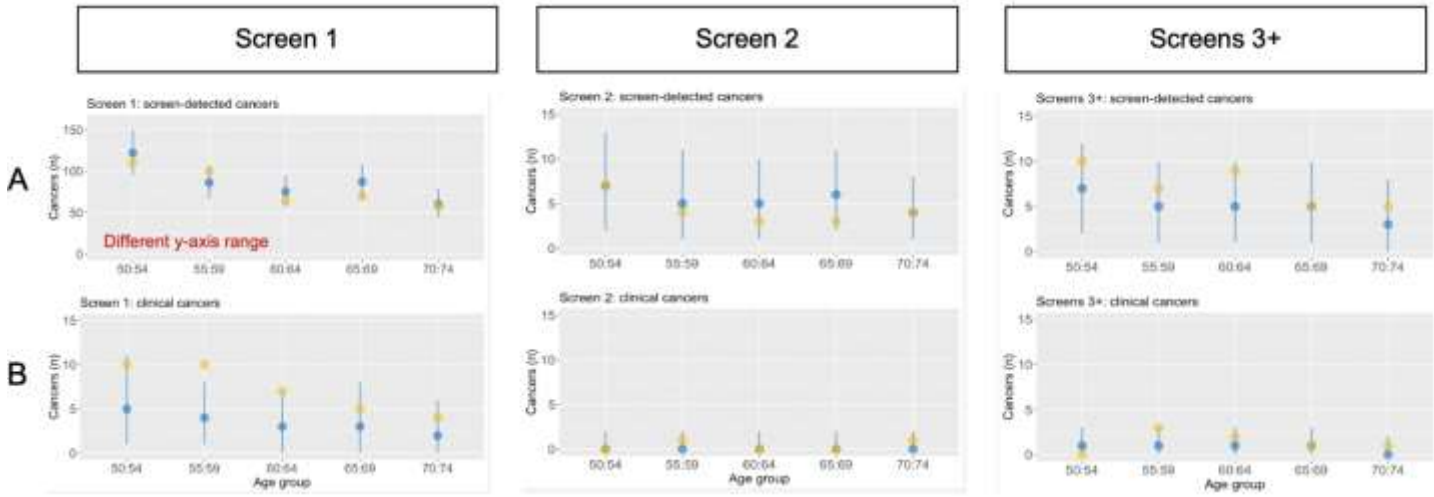
* This supplementary material was provided by the authors to give readers further details on their article. The material was reviewed but not copyedited.



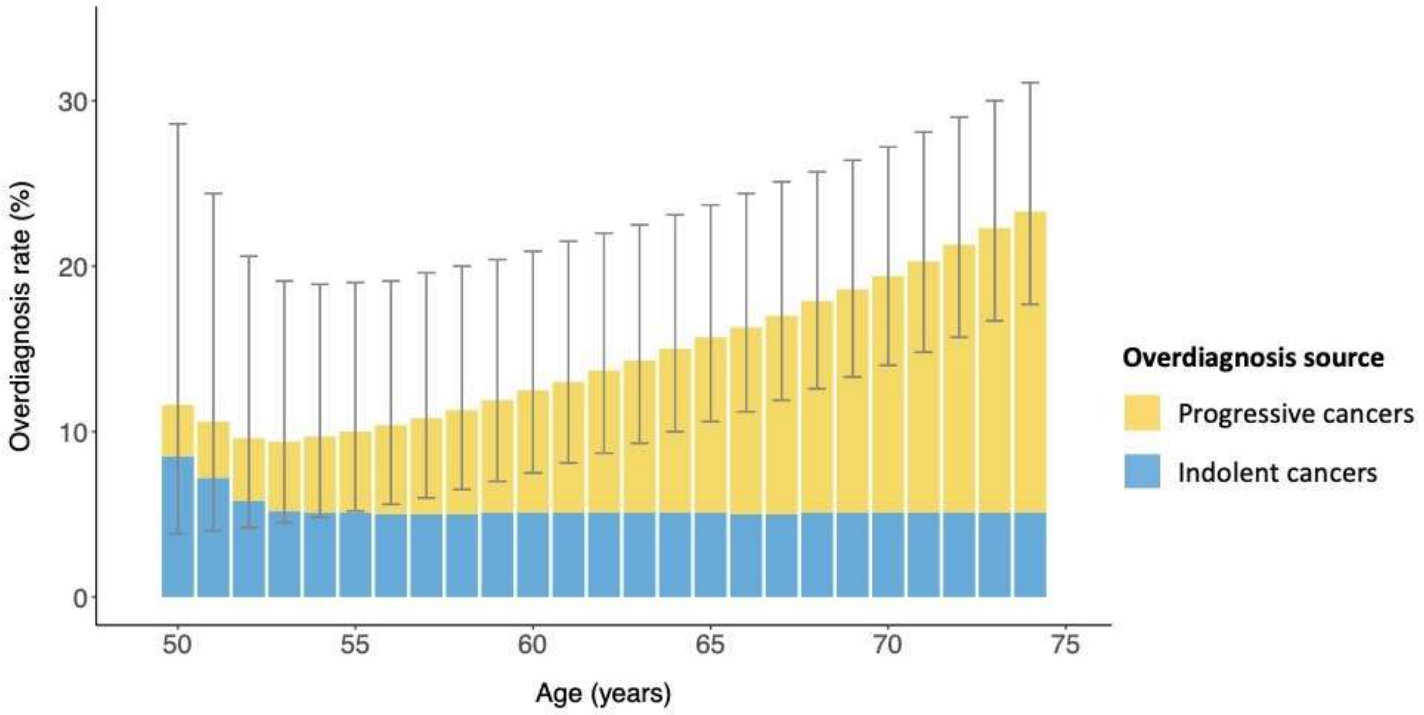
Supplement Figure 1: Sensitivity of parameter estimates to changepoints of onset rate. The impact of different changepoint combinations for the piecewise constant rate of preclinical onset on the posterior parameter distributions. The following changepoint combinations (s_0, s_1, s_2) were compared: G1: (35, 50, 65), G2: (40, 50, 65), G3: (35, 50, 60), G4: (40, 50, 60), G5: (45, 55, 65), G6: (40, 55, 65). Top row: preclinical onset rates h0 on $[s_0, s_1)$, h1 on $[s_1,s_2)$ and h2 on $[s_0,\infty)$. Bottom row: fraction of indolent cancers f_i, screening episode sensitivity, and the rate of clinical progression among progressive preclinical cancers; these parameters were insensitive to the choice of changepoints.



Supplement Figure 2: Parameter identifiability and sampling performance. (A) Marginal univariate (diagonal) and pairwise bivariate (off-diagonal) posterior distributions of the model parameters for the mixture model. Unimodal univariate posteriors and a lack of strong correlations in the pairwise distributions suggest adequate parameter identifiability. **(B)** Trace plots corresponding to the posterior distribution in panel A. Two sampling chains of length 2,000, including a warm-up phase of 200 steps (not shown), were run in parallel. h0, h1, h2: the piecewise constant rates of preclinical onset; psi: fraction indolent cancers: lambda: rate of clinical progression of progressive preclinical cancers; beta: screening period sensitivity. Visual inspection of trace plots, together with standard convergence metrics (not shown) indicate satisfactory chain mixing needed for valid inference. **(C)** Superposition of prior (blue) and posterior (black) parameter distributions. The only informative prior was that for the screening test sensitivity.



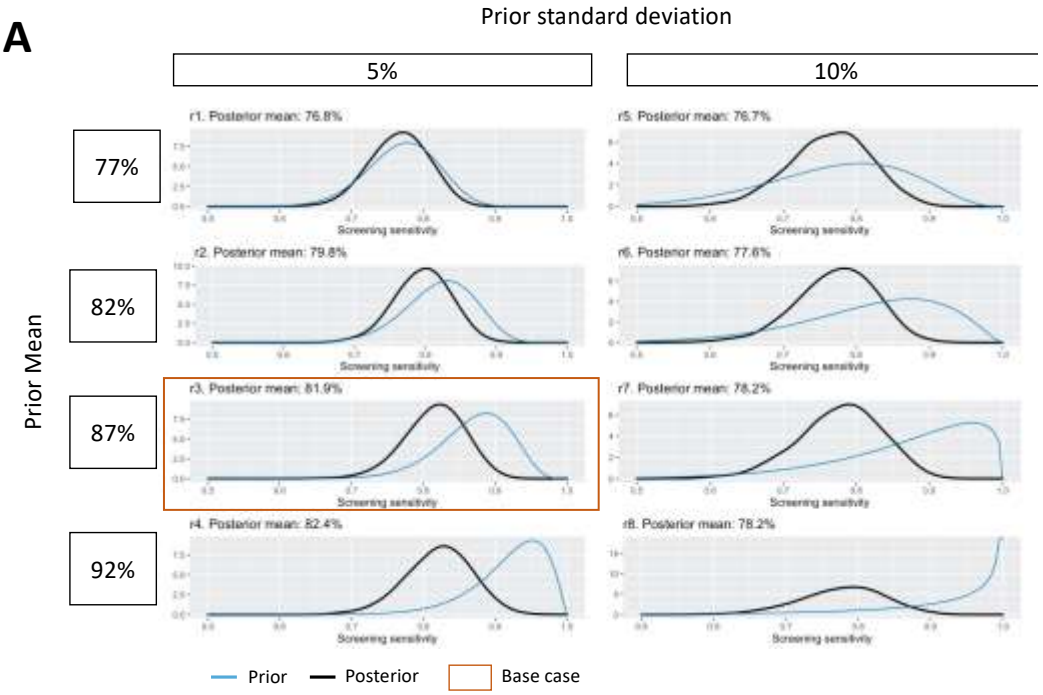
Supplement Figure 3: Posterior predictive model checks. Comparison of posterior predictive median (blue dot) and 95% prediction interval (blue bar) against the observed data (yellow dot) for the number of screen-detected (A) and clinical (B) cancers during the first (*screen 1*), second (*screen 2*) and combined subsequent (*screens 3+*) screening rounds where the latter was defined as the cumulative number of cancers diagnosed in screening rounds 3 to 10. For the purposes of this analysis, we grouped women into 5-year age bins and considered only consecutive annual screens. More precisely, for each woman in the dataset, and starting with the first screen, we included all consecutive screens that occurred approximately one year (9 to 18 months) after the preceding screening mammogram; if the time between two consecutive screens exceeded 18 months, the screening history for that woman was censored 12 months after the last consecutive annual screen. For each age group and each of the 10 quasi-annual screening rounds, we thus obtained the observed numbers of women at risk, screen-detected cancers, and clinical cancers. Using the fitted model, we then computed the expected number of screen-detected and clinical cancer diagnoses in each age group and at each screening round by multiplying the model-predicted probability of a diagnosis by the observed number of women at risk in the dataset. Performing Monte Carlo sampling from the full posterior distribution of model parameters, we thus obtained a posterior predictive distribution of the number of cancer diagnosis which we summarized using the median and 95% prediction interval.



Supplement Figure 4: Overdiagnosis in women undergoing annual screening, ages 50 to 74 years. For women undergoing annual screening starting at age 50, the mean predicted overdiagnosis rate is shown for each screen until age 74 (gray bars represent the 95% prediction intervals). There are two sources of overdiagnosis: the detection of progressive preclinical cancers that would not have progressed to clinical cancer before death due to a breast cancer unrelated cause (yellow), and the detection of indolent preclinical cancers (blue).

Supplement Figure 5: Sensitivity analysis for prior distribution of screening test sensitivity. A Prior and posterior distributions of the screening test sensitivity, varying the prior mean (77%, 82%, 87%, 92%) and the prior standard deviation (5% and 10%). **B** Posterior mean for estimated key model parameters (95% credible interval) predicted overdiagnosis rate (95% prediction interval).

A



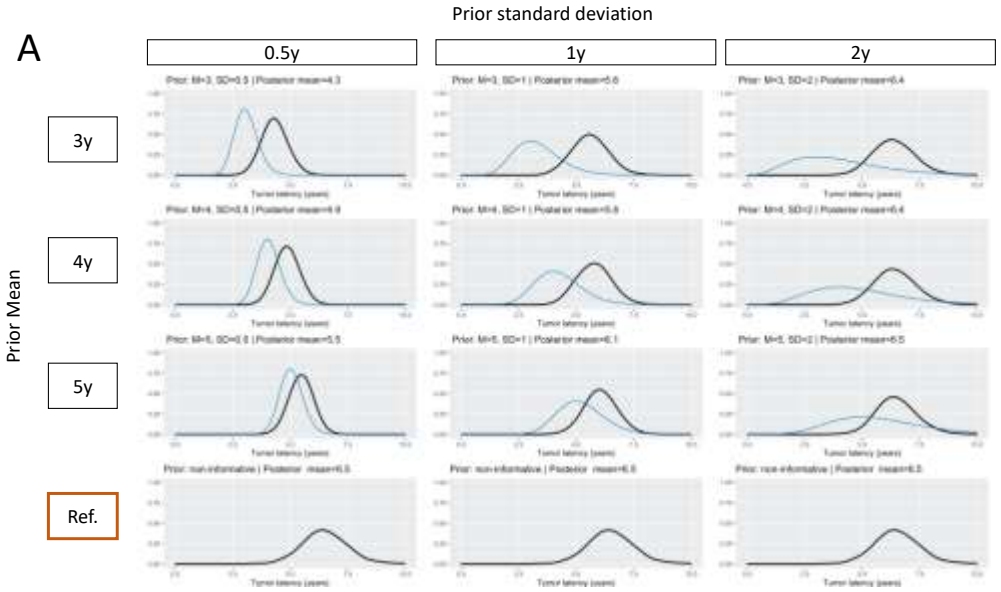
B

Prior standard deviation

Prior mean	5%		10%	
77%	Tumor latency, y	7.1 (5.2 to 9.7)	Tumor latency, y	7.1 (5.1 to 10.2)
	Indolent cancers, %	5.2 (0.2 to 15.6)	Indolent cancers, %	5.4 (0.1 to 17.7)
	Sensitivity, %	76.6 (68.6 to 83.8)	Sensitivity, %	76.3 (64.7 to 86.4)
	Overdiagnosis, %	17.3 (10.7 to 28.2)	Overdiagnosis, %	17.7 (10.1 to 30.7)
82%	Tumor latency, y	6.7 (5.0 to 8.9)	Tumor latency, y	6.9 (5.0 to 9.9)
	Indolent cancers, %	4.8 (0.2 to 15.1)	Indolent cancers, %	5.4 (0.2 to 17.9)
	Sensitivity, %	79.6 (71.6 to 86.7)	Sensitivity, %	77.2 (65.5 to 86.7)
	Overdiagnosis, %	16.0 (9.8 to 27.0)	Overdiagnosis, %	17.5 (10.1 to 31.2)
87%	Tumor latency, y	6.6 (4.9 to 8.6)	Sojourn time	6.9 (5.0 to 9.9)
	Indolent cancers, %	4.5 (0.1 to 14.8)	Indolent cancers	5.2 (0.2 to 16.7)
	Sensitivity, %	81.4 (73.1 to 88.7)	Sensitivity	77.7 (65.8 to 87.7)
	Overdiagnosis, %	15.4 (9.4 to 26.5)	Overdiagnosis	17.2 (10.1 to 30.0)
92%	Tumor latency, y	6.4 (4.8 to 8.8)	Tumor latency, y	6.9 (5.0 to 10.1)
	Indolent cancers, %	4.3 (0.1 to 13.5)	Indolent cancers, %	5.4 (0.2 to 18.0)
	Sensitivity, %	82.1 (72.7 to 90.4)	Sensitivity, %	77.7 (65.4 to 88.1)
	Overdiagnosis, %	15.0 (9.1 to 25.2)	Overdiagnosis, %	17.2 (9.9 to 30.5)

Supplement Figure 6: Sensitivity analysis for prior distribution of the mean sojourn time. A Prior and posterior distributions of the mean tumor latency period, varying the prior mean (3, 4 and 5 years) and the prior standard deviation (0.5, 1, and 2 years). **B** Posterior mean for estimated key model parameters (95% credible interval) predicted overdiagnosis rate (95% prediction interval).

A



B

Prior standard deviation

Prior Mean	0.5y	1y	2y
3y	Tumor latency, y 4.3 (3.4 to 5.4) Indolent cancers, % 13.4 (3.5 to 25.9) Sensitivity, % 84.6 (76.1 to 91.0) Overdiagnosis, % 23.5 (11.3 to 39.3)	Tumor latency, y 5.5 (4.2 to 7.1) Indolent cancers, % 6.8 (0.3 to 18.4) Sensitivity, % 83.4 (75.5 to 90.0) Overdiagnosis, % 16.5 (9.0 to 29.6)	Tumor latency, y 6.3 (4.8 to 8.3) Indolent cancers, % 4.9 (0.1 to 15.5) Sensitivity, % 81.8 (73.7 to 88.8) Overdiagnosis, % 15.5 (9.4 to 27.4)
4y	Tumor latency, y 4.8 (4.0 to 5.8) Indolent cancers, % 9.9 (1.3 to 21.1) Sensitivity, % 84.4 (76.9 to 91.0) Overdiagnosis, % 19.4 (9.0 to 33.2)	Tumor latency, y 5.7 (4.5 to 7.3) Indolent cancers, % 6.0 (0.2 to 17.0) Sensitivity, % 83.3 (75.9 to 90.1) Overdiagnosis, % 16.1 (9.1 to 28.7)	Tumor latency, y 6.3 (4.8 to 8.3) Indolent cancers, % 4.8 (0.2 to 14.4) Sensitivity, % 81.5 (73.2 to 88.8) Overdiagnosis, % 15.4 (9.3 to 25.6)
5y	Tumor latency, y 5.4 (4.6 to 6.4) Indolent cancers, % 7.0 (0.5 to 17.8) Sensitivity, % 83.8 (76.5 to 90.0) Overdiagnosis, % 16.4 (8.8 to 29.1)	Tumor latency, y 6.0 (4.7 to 7.4) Indolent cancers, % 5.4 (0.2 to 15.4) Sensitivity, % 82.7 (75.1 to 89.4) Overdiagnosis, % 15.4 (9.1 to 26.4)	Tumor latency, y 6.4 (4.9 to 8.4) Indolent cancers, % 4.6 (0.1 to 14.5) Sensitivity, % 81.6 (73.5 to 88.7) Overdiagnosis, % 15.2 (9.3 to 26.2)

Reference (flat prior)	
Tumor latency, y	6.6 (4.9 to 8.6)
Indolent cancers, %	4.5 (0.1 to 14.8)
Sensitivity, %	81.4 (73.1 to 88.7)
Overdiagnosis, %	15.4 (9.4 to 26.5)

Supplement Table 1. Selection of changepoint combinations for the preclinical onset rate. Models with different combinations of changepoints (s_0,s_1,s_2) were compared using Bayesian leave-one out cross validation. For each model, the expected log probability density (ELPD) and the ELPD difference to the best fitting model (G6) was computed.

Changepoint combination	ELPD	ELPD difference (SE)
G6: (40, 55, 65)	-4347.0	-
G5: (45, 55, 65)	-4348.9	-1.9 (1.7)
G1: (35, 50, 65)	-4349.8	-2.8 (2.5)
G3: (35, 50, 60)	-4349.8	-2.8 (2.5)
G4: (40, 50, 60)	-4349.834	-2.8 (2.5)
G2: (40, 50, 65)	-4351.6	-4.6 (2.8)

ELPD: expected log probability density; SE: standard error

Supplement Table 2. Purely progressive natural history model. A purely progressive disease model, not allowing for any non-progressive cancers, was fit to the data and the parameter estimates were used to predict the rate of overdiagnosis under biennial screening ages 50 to 74 years (second column). The mixture model results are shown for the sake of comparison (third column).

Model parameter	Posterior mean (95% credible interval)	
	Progressive model	**Mixture model**
Preclinical onset rate for ages [40,55), year^{-1}	0.0018 (0.0015 to 0.0021)	0.0017 (0.0015 to 0.0021)
Preclinical onset rate for ages [55,65), year^{-1}	0.0029 (0.0025 to 0.0034)	0.0029 (0.0024 to 0.0033)
Preclinical onset rate for ages [65,∞), year^{-1}	0.0036 (0.0030 to 0.0043)	0.0035 (0.0029 to 0.0042)
Mean sojourn time, years	7.1 (5.6 to 9.1)	6.6 (4.9 to 8.6)
Fraction indolent cancers, %	NA	4.5 (0.1 to 14.8)
Screening episode sensitivity, %	81.9 (73.8 to 89.2)	81.4 (73.1 to 88.7)
Overdiagnosis (biennial, 50-74), %	10.9 (8.0 to 15.0)	15.4 (9.4 to 26.5)