Supplementary Table 1. Local recurrence rates from publications of the 7 RCTs in Overview trial category (C): "Trials evaluating the need for RT after lumpectomy in low risk patients" (1)

First 4 RCTs are contained in our pooled analysis(1)	Median follow-up (yrs)	Local recurrence rate (n = number of patients)			Breast cancer deaths	
		No-RT #LR/n	RT #LR/n	Relative risk (95%CI)	No-RT	RT
NSABP B-21(2)	7.3	45/334 (13.5%)	9/334 (2.7%)	5.0 (2.7-10.9)	9	5
GBSG V Germany (3)§	10.0	6/80 (7.5%)	5/94 (5.3%)	1.4 (0.5-4.4)	2	1
PMH Toronto (4)	5.6	38*/383 (9.9%)	5*/386 (1.3%)	7.7 (3.1-19.3)	10	10
CALGB 9343(5)	12.6	27/319 (8.5%)	6/317 (1.9%)	4.5 (1.9-10.7)	8#	13#
Subtotal	~8.4	116/1116 (10.4%)	25/1131 (2.2%)	4.7 (3.1-7.2)	29	29
†BASO II (6)	13.9	8/106 (7.5%)	0/98 (0.0%)	∞	NA	NA
Austria BCSG A8 (7)	4.5	19/417 (4.6%)	2/414 (0.5%)	9.4 (2.2-40.2)	2	2
‡PRIME II(8)§	5.0	26/668 (3.9%)	5/658 (0.8%)	5.1 (2.0-13.3)	8	4
Subtotal	~5.6	53/1191 (4.5%)	7/1170 (0.6%)	7.4 (3.4-16.3)	10	6
Subtotal without PRIME II	~6.4	27/523 (5.2%)	2/512 (0.4%)	13.2 (3.2-55.3)	2	2

* Estimated from published Kaplan-Meier curves

[#] These numbers of deaths were reported in the Oxford Overview to be 8 and 14, with order reversed. The Overview used a biased definition of breast cancer death: Death following disease recurrence of any type, including any death following local recurrence.

† Data given are for the two relevant arms of the imbedded factorial design. The publication does not give results separately for those patients who chose to receive tamoxifen and to be randomized to RT.

‡ The indicated data for PRIME II were published in 2015. The Overview was published in 2011 and had limited data from only the PRIME I subset of PRIME II.

§ Includes local or regional recurrence (PRIME II: Majority were local recurrence events. RT arm local recurrence only=18, local-regional=6; no RT arm local recurrence only=4, local-regional recurrence=1)

Supplementary Methods

Model Methods

Two independent models simulated a proposed randomized trial of the omission of radiotherapy in invasive breast cancer patients with biologically low-risk disease. The approach used to develop distributions for model inputs and to run the simulations of the proposed trial are summarized in the manuscript and in additional detail herein for each model. Further information is available from the authors on request.

Since no one source included all the information needed for the simulations, several good partial sources were used by one or both models, including the SEER-Oncotype DX® linked database (9), the Oxford Overview ("Overview") of 17 randomized radiotherapy trials (10), and a pooled individual-level, de-identified database with information from seven trials ("pooled" dataset). The latter were provided under federal data-sharing policies and included four trials from the Overview (2-5) and three that did not randomize radiotherapy (11-13). Data from the SEER registry for patients matching trial eligibility were used to estimate competing non-breast cancer mortality (14). Finally, de-identified proprietary data linking Oncotype DX® scores with data from two NSABP trials (12, 13) were used to model endpoints (Personal communication, S. Shak, Genomic Health Inc., 2016) (15). Specific data were selected for use from these datasets based on model structure.

Model GE

Model GE used an empiric Bayesian analytic approach for this project. New program code was developed for this project since the original CISNET Model GE does not include recurrence endpoints. The simulation using this project-specific model relied primarily on the pooled dataset to derive most inputs. This dataset included individual-level, de-identified

information from seven trials that included radiotherapy data (2-5, 11-13). Four of the trials randomized radiotherapy and were included in the Oxford Overview (2-5) and the other three included data of radiotherapy use as observed within a trial of another modality (11-13). All data were provided under federal data-sharing policies. Other trials included in the Oxford Overview were either not obtained due to administrative issues unrelated to the proposed trial's hypotheses or having radiotherapy protocols that were felt to not reflect current US practice.

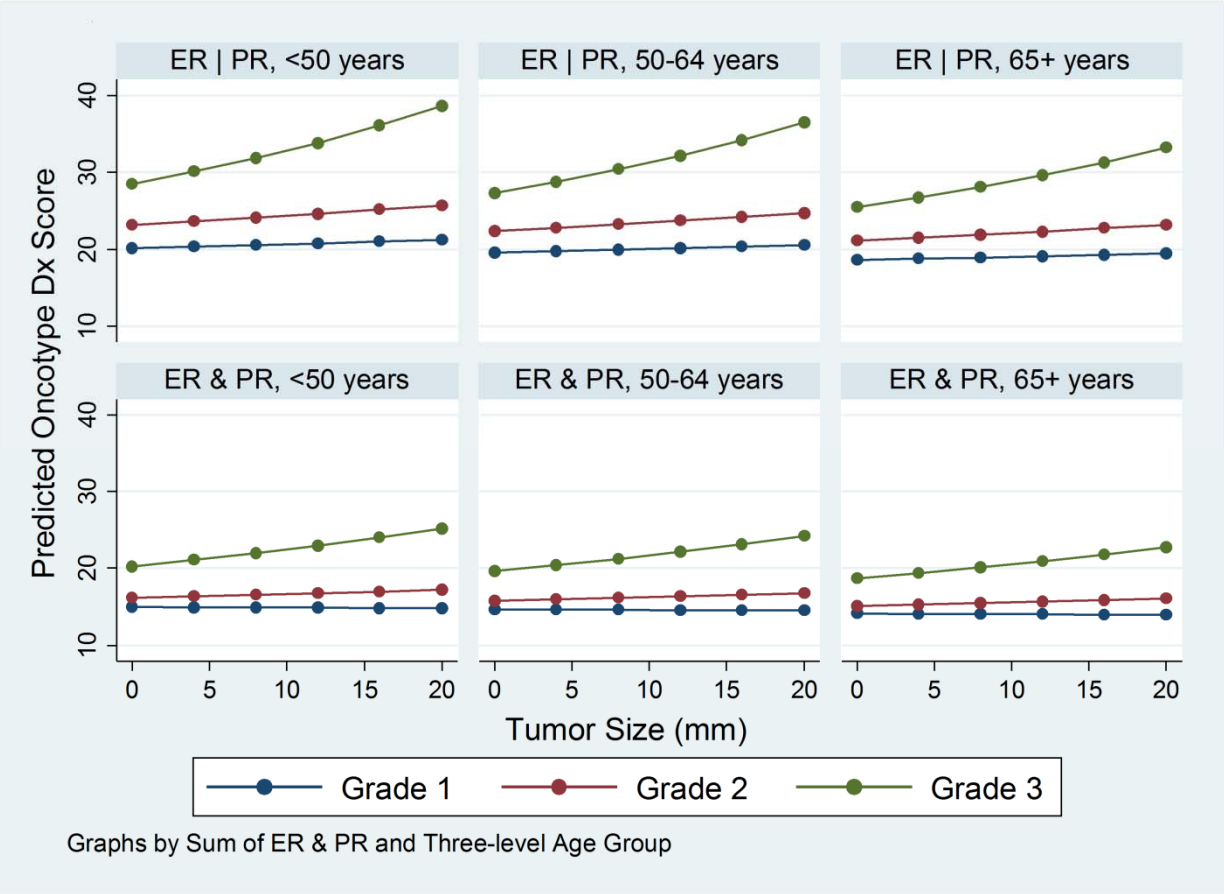
Target Population for the Proposed Trial

The pooled trial data were used to develop joint distributions of most characteristics of the trial eligible target population (e.g., age, tumor grade, tumor size, etc.) For each of the 1,000 trial replicates, a pool of 4,500 potentially eligible patients was created, each having characteristics randomly sampled from this joint distribution.

Oncotype DX®

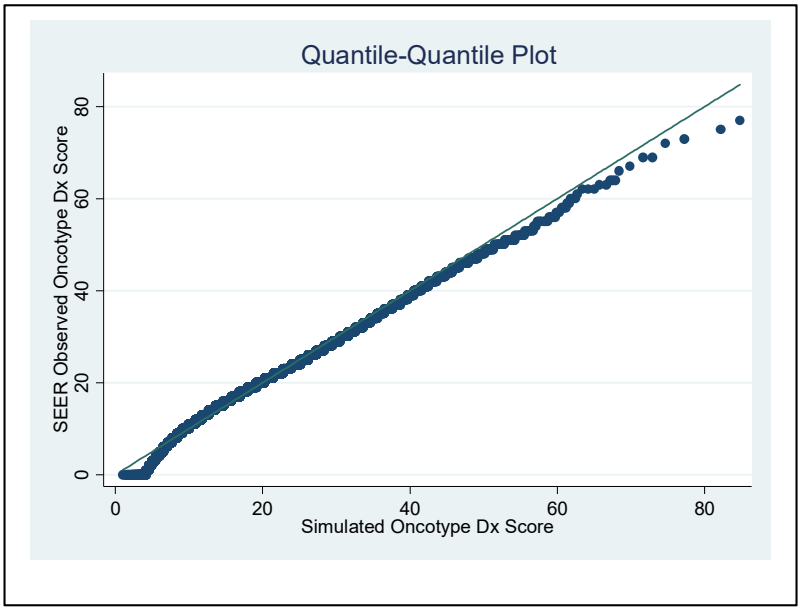
Since six of the seven trials included in the pooled data did not include Oncotype DX®data, SEER-GHI data (9, 16) were used to impute missing Oncotype DX® scores for all patients in the pooled dataset, assuming data were essentially missing at random. A generalized linear model with an inverse link and a gamma distribution was fitted to predict Oncotype DX® score conditional on attributes of the SEER-GHI population that matched the proposed trial eligibility criteria: breast conserving surgery with hormonal therapy for a new primary invasive hormone receptor positive (ER and/or PR positive), HER2 negative (or unknown), lymph node negative breast cancer with ≤ 2cm pathologic size, having age at diagnosis between 40 and 74. The predicted mean Oncotype DX® scores are shown in the figure by age group, grade (grade 1=low; grade 2=intermediate; and grade 3=high), hormone receptors (ER+ or PR+ [top row] or ER+ and PR+ [row 2]), and tumor size. Note that among these predicted mean scores for groups

of patients, there is no combination of age groups, grade, etc. where the *mean* Oncotype DX® score is predicted to be < 10, even though many individuals with mean predicted scores near 10 will have an individual score < 10.



This generalized linear model was used to impute Oncotype DX® scores for the 4,500,000 virtual patients (4,500 each for 1,000 trial replicates). For each virtual patient, the simulated score was sampled from a gamma distribution with shape and scale parameters chosen to match the patient's predicted mean score from the model and a coefficient of variation that was fitted to the dataset from which the model was derived. Virtual patients with simulated Oncotype DX® scores >18 were then removed, and up to 2,194 virtual patients were selected at random from the remaining sample with scores ≤18. The virtual patients in each replicate were assigned a

randomization date in batches of 40 per month and randomized to radiotherapy or no-radiotherapy with equal probability.



Using these methods to impute Oncotype DX® resulted in a pool of virtual patients with a distribution of Oncotype DX® scores in the 1,000 trials that matched the distribution observed in the SEER-GHI population (see quantile-quantile plot). Thus, this method is useful in modeling and simulation and yields valid distributions of results within and across the trial replicates. However, the results are not intended to predict outcomes for individual real patients.

Recurrence- free Interval (and Breast Cancer Deaths)

Competing risk time-to-event semi-parametric models were fitted to the pooled dataset to model recurrence-free interval (RFI) and breast deaths conditional on simulated attributes. The first set of models used age (< 50, 50-69, ≥ 70 years), tumor size (cm), grade (good, intermediate or poor), and radiotherapy (observed vs. randomized).

A second set of semi-parametric models was developed for time to recurrence and breast cancer death conditional on Oncotype DX® scores by fitting to the dataset provided by Genomic Health (personal communication, Steve Shak). Results from these two set of models were synthesized by adjoining the Oncotype DX® term to the coefficient vector from the first regression and using the combined vector to calculate the logarithm of the subhazard ratio for each simulated person. Note that radiation effects from randomized and non-randomized trials in

the pooled dataset were separately estimated by using a radiation by randomization interaction term, and only the randomized effects were used in the simulations.

In this manner, a proportional-hazards cumulative incidence function for recurrences and breast cancer deaths was generated that corresponded to reference levels of the attributes and Oncotype DX® scores, along with estimated subhazard ratios (SHRs) for each attribute and covariances for those SHRs. Some of these distributions were "shrunk" by combining them with an abstract lognormal hyper-prior distribution that gives most of its support to hazard ratios between 0.25 and 4.0.

Each of the 1,000 trial replicates was then assigned its own set of SHRs by random draws from the multivariate normal distribution of SHRs with those means and covariances. The replicate-specific SHRs were applied to the reference level cumulative incidence functions according to each virtual patient's characteristics and Oncotype DX® score to create an incidence function for each event for each virtual patient. The resulting incidence functions were randomly sampled to simulate times to these events. The simulated times to the first events were an excellent match to the observed times in the pooled dataset (not shown).

Virtual patients were sorted in order of time of recurrences (if an event occurred) and time of accrual within their replicate. Patients accrued after the 88th event were removed from the simulation; all remaining patients not having had a recurrence or death up to that time were censored at the date of the trial end.

Note that none of the available data sets provided adequate information for estimating the joint distribution of recurrence times and breast cancer mortality. Therefore, separate models were estimated in the manner described above for recurrence and mortality, based on the marginal distributions of death due to breast cancer.

Non-breast Cancer Mortality

The prior distribution for non-breast cancer-specific survival was based on SEER data for trial eligible patients (14). The summation of time to breast cancer-specific mortality and other cause mortality provided simulated values for all-cause mortality.

Trial Simulation

The simulation is a Monte Carlo microsimulation implemented in Stata version 14.2 (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.). All random sampling utilized the Mersenne Twister pseudo-random number generating algorithm in Stata version 14. Each of the 1,000 trial replicates is analyzed according to the analysis specifications set out in the proposed clinical trial protocol, and a summary of the distribution of the results of the 1,000 trials is generated.

Model M

Model M uses a fully Bayesian analytic approach and embeds the trial simulation within its CISNET model, where the code is adapted to add recurrence events. The prior distribution for model M parameters was derived primarily based on the Oxford Overview (referred to as the Overview) (10). The Overview reports the effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death, based on a meta-analysis of individual patient data for 10,801 women in 17 randomized trials. Corresponding to the definition of recurrence-free interval (RFI) in the Overview, in Model M if breast cancer death occurs without a prior loco-regional recurrence (LRR) or distant recurrence (DR) being recorded, a DR is assumed to immediately precede the breast cancer death. Since the Overview includes only invasive events and does not report LRR and DR separately, Model M uses invasive RFI (IRFI) as the primary outcome, and cannot determine the type of first recurrence.

Given the focus of the proposed trial on low-risk patients, Model M primarily used Overview data for low-risk groups to develop prior distributions. Low-risk was defined in the Overview for node-negative patients as having less than a 10% predicted absolute reduction in 10-year rates of any invasive recurrence with radiotherapy (10). Some data for intermediate- and large-risk groups was also incorporated, due to the modeling of effects of Oncotype DX® based on its correlation with grade (see below). Intermediate- and large-risk were defined in the Overview for node-negative patients as having a less than 10-19% and ≥ 20% predicted absolute reduction in 10-year rates of any invasive recurrence with radiotherapy, respectively.

Target Population for the Proposed Trial

The target population for the trial was based on the subset of patients reported in the SEER and SEER-GHI datasets (9, 14, 16) that met the eligibility criteria of the proposed trial.

Oncotype DX®

The Overview did not contain information on Oncotype DX®, but it did include data by grade and age. Therefore, to select data for a population matching the eligibility of the proposed trial, Model M used the SEER-GHI data (9, 16) to estimate the distribution of grade among patients with Oncotype DX® ≤18, categorized into two groups (0-10 and 11-18), and then examined whether that distribution varied by age in each group. These data were also employed to enable use of the Overview grade- and age group-specific distributions for RFI (and survival outcomes) in simulating these outcomes conditional on an Oncotype DX® score group.

The distribution of grade by Oncotype DX® score varied qualitatively somewhat with age. Therefore, when the Oncotype DX® group for a virtual patient was 0-10, the patient's time-to-event outcomes were simulated according to the corresponding prior distribution in the Overview conditional on both tumor grade and age group, weighted by the proportion of tumor grade for the simulated age group observed in SEER-GHI data. There was less of an age effect for Oncotype DX® 11-18, so events for this group were simulated based on weighted proportions of tumor grade as an average across the age groups.

Recurrence- free Interval

The Overview included data for the estimated 5- and 10-year absolute risks of first recurrence stratified by sets of single factors (e.g., age group, tumor grade, tumor size, ER status), in pathologically node-negative women allocated to breast-conserving surgery (BCS) and BCS with radiotherapy RT (Webtables 5b and 3b) (10). Among patients with ER+ tumors who underwent BCS and tamoxifen treatment, the data suggested potentially different hazard rates for first recurrence in the first and second five years for both RT and no RT arms, and this was the case when RFI was examined for subgroups based on age, tumor grade, or tumor size. Based on

these observations, Model M modeled the RFI using a piecewise exponential distribution with two constant hazard rates for the first five years and five years later, denoted as λ_1 and λ_2, respectively.

Next, the prior mean and standard deviation (SD) for the 5- and 10-year recurrence rates (denoted as p_1 and p_2, respectively) were derived assuming a beta distribution based on the observed recurrence rates in each cell in the no RT and RT arms (Webtables 5b and 3b), and the sample size given in each cell for both arms combined (Webtable 3c), for node-negative, ER+ patients receiving BCS and tamoxifen. Because the sample size in each cell was not given by RT in the Overview, for simplicity an equal split of the sample between arms was assumed when deriving the prior distributions.

Based on the derived mean and SD for p_1 and p_2, and the implied transformation between (λ_1, λ_2) and (p_1, p_2) by the assumed piecewise exponential distribution, the prior mean and SD for log (λ_1) and log (λ_2) was derived using the multivariate delta method. A zero correlation between p_1 and p_2 - p_1 was assumed.

Given the calculated prior mean and SD for log (λ_1) and log (λ_2) in each cell by RT arm, the RFI was simulated assuming a piecewise exponential distribution with a normal prior distribution for each of log (λ_1) and log (λ_2), with the above calculated mean and SD.

Breast Cancer Specific Survival

The estimated absolute 10-year hazards of breast cancer death by RT arm (Figure 5 and Webtable 6 of the Overview) (10) in the Overview-defined low-risk category were used to estimate the prior distribution of the hazard rate for breast cancer death. However, we used an updated version of the results of CALGB 9343 (5). We used the results shown in Appendix Table 2.

An exponential distribution was assumed for breast cancer-specific survival since the Overview did not provide estimates for the absolute risks for breast cancer death during different time periods. As noted above, low-risk was defined in the Overview as having less than a 10% predicted absolute reduction in 10-year rates of any invasive recurrence among ER+, node-negative patients having BCS and tamoxifen (Webtable 3a of the Overview) (10).

However, the estimated absolute 15-year risks of breast cancer death in the Overview-defined low-risk category (Webtable 6 of the Overview) looked considerably higher than that reported for patients with Oncotype DX® scores ≤18 (15), and they were also higher than the risks in the Overview intermediate-risk category, which was counterintuitive. Therefore, the estimated 15-year risks of breast cancer death for the combined no RT and RT arms in the Overview-low risk category were replaced with the corresponding estimates (assuming an exponential distribution) from the pooled dataset for the sub-set of patients eligible for the proposed trial who had originally been randomized to radiotherapy (2-5). Similar to the general approach to the prior derivation for RFI, a normal prior distribution was assumed for the logarithm of the hazard rates, with the prior standard deviation estimated using the multivariate delta method.

Non-breast Cancer Mortality

The prior distribution for non-breast cancer-specific survival was based on SEER data for trial eligible patients (14).

Trial Simulation

The Model M simulations were implemented using Microsoft Visual Studio C# 2010. All random sampling utilized the built-in pseudo-random number generator in C#. Each of the 1,000

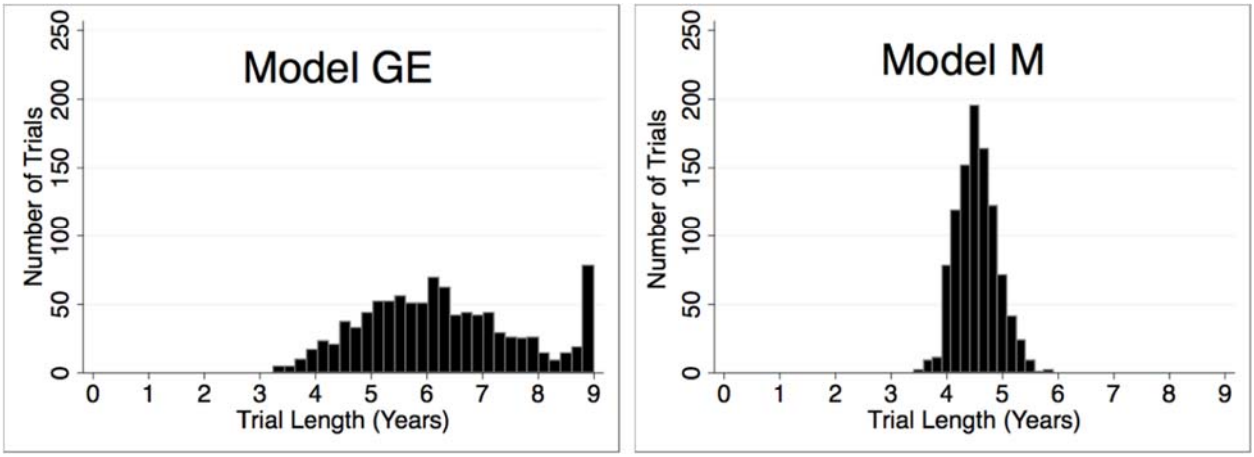
trial replicates was then analyzed according to the specifications described in the proposed clinical trial protocol, and the results of the 1,000 trials were tabulated.

Supplementary Table 2. Grade by Oncotype DX® by Age in SEER-GHI data*

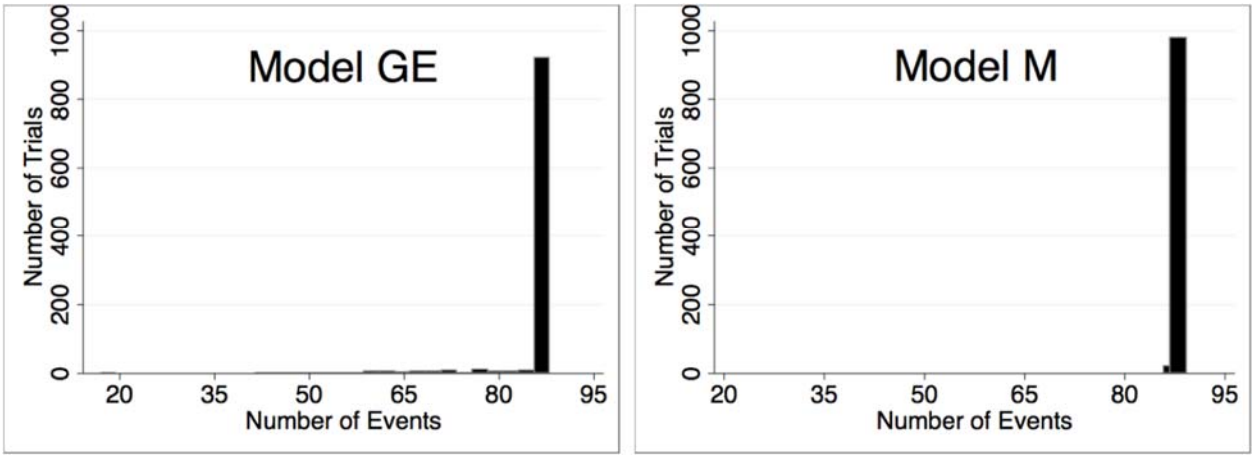
Tumor Grade	Oncotype DX® 0-10		Oncotype DX® 11-18		All
	Age, years		Age, years		
	<50	50+	<50	50+	
Low grade	2.4%	9.5%	6.3%	19.9%	38.1%
Intermediate grade	2.7%	13.5%	9.2%	28.3%	53.7%
High grade	0.4 %	1.7%	1.3%	4.7%	8.1%
All	5.5%	24.7%	16.8%	52.9%	100%

*Data among those with known grade. SEER-GHI, Surveillance Epidemiology and End Results linked with Oncotype data provided by Genomic Health Inc.

Supplementary Figures

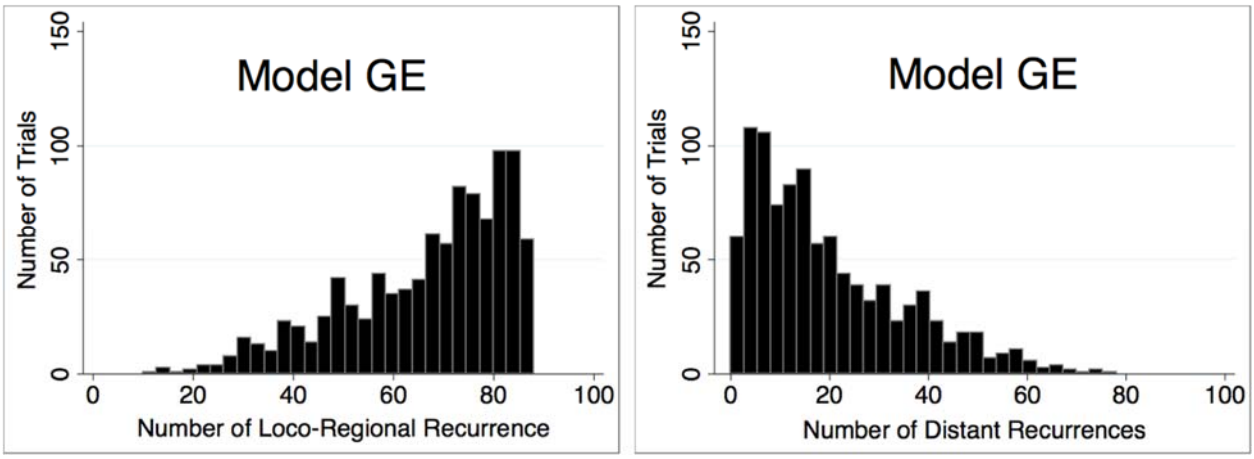


Supplementary Figure 1. Distribution of Trial Length in 1,000 Trial Replications by Model
The histograms demonstrate the differences in total trial length for both models over 1,000 trial simulations. All trials in Model M reached 88 events before nine years; overall trial length for Model M was shorter than GE since events occurred earlier among those with higher-grade tumors, and Model M included more women with intermediate and high grade than Model GE. Most Model GE trials also reached 88 events before 9 years. Results are for patients ages 40-74. Low-risk was defined in the proposed trial as ER+ and/or PR+, human epidermal growth factor 2 (HER2) negative, lymph node negative breast cancers with pathologic tumor size less or equal to 2 cm, Oncotype DX® scores ≤ 18, who were given hormonal therapy following breast conserving surgery, but no adjuvant chemotherapy.



Supplementary Figure 2. Distribution of Numbers of First Recurrence Events in 1,000 Trial Replications by Model

The histograms show the distribution of the numbers of first recurrence events for Models GE and M over 1,000 trial simulations. Almost all of the trials reached the target endpoint of 88 events, although Model GE had several trials with fewer. Results are for patients ages 40-74. Low-risk was defined in the proposed trial as ER+ and/or PR+, human epidermal growth factor 2 (HER2) negative, lymph node negative breast cancers with pathologic tumor size less or equal to 2 cm, Oncotype DX® scores ≤ 18, who were given hormonal therapy following breast conserving surgery, but no adjuvant chemotherapy.



Supplementary Figure 3. Distribution of the Types of First Recurrence Events in 1,000 Trial Replications – Model GE

The histograms show the number of first loco-regional vs. first distant recurrence events over 1,000 trial simulations for Model GE. Most trials had high rates of loco-regional recurrences as the first events and low rates of first distant events. Those in the no radiotherapy arm had higher rates of loco-regional recurrences as first events than those in the radiotherapy arm. Since only the first event is considered in tallying events by arm, and loco-regional recurrences are higher in the no RT than RT arm, the radiotherapy arm has the possibility of having more distant first events than the no radiotherapy arm (data not shown).

Results are for patients ages 40-74. Low risk was defined in the proposed trial as ER+ and/or PR+, human epidermal growth factor 2 (HER2) negative, lymph node negative breast cancers with pathologic tumor size less or equal to 2 cm, Oncotype DX® scores ≤ 18, who were given hormonal therapy following breast conserving surgery, but no adjuvant chemotherapy.

References

1. Darby S, McGale P, Correa C, *et al.* Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. Early Breast Cancer Trialists' Collaborative Group. Lancet 2011;378(9804):1707-16.

2. Fisher B, Bryant J, Dignam JJ, *et al.* Tamoxifen, radiation therapy, or both for prevention of ipsilateral breast tumor recurrence after lumpectomy in women with invasive breast cancers of one centimeter or less. J Clin Oncol 2002;20(20):4141-9.

3. Winzer KJ, Sauerbrei W, Braun M, *et al.* Radiation therapy and tamoxifen after breast-conserving surgery: updated results of a 2 x 2 randomised clinical trial in patients with low risk of recurrence. Eur J Cancer 2010;46(1):95-101.

4. Fyles AW, McCready DR, Manchul LA, *et al.* Tamoxifen with or without breast irradiation in women 50 years of age or older with early breast cancer. N Engl J Med 2004;351(10):963-70.

5. Hughes KS, Schnaper LA, Bellon JR, *et al.* Lumpectomy plus tamoxifen with or without irradiation in women age 70 years or older with early breast cancer: long-term follow-up of CALGB 9343. Journal of Clinical Oncology 2013;31(19):2382-2387.

6. Blamey RW, Bates T, Chetty U, *et al.* Radiotherapy or tamoxifen after conserving surgery for breast cancers of excellent prognosis: British Association of Surgical Oncology (BASO) II trial. Eur J Cancer 2013;49(10):2294-302.

7. Potter R, Gnant M, Kwasny W, *et al.* Lumpectomy plus tamoxifen or anastrozole with or without whole breast irradiation in women with favorable early breast cancer. Int J Radiat Oncol Biol Phys 2007;68(2):334-40.

8. Kunkler IH, Williams LJ, Jack WJ*, et al.* Breast-conserving surgery with or without irradiation in women aged 65 years or older with early breast cancer (PRIME II): a randomised controlled trial. Lancet Oncol 2015;16(3):266-73.

9. Petkov VI, Miller DP, Howlader N*, et al.* Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study. npj Breast Cancer 2016;2:16017.

10. Darby S, McGale P, Correa C*, et al.* Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. Lancet 2011;378(9804):1707-16.

11. Sparano JA, Gray RJ, Makower DF*, et al.* Prospective validation of a 21-gene expression assay in breast cancer. New England Journal of Medicine 2015;373(21):2005-2014.

12. Fisher B, Costantino J, Redmond C*, et al.* A randomized clinical trial evaluating Tamoxifen in the treatment of patients with node-negative breast cancer who have estrogen-receptor–positive tumors. New England Journal of Medicine 1989;320(8):479-484.

13. Fisher B, Dignam J, Wolmark N*, et al.* Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer. J Natl Cancer Inst 1997;89(22):1673-82.

14. Cho H, Mariotto AB, Mann BS*, et al.* Assessing Non–Cancer-Related Health Status of US Cancer Patients: Other-Cause Survival and Comorbidity Prevalence. American Journal of Epidemiology 2013;178(3):339-349.

15. Paik S, Shak S, Tang G*, et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351(27):2817-26.

16. Denberg TD, Lin CT, Myers BA*, et al.* Improving patient care through health-promotion outreach. J Ambul Care Manage 2008;31(1):76-87.