## Supplementary Online Content

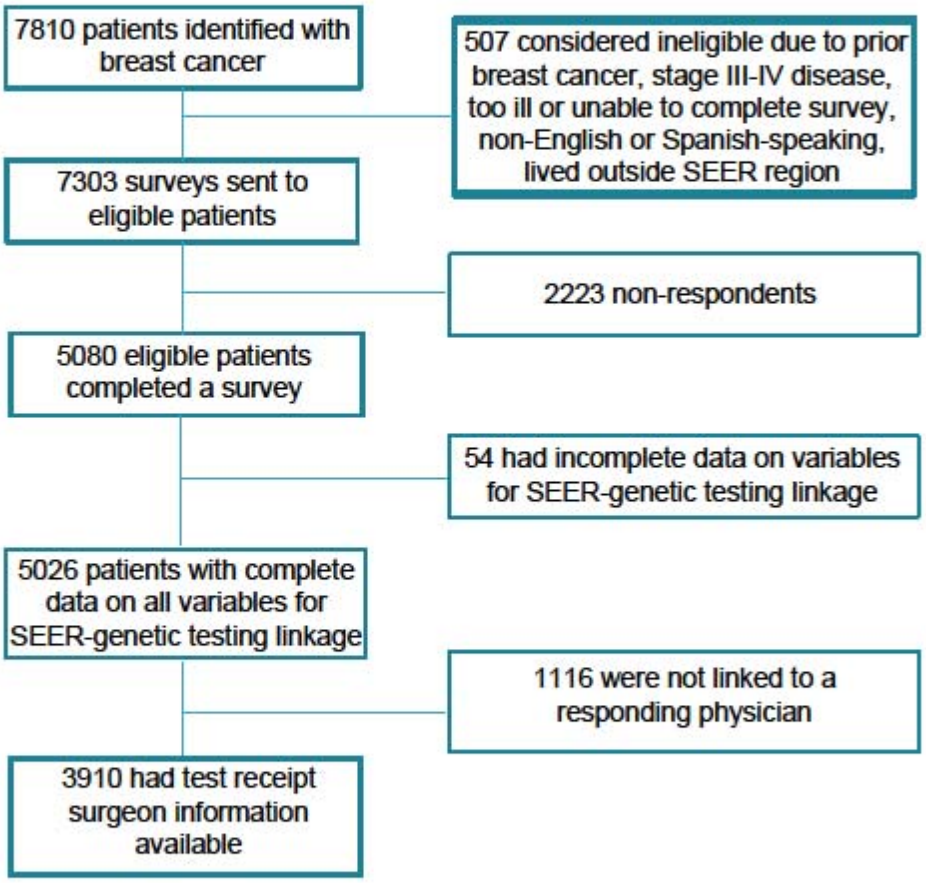
Katz SJ, Bondarenko I, Ward KC, et al. Association of attending surgeon with variation in the receipt of genetic testing after diagnosis of breast cancer. *JAMA Surg*. Published online July 3, 2018. doi:10.1001/jamasurg.2018.2001

**eFigure 1.** Survey Sampling

**eAppendix.** Replication of figure 2 model (in manuscript) that estimated ORs for 3 successive multilevel logistic regression models (MI)

**eFigure 2.** Distribution of Clinical Factors That Define Pretest Risk and Prevalence of Test by Different Groups of Patients

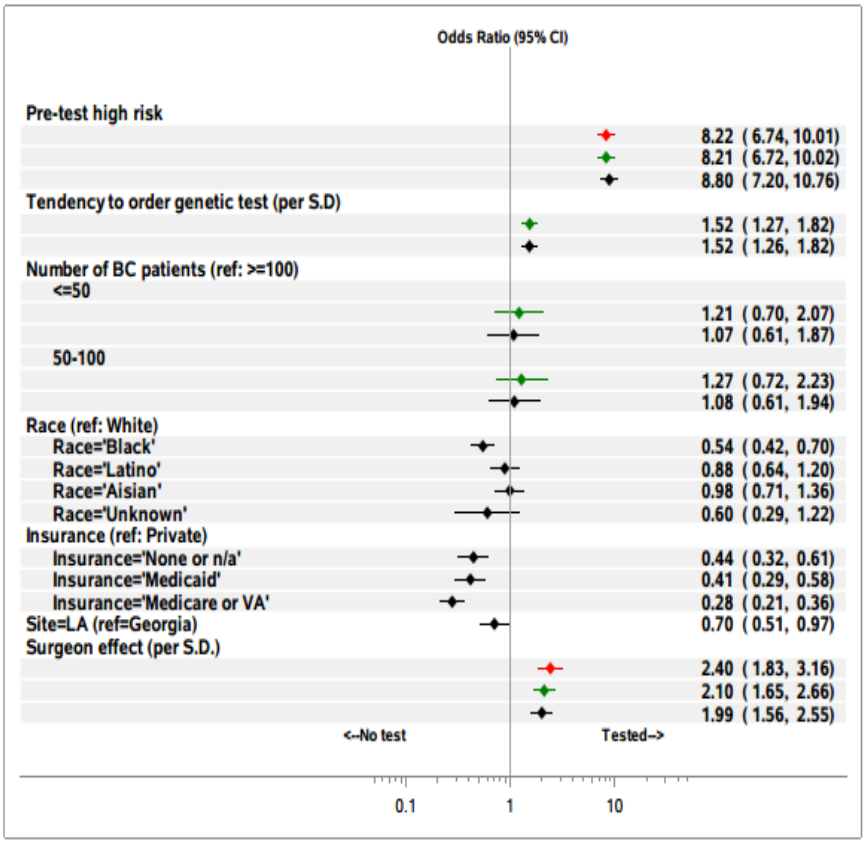
**eFigure 1.** Survey Sampling



**eAppendix. Assessing bias in the multivariate analyses due to missing data:** Though we had a very high surgeon response rate (77%), we had some missing data in the analytic sample due to missing responses from a few surgeons. Seventeen surgeons skipped the question related to the volume of patients. These surgeons were linked to 218 patients. There were no significant differences in the distribution of any of the variables included in models 1-3 between surgeon responders and non-responders. We conducted secondary analyses with multiply imputed surgeon data. Imputed values for volume are drawn from posterior predictive distributions conditional on surgeons' responses about practice, and personal experience, surgeon tendency to test, and number of patients in our survey linked to each surgeon (proxy for volume). Further, to ensure congeniality between predictive and analytic model we conditioned on surgeon-specific rate of testing and prevalence of high-risk patients as well as SES indicators in individual practice. MI yielded 5 imputed data sets. Applying Rubin's rules, we combined estimates from these data **(Figure 2_replication)** and compared them to the results yielded by observed data (Figure 2 manuscript). While there are some small changes in the estimated odds, the direction, magnitude and 95% CI are similar to those in the manuscript.

**Figure 2_replication. The estimated ORs for 3 successive multilevel logistic regression models (MI)**

1



**Analysis of testing by individual components of the NCCN guideline based on pre-test risk of genetic mutation carriage**: **eFigure 2** below shows the distribution of clinical factor that define pre-test risk and prevalence of testing by different groups of patients. Each cell shows a number of women for each combination of these risk factors alonfg with rate of testreceipt in this group. Following the National Comprehensive Cancer Network (NCCN) guidelines, the Pre-test assessment of high risk of pathogenic mutation is a composite of three clinical demographic factors: age, familiar/ancestry history of disease/mutation and triple negative disease (ER-&PR-, and HER2- breast cancer). By these guidelines, a woman is defined as high risk if she meets any of the three criteria listed below:

1. Younger age: age of the diagnosis<=45.
2. Family/ancestry history (FH): Any relatives diagnosed with male breast cancer, ovarian cancer, and/or sarcoma or has two or more relatives diagnosed with breast cancer or Ashkenazi Jewish ancestry.
3. Triple negative disease & age<=60. (TN)

For the 3910 women in the analytic sample 11.7% of women matched the first criterion, 21.0% had family/ancestry history of disease and 2.8% had triple-negative disease and were 60 or younger. Some women matched more than one criterion: 162 (4%) match 2 criteria and 8 women matched all three criteria.

We examined if using the specific criteria independently helped to explain more of the surgeon variation than using the single composite variable representing high risk, by replacing the composite with indicator variables for three criteria and their interactions. We found that the OR associated with the surgeon effect was 2.48 (1.82, 3.38), almost identical to the one yielded by the base model as shown in Figure 2 in the manuscript. However, disaggregating pre-tesk risk into the component parts substantially improved the overall prediction of genetic testing. Being high risk explained 20% of the variation in testing in the base model but a model with the individual components of risk and their interactions explained 33% of the variation. However, the expanded risk variables did not explain any additional amount of the variation in testing attributable to surgeon, which remained at about 17% of the variation in testing in both models.

But the more detailed risk variables can give us some valuable insight about what groups of patients remained systematically under or over-tested relative to the guideline recommendations. **Efigure 2** below shows the predicted prevalence of test receipt across combinations of the three criteria that contribute to the pre-test risk of mutation.

Testing does tend to generally increase with the number of criteria a woman has. Most notably, family/ancestry history tends to be underused as a criteria for testing relative to the guideline recommendations, with only 43% of women receiving testing despite a family history or Jewish ancestry unless a second risk factor is present. Among the group with family history of breast cancer or Jewish ancestry the probability of receiving genetic testing declines with age from 0.68 in women 45 and younger to 0.25 in the elderly. Showing a similar age gradient, within the group of women with triple negative disease <60 felt by NCCN guidelines to be at high risk, women 45-60 with triple negative disease are less likely to be tested than women 45 and younger (52% vs 82%). Among average risk women, for whom the NCCN guidelines would not recommend testing, the highest rate of testing was 23%, which was observed among women 45-60 years old with no other clinical risk factors.

**eFigure 2.** Prevalence of test receipt across combinations of the three criteria that contribute to the pre-test risk of mutation.