Supplementary Online Content

Swisher EM, Rayes N, Bowen D, et al. Remotely delivered cancer genetic testing in the Making Genetic Testing Accessible (MAGENTA) trial. *JAMA Oncol*. Published online September 14, 2023. doi:10.1001/jamaoncol.2023.3748

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Detailed Methods

This trial was generated to facilitate assessment of hereditary breast and ovarian cancer risk with a secondary goal to understand ovarian cancer screening and prevention behaviors. Therefore, eligibility included the presence of one intact ovary. All study interventions, including testing and counseling was provided at no charge to participants. Interested participants completed a stepwise eligibility checklist (Supplementary Table 1). If they were eligible at any step, they were provided that result without answering further questions and offered to complete informed consent. If patients were eligible in both the familial PV and family history cohort, they were included in the familial PV, as those individuals may have had different exposure to genetic information. Race and ethnicity were self-reported by participants and were collected in order to determine diversity of the study population and to identify associations between race and ethnicity and study outcomes.

For the primary endpoint analysis, missing values in partially completed questionnaires were imputed as the mean within the study arm plus the non-inferiority margin of 4. The primary endpoint analysis was based on 1632 impact of event (IES) questionnaires that were at least partially completed. The IES questionnaire included 15 scored items. 1544 subjects (95%) completed all 15 items, 70 subjects (4%) were missing one item, and 18 subjects (1%) were missing two or more items. As a post hoc analysis, we tested for differences in distress three months post-test and differences in the delta from baseline to three months post-test across arms. We also tested for non-inferiority of the experimental arms in the cascade cohort. Rates of high distress (IES scores ≥ 20) were computed for each arm and compared across arms using Fisher's exact test. Continuous scores for depression, anxiety and decisional regret were compared across arms using one-way ANOVA. The interim analysis for efficacy was performed at 50% enrollment (i.e., non-inferiority of experimental arms) and was significant for arm A (one-

sided $p < 0.0002$); since this analysis was non-binding, the decision was made to continue the trial to collect more data for the secondary endpoints and correlatives.

There were two pre-specified hypotheses regarding completion: that the completion rate in the arms requiring pre-test phone counseling would be inferior to those in the arms with only pre-test video by 6% or more, and that the completion rate in the arms requiring post-test phone counseling would be inferior to those in the arms with only a post-test report by 6% or more. For the secondary endpoint analysis, 98.75% one-sided confidence intervals were constructed for the differences in completion rates, and if the upper bound of the confidence interval was less than the non-inferiority margin of 6%, this was determined to be a difference in the completion rate between the family history and cascade cohort using Fisher's exact test and considered grounds to reject the null hypothesis. As a post hoc analysis, we tested for differences in the completion rate between the family history and cascade cohort using Fisher's exact test.

An intent to treat analysis includes data from all randomized patients, but nearly half of randomized patients did not return their three-month question, and approximately half of those who did not return a questionnaire did not receive one as they did not complete testing. We evaluated differences between participants who completed and did not complete their three-month post-test questionnaire (Supplementary Table 1). The pattern of significant differences in subject characteristics for participants that did or did not complete the 3-month questionnaire was similar to the pattern seen in comparing participants who did or did not complete testing, since there was substantial overlap in the group who did not return a questionnaire and did not complete testing (Supplementary Table 5). To take a deeper look at the impact of missing responses on our study results, we performed multiple imputation analysis. [1] We imputed distress scores at 3 months using the mice package in R [2] as follows: we generated 50 imputed data sets using predictive mean matching based on the covariates white, Black, education, and distress score at baseline. As a sensitivity analysis, we constructed an additional version of the

imputed data with the non-inferiority margin of 4 points added to the imputed scores for the non-control arms. We applied the multiple imputation Student's t-test based on the approach of Rubin (1987) [3] with the adjustment of Barnard and Rubin (1999). [4] The multiple imputation t-test results were consistent with those in the primary analysis: each of the three experimental arms was non-inferior to the control arm for distress at 3 months ($p < 0.0082$), for both versions of the imputed data.

Additional statistical tests used: McNemar's test was applied to test for a difference in the proportion of subjects with high distress at 3 months vs. 12 months. Two-way ANOVA was applied to test for differences in the mean change in IES score based on study arm, PV status, and the interaction of these two factors. A Mann-Whitney test was applied to compare baseline distress, anxiety and depression scores for subjects that did vs. did not complete testing.

eReferences

1. Rubin, D. (1976). Inference and missing data. *Biometrika*, 63(3), 581-592.

2. van Buuren, S. and Groothuis-Oudshoorn, K. (2011). “mice: Multivariate Imputation by Chained Equations in R.” *Journal of Statistical Software*, 45(3), 1-67.

3. Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, New York.

4. Barnard, J. and Rubin, D. (1999). Small-Sample Degrees of Freedom with Multiple Imputation. *Biometrika*, 86(4), 948-955.

eTable 1: Eligibility Criteria

Inclusion Criteria Participants must meet **all of the following four criteria**

- Biological female ages 30 or older
- Have access to a healthcare provider and be willing to share genetic results with that provider
- Have at least one ovary
- Have a valid United States mailing address for receipt of saliva kit

And participants must also meet **any one of the following six criteria**

- Diagnosed with breast cancer at age 45 or younger
- Diagnosed with triple negative breast cancer at age 60 or younger
- Have one biological relative with a mutation in BRCA1, BRCA2, BRIP1, PALB2, RAD51C, RAD51D, BARD1, MSH2, MSH6, MLH1, or PMS2
- Have a biological relative with ovarian cancer
- Have at least 2 biological relatives with breast cancer on the same side of the family, one of which is < age 50
- Have one biological male relative with breast cancer

Exclusion Criteria

- Personal history of ovarian cancer
- Unable to speak English
- Unable to provide informed consent
- Unwilling to complete baseline and follow-up questionnaires
- Unable to access the internet
- Previous genetic testing or counseling regarding cancer risk
- Previous bone marrow transplant
- Blood transfusion within 7 days prior to genetic testing
- Active hematologic malignancy

eTable 2: Demographic breakdown by study arm

	A N (% of randomized subjects)	**B** N (% of randomized subjects)	**C** N (% of randomized subjects)	**D** N (% of randomized subjects)	**overall**
Total randomized	960 (25.0%)	958 (25.0%)	960 (25.0%)	961 (25.0%)	3839
Age median (range)	44 (29-83 years)*	44 (25-91 years)*	45 (22-84 years)*	45 (29-78 years)*	44 years (22-91 years)*
Personal history of breast cancer	33 (3.4%)	32 (3.3%)	40 (4.2%)	20 (2.1%)	125 (3.3%)
Race/Ethnicity** Asian/PI/ NH	8 (0.8%)	16 (1.7%)	14 (1.5%)	12 (1.2%)	50 (1.3%)
Black	20 (2.1%)	35 (3.7%)	26 (2.7%)	22 (2.3%)	103 (2.7%)
Hispanic	26 (2.7%)	31 (3.2%)	42 (4.4%)	39 (4.1%)	138 (3.6%)
NatAmer/NatAlaskan	9 (0.9%)	13 (1.4%)	10 (1.0%)	11 (1.1%)	43 (1.1%)
White	925 (96.4%)	902 (94.2%)	912 (95%)	918 (95.5%)	3657 (95.3%)
Education: Grade 12 or below	90 (9.4%)	112 (11.7%)	92 (9.6%)	99 (10.3%)	393 (10.2%)
Any college	613 (63.9%)	605 (63.2%)	602 (62.7%)	595 (61.9%)	2415 (62.9%)
Post College	248 (25.8%)	232 (24.2%)	257 (26.8%)	260 (27.1%)	997 (26.0%)

eTable 3. The means and standard deviations (SD) by arm for continuous outcomes with sample size available per arm at given timepoint. All numbers are for the total cohort unless otherwise specified.

Measure	**Arm A** Mean (SD) (sample size)	**Arm B** Mean (SD) (sample size)	**Arm C** Mean (SD) (sample size)	**Arm D** Mean (SD) (sample size)	**Timepoint**
Total IES score Family History Cohort	9.76 (13.2), (n=464)	7.97 (11.5) (n=388)	9.39 (12.1) (n=361)	9.76 (12.8) (n=419)	3 months
GAD-7 score	4.62 (4.9) (n=536)	4.56 (4.8) (n=453)	4.31 (4.5) (n=413)	4.43 (4.8) (n=484)	3 months
PHQ-9 score	4.21 (4.7) (n=535)	3.99 (4.3) (n=452)	4.16 (4.7) (n=412)	3.98 (4.5) (n=478)	3 months
DRS score	4.21 (4.7) (n=533)	3.99 (4.3) (n=446)	4.16 (4.7) (n=409)	3.98 (4.5) (n=476)	3 months
Total IES score	8.79 (12.7) (n=318)	7.87 (12.3) (n=317)	8.13 (10.9) (n=268)	8.08 (12.0) (n=287)	12 months

eTable 4. Analyses for binary outcomes, along with sample sizes for each arm at given timepoints. All numbers are for the total cohort unless otherwise specified.

Measure	**Arm A** N (%) (sample size)	**Arm B** N (%) (sample size)	**Arm C** N (%) (sample size)	**Arm D** N (%) (sample size)	**Timepoint**
Proportion High Distress	104 (18.7%) (n=556)	79 (16.9%) (n=468)	85 (19.9%) (n=428)	98 (19.6%) (n=501)	3 months
Completion rate	738 (76.9%) (n=960)	753 (78.6%) (n=958)	664 (69.2%) (n=960)	638 (66.4%) (n=961)	Received results at any timepoint
Proportion High Distress	48 (15.1%) (n=318)	47 (14.8%) (n=317)	36 (13.4%) (n=268)	39 (13.6%) (n=287)	12 months

eTable 5. Comparison of participants who completed or did not complete the 3-month questionnaires for the primary outcome.

Characteristic	**Completed**, N = 1,632[1]	**Not completed**, N = 1,493[1]	**p-value**[2]
Age	45 (22, 91)	44 (29, 84)	0.11
Personal history of breast cancer	73 (4.5%)	52 (3.5%)	0.2
Race/Ethnicity			
Asian/PI/NH	24 (1.5%)	19 (1.3%)	0.6
Black	26 (1.6%)	50 (3.3%)	0.001
Hispanic	44 (2.7%)	51 (3.4%)	0.2
NatAmer/NatAlaskan	20 (1.2%)	11 (0.7%)	0.2
White	1,582 (97%)	1,403 (94%)	<.001
Education			<0.001
Grade 12 or below	126 (7.8%)	156 (11%)	
Any college	989 (61%)	952 (64%)	
Post college	506 (31%)	371 (25%)	
Baseline distress	8 (0, 66)	8 (0, 61)	0.069[3]

[1] Median (Range); n (%),

[2] Wilcoxon rank sum test; Pearson's Chi-squared test.

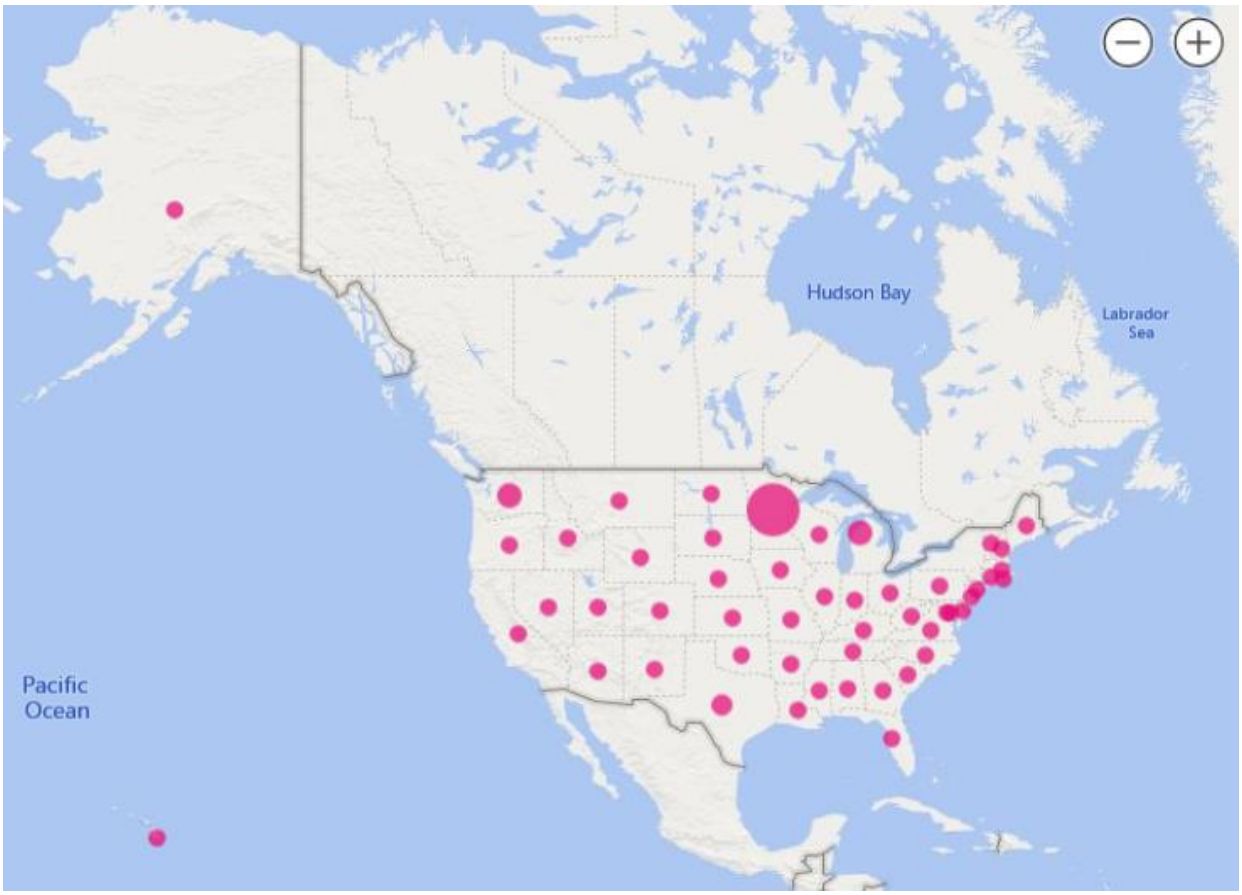
[3] Baseline distress was non-significant but since the p-value was <0.1, we included it as a covariate in the multiple imputation model.

eTable 6. Subject screening, enrollment and completion of each step with breakdown by cohort enrollment.

	N (% of total randomized)	Familial PV Cohort	Family history cohort	P Value*
Completed eligibility questionnaire	14479			
Eligible	6455			
Signed Consent	4130			
Completed baseline questionnaire and randomized	3839	714	3125	
Completed pre-test counseling/video	3363 (87.6%)	636 (89.1%)	2727 (87.3%)	NS
Activated and returned kit	2839 (74.0%)	498 (69.7%)	2341 (74.9%)	.005
Complete post-test counseling/results	2793 (72.8%)	477 (66.8%)	2316 (74.1%)	<.0001
Complete month 3 follow up questionnaire	2072 (53.9%)	341 (47.8%)	1731 (55.5%)	0.002

NS: non-significant
*Chi square analysis

eFigure 1: Geographic distribution of subjects in MAGENTA.



eFigure 2. Psychological outcomes at 3 months post-receipt of genetic testing Results.