



## ARTICLE

Refining the risk for fragile X–associated primary ovarian insufficiency (FXPOI) by *FMR1* CGG repeat size

Emily Graves Allen<sup>1</sup>  , Krista Charen<sup>1</sup>, Heather S. Hipp<sup>2</sup>, Lisa Shubeck<sup>1</sup>, Ashima Amin<sup>1</sup>, Weiya He<sup>1</sup>, Sarah L. Nolin<sup>3</sup>, Anne Glicksman<sup>3</sup>, Nicole Tortora<sup>3</sup>, Bonnie McKinnon<sup>1</sup>, Katharine E. Shelly<sup>1</sup> and Stephanie L. Sherman<sup>1</sup>

**PURPOSE:** Approximately 20–30% of women with an *FMR1* premutation experience fragile X–associated primary ovarian insufficiency (FXPOI); however, current risk estimates based on repeat size only identify women with the midrange of repeats to be at the highest risk.

**METHODS:** To better understand the risk by repeat size, we collected self-reported reproductive histories on 1,668 women and divided them into high-resolution repeat size bins of ~5 CGG repeats to determine a more accurate risk for FXPOI in relation to CGG repeat length.

**RESULTS:** As previously reported, women with 70–100 CGG repeats were at the highest risk for FXPOI using various statistical models to compare average age at menopause and risk of FXPOI, with women with 85–89 repeats being at the highest risk. Importantly, women with <65 repeats or >120 repeats did not have a significantly increased risk for FXPOI compared to women with <45 repeats.

**CONCLUSION:** Using a large cross-section study on 1,668 women, we have provided more personalized risk assessment for FXPOI using high-resolution repeat size bins. Understanding the variability in risk has important implications for family planning and overall health among women with a premutation.

*Genetics in Medicine* (2021) 23:1648–1655; <https://doi.org/10.1038/s41436-021-01177-y>

## INTRODUCTION

The 5' untranslated region of the *FMR1* gene contains a CGG repeat that is responsible for at least three distinct clinical phenotypes: fragile X syndrome (FXS), fragile X–associated tremor/ataxia syndrome (FXTAS), and fragile X–associated primary ovarian insufficiency (FXPOI). The full mutation form of the *FMR1* gene contains more than 200 CGG repeats that are hypermethylated, resulting in transcriptional silencing of the gene.<sup>1</sup> The lack of *FMR1* protein product, FMRP, which is an RNA-binding protein involved in control of translation,<sup>2</sup> leads to FXS.<sup>3</sup> FXTAS and FXPOI are seen in individuals with the premutation form of the *FMR1* gene that contain 55–199 CGG repeats.<sup>4,5</sup> FXTAS is characterized by intention tremor and gait ataxia, with associated features of parkinsonism, neuropsychological dysfunction, autonomic dysfunction, and peripheral neuropathy.<sup>6</sup> The prevalence is approximately 40% and 6–18% among men and women, respectively.<sup>7,8</sup> Women with FXPOI develop hypergonadotropic hypogonadism and have absent or irregular periods prior to age 40.<sup>9</sup> FXPOI occurs in approximately 20% of women with a premutation, and women with FXPOI have an increased risk of infertility and hypoestrogenism, resulting in hot flashes, night sweats, and increased risk for osteoporosis.<sup>10</sup> More recently, an additional classification of FXAND (fragile X–associated neuropsychiatric disorders) was proposed to describe an increased frequency of mental health problems, immune disorders, and various other medical conditions among women with a premutation.<sup>11</sup>

While FXPOI is the most frequent of the fragile X–associated disorders, it is underdiagnosed and biologically the least well understood. Because premutation alleles are seen in about 1 in 151 women,<sup>12</sup> more than 1 million women in the United States

carry a premutation allele. In qualitative interviews with women with a premutation, the diagnosis of FXPOI was delayed in women with known premutation status when presenting with symptoms of POI; this delay in diagnosis was even greater for women who did not know their premutation status.<sup>13</sup> Individuals with premutation alleles and stakeholder groups alike (e.g., health-care providers) have emphasized the importance of understanding risk for FXPOI and identifying predictive markers.<sup>14–17</sup> The need for this effort becomes magnified as fragile X screening studies become more prevalent and identify women with premutation alleles who are at risk for FXPOI.<sup>14,18–21</sup>

Previous studies have consistently identified that the risk for FXPOI is associated with CGG repeat size in a nonlinear fashion: those with ~70 to 100 CGG repeats are at the highest risk.<sup>10,22–26</sup> Because the risk is nonlinear, the molecular mechanism causing FXPOI may be distinct from the mechanism for FXTAS, which shows a linear risk with repeat size. However, because the underlying mechanism of this shortened reproductive lifespan is unknown and there are no biomarkers for prediction, all individuals with premutation alleles are “affected” in the sense that they are at risk and no fertility preserving interventions can be prescribed.

In the current work, we examined fragile X repeat sizes among 1,668 women, using high-resolution bins, to determine a more accurate risk for FXPOI in relation to CGG repeat length. We divided repeat sizes into ranges of ~5 repeats where the sample size was adequate, and statistically modeled the risk for an earlier onset of menopause or increased risk for FXPOI for each repeat size group. In addition, we investigated self-reported cycling and fertility characteristics for women in each of these repeat size

<sup>1</sup>Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA. <sup>2</sup>Department of Gynecology and Obstetrics, Emory University School of Medicine, Atlanta, GA, USA. <sup>3</sup>Department of Human Genetics, New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY, USA. email: emgrave@emory.edu

groups. Our goal is to provide women more personalized risk information based on their repeat size.

## MATERIALS AND METHODS

### Study population

Emory University Institutional Review Board approved protocols and consent forms, and informed consent was obtained from all participants. Participants were identified through two different sampling schemes: (1) a general population survey in metropolitan Atlanta to identify individuals with smaller allele sizes<sup>27</sup> and (2) from fragile X family recruitment efforts to identify individuals with larger alleles (e.g., previous fragile X research projects at Emory, recruitment at scientific conferences, and through collaborations with other research groups). Once a family contact was identified, additional family members were screened for eligibility without respect to phenotype. Eligibility was based on premutation allele status, age, and sex. A blood or saliva sample was collected, and each participant completed a reproductive history questionnaire. Data included general demographics (e.g., age at interview, date of birth, race/ethnicity), lifestyle factors that might affect overall health (e.g., smoking, body mass index [BMI]), and reproductive history (e.g., menstrual history, reason for cessation of menses, pregnancy history). A subset of women were also interviewed by a reproductive endocrinologist ( $n=83$ ).<sup>13</sup> For many women, the questionnaires and interviews were administered at multiple time points. Research personnel curated all data points to determine correct assignments for study variables.

### Variable definitions

**Age at menopause (AAM).** The reproductive history was used to determine whether a woman was still cycling or why her menses had ceased. If the date of last menstrual period was more than two months prior to the interview, we identified the cause of menses cessation. Diagnosis was based on self-reporting. AAM was assigned if menses had been absent for at least 12 months or 4–6 months if a menopausal-level follicle-stimulating hormone level was available.<sup>28</sup> If AAM from the in-depth interview differed from information provided on the reproductive questionnaire, the data from the in-depth interview were used, as the interviewer (H.S.H.) had information from all previous reproductive questionnaires during the interview. Thus, she was able to obtain the necessary information to determine the most accurate AAM. Women whose menses ceased because of surgery, pregnancy, chemotherapy or radiation, eating disorders, or hormone use were not assigned an AAM.

**FXPOI.** Women were defined as having FXPOI if their age at natural menopause (AAM) as defined above was < age 40. Although we recognize that experiencing POI and menopause is not the same, we will use the notation of AAM to reference the onset of these conditions for all women. Women who had menopause or were still having menstrual cycles at age 40 or later were classified as not having FXPOI. For some women, a FXPOI assignment could not be made (e.g., women who were still cycling, but younger than age 40; women who had surgery, such as a hysterectomy, before age 40, etc.).

**Cycle characteristics.** For women with more than one questionnaire completed, the answer from their earliest questionnaire was used for age at menarche. For all other cycling traits, we used the questionnaire done in closest time proximity to when the cycles ended. Variables (e.g., short cycles and bleed length) were categorized as described in our previous study.<sup>10</sup> Short cycles were defined as <27 days from the first day of one menstrual period to the first day of the next menstrual period, and short bleed length was defined as less than five days from the time that bleeding began until it completely stopped within an average cycle during the last year they were naturally cycling.

**Fertility characteristics.** Two questions were analyzed to assess fertility problems. First, each woman reported whether she ever had unprotected intercourse for a year or more without getting pregnant. Second, each woman was asked if she had ever visited a doctor, clinic, or hospital because she was unable to become pregnant.

**Environmental factors.** Smoking characteristics were curated across questionnaires when necessary. Variables for having ever smoked regularly (at least one cigarette per day), current smoker, and pack years were

determined for each woman. Pack years was used in the final models because it showed the strongest association with AAM. BMI was also calculated based on the earliest questionnaire completed. A binary variable for self-reported race/ethnicity was used due to the vast majority self-reporting as Caucasian (1 = Caucasian; 0 = other or not reported).

### Laboratory methods

DNA was extracted from biological samples using Qiagen Qiamp DNA Blood Mini Kit, Gentra Puregene extraction kit, or prepIT-L2P protocol from Oragene.

FRAXA CGG repeat numbers were determined by a fluorescent sequencer method.<sup>29</sup> For females with only one allele, a second polymerase chain reaction (PCR) protocol was used to look for larger allele.<sup>30</sup> If only one allele was seen using this method, the woman was classified as homozygous for that repeat size. The PCRs for FRAXA consisted of 1X PCR Buffer (Gibco/BRL), 10% dimethyl sulfoxide (DMSO), 370  $\mu$ M deazaG, 500  $\mu$ M d(ACT), 0.3  $\mu$ M each primer, 15 ng T4 gene 32, and 1.05 U Roche Expand Long Taq. Primers for the FMR1 gene were C: 5'GCTCAGCTCCGTTTCGGTTTCACTCCGGT3', and F: 5'AGCCCCGCACTCCA CCAGTCTCTCA3'.<sup>31</sup>

For a subset of samples (see Supplementary Figure 1), *FMR1* repeat sizing was performed by PCR analysis using AmpliDx PCR CE/*FMR1* (Asuragen, Inc.) assay. Repeat size and the presence of repeat size mosaicism (PM/PM: more than one premutation size allele detected; PM/FM: both PM and FM alleles detected) were determined using AmpliDx PCR CE/*FMR1* at the New York State Institute for Basic Research in Developmental Disabilities or Xpansion Interpreter (Asuragen, Inc.) at Asuragen.

### Statistical analysis

Standard statistical methods were used. Age, BMI, and smoking pack years are reported as the mean  $\pm$  standard error. Differences in demographics among repeat size groups were examined using an analysis of variance (ANOVA) model. Differences for binary variables (e.g., race/ethnicity and smoking) among groups were tested using a  $\chi^2$  analysis. ANOVA models to compare means of continuous measures (e.g., AAM) between repeat groups and significant differences between groups were determined using Tukey's post hoc analysis to control for multiple testing. Logistic regression models were used to test for associations between repeat groups and risk of FXPOI, cycling, and fertility traits. Models included all repeat classes as predictor variables, along with covariates. Women with <45 repeats were the referent group. To compare the AAM across repeat size groups, survival analyses were used. For this analysis, women who had experienced AAM were included as those with the "event," and the AAM was used as the "event" age. Women who did not have an assigned AAM were censored at the following time points: women who were still cycling and not currently on hormone medication were censored at their age of interview, women who had a surgery that led to cessation of menses (e.g., hysterectomy) were censored at the age of surgery, women who were currently on hormone medication were censored at the last age they were not on hormone medications, and women who reported their periods had stopped for other reasons (e.g., chemotherapy) were censored at the age of their last period. All models were adjusted for age at interview, race/ethnicity, smoking pack years, and BMI.

Linear regression models were used to test for the association of repeat size mosaicism with AAM. In this model the outcome measure was AAM, and the model included repeat size, the quadratic of repeat size, binary variables for the presence or absence of each type of mosaicism (i.e., PM/PM or PM/FM (1 = mosaicism detected; 0 = no mosaicism detected)), age at interview, race/ethnicity, smoking pack years, and BMI.

Because some of the women were related, we confirmed our findings using generalized estimating equations (GEE) and frailty analysis to adjust for the dependency of related individuals.

All analyses were done using SAS 9.4 and R.

## RESULTS

Genotyping results for the *FMR1* CGG repeat were used to group the 1,668 women into repeat size groups. Most repeat size groups included a range of five repeats, with the exception of the <45 repeats group, the 45–54 repeats group, and the higher end of the premutation size range. We collapsed these groups due to

**Table 1.** Demographic, environmental, and descriptive characteristics for repeat size groups.

Repeat group	N	Age at interview mean $\pm$ S.E.	Race			BMI <sup>a</sup> mean $\pm$ S.E.	% Ever smoke	Pack years mean $\pm$ S.E.
			% Caucasian	% African American	% Other/ not stated			
<45 repeats	634	43.9 $\pm$ 0.6	77.8	13.9	8.3	27.0 $\pm$ 0.3	34.0	4.9 $\pm$ 0.5
45–54 repeats	66	40.9 $\pm$ 1.8	75.8	19.7	4.5	26.8 $\pm$ 0.8	37.9	4.6 $\pm$ 1.4
55–59 repeats	50	42.8 $\pm$ 2.1	88.0	8.0	4.0	26.1 $\pm$ 0.9	32.0	2.4 $\pm$ 0.7
60–64 repeats	38	47.6 $\pm$ 2.5	89.5	0	10.5	26.9 $\pm$ 0.9	42.1	6.5 $\pm$ 2.1
65–69 repeats	53	51.1 $\pm$ 1.8	92.4	3.8	3.8	27.5 $\pm$ 0.8	46.1	7.5 $\pm$ 1.6
70–74 repeats	73	49.5 $\pm$ 1.4	90.4	1.4	8.2	28.8 $\pm$ 0.8	31.5	7.3 $\pm$ 2.0
75–79 repeats	112	47.4 $\pm$ 1.3	88.4	4.5	7.1	26.3 $\pm$ 0.5	30.6	3.9 $\pm$ 0.9
80–84 repeats	110	48.4 $\pm$ 1.3	88.2	3.6	8.2	25.7 $\pm$ 0.5	37.4	6.6 $\pm$ 1.5
85–89 repeats	117	44.1 $\pm$ 1.1	85.6	5.1	9.3	27.8 $\pm$ 0.7	34.2	5.3 $\pm$ 1.1
90–94 repeats	119	46.3 $\pm$ 1.1	89.1	3.4	7.5	26.9 $\pm$ 0.6	30.8	4.9 $\pm$ 1.1
95–99 repeats	89	44.1 $\pm$ 1.3	92.1	2.2	5.7	28.2 $\pm$ 0.9	37.9	4.4 $\pm$ 1.1
100–104 repeats	62	44.1 $\pm$ 1.5	93.5	3.2	3.3	26.6 $\pm$ 0.8	29.0	2.4 $\pm$ 0.7
105–109 repeats	28	45.5 $\pm$ 2.5	92.9	3.6	3.5	27.4 $\pm$ 1.3	25.9	1.3 $\pm$ 0.6
110–119 repeats	45	43.4 $\pm$ 1.9	88.9	4.4	6.7	27.2 $\pm$ 1.0	50.0	5.6 $\pm$ 1.4
120–129 repeats	35	44.0 $\pm$ 2.2	91.4	2.9	5.7	26.5 $\pm$ 1.2	31.4	2.5 $\pm$ 1.0
130–199 repeats	37	39.0 $\pm$ 2.1	94.6	0	5.4	26.9 $\pm$ 0.9	29.7	2.4 $\pm$ 0.9
Total	1668	45.0 $\pm$ 0.3	84.6	8.1	7.3	27.0 $\pm$ 0.2	34.5	4.9 $\pm$ 0.3

BMI body mass index.

<sup>a</sup>For women with multiple questionnaires, answer from first questionnaire was used.

decreased sample sizes of either the number of women with known AAM (45–54 repeats group) or overall sample size (higher premenstrual groups). Table 1 shows the sample size for each of the groups as well as demographic, environmental, and descriptive variables. There were significant differences in the age at interview and race/ethnicity across groups; therefore, all statistical models were adjusted for these variables. Although we did not identify a significant difference between groups for smoking or BMI, these variables were also included in all statistical models based on their known association with AAM.<sup>32,33</sup>

#### Risk for AAM/FXPOI based on repeat size group

**Adjusted mean AAM per repeat size group (ANOVA).** To test for differences in average AAM across repeat groups, we first used an ANOVA model that was adjusted for age at interview ( $p < 0.0001$ ), race/ethnicity ( $p = 0.88$ ), smoking pack years ( $p = 0.0006$ ), and BMI (0.99). Only the 529 women with an assigned AAM were included in this analysis. Table 2 shows the adjusted least squares means with the 95% confidence limits. Figure 1 shows the distribution of raw AAM for each of the repeat size groups. The sample size for each repeat size group in the model is also shown. Groups shown in bold in Table 2 were significantly different from the referent group (<45 repeats) using Tukey's Studentized range test that adjusts for multiple testing. Importantly, repeat size groups with <70 repeats and  $\geq 120$  repeats were not significantly different from the referent group.

**Adjusted odds ratios (OR) for FXPOI per repeat size group (logistic regression).** Using a logistic regression model with a binary variable for POI included a total of 903 women. This model was also adjusted for age at interview ( $p < 0.0001$ ), race/ethnicity ( $p = 0.10$ ), smoking pack years ( $p = 0.18$ ), and BMI ( $p = 0.68$ ). Figure 2 shows the raw percentages of POI for each repeat size group, and

Table 2 shows the odds ratios and 95% confidence interval for each repeat size group from the adjusted model. The odds ratios that are significantly different from referent group are shown in bold in Table 2. For this analysis, repeat groups with <65 repeats and  $\geq 120$  repeats were not significantly different from the referent group. Odds ratios could not be calculated for repeat groups with 45–64 repeats because there were no women with POI in that repeat size range.

**Adjusted hazard ratios (HR) for FXPOI per repeat size group (survival analysis).** Using survival analysis enables us to include the information from the maximum number of women: a total of 1,607 women were included in this model. As described above for the ANOVA model, 529 women had an assigned AAM, and the remaining 1,078 women were censored from the analysis at the age described in statistical methods. The model was adjusted for age at interview ( $p = 0.001$ ), race/ethnicity ( $p = 0.61$ ), smoking pack years ( $p < 0.0001$ ), and BMI (0.24). Consistent with the ANOVA and logistic regression models, repeat groups with <65 repeats and  $\geq 120$  repeats were not significantly different from the referent group (Table 2). Figure 3 shows survival curves for each repeat size group (shown in black) compared to the low-risk <45 repeats group (shown in green) and the highest risk group (85–89 repeats; shown in red). The 85–89 repeat size group was chosen as the high-risk group because this group showed the lowest average AAM. The difference in survival curves reflect the findings for the hazard ratios presented in Table 2.

#### Cycling traits based on repeat size groups

We also examined self-reported cycling traits for each repeat size group. Women were asked to describe these traits for the last year they were naturally cycling. No significant difference in reported age at menarche was detected between groups in an ANOVA

**Table 2.** Outcome measures by repeat size group for the three statistical models used to evaluate age at menopause (AAM) or risk of FXPOI.

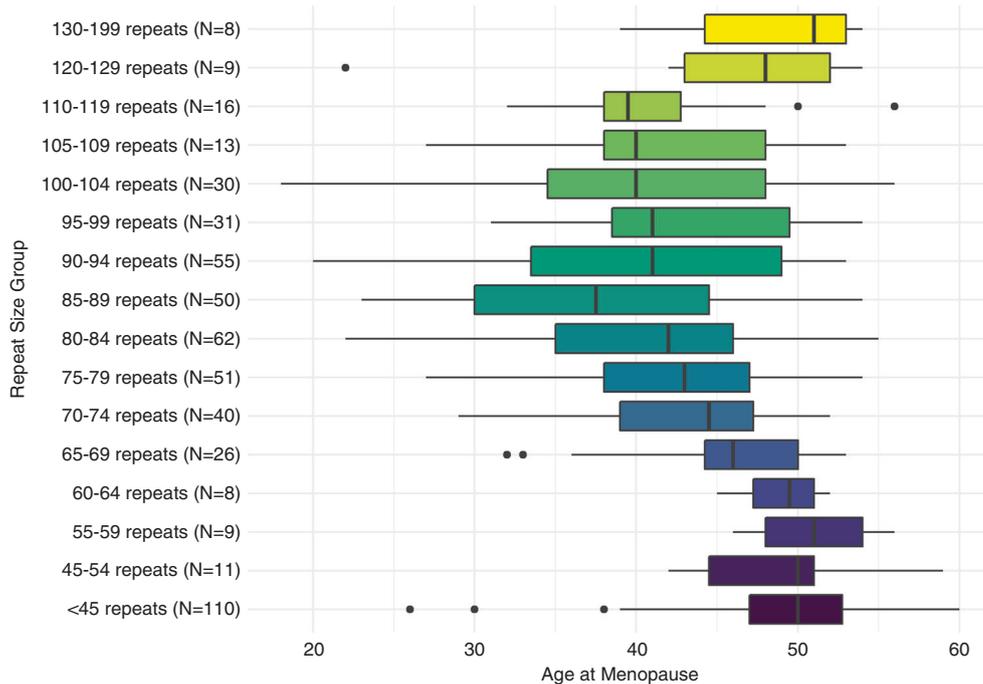
	ANOVA (N = 529) <sup>a,b</sup>	Logistic Regression (N = 903) <sup>a</sup>	Survival Analysis (N = 1,607) <sup>b</sup>
Repeat Group	Least squares mean for AAM (95% CI)	OR for FXPOI (95% CI)	HR for Menopause (95% CI)
<45 repeats (ref)	47.8 (46.6–49.1)	-	-
45–54 repeats	48.2 (44.4–52.1)	N/A <sup>c</sup>	1.2 (0.7–2.3)
55–59 repeats	50.2 (46.0–54.5)	N/A <sup>c</sup>	1.0 (0.5–1.9)
60–64 repeats	47.2 (42.7–51.7)	N/A <sup>c</sup>	1.5 (0.7–3.2)
65–69 repeats	44.5 (42.0–47.0)	<b>6.4 (1.3–30.1)</b>	<b>3.6 (2.3–5.6)</b>
70–74 repeats	<b>43.1 (41.1–45.1)</b>	<b>20.2 (6.1–67.0)</b>	<b>4.6 (3.2–6.6)</b>
75–79 repeats	<b>42.7 (41.0–44.5)</b>	<b>23.1 (7.4–72.3)</b>	<b>4.0 (2.8–5.6)</b>
80–84 repeats	<b>40.6 (39.0–42.2)</b>	<b>41.7 (13.7–126.8)</b>	<b>4.6 (3.3–6.3)</b>
85–89 repeats	<b>38.7 (36.9–40.5)</b>	<b>37.7 (12.4–114.5)</b>	<b>5.2 (3.7–7.3)</b>
90–94 repeats	<b>41.4 (39.7–43.1)</b>	<b>29.9 (9.8–91.0)</b>	<b>4.4 (3.2–6.2)</b>
95–99 repeats	<b>43.3 (41.0–45.5)</b>	<b>14.2 (4.2–48.4)</b>	<b>3.0 (2.0–4.6)</b>
100–104 repeats	<b>41.7 (39.3–44.0)</b>	<b>30.3 (9.2–100.0)</b>	<b>2.8 (1.9–4.3)</b>
105–109 repeats	<b>41.2 (37.7–44.7)</b>	<b>32.2 (7.5–139.1)</b>	<b>4.6 (2.6–8.2)</b>
110–119 repeats	<b>41.9 (38.7–45.1)</b>	<b>24.2 (6.5–90.5)</b>	<b>3.9 (2.3–6.5)</b>
120–129 repeats	44.6 (40.4–48.8)	2.9 (0.3–28.2)	2.0 (1.0–4.0)
130–199 repeats	48.4 (43.9–52.8)	3.7 (0.4–36.4)	1.6 (0.8–3.4)

All models were adjusted for age at interview, smoking pack years, race, and BMI. Bold indicates a significant difference (<0.05) from the <45 repeats (referent) group. Deepness of color indicates variation of the test statistic across repeat groups.

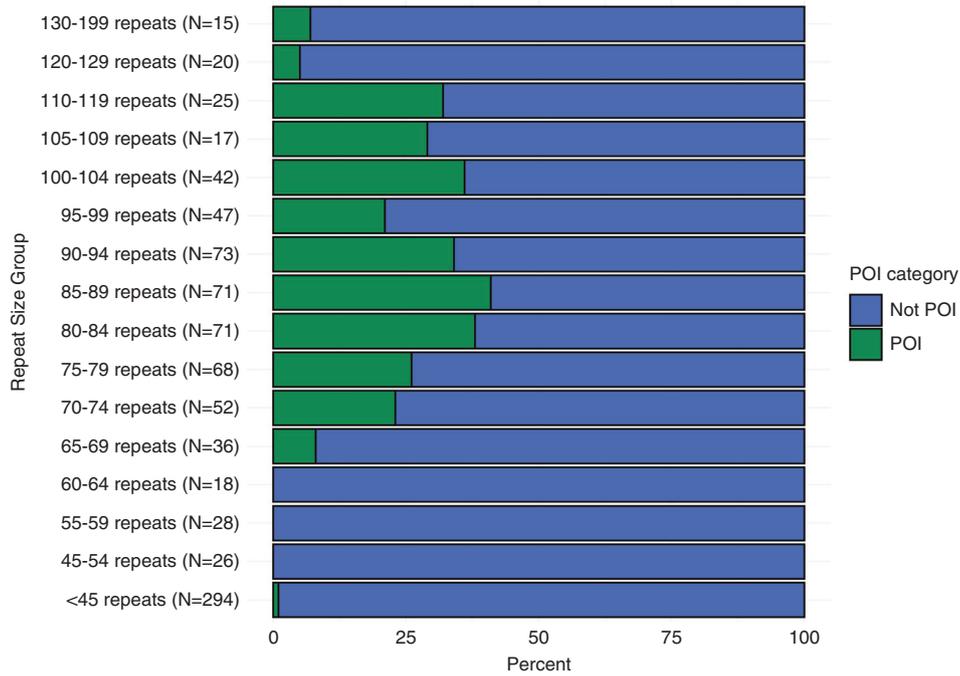
<sup>a</sup>Age at interview ( $p < 0.0001$ ).

<sup>b</sup>Smoking pack years ( $p < 0.001$ ) significant in model.

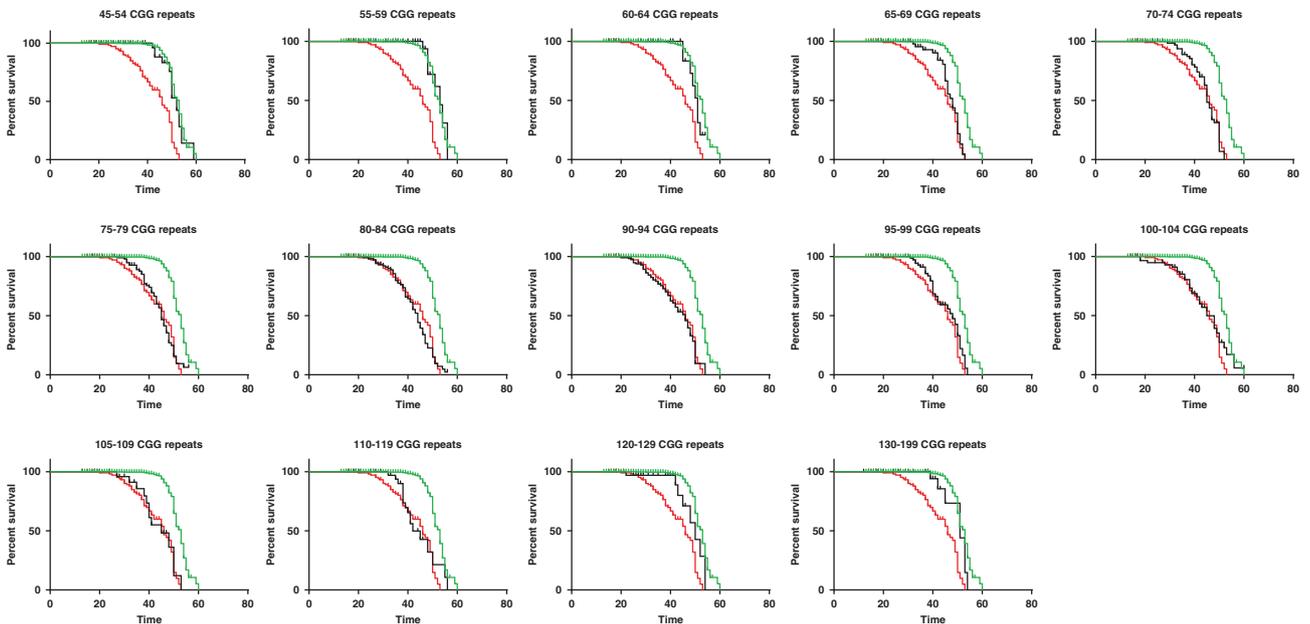
<sup>c</sup>Odds ratio could not be calculated because there were no women with POI in that repeat size range.



**Fig. 1** Box plot of age at menopause distribution by repeat size group. Vertical lines within the box from left to right represent the lower quartile, the median, and the upper quartile, respectively. The horizontal lines represent the 5th and 95th percentiles, and the values beyond these lines, marked as dots, are considered outliers.



**Fig. 2 Risk of POI by repeat size group.** Raw percentages of women with fragile X-associated primary ovarian insufficiency (FXPOI) (age at menopause [AAM] < age 40) by repeat size group.



**Fig. 3 Survival analysis plots for each repeat size group.** Survival analysis plots are shown for each repeat size group (shown in black) compared to the low-risk <45 repeats group (shown in green) and the high-risk group of 85–90 repeats (shown in red).

model that was adjusted for age at interview and race/ethnicity (Supplementary Table 1). The remaining cycling traits were each examined using logistic regression models that were adjusted for age at interview, race/ethnicity, smoking pack years, and BMI. Reporting of short cycle length, defined as <27 days, was significantly higher in the 70–74 repeat size group (OR = 1.77; 95% CI = 1.04–3.02) and the 90–94 repeat size group (OR = 1.57; 95% CI = 1.02–2.42). Short bleed length was defined as <5 days of bleeding, and the 80–84, 85–89, 95–99, and the 105–109 repeat group showed an increased likelihood of reporting this (OR = 2.22,

95% CI = 1.42–3.45; OR = 1.83, 95% CI = 1.17–2.82; OR = 1.68, 95% CI = 1.01–2.74; and OR = 3.19, 95% CI = 1.38–7.35, respectively). Two measures of cycle irregularity were also included. For the first variable of irregular cycles, four repeat groups reported significantly increased frequencies compared to the referent group: 70–74 repeats (OR = 1.70, 95% CI = 1.01–2.85), 85–89 repeats (OR = 1.94, 95% CI = 1.28–2.93), 90–94 repeats (OR = 1.88, 95% CI = 1.24–2.83), and 120–129 repeats (OR = 2.26, 95% CI = 1.09–4.77). The second variable for irregularity asked if women ever went six weeks or more without having a period, and the

majority of repeat groups between 75 and 104 repeats reported a significantly increased likelihood of this compared to the referent group (Supplementary Table 1).

#### Fertility traits based on repeat size groups

Two variables were analyzed to investigate the presence of fertility issues: (1) going a year or more having unprotected intercourse without getting pregnant and (2) visiting a doctor, clinic, or hospital because they could not get pregnant. Logistic regression was used to test for the likelihood of reporting these variables compared to the referent group. Four repeat size groups showed an increased likelihood of ever going a year or more having unprotected intercourse without getting pregnant compared to the <45 repeats group: 75–79, 85–89, 90–94, and 110–119 repeats (Supplementary Table 2). The second question, if they had ever visited a doctor, clinic, or hospital because they could not get pregnant, indicated a significantly increased OR for repeat groups from 75–94 repeats (Supplementary Table 2).

#### Repeat size mosaicism as a possible modifier of risk

Mailick et al.<sup>34</sup> found that women with a premutation allele with PM/FM mosaicism of their repeat size, as measured in blood, were associated with better health, mental health, and executive function compared with women who were nonmosaic for a premutation allele. Based on these findings, we evaluated whether repeat size mosaicism altered the risk of FXPOI in our population. We had information on the presence of mosaicism on 320 women with a premutation (Asuragen assay conducted; see “Materials and Methods”). Based on this assay, 91 women were identified to have PM/PM mosaicism and 7 women had PM/FM mosaicism. We performed a survival analysis to identify whether either type of mosaicism had a significant impact on AAM. This model included the presence of mosaicism as the predictor variable, adjusted for repeat size and the quadratic of repeat size. The presence of PM/PM mosaicism and of PM/FM mosaicism were each tested separately using a binary variable (1 = mosaicism was present, 0 = no mosaicism detected). Neither type of mosaicism was significant in the models tested; although, the limited sample size for PM/FM mosaicism was clearly a limitation. The distributions for repeat size by AAM on the 131 women with an assigned AAM by presence of type of mosaicism as assessed by the Asuragen assay are shown in Supplementary Figure 1.

## DISCUSSION

Using a cross-sectional study of reproductive histories from 1,668 women, we have detailed our findings on risk for FXPOI and associated traits, including decreased AAM, stratified by high-resolution repeat size groups to provide more personalized risk information for women with a premutation. Using multiple statistical methods, we found consistent results with previous studies:<sup>10,22–26</sup> women with midrange premutation repeats were at the highest risk for FXPOI and earlier AAM. Because of this large sample, we were able to garner additional risk information using more detailed assessments. In general, women with a repeat size of 70–120 experience a 5-year reduction in AAM—a difference of 5 years in this window of fertility has important implications for family planning and overall health. Earlier onset for ovarian insufficiency has a strong negative impact on a woman's overall reproductive, physical, and mental health. Beyond the significant consequences of infertility and not being able to have children, women with POI also have earlier onset of bone loss, cardiovascular disease, and menopausal symptoms of hot flashes and night sweats, which can severely impact quality of life. Significantly, women with 85–89 repeats have the highest risk for ovarian insufficiency: they had an average AAM ~10 years earlier than the <45 repeats group. This compares to women in the surrounding

repeat intervals of 80–84 repeats and 90–94 repeats who had a 7-year reduction, and, moving out an additional repeat interval, women with 75–79 or with 95–99 repeats who had an average AAM that was 5 years earlier than the referent group (Supplementary Table 3). The percentage of women reporting POI (AAM <40) or early menopause (AAM <45) shows a similar pattern (Supplementary Table 3). Consistent with our previous work,<sup>10</sup> we were also able to detect altered cycling characteristics and fertility traits among women with the midrange of repeats during their last year of natural cycling. Understanding why this narrow peak of repeat sizes imposes such a high risk should provide insight into the etiological mechanism.

Importantly, women with 55–64 repeats did not show an increased risk in any of the models that were run, and in fact, there were not any women in this repeat range that reported their AAM as less than age 45 (Supplementary Table 3). To ensure the lack of significance was not related to the reduced sample size in these groups, we combined the two categories for a total of 88 women, and the findings were consistent: women with 55 to 64 CGG repeats do not show an increased risk for an earlier AAM or FXPOI compared to the referent group. Further, women within this repeat size range did not report an increase in altered cycle traits or fertility issues. This finding has important implications for women who are identified as having a premutation allele through population screening, as the majority of these women fall within this smaller premutation repeat size range.<sup>14,19–21,35</sup>

Of note, the repeat group with 45–54 repeats, those with intermediate alleles, did not have any significantly different findings compared to the <45 repeat referent group (Table 2, Figs. 1–3). Previous studies on this repeat range have reported conflicting results;<sup>36–38</sup> however, in our population we do not see evidence for an earlier age at menopause in the 45–54 repeat group.

Based on the work by Mailick et al.,<sup>34</sup> we hypothesized that presence of a PM/FM mosaic pattern in women may explain why women with higher premutation repeats have a decreased risk for FXPOI. Unfortunately, our sample size of women with PM/FM was not sufficient to test this hypothesis, as only two of the seven PM/FM mosaic women had an assigned AAM. Both women with PM/FM mosaicism had AAM over age 50; however, this was similar to the others in the same high premutation range (>120 repeats) (Supplemental Fig. 1). A larger cohort or a collaborative effort to target women with a PM/FM mosaic profile would be needed to test this hypothesis.

The primary limitation of this work is that the data are largely based on self-report through a structured questionnaire; although, for a subset of the data, a clinician (H.S.H.) conducted qualitative interviews. Further, women completed interviews at different time points in their reproductive lifespan: some women were still cycling while others had gone through menopause years before they completed the interview. In addition, women who have experienced POI or infertility may have a greater motivation to participate in research. Importantly, these limitations of the data apply to all repeat size groups.

In conclusion, we have used our cohort of women to determine more accurate risk guidelines for FXPOI based on repeat size groups. An important next step for this work is to ensure this information is available to women with a premutation. Providing more specific risk rates for FXPOI for women who are identified through screening studies has the potential to lessen the psychological burden that comes with this diagnosis. Qualitative data from women with a premutation have identified a desire for women to understand their risk for FXPOI.<sup>14,15,17,39</sup> A lack of educational materials and clinician training might be contributing to the difficulty women experience when they receive a FXPOI diagnosis.<sup>13,39,40</sup> In many cases, a genetic counselor is the essential health-care professional to work with these families from precounseling to diagnosis to postcounseling.<sup>16</sup> Thus, for our next steps, we will work with genetic counselors to study the best

methods for distributing this more accurate FXPOI risk information to families as well as health-care professionals.

## DATA AVAILABILITY

All data described here will be made available upon request from the corresponding author.

Received: 10 February 2021; Revised: 1 April 2021; Accepted: 1 April 2021;

Published online: 29 April 2021

## REFERENCES

- Sutcliffe, J. S. et al. DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum. Mol. Genet.* **1**, 397–400 (1992).
- Greenblatt, E. J. & Spradling, A. C. Fragile X mental retardation 1 gene enhances the translation of large autism-related proteins. *Science*. **361**, 709–712 (2018).
- Ashley, C. T. Jr, Wilkinson, K. D., Reines, D. & Warren, S. T. FMR1 protein: conserved RNP family domains and selective RNA binding. *Science*. **262**, 563–566 (1993).
- Cronister, A. et al. Heterozygous fragile X female: historical, physical, cognitive, and cytogenetic features. *Am. J. Med. Genet.* **38**, 269–274 (1991).
- Hagerman, P. J. & Hagerman, R. J. Fragile X-associated tremor/ataxia syndrome (FXTAS). *Ment. Retard. Dev. Disabil. Res. Rev.* **10**, 25–30 (2004).
- Jacquemont, S. et al. Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. *Am. J. Hum. Genet.* **72**, 869–878 (2003).
- Berry-Kravis, E. et al. Fragile X-associated tremor/ataxia syndrome: clinical features, genetics, and testing guidelines. *Mov. Disord.* **22**, 2018–2030 (2007). quiz 2140.
- Hagerman, R. J. & Hagerman, P. fragile X-associated tremor/ataxia syndrome - features, mechanisms and management. *Nat. Rev. Neurol.* **12**, 403–412 (2016).
- Sherman, S. L. Premature ovarian failure in the fragile X syndrome. *Am. J. Med. Genet.* **97**, 189–194 (2000).
- Allen, E. G. et al. Examination of reproductive aging milestones among women who carry the FMR1 premutation. *Hum. Reprod.* **22**, 2142–2152 (2007).
- Hagerman, R. J. et al. Fragile X-associated neuropsychiatric disorders (FXAND). *Front. Psychiatry.* **9**, 564 (2018).
- Seltzer, M. M. et al. Prevalence of CGG expansions of the FMR1 gene in a US population-based sample. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **159B**, 589–597 (2012).
- Hipp, H. S., Charen, K. H., Spencer, J. B., Allen, E. G. & Sherman, S. L. Reproductive and gynecologic care of women with fragile X primary ovarian insufficiency (FXPOI). *Menopause.* **23**, 993–999 (2016).
- Anido, A., Carlson, L. M. & Sherman, S. L. Attitudes toward fragile X mutation carrier testing from women identified in a general population survey. *J. Genet. Couns.* **16**, 97–104 (2007).
- Archibald, A. D. et al. "It's about having the choice": stakeholder perceptions of population-based genetic carrier screening for fragile X syndrome. *Am. J. Med. Genet. A.* **161A**, 48–58 (2013).
- Archibald, A. D. et al. "It gives them more options": preferences for preconception genetic carrier screening for fragile X syndrome in primary healthcare. *J. Community Genet.* **7**, 159–171 (2016).
- Archibald, A. D. et al. "It's something I need to consider": decisions about carrier screening for fragile X syndrome in a population of non-pregnant women. *Am. J. Med. Genet. A.* **149A**, 2731–2738 (2009).
- Hantash, F. M. et al. FMR1 premutation carrier frequency in patients undergoing routine population-based carrier screening: insights into the prevalence of fragile X syndrome, fragile X-associated tremor/ataxia syndrome, and fragile X-associated primary ovarian insufficiency in the United States. *Genet. Med.* **13**, 39–45 (2011).
- Cronister, A., Teicher, J., Rohlf, E. M., Donnenfeld, A. & Hallam, S. Prevalence and instability of fragile X alleles: implications for offering fragile X prenatal diagnosis. *Obstet. Gynecol.* **111**, 596–601 (2008).
- Tassone, F. et al. FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States. *Genome Med.* **4**, 100 (2012).
- Lee, S. et al. Validation of fragile X screening in the newborn population using a fit-for-purpose FMR1 PCR assay system. *J. Mol. Diagn.* **22**, 346–354 (2020).
- Sullivan, A. K. et al. Association of FMR1 repeat size with ovarian dysfunction. *Hum. Reprod.* **20**, 402–412 (2005).
- Ennis, S., Ward, D. & Murray, A. Nonlinear association between CGG repeat number and age of menopause in FMR1 premutation carriers. *Eur. J. Hum. Genet.* **14**, 253–255 (2006).
- Tejada, M. I. et al. Analysis of the molecular parameters that could predict the risk of manifesting premature ovarian failure in female premutation carriers of fragile X syndrome. *Menopause.* **15**, 945–949 (2008).
- Spath, M. A. et al. Predictors and risk model development for menopausal age in fragile X premutation carriers. *Genet. Med.* **13**, 643–650 (2011).
- Mailick, M. R., Hong, J., Greenberg, J., Smith, L. & Sherman, S. Curvilinear association of CGG repeats and age at menopause in women with FMR1 premutation expansions. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **165B**, 705–711 (2014).
- Allen, E. G. et al. Examination of the effect of the polymorphic CGG repeat in the FMR1 gene on cognitive performance. *Behav. Genet.* **35**, 435–445 (2005).
- Nelson, L. M., Covington, S. N. & Rebar, R. W. An update: spontaneous premature ovarian failure is not an early menopause. *Fertil. Steril.* **83**, 1327–1332 (2005).
- Meadows, K. L. et al. Survey of the fragile X syndrome and the fragile X E syndrome in a special education needs population. *Am. J. Med. Genet.* **64**, 428–433 (1996).
- Brown, W. T. et al. Rapid fragile X carrier screening and prenatal diagnosis using a nonradioactive PCR test. *JAMA.* **270**, 1569–1575 (1993).
- Fu, Y. H. et al. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell.* **67**, 1047–1058 (1991).
- Cooper, G. S., Sandler, D. P. & Bohlig, M. Active and passive smoking and the occurrence of natural menopause. *Epidemiology.* **10**, 771–773 (1999).
- Akahoshi, M. et al. The effects of body mass index on age at menopause. *Int. J. Obes. Relat. Metab. Disord.* **26**, 961–968 (2002).
- Mailick, M. R. et al. Health profiles of mosaic versus non-mosaic FMR1 premutation carrier mothers of children with fragile X syndrome. *Front. Genet.* **9**, 173 (2018).
- Hantash, F. M. et al. Qualitative assessment of FMR1 (CGG)n triplet repeat status in normal, intermediate, premutation, full mutation, and mosaic carriers in both sexes: implications for fragile X syndrome carrier and newborn screening. *Genet. Med.* **12**, 162–173 (2010).
- Bretherick, K. L., Fluker, M. R. & Robinson, W. P. FMR1 repeat sizes in the gray zone and high end of the normal range are associated with premature ovarian failure. *Hum. Genet.* **117**, 376–382 (2005).
- Bodega, B. et al. Influence of intermediate and uninterrupted FMR1 CGG expansions in premature ovarian failure manifestation. *Hum. Reprod.* **21**, 952–957 (2006).
- Gleicher, N., Weghofer, A. & Barad, D. H. A pilot study of premature ovarian senescence: I. Correlation of triple CGG repeats on the FMR1 gene to ovarian reserve parameters FSH and anti-Mullerian hormone. *Fertil. Steril.* **91**, 1700–1706 (2009).
- Espinell, W., Charen, K., Huddleston, L., Visootsak, J. & Sherman, S. Improving health education for women who carry an FMR1 premutation. *J. Genet. Couns.* **25**, 228–238 (2016).
- Smolich, L., Charen, K. & Sherman, S. L. Health knowledge of women with a fragile X premutation: improving understanding with targeted educational material. *J. Genet. Couns.* **29**, 983–991 (2020).

## ACKNOWLEDGEMENTS

First and foremost, we thank the families who participated in this project. Without their contribution and encouragement, this work could not be done. We would also like to thank the Fragile X Research Participant Registry of the Carolina Institute for Developmental Disabilities (P50 HD103573) at the University of North Carolina at Chapel Hill who helped with recruitment. Lastly, we want to thank the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) and the National Institute of Neurological Disorders and Stroke (NINDS) for supporting our National Fragile X Center (U54NS091859 and P50HD104463) in which this work was conducted. This study was supported in part by the Emory Integrated Genomics Core (EIGC), which is subsidized by the Emory University School of Medicine and is one of the Emory Integrated Core Facilities. This study also utilized a REDCap database supported by Emory's Library and Information Technology Services (UL1TR000424).

## AUTHOR CONTRIBUTIONS

E.G.A., S.L.S., and H.S.H. were involved in the conceptualization of the project. Formal analyses and writing of the original draft was performed by E.G.A. Data curation was performed by K.C., L.S., and H.S.H. Laboratory analyses were performed by A.A., W.H., S.L.N., A.G., and N.T. Writing and methodology were completed by E.G.A., K.C., H.S.H., L.S., S.L.N., B.M., K.E.S., and S.L.S. Funding acquisition was done by E.G.A. and S.L.S.

## ETHICS DECLARATION

Emory University Institutional Review Board approved protocols and consent forms, and informed consent was obtained on all participants.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41436-021-01177-y>.

**Correspondence** and requests for materials should be addressed to E.G.A.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021