Supplementary Online Content

eAppendix

This supplementary material has been provided by the authors to give readers additional information about their work.
We present 3 distinct models (discussed and presented as Tables 2, 3, and 4) to address specific hypotheses regarding the prevalence of high CES-D scores (CESD>16 vs < 16) as a function of time. All models are versions of a generalized linear mixed effects regression model. For all models, the outcome is the prevalence of high CESD for each participant i, at each time t. We denote this as:

\[ D_{it} = \begin{cases} 
1 & \text{if } \text{CESD} \geq 16 \\
0 & \text{otherwise.} 
\end{cases} \]

The goal of the model presented in Table 2 is to compare the prevalence of high CESD scores relative to FMP, time 0, for each year, \(-10 < t < 8\). For each participant, \(i\), for each observed time the model is:

\[
\log[\Pr(D_{it} = 1)/(1 - \Pr(D_{it} = 1))] = \beta_0 + \beta_1 t_i + \epsilon_i 
\]

Regression coefficients were estimated separately for women with and without a history of depression. Models were fit using version 9.3 of SAS, Proc Genmod assuming that time in year was a class variable. For this procedure we assume that the repeated measurements of CESD over time for each woman are correlated, and estimated standard deviations needed for the confidence intervals for the Odds Ratio estimates using robust methods (GEE)\(^1\). Estimates of the regression parameters are then transformed as \(\exp(\beta_i)\) to represent the odds ratio for prevalence of high CESD in year \(t\) compared to \(t=0,\) the FMP.

Next, the goal of the model presented as Table 3 was to evaluate the within-woman change over time in prevalence of high CES-D scores and ascertain whether covariates influences the rate of change over time. SAS Proc NLMIXED was used to fit these models, and the appropriate number of variance components (i.e., random slope and random intercept) were evaluated by comparing Akaike information criteria (AIC).\(^2\) The final model included random intercept and slope for time. For each covariate, \(X\), the following regression model was fit:

\[
\logit[\Pr(D_{it} = 1)] = \beta_0 + \beta_1 t_i + \beta_2 X_i + \epsilon_i 
\]

where \(\beta_0\) and \(\beta_1\) are the random intercept and slope. In table 3 as above, we present transformed regression coefficients \(\exp(\beta_i)\) which are odds ratios of interest.

Models presented in Tables 4 and 5 are restricted to postmenopausal times \(0 < t < 12\). For these 2 models we wish to evaluate the impact of covariates defined prior to FMP on depressive symptoms post FMP. Similar to the model presented in Table 2, SAS Proc GENMOD was used to estimate:

\[
\logit[\Pr(D_{it} = 1)] = \beta_0 + \beta_1 t_i + \beta_2 hxDep + \beta_3 PreMDep + \beta_4 TranDep, 
\]

presented in Table 4, where hxDep is an indicator variable representing women who reported a
depression history at baseline or prior to study entry. Similarly, PreMDep was used to indicate the subset of women who reported their first incidence of CESD>16 while pre-menopausal, and TranDep indicates those whose first episode was during the menopausal transition. For this model \( \exp(\beta_1) \) is the odds ratio comparing prevalence of high CESD for women with > 2 years post menopause compared to that for times <2. Similarly, \( \exp(\beta_2) \) is the estimated odds of high CESD symptoms for women with a history of depression prior to study enrollment, compared to the odds of high CESD symptoms in women who experienced no depressive symptoms prior to FMP.

The results presented in Table 5 extend the Table 4 model to include hormone summaries as an additional covariate. Define a subset of time prior to FMP where the rate of change in the hormones is approximately linear (i.e. linear range). These linear ranges were identified by inspection of the mean slope for the log hormone values for each year from baseline to FMP; linear ranges were 3 years for estradiol, 4 years for FSH, and 6 years for inhibin b.

Our identified ranges were consistent with those previously identified by Zheng and colleagues\(^3\) who utilized a piecewise linear mixed effects model to identify the timing of changes in hormone trajectories over time in the SWAN (Study of Women Across the Nation). Below is a graph of individual level data for log transformed FSH for 50 participants, along with the population average curve which included linear and quadratic terms for time, along with random effects. Due to space constraints this graph has not been incorporated into the manuscript.

Given these specific linear ranges for time, the rate of hormone change prior to menopause (the slope) was calculated for each hormone for each woman. This was done by using the data for
each woman, separately (i.e. no averaging across women, as with a mixed model, was conducted), therefore, each estimate is uninfluenced by the data from other study subjects as follows:

\[ \log(Y_{it}) = \alpha + \alpha_i t + \epsilon_i. \]

The estimated slope, \( \alpha_i \), for the rate of log transformed hormone per year was then included in the model for high post-menopausal CESD scores

\[
\text{logit}[\text{Pr}(D_{it} = 1)] = \beta_0 + \beta_1 (t_i - 2) + \beta_2 \text{hxDep} + \beta_3 \alpha_i.
\]

The odds ratio for FSH as presented in Table 5 is computed as \( \exp(\beta_3 \times 0.33) \), and represents the change in odds of post-menopausal odds high CESD for a 1 SD higher slope for \( \log(\text{FSH}) \).

**References**

