**Intra- and inter-cycle variability of anti-Müllerian hormone (AMH) levels in healthy women during non-consecutive menstrual cycles: the BICYCLE study**

Marieke Biniasch, Ruediger Paul Laubender, Martin Hund, Katharina Buck, Christian De Geyter

**Supplementary Materials and Methods**

**Linear mixed-effects model**

Since a single menstrual cycle of a woman is a periodic phenomenon, it is reasonable to model the menstrual cycle by a periodic model which allows non-linear trends to be captured. This model is given by

$$log\left(y\_{ijt}\right)= \left(β\_{0}+ b\_{0ij}\right)+\left(β\_{1}+ b\_{1ij}\right)\cos(\left(2π\frac{t}{28}\right))+\left(β\_{2}+ b\_{2ij}\right)\sin(\left(2π\frac{t}{28}\right))+\left(β\_{3}+ b\_{3ij}\right)\cos(\left(2π\frac{t}{14}\right))+\left(β\_{4}+ b\_{4ij}\right)\sin(\left(2π\frac{t}{14}\right))+e\_{ijt}$$

where $y\_{ijt}$ denotes the AMH result measured at University Hospital Basel for woman $i$ of cycle $j$ at standardized time $t$, and $log\left(y\_{ijt}\right)$ denotes the natural logarithm of the corresponding AMH result. The fixed effects in this model are the regression coefficients $β\_{0}, β\_{1}, β\_{2}, β\_{3}, β\_{4}$ and the random effects are the coefficients $b\_{0ij}, b\_{1ij}, b\_{2ij},b\_{3ij},b\_{4ij}$, which are linearly related. Further, the standardized time $t$ is transformed by sine and cosine functions with $π$ denoting a mathematical constant (approx. 3.14159); the frequency $\frac{t}{28}$ means that the series of AMH results of a woman $i$ within her cycle $j$ makes a sine (cosine) cycle every 28th standardized day and the frequency $\frac{t}{14}$ makes a sine (cosine) cycle every 14th standardized day. The combination of these sine and cosine transformations (sinusoids) describe a Fourier series of second degree and allow the modelling of a non-linear periodic phenomenon like the variation of (logarithmic) AMH results within a menstrual cycle. The choice of a Fourier series of second degree gives enough flexibility for modelling the different AMH results series within the cycle of a woman. For more details about periodic regression, see Shumway and Stoffer 2017, introductory Chapter 4.1. The term $\left(β\_{0}+ b\_{0ij}\right)$ describes the mesor (on a logarithmic scale) for cycle $j$ of woman $i$ (shown in Figure 3). The amplitude of cycle $j$ of woman $i$ (shown in Figure 3) can be derived by numerically finding the maximum and minimum of the cycle using the regression model based on the estimated fixed and random regression coefficients of cycle $j$ of woman $i$. Further, it is assumed that the random coefficients follow a multivariate normal distribution with zero mean estimated by restricted maximum likelihood techniques. The residuals are denoted by $e\_{ijt}$ and are assumed to follow a normal distribution with mean 0 and variance $σ\_{e}^{2}$. The residuals quantify variation not explained by the sine and cosine transformed standardized time variables, like biological influences or measurement uncertainty of the AMH results (analytical variability). The variance $σ\_{e}^{2}$ can be transformed back to the original AMH unit in form of a coefficient of variation (CV) in percentage, which is given by

$$\sqrt{exp\left(σ\_{e}^{2}-1\right)}×100$$

In order to separate the biological variability from the analytical variability, the duplicate (remeasurements) and triplicate (original measurement and duplicate remeasurements) AMH results can be used and the following mixed effects model can be fitted:

$$log\left(y\_{ijtr}\right)= \left(β\_{0}+ b\_{0ij}\right)+\left(β\_{1}+ b\_{1ij}\right)\cos(\left(2π\frac{t}{28}\right))+\left(β\_{2}+ b\_{2ij}\right)\sin(\left(2π\frac{t}{28}\right))+\left(β\_{3}+ b\_{3ij}\right)\cos(\left(2π\frac{t}{14}\right))+\left(β\_{4}+ b\_{4ij}\right)\sin(\left(2π\frac{t}{14}\right))+ε\_{ijt}+ϵ\_{ijtr} $$

where $r$ denotes the replicate. Note that the residual variability $e\_{ijt}$ of the above model is now split into $ε\_{ijt}$, which estimates the unexplained biological variability (variability not explained by the non-linearly transformed variables), and $ϵ\_{ijtr}$, which denotes the residuals of the replicates, thus quantifying the analytical variability. It is assumed that both $ε\_{ijt}$ and $ϵ\_{ijtr}$ follow a normal distribution each with mean 0 and variances $σ\_{ε}^{2}$ and $σ\_{ϵ}^{2}$. By analogy, the variances can be transformed back to the original AMH unit in form of a CV, which are given by

$$CV\_{ε}=\sqrt{exp\left(σ\_{ε}^{2}-1\right)}×100$$

and

$$CV\_{ϵ}=\sqrt{exp\left(σ\_{ϵ}^{2}-1\right)}×100$$

In the case of using triplicates, the coefficient of variation $CV\_{ϵ}$ represents reproducibility since several days, instrument types, and sites are involved as sources of analytical variability. In the case of using duplicates from the same laboratory, the coefficient of variation $CV\_{ϵ}$ represents the repeatability estimate since two sequential determinations of AMH results are involved as a source of analytical variability.

**Reference**

Shumway RH, Stoffer DS. Time Series Analysis and Its Applications (Fourth Edition). New York; Springer International Publishing; 2017.