

Optimal sampling strategies for the baseline characterization of testosterone in hypogonadal men

Olivier Barrière, Ph.D¹, Mario González-Sales, Ph.D^{1,2}, Julie Desrochers, M.Sc¹, Fahima Nekka, Ph.D,² Pierre-Olivier Tremblay, M.Sc¹, Mario Tanguay, Ph.D¹
¹inVentiv Health Clinical, Montréal, Canada ²Université de Montréal, Montréal, Canada

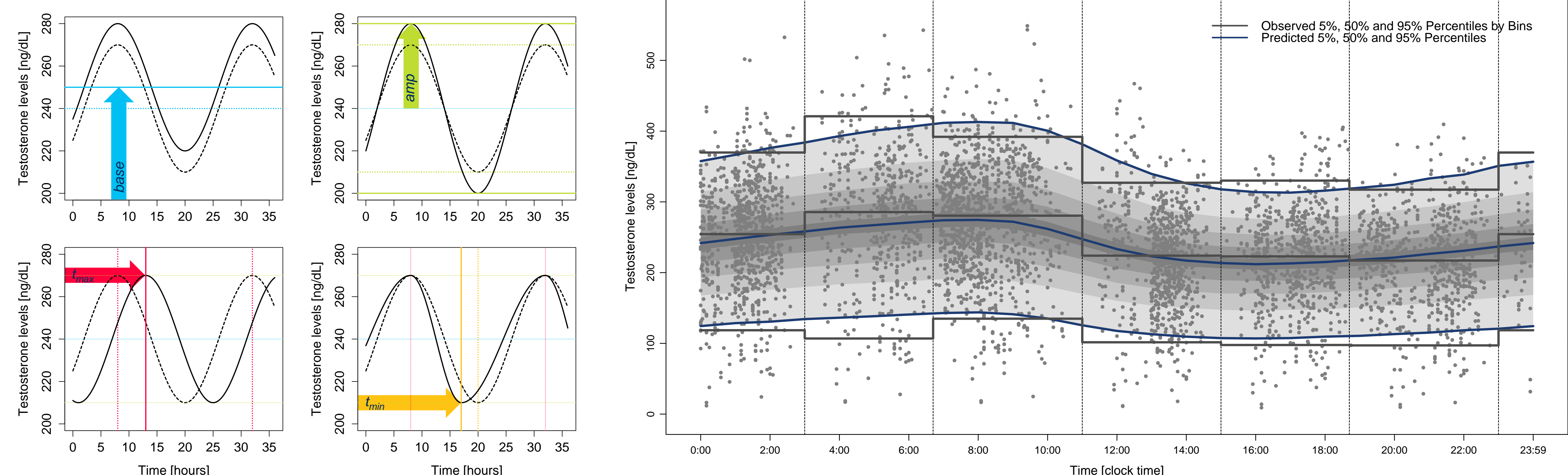
Purpose

The FDA testosterone bioequivalence guidances recommend that the serum concentrations of testosterone should be corrected for baseline endogenous levels by subtracting the mean pre-dose baseline value. The timepoints for the baseline computation are specified in 8 out of the 10 testosterone related guidances currently available on the FDA website: they consist of *at least three pre-dose values, e.g. -1.0, -0.5, and 0 hours* in 7 of them, and they were recently updated to *-12 and 0 hours before dosing* in the most recent guidance. The former approach does not take into account the circadian rhythm, and can over or underestimate the baseline when samples are taken close to the peak or the nadir. The latter approach is more balanced, so it is less prone to bias but the reduced number of points used for the baseline computation can increase its variance.

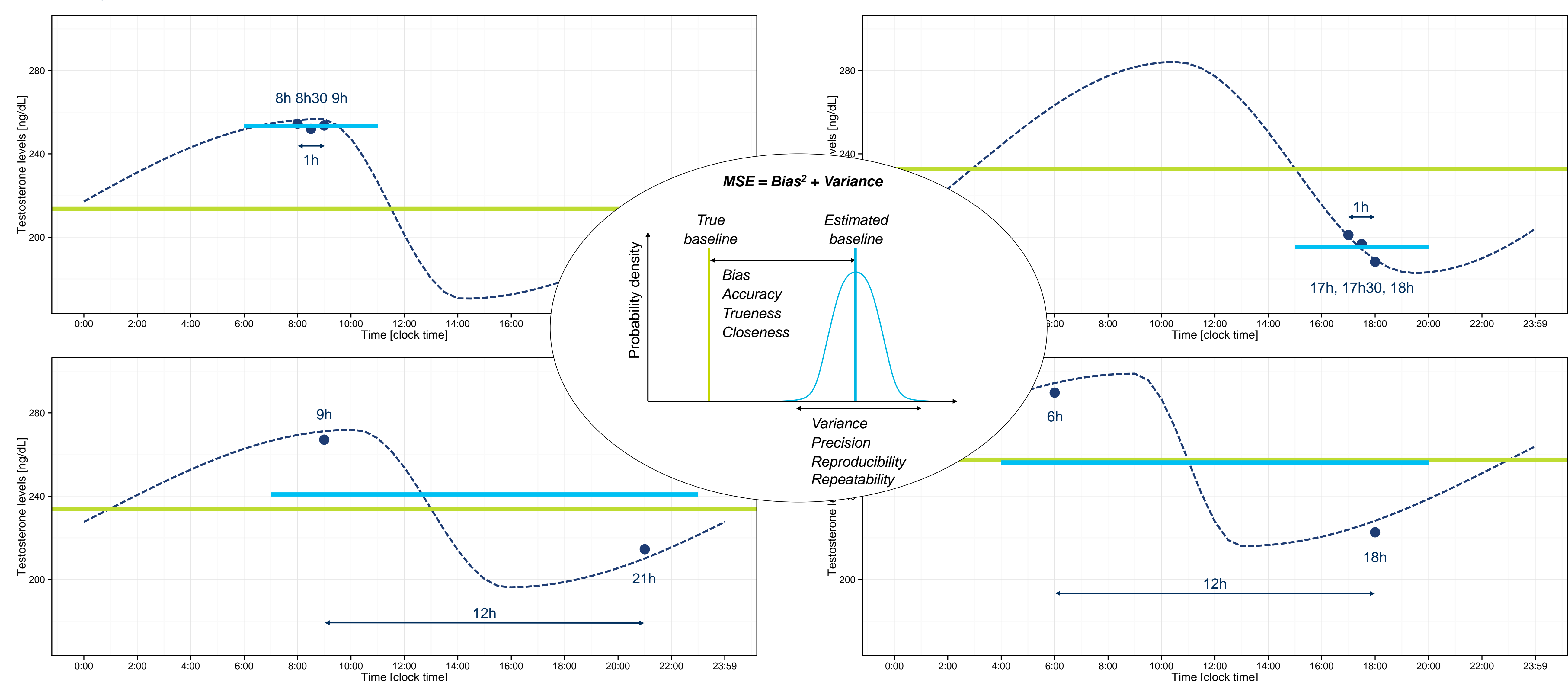
The objective of this work is to formally test these sampling strategies and provide optimal sampling times accounting for the circadian behavior for the estimation of the testosterone baseline, here defined as the average concentration over 24h.

Methods

Based on 859 pre-dose PK profiles from 239 hypogonadal men (4556 observations), a kinetic model describing the circadian rhythm of testosterone in hypogonadal men was previously built. The circadian rhythm is usually described by a standard cosine function, which implies that the increasing and the decreasing behaviors are symmetric. A stretched cosine function parameterized by *base*, *amp*, *t_{max}*, and *t_{min}*, within a 24 hours cycle, was here defined. It allows the time between *t_{max}* and *t_{min}* to be different from 12 hours in order to model the observed slow increase and fast decrease.

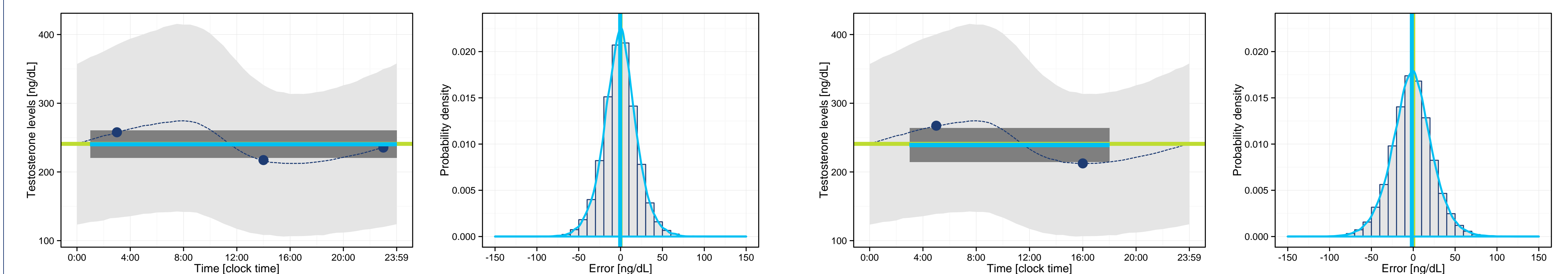


Testosterone profiles were then simulated using this model to determine the optimal sampling times for the baseline estimation with the best trade-off between bias and variance, by minimizing the mean squared error (MSE) of the *base* parameter. MSE is the sum of the squared bias and the variance, and thus incorporates both components at the same level.

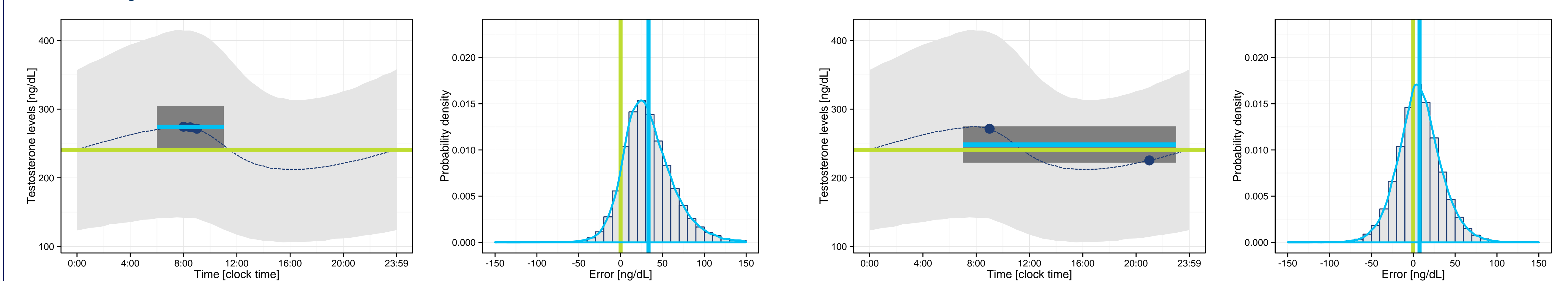


Results

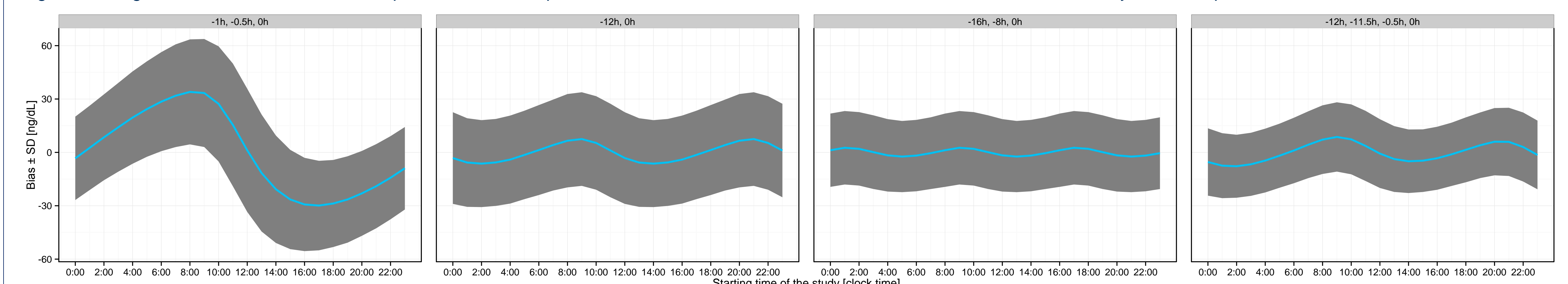
Without any constraints, the best estimator based on 3 timepoints (3:00AM, 2:00PM and 11:00PM) has a bias lower than 1 ng/dL with a standard deviation of 20 ng/dL. The best estimator based on 2 timepoints (5:00AM and 4:00PM) has a bias lower than 1.5 ng/dL with a standard deviation of 24.7 ng/dL.



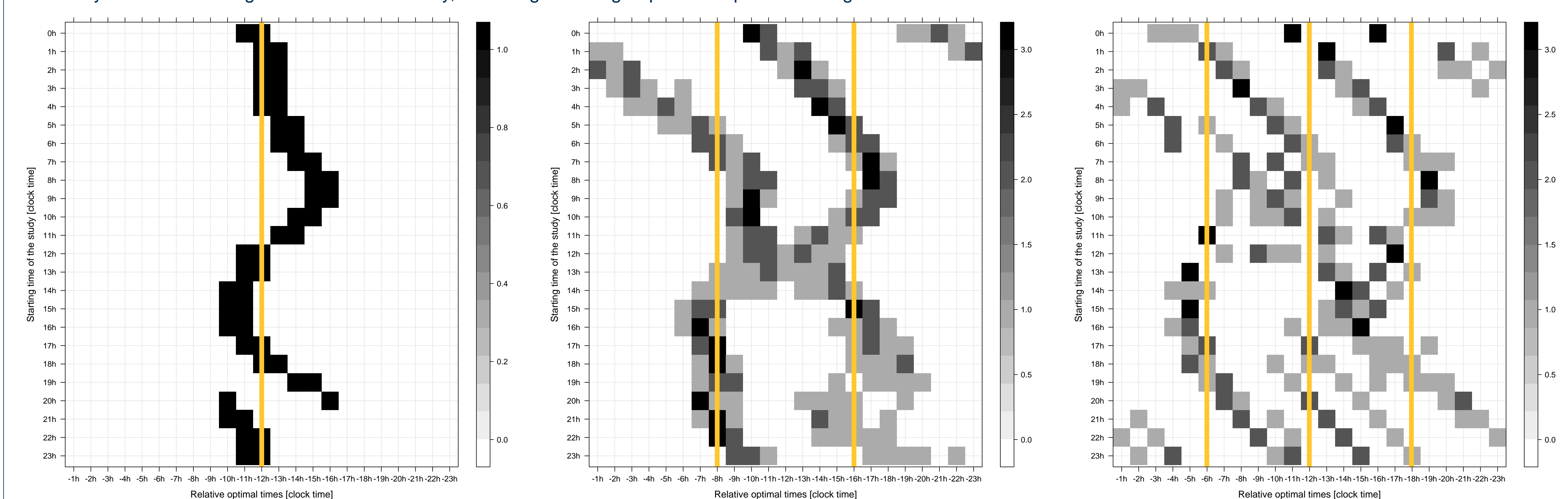
Applying the former FDA design to a study that starts at 9:00AM can lead to an over-prediction of the baseline with a bias of 33.5 ± 30.3 ng/dL. The latter FDA design has a bias of 7.6 ± 26.4 ng/dL.



The pattern *-16h, -8h, 0h* adequately spans the 24 hours of a day and provide a bias lower than 2.5 ng/dL whatever the starting time of the study, with a standard deviation around 20 ng/dL. If the goal is to limit the confinement period to 12h, the pattern *-12h, -11.5h, -0.5h, 0h* has a standard deviation decreased by 27% compared to *-12h, 0h* with a similar bias.



The bias can always be reduced to almost zero with an optimal choice of sampling times, but the standard deviation cannot be smaller than 24 ng/dL, 20 ng/dL and 16 ng/dL, when 2, 3 and 4 samples are used, respectively. Assuming that a sample is always taken just before the dose, the times of the other samples can be optimized given the starting time of the study. Instead selecting the best estimator only, collecting a small group of the top estimators gives a more flexible and robust view.



Conclusions

Sampling strategies for the baseline estimation of testosterone in hypogonadal males were extensively tested in terms of bias and variance. The FDA recommendations were compared to the optimal timepoints and improved patterns were suggested.