Supplementary Text 3: Data preprocessing of dihydrotestosterone (DHT) and total testosterone (TT) measurements

The steps listed below were undertaken to normalize the measurements to their plate effects as well as NA imputation and NA removal, where necessary. The steps are ordered in the same way they were applied to the data set.

NA removal and imputation

NA detection was conducted and removal was performed if more than 40% of steroids of all samples were labeled "NA" or "0" in the raw data set. For the raw data set, only estrone was detected to fulfill this criteria and was therefore subsequently removed from the data set. All other steroids were below 40% NA. These remaining NA or zero values were imputed by a minimum values replacement algorithm, which is implemented as follows:

The minimum value - which could be measured for each metabolite from all plates combined - is gathered. This value is not allowed to be equal to zero as the export from the MetIDQ-System, in which the raw measurement data is translated to quantitative data, unfortunately uses "NA" and "0" synonymously. In order to mitigate the effects which could be incurred by using these minimum values directly as imputation values, further steps are to be undertaken.

1. The minimum values are not used directly: they are divided by √2 to emulate their real concentrations being well below the minimally detected ones.

2. These newly calculated values, called “replacers”, are then permuted randomly in a range of 0.75* replacer up to 1.25* replacer in order to mitigate the statistical effects of repeating the same number over and over again.

Plate normalization

For each plate and each metabolite, plate specific mean values ("plate means") are calculated. For this, the concentrations of each metabolite of the QC-2 samples (N = 5) are used. Then the plate means are used to calculate an overall mean of all plates. These steps have to be performed for all metabolites (x_1, x_2, x_3, ..., x_n) on all plates (n_1, n_2, n_3, ..., n_j) of the data set.

$$Platemean[metabolite(x)] = \frac{\sum C\ [metabolite(x)]}{N}$$

$$\begin{gathered} where: \\ N = numberofreferencesamples \\ C = concentration \end{gathered}$$

These plate means are then used to calculate the overall "means of all plates". This has to be calculated, again, for all metabolites(x_1 , x_2 ,x_3 ... x_n) for each plate in the data set.

$$Overallmean[X] = \frac{\sum Means\ [X]}{N}$$

$$\begin{gathered} where: \\ X = metabolite(x)ofPlates\big(n\ldots n_j\big) \\ N = numberofPlates\big(n\ldots n_j\big) \end{gathered}$$

The plate factors are calculated for each plates separately, e.g. for plate 1 as follows:

$$Factor[Y] = \frac{Overallmean[X]}{Mean[Y]}$$

$$\begin{gathered} where: \\ X = metabolite(x)ofPlates\big(n\ldots n_j\big) \\ Y = Plate(n)[metabolite(x)] \end{gathered}$$

In the final step of normalization, each and all metabolite concentrations are multiplied with their corresponding plate factors, e.g. metabolite concentrations of each sample on plate 1 are multiplied with plate factor 1.

$$Normalized[Y] = Factor[Y] * Concentration[Y]$$

$$where:$$
$$Y = Plate(n)[metabolite(x)]$$

In a last step, the duplicates from Batch 1 of the measurements are removed by averaging them.

Calibration of batch effects of dihydrotestosterone and total testosterone

Regarding TT and DHT measurements, serum samples from all KORA F4 participants were initially measured between January and November 2013 (batch 1). Due to measurement problems, measurements had to be repeated for 980 serum samples with the same method described above between July 2017 and March 2018 (batch 2). To avoid batch effects, we used 175 duplicate measurements from the same participants of batch 1 and 2 to develop calibration formulas. Intercept and the slope of Passing-Bablok regressions were used for calibration of batch 2 measures with the batch 1 measures.