

Combining ADRD risk factors of brain, lifestyle, and genetics into a multimodal signature for menopause classification



Elizabeth Haddad^{1*}, Ravi R. Bhatt¹, Claudia Barth², Talia M. Nir¹, Neda Jahanshad¹

¹Mark and Mary Stevens Neuroimaging and Informatics Institute Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

²Department of Psychiatry and Neuroscience (CCM) & Institute of Gender in Medicine, Charité - Universitätsmedizin Berlin, Berlin, Germany

*haddade@usc.edu @liz_haddad @USC_LoBeS

Keck School of
Medicine of USC

USC Stevens Neuroimaging
and Informatics Institute

Introduction

- During the transition to menopause, women often experience brain-based vasomotor symptoms, mood, sleep, and cognitive disturbances, alongside systemic metabolic and vascular consequences.
- Symptoms of menopause and brain-based vulnerabilities may indicate increased risk of neurodegeneration and ADRD later in life.
- The transitional years may be a window of opportunity for ADRD prevention, yet specific brain-based biomarkers of this transition and their lifestyle correlates have not been well-established.
- Here, in an age-matched sample of midlife women, we used multimodal data integration to identify networks of brain-imaging based, clinical, lifestyle, and genetic markers for classifying before and after the menopause transition.

Methods

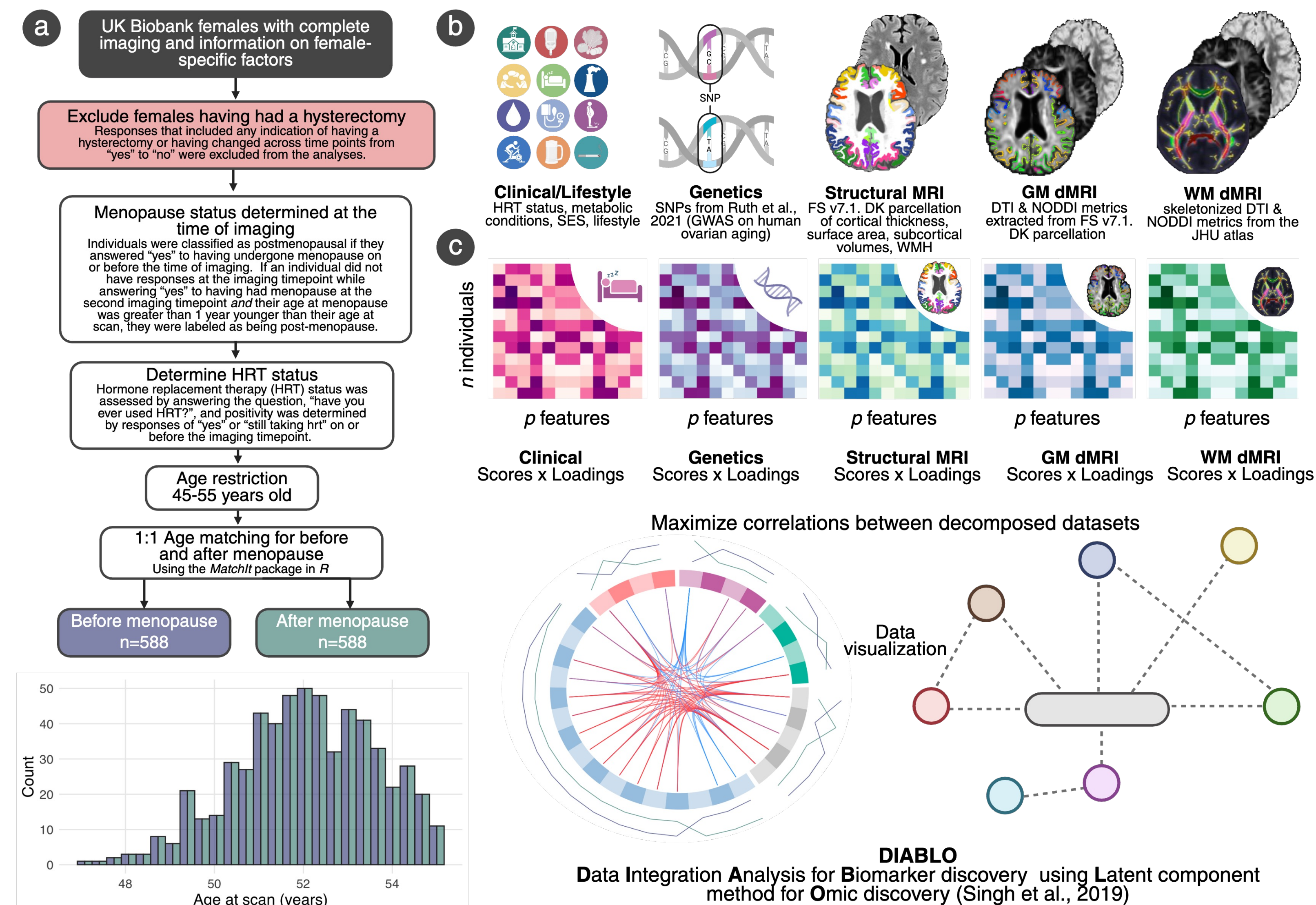


Figure 1. Methods. **a.** Flow diagram detailing subject selection process. **b.** 5 categories or "blocks" used in the analysis. **c.** DIABLO analysis schematic. Details are as follows: data were split 80%/20% using stratified sampling to ensure equal representation of pre- and postmenopausal women in both partitions. We employed DIABLO, a multi-block discriminant analysis method that identifies sparse sets of correlated variables across multiple high-dimensional datasets to predict a categorical outcome. DIABLO extends sparse generalized canonical correlation analysis (sgCCA) to a supervised learning framework, simultaneously maximizing between-block correlations while discriminating between outcome groups. An initial DIABLO model with five latent components was fitted without variable selection to assess the maximum achievable performance. The optimal number of components (three) was determined by evaluating global classification performance on the training data using 50 repeats of 10-fold cross-validation, with balanced error rate (BER) as the primary metric to account for class imbalance. Repeated cross-validation provides more robust performance estimates by reducing variance due to random fold assignments. The design matrix, which specifies the expected correlations between data blocks, was set to 0.1 for all pairwise block connections. This weak correlation constraint prioritizes outcome prediction over inter-block associations, allowing the model to identify discriminative features that may not be strongly correlated across modalities while still permitting biologically relevant cross-block relationships to emerge. Initial models were fit with 20 features per data block per component; features were subsequently refined by examining correlation circle plots and retaining only variables with absolute loading weights ≥ 0.7 to prioritize variables with stronger contributions to each component. Final model performance was evaluated on the independent test set using BER, along with sensitivity, specificity, positive predictive value, negative predictive value, and F1 score derived from the confusion matrix. Network analysis of selected variables was performed to visualize inter-block correlations exceeding $|r| > 0.7$.

Results

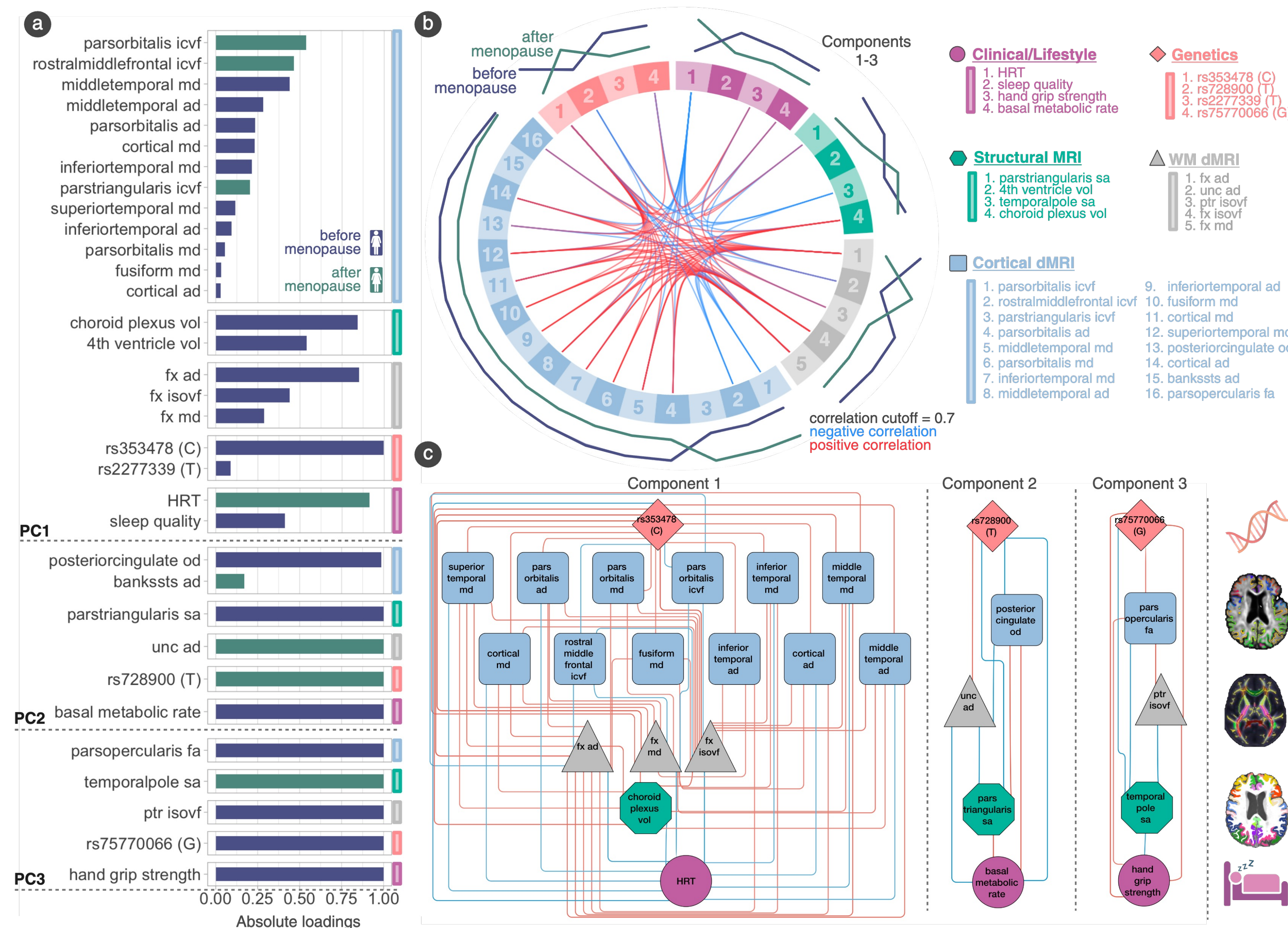


Figure 2. Results of the DIABLO model with HRT included. **a.** Absolute loadings on components. The factors with the highest loading on C1 for classifying menopause status included variants rs353478 (C), rs2277339 (T), sleep scores, cortical MD, cortical AD, choroid plexus volume, ventricular volume, and AD, MD, and ISOVF in the fornix. Higher BMR, surface area of the parstriangularis, OD of the posterior cingulate were predictive of (before) menopause and higher AD of the banks of the superior temporal sulcus, AD of the uncinate fasciculus, and frequency of rs728900 (T) was predictive of postmenopause on C2. Higher grip strength, frequency of rs75770066 (G), FA of the parsopercularis, and ISOVF of the posterior thalamic radiation were predictive of (before) menopause and higher surface area of the temporal pole was predictive of postmenopause on C3. None of the cortical thickness measures included had absolute loadings greater than 0.01. Colors indicate which group showed higher average values for each feature. **b.** Circos diagram showing factors on components 1-3 contributing to menopause transition classification. The inner circle represents positive (red) and negative (blue) correlated features with a correlation value of 0.7 and above. Outer lines represent whether the before and after menopause transition groups were higher (most outer = higher). **c.** Network diagram showing results from (b) separated by component.

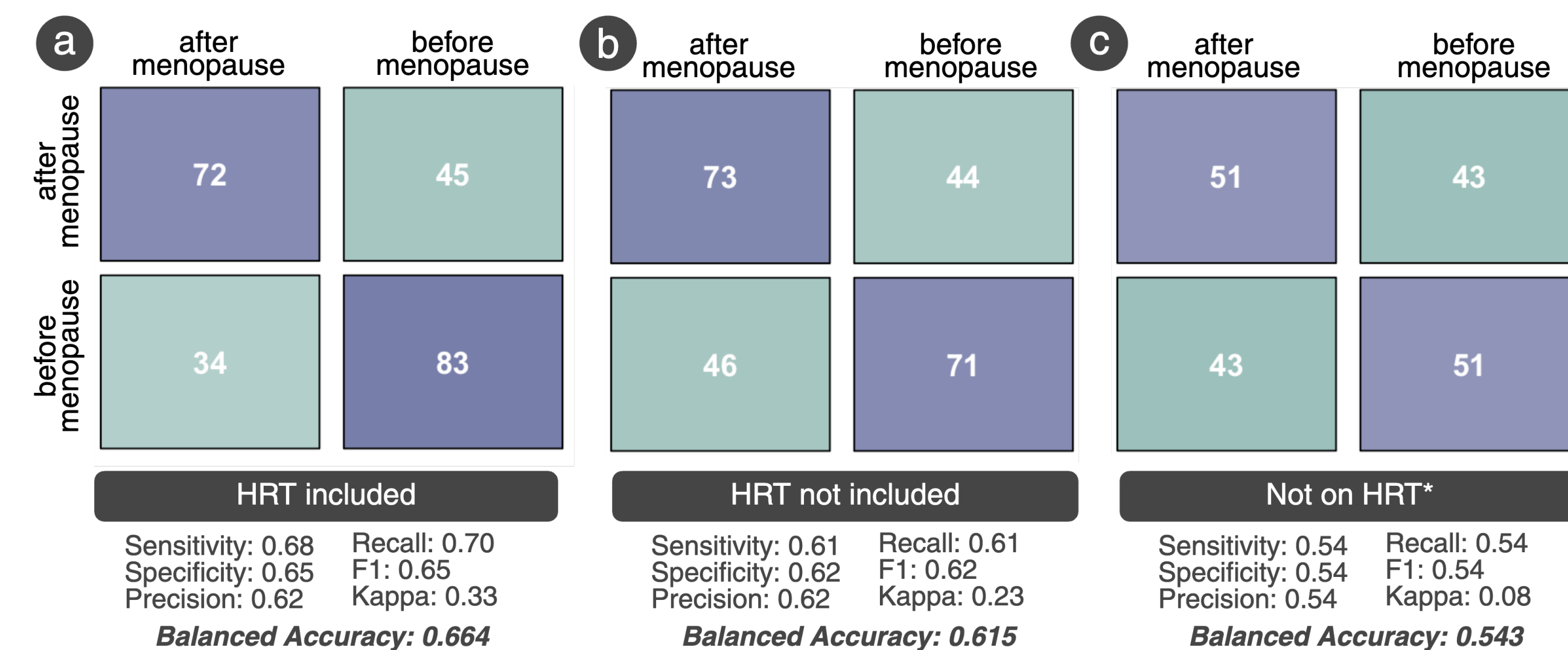


Figure 3. DIABLO classification model performances across runs with **a.)** HRT included in the model; **b.)** HRT not included in the model; **c.)** a subgroup of individuals not on HRT. "When excluding subjects who were on HRT, the remaining sample no longer had 1:1 age matching; therefore we rematched individuals not on HRT and our sample size per group was $n=473$."

Main Takeaways

- There are complex networks of factors that are altered in the menopause transition, yet prediction accuracies for classifying before and after menopause in an age-matched sample are modest.
- Identifying biological alterations and how genetic and lifestyle factors relate to their variability may identify strategies for mitigating symptoms and reducing risk for decline.

Funding

- The work is supported by Wellcome LEAP's CARE initiative and NIH grants: T32AG058507, R01AG087513, AARG-23-1149996 and S10OD032285.
- UK Biobank Resource under application number 11559.

