9. Statistics

9.1 Sample Size Calculation

Based on the sample size calculation, a minimum of 3,500 participants is required to detect statistically significant differences between the intervention groups for the primary outcomes. To account for an anticipated attrition rate of approximately 10%, a total of 4,000 participants will be recruited to ensure adequate statistical power at study completion. Eligible participants will be recruited from individuals receiving care at affiliated Primary Care Centers.

9.2. Data Management and Statistical Software

All study data will be entered into a secure, permanent, interactive database. Data entry will include verification procedures, including double data checks, to minimize transcription and input errors. Statistical analyses will be performed using SPSS software for Windows (IBM Corp., Armonk, NY, USA).

9.4. Statistical Methods

Continuous variables will be summarized as means ± standard deviations or medians with interquartile ranges, as appropriate. Categorical variables will be described using frequencies and percentages. Between-group comparisons will be performed using Student's t-test, Mann–Whitney U test, χ^2 test, or Fisher's exact test, depending on data distribution and variable type. Longitudinal changes over time will be evaluated using repeated-measures ANOVA or mixed-effects regression models, with adjustment for baseline values and relevant covariates, including age, sex, BMI, medication use. Multivariate regression analyses will be conducted to explore associations between biological aging markers and clinical outcomes, as well as to assess gene-diet interaction were appropriate. All statistical tests will be two-sided, and P value < 0.05 will be considered statistically significant.

Descriptive statistics will be summarized as means ± standard deviations (SD) for continuous variables and as appropriate measures for categorical variables. Because most study variables will be assessed at two or more time points, participant-level values will be summarized using the mean of all available measurements for each participant in the primary analyses. Variables exhibiting non-normal distributions will be log-transformed (natural logarithm) prior to analysis to approximate normality. Within-group changes from baseline to follow-up differences will be evaluated using paired Student's t-test, whereas between-group differences will be assessed using analysis of covariance (ANCOVA), with baseline values included as covariates. Multivariate models will be adjusted for age, sex, and any clinical parameters showing significant baseline imbalances between to control for potential confounding. Post-hoc comparisons will be corrected for multiple testing using Bonferroni method. Effect estimates will be expressed as mean differences with 95% confidence intervals (CIs).