Modeling Cardiometabolic Multimorbidity Over Time: Integrating Subject-Specific and Population-Averaged Ordinal Approaches

Qikai Jiang

Introduction

Over the past several decades, the prevalence of cardiometabolic conditions, including type 2 diabetes, hypertension, and dyslipidemia, has increased markedly across the United States. These interrelated disorders are now well-established contributors to cardiovascular disease (CVD), chronic kidney disease, and premature mortality (Benjamin et al., 2017; Ford et al., 2007; Gregg et al., 2014). Despite expanded access to preventive care and pharmacologic therapies, surveillance data continue to document not only rising rates of diagnosis, but also earlier onset, trends that have been particularly pronounced across disadvantaged sociodemographic groups. For example, nearly half of U.S. adults currently meet criteria for hypertension (defined as ≥130/80 mmHg), with prevalence climbing above 60% among individuals over age 40 (Muntner et al., 2018). Type 2 diabetes affects more than 14% of adults, and another 38% are considered prediabetic (CDC, 2014), with incidence increasingly concentrated among younger individuals, a shift attributed to the widespread rise in obesity and declining physical activity (Mayer-Davis et al., 2017; Menke et al., 2015). Dyslipidemia remains similarly common, affecting over 40% of the adult population, and continues to be underrecognized and undertreated in the absence of clinical CVD (Gao et al., 2023; Dimitriadis et al., 2024).

Notably, these conditions rarely develop in isolation. Their co-occurrence reflects a common biological foundation rooted in chronic low-grade inflammation, insulin resistance, and vascular dysfunction. Hyperglycemia disrupts endothelial homeostasis via oxidative stress and the formation of advanced glycation end-products; elevated blood pressure promotes vascular remodeling; and dyslipidemia fosters atherogenesis through oxidation of low-density lipoproteins and macrophage activation (Libby et al., 2002; Grundy et al., 2005). Together, these mechanisms contribute to progressive arterial damage and organ dysfunction. The Framingham Heart Study was among the first to show that individuals with

multiple metabolic risk factors faced significantly higher 10-year risks of myocardial infarction and cardiovascular death, compared to those with none or only one risk factor (Wilson et al., 1998). These findings helped inform the construction of metabolic syndrome, which emphasizes the clinical importance of such clustering (Alberti et al., 2009).

To design effective interventions, it is critical to understand how these conditions evolve together over time. While individual behaviors and biological traits certainly play important roles, broader social and structural forces, including neighborhood disadvantages, and unequal access to preventive care, also shape cardiometabolic risk trajectories. These upstream influences contribute to earlier disease onset and faster progression in low-income and racially marginalized populations (Thorpe et al., 2016). In this context, longitudinal modeling strategies offer valuable tools for studying cardiometabolic burden as it accumulates within individuals. Approaches such as mixed-effects and multi-state models allow researchers to capture within-person variation, accommodate time-varying exposures, and trace how intersecting risk factors unfold over time.

These methods have proven useful in large cohort studies. The Coronary Artery Risk Development in Young Adults (CARDIA) study, which has followed a diverse sample from early adulthood into middle age, has shown that subtle early elevations in blood pressure, glucose, and lipid levels are associated with later subclinical heart disease, including coronary artery calcification and ventricular dysfunction (Allen et al., 2014; Perak et al., 2018). Similarly, findings from the Atherosclerosis Risk in Communities (ARIC) study indicate that prolonged exposure to high triglycerides, abdominal adiposity, and low HDL-C is strongly associated with cardiovascular events later in life, further underscoring the value of longitudinal measurement in risk stratification (Liao et al., 1996; Norby et al., 2016).

Building on this foundation, the present study examines the progression of three cardiometabolic conditions, diabetes, high cholesterol, and high blood pressure, over time, using data from a nationally representative U.S. cohort followed across three waves. To reflect the multifaceted drivers of cardiometabolic health, we include a broad array of demographic, behavioral, and clinical predictors.

Body mass index (BMI), a key indicator of excess adiposity, plays a central role. Elevated BMI is linked to insulin resistance, heightened sympathetic tone, systemic inflammation, and lipid abnormalities. Studies have shown that individuals in the highest BMI quintiles have two- to fivefold increased odds of developing diabetes or hypertension (Kopelman, 2000; Field et al., 2001; Wilkins et al., 2012).

Other important risk factors include sex and age. Premenopausal women tend to have lower cardiometabolic risk than men, but this protective effect narrows with age. Moreover, women may face distinct metabolic challenges, such as visceral adiposity or gestational diabetes, and are more likely to encounter barriers to diagnosis and treatment due to gender-related disparities in care (Regitz-Zagrosek et al., 2016; Maas & Appelman, 2010). Age itself is a key determinant, magnifying the effects of both behavioral and physiological exposures across the lifespan (Ford et al., 2010; Mozaffarian et al., 2008).

Socioeconomic context further shapes cardiometabolic risk. Educational attainment, a widely used proxy for socioeconomic status, is strongly associated with healthier behaviors, better access to care, and lower disease burden (Cutler & Lleras-Muney, 2010; Winkleby et al., 1992). Diet quality, as measured by the Healthy Eating Index–2010 (HEI-2010), has also been linked to lower prevalence of metabolic syndrome and better overall cardiometabolic profiles. Higher HEI scores have been associated with 20–30% lower risk of developing multiple metabolic abnormalities (Reedy et al., 2014). Adherence to physical activity guidelines (PAG2008) plays a similarly protective role, with regular activity linked to 30–40% reductions in incidence of diabetes and hypertension (Myers et al., 2002).

Psychological and behavioral health factors matter as well. Depression has been identified both as a contributor to and a consequence of cardiometabolic disease, acting through behavioral pathways (e.g., inactivity, poor diet) as well as physiological mechanisms such as HPA-axis dysregulation and inflammatory signaling. Meta-analyses estimate that depressive symptoms increase risk of diabetes and hypertension by 30–60% (Pan et al., 2012). Alcohol use presents a more complex picture: while light-to-moderate drinking may offer some metabolic benefit, excessive consumption raises blood pressure and triglyceride levels (Bell et al., 2017). Smoking, meanwhile, remains a consistent and well-documented

risk factor across all stages of metabolic dysfunction, linked to increased insulin resistance, dyslipidemia, and central adiposity (Chiolero et al., 2008).

While much has been learned from past research, several analytical challenges remain. Most notably, many studies examine diabetes, hypertension, and dyslipidemia as separate outcomes, modeling them independently and failing to account for their combined burden. This approach misses important insights into the joint development of multimorbidity and the cumulative impact of these conditions over time (Grundy et al., 2005; Libby et al., 2002). Even large-scale analyses from NHANES (Ford et al., 2010), CARDIA (Pedamallu et al., 2023), and ARIC (Norby et al., 2016) have often focused on single outcomes in isolation. Relatively few adopt composite or ordinal measures that capture disease accumulation (Tian et al., 2024).

Another limitation is the reliance on cross-sectional designs, which offer only snapshots of risk and cannot distinguish within-person changes from between-person differences. While cohort studies like CARDIA (Allen et al., 2014) and Framingham (Wilson et al., 1998) have offered valuable longitudinal data, much of the resulting analysis has focused on incident events rather than chronic multimorbidity or its accumulation across repeated observations (Twisk, 2013).

In the subset of studies that do use ordinal outcomes, most rely on proportional odds (PO) models. Yet the assumption that covariate effects remain constant across outcome thresholds is often unrealistic. For example, the effect of BMI may differ when moving from no conditions to one, compared to the transition from two to three conditions (Peterson & Harrell, 1990; Brant, 1990). Partial proportional odds models allow for more flexibility, but remain underutilized in applied research despite their potential to better reflect variation across the disease burden spectrum (Williams, 2006).

Finally, integrated models that combine behavioral, clinical, and structural predictors, particularly in longitudinal ordinal frameworks, are rare. Many existing studies focus narrowly on a few behavioral risk factors or limit structural variables to cross-sectional assessments (Diez Roux, 2012; Thorpe et al., 2016).

Approaches that allow for repeated measurement, time-varying predictors, and within-person change, using tools such as generalized estimating equations (GEE), mixed-effects models, or Bayesian cumulative logit models, are infrequently applied to nationally representative datasets like NHANES, MIDUS, or HRS.

This study seeks to address these gaps through a comprehensive, multilevel analysis of cardiometabolic disease progression. We construct an ordinal disease burden score, ranging from 0 to 3, based on the number of concurrent diagnoses of type 2 diabetes, high blood pressure, and high cholesterol. This approach allows us to move beyond binary disease classification and reflect how risk accumulates across time and across domains.

Our primary goal is to examine how behavioral (e.g., diet, physical activity, smoking, alcohol use), clinical (e.g., BMI, depressive symptoms), and structural (e.g., education, age, sex) factors jointly influence the progression of cardiometabolic burden. Using three waves of data from a nationally representative U.S. sample, we apply a suite of complementary modeling strategies: (1) cross-sectional cumulative logit models to examine baseline associations; (2) generalized linear mixed models (GLMMs) to assess within-person trajectories and interactions with time; (3) GEE models to produce population-level inference accounting for repeated ordinal responses; and (4) Bayesian partial proportional odds models to relax the PO assumption and quantify uncertainty across the disease continuum. Taken together, these methods allow us to study the accumulation of cardiometabolic conditions as a dynamic, longitudinal process shaped by individual behaviors and broader structural influences, laying the groundwork for more precise and equitable prevention strategies in public health.

Methods

Study Variables and Data Structure

The analytic dataset comprised a well-characterized mix of static baseline attributes and dynamic, time-varying covariates collected across three consecutive assessment periods. Together, these variables

captured a comprehensive range of sociodemographic, behavioral, and clinical factors relevant to understanding both baseline risk and longitudinal progression of cardiometabolic disease. The data structure was specifically designed to support repeated-measures modeling, allowing for the investigation of changes in exposures and outcomes over time within individuals, while also accommodating between-person differences in baseline characteristics.

Baseline covariates were recorded once and treated as fixed throughout the follow-up period. These included sex (female vs. male), highest level of educational attainment (categorical), and dietary quality, assessed via the Healthy Eating Index–2010 (HEI-2010). The HEI-2010 is a continuous index score, with higher values indicating closer adherence to established dietary guidelines. Physical activity was dichotomized according to whether participants met the 2008 Physical Activity Guidelines for Americans (PAG2008), while depressive symptoms were assessed using the Center for Epidemiologic Studies Depression Scale (CES-D). A binary indicator was constructed to reflect the presence or absence of clinically significant depressive symptoms at baseline, in accordance with established CES-D thresholds.

Time-varying covariates, which were measured at each of the three visits, reflected behavioral and physiological characteristics that could evolve over the study period. These included age (treated as a continuous variable in years), current alcohol use (binary: current use vs. no use), and body mass index (BMI), calculated as weight in kilograms divided by height in meters squared (kg/m^2). BMI was modeled both continuously and dichotomously, with obesity defined as $BMI \geq 30$. Smoking status was similarly treated as a dynamic binary variable, indicating whether a participant reported current smoking at each time point.

In addition to these predictors, participants reported physician-diagnosed cardiometabolic conditions at each assessment, including diabetes mellitus, high blood pressure (hypertension), and high cholesterol. Each of these diagnoses was coded as a binary indicator (yes/no) at each visit. These three condition-specific outcomes were then combined into a composite ordinal score ranging from 0 to 3, indicating the cumulative number of cardiometabolic conditions present at a given time. This cumulative score served as

the primary dependent variable in the longitudinal analyses, providing a clinically meaningful, easily interpretable measure of overall cardiometabolic disease burden and its potential progression across time.

Table 1 summarizes the distribution of all study variables, organized by type (categorical or continuous) and temporal stability (static vs. time-varying). Static variables of sex, education, HEI-2010 score, physical activity adherence, and baseline depressive symptoms, were assessed once at baseline and treated as invariant across visits. In contrast, time-varying predictors such as age, alcohol use, BMI, obesity status, and smoking status, were recorded at each visit and allowed to vary within individuals over time. Each of the three primary binary outcomes (diabetes, hypertension, and high cholesterol) was measured longitudinally and then aggregated into the composite ordinal outcome score used throughout the modeling. Descriptive summaries include means and standard deviations for continuous variables and counts and proportions for categorical or binary variables, providing a clear depiction of the sample's baseline profile and observed variability over the course of the study.

This detailed structure, integrating both fixed and time-varying covariates alongside an interpretable ordinal outcome, forms the foundation for the subsequent longitudinal modeling framework, which is designed to capture both within-person changes and between-person differences in the evolution of cardiometabolic multimorbidity.

Exploratory Data Analysis and Descriptive Visualizations

Building upon the structured variable framework described above, we conducted a series of exploratory analyses to examine the distributional characteristics, temporal patterns, and interrelationships among key predictors and outcomes. These initial assessments were essential for guiding appropriate model specification, identifying potential data quality issues, and understanding how covariates and outcomes evolve over time.

To evaluate the empirical distribution of continuous covariates which specifically are age, body mass index (BMI), and dietary quality (HEI-2010), we constructed pooled histograms across all three study

visits (Figure 1). These visualizations served as a diagnostic tool for assessing approximate normality, identifying skew, and detecting outliers. While age and HEI-2010 exhibited reasonably symmetric distributions, the BMI distribution showed evidence of heavy upper-tail values. To further investigate the influence of extreme BMI observations, we examined the top three highest BMI values across visits, as reported in Table 2. These records were scrutinized for data entry anomalies or clinical implausibility. In cases where values appeared to be spurious and could not be defended as legitimate, they were excluded to safeguard model stability and reduce undue leverage on estimation procedures.

Given that both age and BMI were modeled as time-varying continuous covariates in longitudinal analyses, we next assessed the extent of within-subject variability for each. While age increased deterministically with time and required no additional reliability evaluation, BMI was expected to fluctuate across visits due to both biological changes and measurement variability. To quantify this variation and determine the degree of consistency within individuals, we computed within-subject variance and the intraclass correlation coefficient (ICC), as summarized in Table 4. The ICC, which captures the proportion of total variance attributable to between-subject differences, provided a formal measure of measurement reliability for BMI across the study period. Higher ICC values suggested that most of the variation in BMI arose between individuals rather than within, thereby justifying its inclusion as a stable yet time-varying covariate in longitudinal models.

In parallel, we examined temporal shifts in categorical variables to gain insight into how health status and behaviors evolved within the cohort. Table 3 summarizes the number and percentage of participants who experienced either an increase (e.g., disease onset, behavioral adoption) or decrease (e.g., remission, cessation) in each binary variable between consecutive visits. These included physician-reported diagnoses of diabetes, hypertension, and high cholesterol, as well as behavioral variables such as alcohol use, smoking, and obesity (BMI ≥ 30). Changes in the composite comorbidity score, reflecting accumulation or reduction of disease burden from 0 to 3 conditions, were also tracked. This longitudinal characterization of transitions highlighted the dynamic nature of both health and behavior over time,

underscoring the importance of modeling strategies capable of capturing within-subject change and population-level heterogeneity.

To further explore how covariates relate to the ordinal comorbidity outcome, we generated a series of stratified visualizations, disaggregated by comorbidity score and study visit. Figure 2 presents boxplots of time-varying continuous predictors includes age and BMI across levels of cumulative disease burden. These plots provided intuitive visual evidence of gradient effects, suggesting that higher comorbidity scores were associated with older age and greater BMI. HEI-2010, which was assessed only at baseline and treated as time-invariant, was also summarized to assess its cross-sectional association with cardiometabolic burden.

In addition, we visualized distributions of categorical predictors using barplots to compare prevalence across outcome categories. Figure 3 focuses on time-varying behavioral factors of alcohol use, obesity status, and smoking status, illustrated separately at visits 1, 2, and 3 to highlight temporal dynamics. Figure 4 complements this by presenting stable baseline predictors of sex, education level, and physical activity adherence (PAG2008), also separately for those 3 visits. This dual approach, combining time-resolved and static summaries, allowed us to contrast how enduring baseline traits and modifiable behaviors interact with disease burden over time.

These descriptive visualizations are a critical component of any longitudinal analytic workflow. They not only reveal patterns of change and stability across the cohort but also inform decisions around model complexity, variable coding, and the potential need for interaction terms or non-linear specifications. Moreover, by stratifying covariate distributions by levels of comorbidity burden, we obtained preliminary evidence of dose-response relationships and directional associations that would later be evaluated formally through multivariable modeling.

Predictor Collinearity and Associations, both Cross-sectionally and Longitudinally

To guide model specification and assess potential overlap among explanatory variables, we examined associations between covariates using both cross-sectional and longitudinal approaches. Understanding the extent of collinearity among predictors is especially important in longitudinal settings, where repeated measures and time-varying exposures can complicate interpretation and model stability. High correlations between covariates may obscure the independent contributions of each factor, inflate standard errors, and challenge the interpretability of effect estimates in multivariable frameworks.

Our initial assessment focused on cross-sectional diagnostics, applying methods appropriate to the scale and type of each variable pair. For associations between continuous variables, we used Spearman rank correlations; continuous–binary pairs were evaluated using point-biserial correlations; and associations between binary or categorical variables were assessed using chi-squared tests. These analyses were conducted separately for each visit when applied to time-varying covariates, while static variables were analyzed using pooled data from all time points. Among the 45 unique predictor pairs evaluated, Table 5 highlights the ten strongest or most policy-relevant associations. For each pair, we report the corresponding test statistic or correlation coefficient, along with the associated p-value. These results offer a practical starting point for identifying pairs of variables that may introduce redundancy or complicate interpretation in multivariable models.

While cross-sectional associations provide useful snapshots, they do not capture the temporal dynamics that characterize longitudinal data. In studies with repeated measures, it is essential to distinguish between variation occurring within individuals over time and differences observed between individuals. This distinction has important implications for model selection, particularly when deciding between marginal and subject-specific approaches. Properly accounting for within-subject correlation is fundamental to avoiding biased inference and ensuring that parameter estimates remain interpretable.

From a statistical modeling standpoint, unaddressed time-dependent correlations between covariates can lead to misspecified models, inefficient estimation, and misleading inference. For example, strong within-subject correlations may violate assumptions underlying standard error estimates in generalized

estimating equations (GEE), while persistent collinearity among time-varying variables can obscure causal interpretation and destabilize parameter estimation in mixed-effects or Bayesian hierarchical models. Thus, a careful evaluation of how covariates relate to each other over time is critical for building models that are both valid and parsimonious.

To capture these dynamic relationships, we conducted a series of longitudinal association diagnostics designed to quantify both the strength and structure of correlations among key predictors across the study period. Table 6 presents repeated-measures point-biserial correlations between continuous and binary covariates, such as BMI and obesity status, with random-intercept models used to account for intra-individual variability. Table 7 focuses on continuous–continuous pairs, such as age and BMI, using linear mixed-effects models to partition the total variance into within- and between-subject components. For each pair, we report the estimated slope, standard error, and intraclass correlation coefficient (ICC), offering insight into how much of the variability is attributable to differences between individuals versus within-person change. Finally, Table 8 summarizes associations between binary time-varying predictors, such as alcohol use and smoking status, estimated using GEE models to capture population-averaged relationships while accounting for the repeated nature of the data.

Taken together, these diagnostics provide a comprehensive picture of the evolving correlation structure among the study's predictors. By identifying covariates that are highly correlated either cross-sectionally or longitudinally, and distinguishing stable associations from those that vary over time, these results directly inform model building. They help ensure that covariates are included judiciously, that redundancy is minimized, and that the modeling framework appropriately reflects the temporal and structural complexity of the data.

Diagnostic Testing and Inference in Cross-Sectional Data

Building on the preceding diagnostic evaluations of covariate structure and interrelationships, we next turned to modeling the composite ordinal outcome representing cumulative cardiometabolic burden.

Given the outcome's ordered categorical nature which ranging from zero to three diagnosed conditions (diabetes, high cholesterol, and hypertension), we initially considered cumulative logit models under the proportional odds (PO) framework. However, given prior indications of complex, time-varying relationships among predictors, it was essential to formally assess whether the PO assumption held, both at the level of individual predictors and across the full model structure.

To this end, we began by fitting separate cumulative logit models at each of the three study visits and conducted a series of diagnostic procedures to evaluate whether each predictor exerted a consistent effect across all outcome thresholds. The proportional odds assumption requires that the log-odds of exceeding each threshold (e.g., ≥1, ≥2, =3) differ only by a constant intercept shift, implying a single effect estimate per predictor. Violations of this assumption would suggest the need for a more flexible modeling approach that allows predictor effects to vary with the severity of cardiometabolic burden.

As a first step, we applied the Brant test separately at each visit to assess proportionality at the predictor level. The results, presented in Table 9, highlight several covariates for which the PO assumption was statistically rejected at the $\alpha = 0.05$ level. Notably, violations were more frequent at visits 2 and 3, particularly for behavioral and clinical variables such as BMI, smoking status, and alcohol use, suggesting that these predictors may exert threshold-dependent effects on disease burden as individuals age or experience changes in health behavior over time.

To complement these predictor-specific tests, we evaluated the overall adequacy of the PO model using likelihood ratio tests. Specifically, we compared the cumulative logit specification to a fully generalized multinomial model that places no constraints on the relationship between predictors and outcome categories. As reported in Table 10, the nominal models offered significantly improved fit over their proportional counterparts at each visit (all p-values < 0.01), reinforcing the concern that the PO assumption may not be tenable in this context.

Graphical diagnostics provided further insight into the nature of these violations. Figure 5 displays log-odds estimates and 95% confidence intervals from separate binary logistic models comparing outcome levels ≥1, ≥2, and =3 for each predictor. These threshold-specific estimates make apparent the varying direction and magnitude of several effects. For instance, the association between BMI and disease burden increased markedly at higher outcome levels, while the impact of smoking and alcohol use varied both in sign and strength depending on the severity threshold. Such patterns visually confirmed the statistical evidence for non-proportionality.

In tandem, we assessed whether the assumption of logit-linearity held for continuous predictors that namely, age, BMI, and HEI-2010 diet quality scores. Figure 6 plots predicted log-odds against each of these predictors, overlaid with LOESS-smoothed curves derived from separate binary logistic models for each cumulative threshold. In all three cases, the observed trends supported approximate linearity across thresholds, justifying their treatment as continuous linear terms in subsequent models.

Given the consistent evidence of non-proportional effects, we specified partial proportional odds (PPO) models at each visit. These models allow covariates that violate the PO assumption to vary flexibly across thresholds, while retaining the proportionality constraint for others, balancing flexibility with interpretability. Tables 11, 12, and 13 present the estimated coefficients, standard errors, z-values, and p-values from these PPO models for visits 1, 2, and 3, respectively. Threshold-specific estimates are reported for covariates violating the PO assumption, providing a more accurate representation of how the effects of key variables shift across levels of disease burden.

To evaluate model adequacy and identify potential sources of poor fit, we conducted residual diagnostics summarized in Table 14. For each visit, we report the residual deviance, log-likelihood, AIC, and BIC, along with the number of observations exhibiting large Pearson residuals ($|\text{residual}| > 4$). We also identify the maximum residual at each visit, including the observation ID and associated outcome threshold, offering a targeted lens for sensitivity checks or influential case analysis.

Graphical residual diagnostics are presented in Figure 7, which includes (1) Pearson residuals plotted against fitted values to assess model dispersion, (2) Q–Q plots of residuals to evaluate distributional assumptions, and (3) residuals indexed by observation number to detect temporal or cluster-specific patterns. Dashed reference lines at ±4 highlight thresholds for potential outliers or influential observations. Across all three panels and visits, these visuals helped confirm the general adequacy of model fit while flagging a small number of atypical cases meriting closer examination.

Taken together, the combination of formal tests, graphical diagnostics, and flexible modeling via partial proportional odds structures provided a principled approach to addressing the complex, threshold-dependent effects observed in this study. By relaxing the proportionality assumption where appropriate, the final modeling framework more accurately captured the evolving and heterogeneous nature of cardiometabolic risk accumulation across individuals and time.

Longitudinal Threshold-Specific Mixed-Effects Models for Evaluating Subject-Level Heterogeneity and Proportional Odds

Building on the preceding correlation diagnostics and cross-sectional analyses, we advanced to a longitudinal modeling framework to capture within-subject dynamics in cardiometabolic burden across repeated assessments. As previously defined, the outcome represents a cumulative ordinal score ranging from 0 to 3, corresponding to the number of physician-diagnosed conditions of diabetes mellitus, hypertension, and high cholesterol, reported at each visit. The ordinal nature of this outcome, coupled with the longitudinal design, motivated the use of cumulative logit generalized linear mixed models (GLMMs) incorporating random intercepts to account for intra-individual correlation over time.

Before fitting the full longitudinal models, we formally evaluated the proportional odds (PO) assumption, a central requirement of cumulative logit models. To this end, we estimated threshold-specific fixed effects using a series of cumulative binary logistic mixed models, summarized in Table 15. This approach involved fitting three separate mixed-effects binary logistic regressions corresponding to each ordinal

threshold: (i) presence of at least one condition, (ii) at least two conditions, and (iii) all three conditions. Each threshold-specific model included the same set of covariates: indicators for Visit 2 and Visit 3 (with Visit 1 as the reference), sex (female), education categories (levels 2 and 3, relative to level 1), physical activity adherence (PAG2008), depressive symptoms, alcohol use, smoking status, diet quality (HEI-2010), age, and body mass index (BMI). By comparing estimated coefficients across these models, we evaluated the extent to which the proportional odds assumption held. Substantial variation across thresholds would suggest the need for more flexible alternatives, such as partial proportional odds or nominal outcome models.

To further assess model adequacy in the context of repeated ordinal measurements, we compared specifications with and without subject-specific random intercepts. These intercepts are designed to capture unobserved heterogeneity across individuals and to account for within-subject correlation over time. Table 16 presents comparative model fit statistics, including the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), for both the random-intercept and fixed-effects-only models. These metrics provide a basis for evaluating whether accounting for intra-individual dependence improves model parsimony and adequacy. While lower AIC and BIC values are typically interpreted as evidence of better fit, the direction and magnitude of any differences must be interpreted in context, particularly in relation to model complexity and sample size. This comparison informs whether the inclusion of random effects meaningfully captures latent subject-level variability not explained by the observed covariates.

Figure 8 provides important diagnostic visualizations for the random intercepts estimated in the multinomial generalized linear mixed models. The histogram illustrates the empirical distribution of the empirical Bayes estimates, allowing a visual check for symmetry and potential deviations from normality. The accompanying normal Q–Q plot further evaluates the assumption that these random intercepts follow a Gaussian distribution, which is fundamental to the mixed-effects modeling framework. Together, these

plots support the appropriateness of modeling subject-specific heterogeneity through normally distributed random effects.

With these diagnostics in place, we proceeded to estimate the primary longitudinal model using a multinomial GLMM with a cumulative logit link. The initial model excluded interaction terms, allowing for estimation of average covariate effects across all three visits. Table 18 presents fixed effects estimates from this model, including coefficients, standard errors, odds ratios, and corresponding t-statistics with degrees of freedom for each predictor and cumulative threshold. This specification captures stable associations of visit, sex, education, depressive symptoms, alcohol and smoking behaviors, diet quality, age, and BMI with the accumulation of cardiometabolic conditions over time.

Table 19 reports the Type III tests of fixed effects for this no-interaction model, summarizing joint significance of each predictor using F-statistics and associated numerator and denominator degrees of freedom. These omnibus tests clarify which variables exert consistent longitudinal influence on comorbidity burden, serving as a complement to the threshold-specific estimates and model diagnostics.

To explore whether covariate effects changed over time, we fit an extended multinomial GLMM that included interaction terms between visit indicators and all predictors. This model accommodates temporal heterogeneity in predictor-outcome relationships, allowing associations to vary by assessment period. Table 20 summarizes fixed effects from this interaction model, enabling direct comparison of effect sizes across visits. Table 21 reports Type III tests of these interaction terms, quantifying the additional explanatory value gained by allowing covariate effects to evolve longitudinally.

Following the examination of fixed effects with and without visit-by-covariate interactions, Table 17 reports the estimated variance components and associated standard errors for the subject-specific random intercepts from multinomial generalized linear mixed models. Variance components are presented for models both excluding and including interaction terms between visit indicators and covariates, reflecting

the modeling of between-subject variability in baseline cardiometabolic burden under different specifications.

Together, these models, variance estimates, and diagnostics offer a comprehensive and rigorous framework for modeling an ordinal longitudinal outcome that captures the complexity of cumulative cardiometabolic burden. The use of threshold-specific models to test the proportional odds assumption, incorporation of random intercepts to account for subject-level correlation, and consideration of interaction terms to address time-varying covariate effects collectively enhance the robustness, interpretability, and clinical relevance of our findings. This multilevel approach supports principled inference regarding the predictors and progression of multimorbidity over time, balancing statistical parsimony with the structural intricacies inherent in repeated ordinal data.

Bayesian Cumulative Logit Mixed-Effects Model with Partial Proportional Odds

To address violations of the proportional odds assumption identified in threshold-specific cumulative binary logistic models (Table 15), we implemented a Bayesian cumulative logit mixed-effects model with a partial proportional odds (PPO) structure. This framework accommodates longitudinal ordinal outcomes by permitting covariate effects to vary across cumulative thresholds of the ordinal scale, while simultaneously accounting for within-subject correlation through the inclusion of subject-level random intercepts.

The outcome variable was a cumulative ordinal score indicating cardiometabolic multimorbidity, defined by the total number of conditions (diabetes, hypertension, high cholesterol) present at each study visit. The response ranged from 0 to 3 and was modeled using a cumulative logit link. The PPO structure distinguishes between predictors whose effects are assumed constant across thresholds (proportional odds effects) and those whose effects vary across thresholds (non-proportional odds effects). This formulation allows the model to maintain interpretability for predictors conforming to the proportional odds assumption, while providing flexibility for those violating it, as determined by prior diagnostics.

The model was estimated using Bayesian hierarchical methods via Hamiltonian Monte Carlo (HMC) implemented in the brms package. The formulation included a random intercept for each subject to model intra-individual correlation across repeated measures. Proportional effects were specified for predictors that did not exhibit evidence of threshold-dependent effects in prior analyses, while the cs() operator was used to model non-proportional effects, allowing these covariates to have distinct coefficients at each threshold. This partial specification strikes a balance between model parsimony and fidelity to the ordinal data structure.

The full model was estimated using adaptive HMC sampling with Gaussian quadrature, configured to ensure accurate integration over the random effects distribution. Convergence and mixing of the Markov chains were assessed using standard diagnostics, including potential scale reduction statistics, effective sample sizes, and trace plots for all parameters. Table 22 summarizes posterior estimates for the fixed effects, threshold-specific intercepts, and the standard deviation of the random intercept. Each parameter's posterior distribution was characterized by its mean, standard error, posterior odds ratio transformation (where applicable), and 95% credible intervals. Additional convergence diagnostics were reported, including bulk and tail effective sample sizes and $\hat{R}$ values.

Model adequacy and out-of-sample predictive performance for the Bayesian cumulative logit mixed-effects model with partial proportional odds (PPO) specification were assessed using approximate leave-one-out cross-validation (LOO-CV) and the widely applicable information criterion (WAIC), as summarized in Table 23. Both criteria estimate the expected log predictive density (ELPD) of unseen data, reflecting how well the model generalizes beyond the observed sample. In the context of longitudinal ordinal regression, these metrics provide an essential check on the model's ability to capture the underlying distribution of outcomes across repeated measurements while accounting for both fixed and random effects.

LOO-CV is implemented in a fully Bayesian framework via Pareto-smoothed importance sampling (PSIS), which allows efficient estimation of pointwise out-of-sample prediction error without explicitly

refitting the model for each left-out observation. This procedure adjusts for model complexity through the effective number of parameters, thereby balancing fit with parsimony. WAIC, a related fully Bayesian metric derived from the log pointwise posterior predictive density, also penalizes complexity via variance-based estimates of model flexibility. While both metrics converge asymptotically under regularity conditions, LOO-CV is generally preferred in applied settings involving hierarchical structure or latent variables due to its greater numerical stability and lower sensitivity to misspecification of the likelihood.

In this analysis, LOO-CV was prioritized as the primary measure of predictive adequacy. The Pareto-k diagnostic values were used to evaluate the reliability of PSIS approximations, with a majority of observations falling below the conservative $k \leq 0.7$ threshold, indicating minimal influence of any single observation on the overall predictive distribution. In addition, the minimum effective sample size was computed to assess the stability of tail behavior in the importance weights, ensuring robustness in posterior summaries. These diagnostics collectively validate the use of LOO-CV for this model and support its role in selecting among alternative specifications.

Together, the LOO and WAIC results provided a principled, fully Bayesian framework for model comparison and validation, accounting not only for predictive accuracy but also for complexity introduced by random intercepts and non-proportional covariate effects. This approach is particularly important in the longitudinal setting, where repeated measures and individual-level heterogeneity pose challenges to conventional model evaluation techniques.

Figure 9 displays diagnostic plots, posterior density estimates and MCMC trace plots, for a subset of model parameters from the Bayesian cumulative logit mixed-effects model with partial proportional odds specification. The top three rows correspond to the threshold-specific intercepts, which define the cumulative logit cut-points across the ordinal outcome scale. These intercepts are essential for determining the position of each cumulative category and serve as benchmarks for interpreting covariate effects. Rows 4 through 6 show selected proportional odds covariates, whose coefficients are constrained

to be constant across thresholds. The remaining rows illustrate non-proportional covariates, where threshold-specific estimates were modeled using the cs() operator within the brms syntax.

The posterior density plots visualize the estimated posterior distributions for each parameter, allowing assessment of distributional shape, central tendency, and tail behavior. Smooth, unimodal densities with symmetric credible intervals support adequate exploration of the posterior space. Accompanying trace plots show the iteration-wise sampling paths for multiple chains, providing a dynamic visualization of MCMC mixing and convergence. Well-mixed chains that exhibit rapid traversal, stationarity, and lack of autocorrelation (i.e., “hairy caterpillar” appearance) are indicative of stable posterior sampling and adequate tuning.

Figure 10 continues this diagnostic suite, presenting posterior densities and trace plots for the remaining non-proportional predictors which includes alcohol use, BMI, education levels, depression, and smoking status, each modeled with three distinct coefficients for the ordinal thresholds. The final row of the figure includes the posterior distribution and trace plot for the subject-level random intercept standard deviation, a key parameter capturing between-subject heterogeneity in baseline ordinal risk.

These diagnostic visualizations align with best practices for Bayesian inference implemented via the brms package and Stan backend. They enable qualitative verification of model performance, complementing numerical convergence diagnostics such as $\hat{R}$ (Gelman-Rubin statistic), effective sample size (ESS), and autocorrelation estimates. Together, Figures 9 and 10 provide critical evidence for the validity of posterior estimation, the stability of model specification, and the plausibility of threshold-specific and random effect assumptions central to the partial proportional odds framework.

In sum, the Bayesian cumulative logit mixed-effects model with a partial proportional odds structure offers a principled and flexible approach to modeling longitudinal ordinal outcomes defined by cumulative cardiometabolic burden. By jointly capturing threshold-specific covariate effects and subject-level random heterogeneity, the framework effectively reconciles the ordinal nature of multimorbidity

progression with the temporal structure of repeated measures data. The integration of proportional and non-proportional components accommodates variation in covariate associations across levels of the outcome, ranging from the absence of any conditions to the presence of all three (diabetes, hypertension, and high cholesterol), while preserving parsimony where the proportional odds assumption remains valid. Bayesian estimation via Hamiltonian Monte Carlo further enables coherent uncertainty quantification and model evaluation, supported by posterior diagnostics, predictive accuracy metrics, and convergence checks. As demonstrated through posterior summaries, information criteria, and visual diagnostics, this modeling strategy affords a robust inferential framework for studying complex, temporally evolving ordinal health outcomes in population-based settings.

Generalized Estimating Equations (GEE) for Marginal Analysis

To characterize the longitudinal relationship between covariates and cumulative cardiometabolic multimorbidity, a composite ordinal outcome representing the number of concurrent diagnoses of diabetes, hypertension, and high cholesterol at each study visit (ranging from 0 to 3), we employed generalized estimating equations (GEE) as a marginal modeling strategy. Unlike subject-specific mixed-effects models that target individual-level conditional means, GEE estimates the population-averaged effects of predictors over time. This approach accommodates repeated ordinal measurements by adjusting for within-subject correlation using a prespecified working correlation matrix (e.g., exchangeable, autoregressive), while making minimal assumptions about the distributional form of subject-specific effects. Consequently, GEE is particularly advantageous for large-sample settings where inference about average trends across the population is of primary interest.

To examine whether the proportional odds assumption, central to cumulative logit models, was tenable within the GEE framework, we began with threshold-specific marginal binary logistic models. This diagnostic strategy involves fitting separate GEE models at each cutpoint of the ordinal scale: (i) comparing individuals with outcome ≥ 1 vs <1, (ii) ≥ 2 vs <2, and (iii) $=3$ vs <3. Table 24 summarizes the coefficient estimates, standard errors, Wald z-statistics, and associated p-values for each covariate across

these thresholds. Discrepancies in effect magnitudes or significance across thresholds suggest violation of the proportional odds assumption and indicate that a more flexible model, such as one allowing threshold-specific effects, may be warranted.

Motivated by these findings, we next implemented a cumulative logit GEE model assuming proportional odds across thresholds to estimate the marginal association between covariates and the ordinal outcome. Table 25 presents the parameter estimates for this model, which included sociodemographic and behavioral predictors, such as visit (time), sex, education, smoking, alcohol use, adherence to physical activity guidelines (PAG2008), Healthy Eating Index (HEI2010), age, and body mass index (BMI), but excluded interaction terms. This model constrains the log-odds of being in a higher outcome category to change uniformly across thresholds, thereby providing a parsimonious summary of covariate effects across the ordinal scale. Each parameter is accompanied by its standard error, Wald z-statistic, odds ratio transformation, and 95% confidence interval, facilitating inference about the average effect of each predictor across all levels of multimorbidity burden.

To relax the assumption of temporally homogeneous effects and capture potential time-varying covariate influences, a second cumulative logit GEE model was specified with visit-by-covariate interaction terms. Table 26 reports the corresponding parameter estimates for this interaction model, allowing each covariate's effect to vary across study visits. This model accommodates dynamic changes in the strength or direction of associations over time, such as those that might arise due to aging, treatment exposure, lifestyle modification, or disease progression. Modeling such interactions is critical in longitudinal contexts where covariate effects are hypothesized to evolve as individuals accumulate additional risk or respond to clinical interventions.

To formally assess the joint contribution of each predictor (and its interactions) to model fit, Type 3 score statistics were computed. These global tests, which are robust to the working correlation specification, test whether a given predictor or set of predictors has any statistically significant effect on the outcome. Table 27 presents the score statistics from the GEE model without interactions, allowing assessment of

the overall impact of each covariate across the study period under the proportional odds constraint. Table 28 provides the analogous tests from the interaction model, evaluating the added contribution of time-varying effects. These score statistics, comprising degrees of freedom, chi-square values, and p-values, offer a model-based framework for hypothesis testing under the GEE paradigm, informing decisions about the inclusion or exclusion of interaction terms based on their collective explanatory value.

Taken together, the GEE modeling framework provides a robust, semi-parametric approach for estimating marginal associations between time-varying covariates and repeated ordinal outcomes. By incorporating both threshold-specific diagnostics (Table 24) and flexible interaction models (Tables 25–28), we evaluated not only the stability of covariate effects over time but also the adequacy of the proportional odds assumption in a longitudinal setting. These models serve as a complementary analytic tool to mixed-effects models, offering insight into average population-level trends in cumulative cardiometabolic burden while adjusting for within-subject correlation without requiring specification of individual-level random effects. This dual approach, combining both marginal and conditional perspectives, enhances the robustness and interpretability of longitudinal inferences in studies of multimorbidity progression.

Results

Descriptive Characteristics of the Study Sample

Table 1 summarizes baseline and longitudinal characteristics of the 2,000 study participants across three visits. The sample was 63.9% female, with education levels relatively balanced across categories: 36.7% had less than a high school education, 26.9% completed high school, and 36.5% reported some college or higher. At baseline, 25.6% screened positive for depressive symptoms based on the CES-D, and 62.1% met 2008 physical activity guidelines.

Cardiometabolic outcomes showed marked progression over time. Diabetes prevalence increased from 19.9% at Visit 1 to 36.7% at Visit 3; similarly, high cholesterol rose from 43.6% to 47.9%, and hypertension from 29.9% to 40.8% over the same period. This progression was reflected in the composite

comorbidity score: the proportion of participants reporting all three conditions increased from 7.2% at Visit 1 to 13.9% at Visit 2, before declining slightly to 7.5% at Visit 3, while those reporting zero conditions dropped steadily from 39.5% to 19.9%.

Risk factor profiles also evolved longitudinally. Obesity prevalence increased from 42.4% to 49.1%, and mean BMI rose from 29.84 to 30.69 kg/m² over the three visits. Alcohol use was reported by 46.5% of participants at baseline, peaking at 56.0% at Visit 2 before declining slightly. Smoking prevalence declined across visits, from 17.1% at Visit 1 to 12.8% at Visit 3. The average participant age increased from 47.8 to 58.8 years across the study period, while average dietary quality, measured via the HEI-2010, was 60.6 (SD = 13.8) and remained stable.

These descriptive statistics provide context for the longitudinal modeling of cardiometabolic burden, revealing temporal trends in multimorbidity and associated risk exposures.

Exploratory Data Analysis and Descriptive Visualizations

Figure 1 depicts the empirical distributions of three continuous predictors fundamental to modeling cardiometabolic multimorbidity longitudinally: age, body mass index (BMI), and the Healthy Eating Index–2010 (HEI-2010). These variables capture essential demographic, anthropometric, and behavioral risk domains and were pooled across visits to evaluate their range, shape, and potential outliers. A clear understanding of their empirical characteristics is critical for guiding model development, validating assumptions, and interpreting their roles in the cumulative cardiometabolic burden.

Age exhibited an approximately normal distribution with a slight right skew, spanning from 20 to 100 years. The majority of participants were clustered between 40 and 70 years, with a median of 55 and mean of 53.5 years. This age range aligns well with established epidemiology, where cardiometabolic conditions become more prevalent in midlife and older adulthood due to accumulating vascular and metabolic changes. The incremental age increases observed across study visits reflect the deterministic

nature of age progression and highlight its dual role as a time-varying risk marker and potential effect modifier in longitudinal disease models.

The distribution of HEI-2010 scores, a validated measure of diet quality relative to federal dietary guidelines, was also approximately normal, with a mean of 60.6 and a broad range from 10.9 to 97.7. While most participants reported moderate to high diet quality, a noticeable left tail indicated a subset with substantially poorer dietary habits, which are linked to elevated cardiometabolic risk. As a modifiable behavioral risk factor, HEI-2010 provides important insight into lifestyle influences on multimorbidity development, especially when considered alongside other health behaviors like smoking and physical activity.

BMI showed the greatest skewness, characterized by a pronounced right tail extending beyond 60 kg/m^2. The modal value was near 30 kg/m^2, consistent with the clinical cutoff for obesity, and a large proportion of participants fell between 25 and 40 kg/m^2. This pattern reflects a high baseline prevalence of overweight and obesity, both established risk factors for insulin resistance, systemic inflammation, and hypertension. Table 2 highlights the highest BMI values per visit, including one participant with repeated measures exceeding 70 kg/m^2. Although biologically plausible, such extreme values can unduly influence model estimates; therefore, the single maximum value (70.9) was excluded from multivariable analyses to improve estimation stability without distorting the overall exposure distribution. Remaining high values were retained to preserve clinically relevant variability and capture obesity-related risk effects adequately.

Taken together, the distributions of age, BMI, and diet quality illustrate the heterogeneous risk landscape within the cohort, underscoring their analytic importance as core predictors in cumulative cardiometabolic burden models.

Because both age and BMI were collected at each visit, they qualify as time-varying variables, though only BMI necessitated a formal assessment of longitudinal variability and reliability. Age increases deterministically over time, serving primarily as a control variable to reflect cohort aging. Validation

confirmed monotonic age progression without coding errors or temporal inconsistencies. In contrast, HEI-2010 was assessed only at baseline and treated as a time-invariant covariate. BMI, by contrast, represents a dynamic, modifiable risk factor influenced by physiological changes and behavioral factors, making it critical to quantify its temporal stability before repeated-measures modeling.

Table 4 presents BMI variability and reliability metrics, including the intraclass correlation coefficient (ICC) of 0.92, which indicates that most variation in BMI arises from differences between individuals rather than within-subject fluctuations. The observed within-subject variance (2.73 kg/m^2) confirms that BMI trajectories are not static, supporting its inclusion as a time-varying predictor with random intercepts to account for baseline heterogeneity. This quantification informs appropriate longitudinal model specification by capturing both stable individual differences and evolving risk contributions over time.

Table 3 further characterizes longitudinal dynamics by summarizing within-subject categorical transitions for cardiometabolic conditions and health behaviors between consecutive visits. These transition frequencies elucidate evolving comorbidity and exposure profiles that shape cumulative cardiometabolic burden.

Diabetes prevalence rose steadily: 10.6% of participants developed diabetes between Visits 1 and 2, with an additional 8.2% incident cases by Visit 3. Remission was rare (1.9%) and absent in the later interval, consistent with diabetes's progressive nature and reinforcing its role in cumulative disease accumulation. In contrast, high cholesterol exhibited substantial bidirectional movement: 14.4% of participants newly developed high cholesterol between Visits 1 and 2, while 10.4% improved. Between Visits 2 and 3, nearly one-quarter changed status in either direction, likely reflecting pharmacological interventions, lifestyle changes, or diagnostic reclassification.

Hypertension followed a similar but less volatile pattern, with 14.4% transitioning to hypertensive status and 3.3% improving between Visits 1 and 2, followed by nearly symmetrical transitions (~22%) in both

directions between Visits 2 and 3. This dynamic behavior highlights the need to treat these conditions as time-varying and potentially reversible factors influencing cardiometabolic burden.

Health-related behaviors also shifted notably over time. Alcohol use increased sharply by 18.0% between Visits 1 and 2, with 8.6% decreasing use; between Visits 2 and 3, 28.3% reduced consumption while 20.9% increased, illustrating greater behavioral fluctuation later in follow-up. Smoking status was more stable but dynamic: net declines in smoking narrowed from 3.6% between Visits 1 and 2 to 0.8% thereafter. These behavioral patterns carry clinical importance, given their established links to cardiometabolic risk and underscore their relevance as time-varying predictors.

BMI-based obesity classification showed modest but consistent increases over time. New obesity incidence rose from 7.2% to 7.4% across intervals, while remission decreased from 6.0% to 1.9%. This net upward trend aligns with known age-related weight gain and emphasizes obesity's growing contribution to worsening multimorbidity.

The composite cardiometabolic outcome, aggregating conditions from zero to three, demonstrated substantial turnover. Between Visits 1 and 2, nearly 30% of participants experienced increased comorbidity, while 11.6% improved; between Visits 2 and 3, worsening increased to 35.9%, alongside a notable 28.7% showing improvement. This bidirectional pattern highlights the need for flexible models that capture both progression and remission rather than assuming unidirectional disease accumulation.

Figure 2 complements these results by illustrating the distribution of BMI, age, and diet quality stratified by comorbidity burden. BMI distributions showed modest differences across comorbidity levels, with median values clustering between 29 and 32 kg/m² and overlapping interquartile ranges. Greater outlier dispersion appeared in higher comorbidity groups, suggesting a complex, potentially nonlinear relationship between adiposity and multimorbidity that longitudinal modeling must accommodate.

Age displayed a pronounced gradient at baseline, with median ages increasing steadily from the zero- to three-condition groups. This gradient diminished over time as aging cohorts converged in disease burden,

reflecting both temporal dynamics and the influence of increasing multimorbidity prevalence in older participants.

Diet quality distributions remained stable across comorbidity strata, with a modest decrease in median HEI-2010 among those with two or three conditions. This subtle pattern indicates a diffuse but potentially important association between diet and cardiometabolic burden, emphasizing the need to consider diet alongside clinical and behavioral factors.

Figure 3 examines three categorical predictors includes alcohol use, obesity status, and smoking status, stratified by comorbidity across visits. Alcohol users consistently had higher proportions in lower burden categories, though this may partly reflect reverse causation or selection effects. Obese participants showed a persistently higher cardiometabolic burden, with increasing representation in two- and three-condition groups over time, affirming obesity's central role in multimorbidity progression. Smoking status differences were minimal at baseline but became pronounced by Visit 3, with smokers more likely to have multiple conditions, consistent with cumulative adverse effects.

Figure 4 presents comorbidity distributions by education, physical activity adherence (PAG2008), depression status, and sex. Education revealed a steep gradient at Visit 1: over 40% of those with college or higher education were free of conditions, whereas participants with less than high school education showed more dispersed and heavier comorbidity burdens. Although disparities narrowed over time, lower-education groups retained higher proportions with multiple conditions, highlighting education's role as a persistent social determinant of cardiometabolic health.

Physical activity adherence consistently associated with more favorable profiles. Active individuals had higher zero-condition proportions and fewer with multiple conditions at all visits, reinforcing physical activity's protective effect and supporting its inclusion as a time-varying behavioral predictor.

Depression status showed pronounced differences, with non-depressed participants consistently exhibiting lower comorbidity burden. At Visit 1, over 40% of non-depressed individuals reported no conditions,

whereas those with depression had higher proportions in one- and two-condition categories. This divergence widened by Visit 3, where depressed participants had only 15% with zero conditions but 35% with two and about 10% with three conditions, emphasizing depression's strong association with increased multimorbidity.

Sex differences were also evident early: males had higher prevalence of one and two conditions, while females were more likely to be free of conditions. However, distributions converged by Visit 3, suggesting initial sex-based disparities diminish over time, underscoring the importance of treating sex as a dynamic factor in longitudinal cardiometabolic modeling.

Predictor Collinearity and Associations, both Cross-sectionally and Longitudinally

Before fitting multivariable longitudinal models, assessing collinearity among key predictors is crucial, as strong correlations can affect model stability and interpretability. Table 5 reports cross-sectional associations for ten selected predictor pairs across all three visits, chosen based on prior knowledge. For dynamic pairs measured repeatedly (Visits 1, 2, and 3), statistics and p-values are presented separately by visit; for stable or time-invariant pairs, a single overall statistic is reported.

Appropriate association measures were applied according to variable type: Spearman rank correlations for continuous-continuous pairs (e.g., BMI and HEI-2010) to capture monotonic relationships; point-biserial correlations for binary-continuous pairs (e.g., obesity status and BMI); chi-squared tests of independence for binary-binary or ordinal-binary pairs (e.g., smoking status and alcohol use, education level and physical activity adherence); and Spearman correlations for ordinal-continuous pairs (e.g., education level and HEI-2010). This tailored approach ensures accurate characterization of predictor interrelations critical for longitudinal modeling.

The results confirm a very strong positive correlation between BMI and obesity status at all visits (ranging from 0.75 to 0.77, all $p < 0.0001$), reflecting the clinical cutoff definition of obesity based on BMI thresholds. This near-deterministic relationship suggests that including both predictors concurrently

could introduce multicollinearity and warrants careful consideration in model specification. Conversely, HEI-2010 demonstrated a consistently weak but statistically significant negative correlation with BMI at Visits 2 and 3 (approximately –0.085; $p = 0.0001$), and a borderline significant negative association at Visit 1, indicating that poorer diet quality modestly corresponds to higher adiposity in this cohort. This supports the conceptualization of diet quality and BMI as related but distinct contributors to cardiometabolic risk.

Behavioral risk factors exhibited notable, temporally dynamic associations. Alcohol use and smoking status were strongly associated at baseline and Visit 2 ($\chi^2 = 46.0$ and 25.5; $p < 0.0001$), demonstrating behavioral clustering common in epidemiologic studies, but this association diminished and lost significance by Visit 3, signaling divergence in these behaviors over time. Similarly, sex was significantly linked with smoking and alcohol use at early visits ($p < 0.0001$), reflecting established sex differences in these behaviors, but these associations attenuated with follow-up, suggesting evolving behavioral patterns or cohort effects.

Age showed minimal correlation with BMI across visits, but consistently demonstrated a small positive association with HEI-2010 ($\rho \sim 0.15$, $p < 0.0001$), implying that older participants tend to report marginally better diet quality. Depression status was associated with smoking and alcohol use at baseline, though only the relationship with alcohol use persisted across visits. Lastly, education level exhibited weak negative correlation with HEI-2010 and a modest positive association with physical activity adherence, highlighting the influence of socioeconomic factors on health behaviors.

Collectively, Table 5 indicates that while certain predictors exhibit strong, potentially problematic correlations (notably BMI and obesity), most associations are moderate or weak and vary over time. These findings underscore the importance of modeling these variables as distinct but potentially interrelated contributors, employing longitudinal approaches that accommodate their dynamic relationships while mitigating multicollinearity risks.

To complement the cross-sectional diagnostics in Table 5, Tables 6 through 8 summarize longitudinal associations among key time-varying predictors. These repeated-measures correlations capture both within-subject and between-subject dependencies over time, providing essential input for specifying multivariable GEE and mixed-effects models. Longitudinal correlation estimates not only clarify which variables may co-vary over time but also guide decisions about model structure, inclusion of interaction terms, and multicollinearity risk in repeated-outcomes frameworks.

Table 6 reports the point-biserial correlation between BMI and obesity status, calculated across repeated measures. The estimated mean BMI difference between obese and non-obese participants was 4.07 kg/m^2, yielding a high correlation of 0.90 due to the definitional link between the two measures. Given the redundancy, including both in the same longitudinal model would introduce collinearity. As BMI offers greater granularity and reflects dynamic physiological change, it is preferred over the binary obesity indicator in time-varying modeling.

Table 7 summarizes repeated-measures associations between continuous variables using linear mixed-effects models, providing estimates of subject-level variance, residual variance, and intraclass correlation coefficients (ICCs). The association between BMI and HEI-2010 was modestly negative ($\beta = -0.0384$, $p < 0.0001$), suggesting slightly lower BMI among participants with better diet quality. The high ICC for BMI (0.921) confirms that most variance is between individuals, though meaningful within-person variation supports treating BMI as a dynamic covariate.

The relationships between age and HEI-2010 ($\beta = 0.1418$, $p < 0.0001$) and age and BMI ($\beta = 0.3426$, $p < 0.0001$) indicate that older individuals tend to report better diet and higher BMI, respectively. The reciprocal age–BMI model also reached significance ($\beta = 0.0622$), though age is deterministic over time. These results reinforce age as a key time-varying control and suggest plausible cross-time influence between diet, weight, and aging.

Table 8 evaluates longitudinal associations between binary predictors using GEE models, reporting both working correlations and Z-statistics. Since both variables in each pair may vary across visits, bidirectional associations were computed. However, in several cases, most notably where stable variables like sex or depression were modeled as outcomes estimated working correlations reached 0.9999, a clear artifact of modeling a time-invariant variable as if it varied over time. These inflated values are not interpretable and are excluded from substantive inference.

Instead, focus is placed on associations with plausible temporal dynamics and realistic correlation estimates. The relationship between smoking and alcohol use was significant in both directions ($Z = –6.33$ and $–6.49$, $p < 0.0001$), with working correlations of 0.2974 (smoking on alcohol) and 0.1626 (alcohol on smoking). The stronger association when smoking is treated as the outcome suggests that alcohol use may be a more temporally proximal or influential behavior in shaping tobacco use trajectories.

Similarly, when modeling smoking status as a function of sex, the working correlation was 0.2910, and for alcohol use on sex, it was 0.1543, both consistent with higher risk behavior prevalence among males. These correlations are interpretable and highlight sex as a useful static predictor of time-varying exposures.

The relationship between depression and smoking yielded a working correlation of 0.2709 when smoking was modeled as the outcome. This indicates a moderate association and supports inclusion of depression status as a potential time-varying risk factor in models of behavioral trajectories. Similarly, modeling alcohol use as a function of depression yielded a correlation of 0.2974, reinforcing this directional dependency. In contrast, modeling depression as the response again resulted in a non-interpretable correlation of 0.9999 due to its limited variation over time.

In summary, Tables 6–8 collectively underscore the importance of aligning model directionality with the temporal characteristics of each variable. Time-varying exposures (e.g., smoking, alcohol use, BMI) are most appropriately modeled as outcomes when paired with more stable or exogenous predictors (e.g., sex,

depression), avoiding artificial inflation of working correlation estimates. The interpretable correlations in the 0.2–0.3 range provide empirical justification for their joint inclusion in longitudinal modeling frameworks, while also guiding decisions on controlling for interdependence and assessing potential effect modification in mixed-effects and GEE models.

Diagnostic Testing and Inference in Cross-Sectional Data

After detailed correlation analysis both in cross-sectional and longitudinal structure, we next present a cross-sectional analysis of cumulative comorbidity burden at each visit, treating the data from each time point independently to establish baseline associations. The outcome of interest is an ordinal variable reflecting the total number of diagnosed conditions of diabetes, high cholesterol, and high blood pressure, with possible scores ranging from 0 to 3. At each visit, the cumulative comorbidity score was modeled using a cumulative logit (ordinal logistic) model, incorporating a consistent set of predictors: sex; education level 2 and education level 3 (with level 1 as the reference); continuous HEI2010 (Healthy Eating Index); binary PAG2008 (meeting physical activity guidelines); binary depression; continuous age; binary alcohol use; continuous BMI (not dichotomized into obesity); and smoking status. A key assumption of the ordinal logistic model is the proportional odds (PO) assumption, which posits that the effect of each predictor remains constant across the cumulative logits of the outcome. To ensure the validity of this assumption and guide model specification, we evaluated both graphical and statistical diagnostics at each visit.

Figure 5 presents the threshold-specific log-odds coefficients and corresponding 95% confidence intervals for a set of predictors modeled across three cumulative thresholds (≥1, ≥2, =3) of an ordinal outcome variable. Each panel corresponds to a distinct cross-sectional visit (from left to right: Visit 1, Visit 2, and Visit 3), and each predictor is estimated using separate binary logistic regressions to visualize how its effect varies across increasing outcome severity. Under the proportional odds (PO) assumption, the effect of each predictor should remain constant across thresholds, and thus all three colored points (and their

intervals) should align horizontally. Deviations from this pattern suggest non-proportionality, where the predictor's influence on the odds of higher outcomes depends on the threshold considered.

To statistically evaluate this assumption, Table 9 reports results from the Brant test, a score-based diagnostic specifically designed for ordinal logistic regression models. For each predictor, the Brant test compares the estimated coefficients from multiple binary logistic regressions corresponding to each threshold, testing whether they differ significantly from one another. Formally, the null hypothesis of the Brant test states that the regression coefficients for a given predictor are equal across all cumulative logits, that is, the PO assumption holds. A statistically significant test result ($p < 0.05$) indicates that this null hypothesis is rejected, and that the PO assumption is violated for that variable. The omnibus test assesses this hypothesis jointly across all predictors in the model. Brant test results are particularly useful for identifying which specific predictors may require relaxation of the PO assumption in a partial proportional odds model, allowing some coefficients to vary while keeping others constrained.

Complementing the Brant test, Table 10 provides results from a Likelihood Ratio Test (LRT) comparing two nested models: (1) a proportional odds ordinal logistic regression model using the cumulative logit link, and (2) a non-proportional alternative using the generalized (nominal) logit link that allows each category to have its own set of coefficients. The null hypothesis in this context is that the more constrained proportional odds model fits the data as well as the fully unconstrained nominal model. A significant LRT ($p < 0.05$) suggests that relaxing the PO assumption improves model fit, indicating the presence of global non-proportionality. The LRT thus provides model-level evidence of whether the proportional structure is tenable, while the Brant test offers variable-specific guidance.

At Visit 1, visual inspection of Figure 5 indicates clear violations of the PO assumption for age, sex, and both levels of education, all of which show substantial shifts in effect sizes or direction across thresholds. Depression and alcohol use also show varying slopes and non-overlapping confidence intervals, suggesting potential non-proportionality. These findings are strongly supported by the Brant test results in Table 9, where age ($p < 0.01$) and sex ($p < 0.01$) significantly violate the PO assumption, and depression

and alcohol use present marginal p-values (< 0.3), indicating possible partial violations. In contrast, variables such as HEI2010, PAG2008, BMI, and smoking status exhibit relatively stable coefficients across thresholds with overlapping confidence intervals, and none show signs of violation in the Brant test (all $p > 0.4$). Therefore, we treat HEI2010, PAG2008, BMI, and smoking status as proportional in the partial PO model for Visit 1.

At Visit 2, Figure 5 reveals strong non-proportionality for BMI and age, with visibly diverging coefficient patterns across thresholds of particularly a steep increase for BMI at the third threshold. Education level 3 also demonstrates a notable decrease in effect at the highest threshold, and alcohol use shows a gradient decline in coefficients across thresholds. Table 9 supports these observations with age ($p < 0.01$) being statistically significant, while BMI ($p = 0.13$) and alcohol use ($p = 0.32$) show suggestive but non-significant evidence of violation. Meanwhile, sex, HEI2010, PAG2008, depression, and smoking status display consistent log-odds estimates across thresholds with overlapping intervals and Brant p-values well above 0.4, justifying their inclusion as proportional variables in the partial PO model for Visit 2.

At Visit 3, while Table 9 suggests no significant Brant test violations (omnibus $p = 0.59$), Figure 5 reveals nuanced departures from proportionality that support partial modeling. Smoking status exhibits a pronounced jump in effect size at the third threshold, sex shows a gradual increase in log-odds across thresholds, and BMI demonstrates clear variation, all indicating violations despite their Brant p-values not reaching significance. Physical activity (PAG2008) also has a confidence interval fully below zero at one threshold, suggesting potential non-proportionality. Conversely, education, HEI2010, depression, age, and alcohol use all show relatively flat or overlapping log-odds coefficients across thresholds, with consistent direction and magnitude. Given this graphical evidence, these five variables were selected as proportional in the partial PO model for Visit 3.

Lastly, Table 10 further confirms these modeling decisions through Likelihood Ratio Tests comparing ordinal (proportional odds) and nominal (non-proportional) logistic models. For Visit 1 ($p < .0001$) and Visit 2 ($p = 0.0011$), the significant LRT results indicate that relaxing the PO assumption significantly

improves model fit, justifying the use of partial PO models. For Visit 3, however, the non-significant LRT ($p = 0.5974$) suggests that the fully proportional model may be adequate in a global sense, but our variable-level assessment via Figure 5 still supports partial modeling to account for specific deviations.

Following the evaluation of the proportional odds assumption, we further examined whether the continuous covariates of age, HEI2010, and BMI satisfy the logit-linearity assumption, which underlies cumulative logit models. Figure 6 displays LOESS-smoothed curves of predicted log-odds against each continuous variable, estimated separately for each threshold of the ordinal outcome (≥ 1, ≥ 2, $=3$). The top panel demonstrates that age maintains a clear linear relationship with log-odds across all three visits, with nearly parallel and smooth curves. One extreme value, a participant aged 100 at Visit 3, appears as a non-clustered outlier, but does not substantially distort the overall linear pattern. In the middle panel, HEI2010 exhibits near-flat LOESS trends at all thresholds, suggesting a minimal overall effect and a potential threshold relationship where only extreme values may be influential. Predicted log-odds are highly overlapped across HEI2010 values in Visits 1 and 2, especially between 30 and 80, slightly reducing interpretability in that range, while Visit 3 displays more separation between thresholds. Despite this, no strong evidence of non-linearity is observed, and HEI2010 is retained as a continuous covariate in this preliminary analysis. The bottom panel shows BMI with a generally increasing and monotonic LOESS curves across all visits, supporting approximate linearity. Some mild curvature is observed below 25 and above 45 at Visit 1, and predicted log-odds display greater spread as BMI increases. These trends are smoother in Visit 2 and nearly linear in Visit 3. Although minor deviations from linearity are present, particularly for BMI at extreme values, all three continuous predictors reasonably satisfy the linearity assumption and are treated as continuous variables in the final models.

Having established the suitability of the partial proportional odds framework and verified the logit-linearity of continuous covariates, we now turn to cross-sectional model inferences across all three visits, using the fitted partial proportional odds logistic regression models presented in Tables 11–13. These models accommodate threshold-specific effects for predictors that violated the proportional odds

assumption while retaining the proportional structure for covariates demonstrating consistent effects across outcome thresholds. The resulting models provide flexible yet interpretable estimates of how demographic, behavioral, and psychosocial factors relate to increasing levels of cumulative cardiometabolic burden.

Across Tables 11 to 13, partial proportional odds models identified both consistent and changing predictors of cumulative cardiometabolic comorbidity. Male sex was significantly associated with higher comorbidity at Visits 1 and 2 (Visit 1 ORs: 2.08 to 1.67; Visit 2 OR = 1.67; all $p < 0.01$), but this association was not significant at Visit 3. Age showed a consistent inverse association at all visits, with each additional year reducing odds of greater comorbidity by 3% to 8% across thresholds (all $p < 0.001$).

Alcohol use exhibited threshold-specific effects at Visits 1 and 2, increasing odds of moderate and severe comorbidity (Visit 1 ORs: 1.48 and 2.31; Visit 2 ORs: 1.33 and 1.62; $p \leq 0.005$), but showed no significant association at Visit 3. Education demonstrated variable effects: higher education was associated with increased odds of severe comorbidity at Visit 1 (ORs 1.61 and 1.71; $p \leq 0.01$), while at Visit 2, high school education was associated with lower odds of any comorbidity (OR = 0.71; $p = 0.013$); education effects were not significant at Visit 3.

Depression was inversely associated with comorbidity at Visit 1 (OR = 0.70; $p = 0.0035$), marginally so at Visit 2, and more strongly at Visit 3 (OR = 0.68; $p < 0.001$). Physical activity (PAG2008) was not significant at Visits 1 or 2 but was positively associated with higher comorbidity at Visit 3 thresholds ≥2 and ≥3 (ORs 1.25 and 1.44; $p < 0.05$). BMI consistently showed inverse associations at lower thresholds across visits (7–10% reduced odds per unit increase; $p < 0.001$), but at Visit 3 demonstrated a U-shaped relationship with increased odds at the highest comorbidity level (OR = 1.04; $p < 0.001$). Smoking status was only significant at Visit 3, inversely associated with highest comorbidity (OR = 0.62; $p = 0.025$). Diet quality (HEI2010) showed no significant associations at any visit.

These results indicate that sex and age are stable predictors of cardiometabolic burden early in follow-up, while other factors, including behavioral and psychosocial variables, exhibit threshold- and visit-specific effects. The evolving patterns underscore the need for longitudinal modeling to capture dynamic risk factor relationships relevant to cardiometabolic comorbidity.

In support of this, model diagnostics presented in Table 14 affirm the adequacy of the partial proportional odds logistic regression models at each visit. Residual deviance values of 4513.15 (df = 5974), 4813.35 (df = 5976), and 4922.04 (df = 5978) for Visits 1, 2, and 3, respectively, were each substantially lower than their corresponding degrees of freedom, suggesting no evidence of gross model misfit. Modest increases in model selection criteria over time (AIC: 4551.15 to 4960.04; BIC: 4657.56 to 5066.45) reflect growing model complexity and variation in covariate structure across visits.

Outlier detection, based on Pearson residuals exceeding an absolute value of 4, identified a relatively small number of influential observations: 29 at Visit 1, 11 at Visit 2, and 40 at Visit 3, out of 1,999 total participants. The most extreme residuals of 6.38 at Visit 1, 5.23 at Visit 2, and 5.99 at Visit 3, were concentrated in higher comorbidity burden categories but represented only a minor portion of the dataset, with negligible impact on overall model validity.

Figure 7 provides complementary graphical diagnostics of model fit across visits, including plots of Pearson residuals versus fitted values, normal Q-Q plots, and residuals by observation index. For Visits 1 and 2, the residual-versus-fitted plots reveal modest curvature and heteroscedasticity, suggesting some departures from model assumptions. By Visit 3, the residual distribution appears more stable and homoscedastic, indicating improved calibration. Q-Q plots reveal right-skewed residuals across visits, particularly at Visit 1, but alignment improves markedly by Visit 3. The residuals plotted against observation index appear randomly scattered, without visible dependence or temporal structure, consistent with the assumption of independence.

Together, these diagnostics reinforce the robustness of the model fits while highlighting subtle shifts in covariate behavior and residual structure across time. These patterns further justify the need for longitudinal modeling approaches that account for within-subject correlation, time-varying effects, and dynamic risk profiles, analytical capabilities that extend beyond cross-sectional or visit-specific models and are essential for capturing the full complexity of cardiometabolic comorbidity trajectories.

Longitudinal Threshold-Specific Mixed-Effects Models for Evaluating Subject-Level Heterogeneity and Proportional Odds

To explore longitudinal changes in cardiometabolic burden while accounting for subject-level heterogeneity, we estimated a series of multinomial generalized linear mixed models (GLMMs) using a cumulative logit link. These models incorporate random intercepts to capture within-subject correlation across repeated measures and allow for the assessment of fixed effects on the ordinal outcome scale.

For this section, we assume the proportional odds assumption holds, meaning the effect of each covariate is constrained to be constant across all cumulative thresholds of the outcome. We first present a model excluding interaction terms between visit and covariates (Table 18), which serves as the baseline specification under this assumption. To further examine potential effect modification over time, we then estimate an expanded model that includes interaction terms between visit and key covariates (Table 20), while still maintaining the proportional odds structure within each visit. This approach provides a structured framework to assess temporal dynamics and individual variation without relaxing the PO assumption.

Before drawing inference from the proportional odds cumulative logit mixed-effects models, we conducted diagnostic checks to evaluate the appropriateness of the proportional odds (PO) assumption across covariates. While the main models presented below assume proportional odds for parsimony and interpretability, these diagnostics serve to assess whether that assumption holds uniformly across predictors, and whether caution is warranted when interpreting parameter estimates. Table 15 presents

threshold-specific fixed effects estimates from cumulative binary logistic mixed models, one for each dichotomization of the ordinal outcome. Consistency of covariate effects across thresholds supports the PO assumption; in contrast, meaningful variation indicates possible violation. Though we proceed with models assuming proportionality, these results offer valuable context for evaluating the robustness of our findings.

Visit demonstrates clear deviations from proportionality. At threshold 1, Visit 2 has a negligible effect (Estimate = 0.032, $p = 0.7110$), but its association strengthens significantly at thresholds 2 (Estimate = 0.234, $p = 0.0120$) and 3 (Estimate = 0.541, $p < 0.0001$). Visit 3 displays an even stronger non-proportional pattern: a large positive effect at threshold 1 (Estimate = 0.550, $p < 0.0001$), a null effect at threshold 2 (Estimate = –0.082, $p = 0.4234$), and a strong negative effect at threshold 3 (Estimate = –1.023, $p < 0.0001$). This reversal in directionality across thresholds suggests visit-specific changes in risk that violate the PO assumption. Similarly, sex (female) shows diminishing protective effects across increasing thresholds, with estimates decreasing in magnitude from –0.763 (threshold 1, $p < 0.0001$) to –0.453 (threshold 2, $p < 0.0001$) and –0.433 (threshold 3, $p = 0.0074$).

Other covariates also exhibit evidence of threshold-specific behavior. For alcohol use, the association with lower comorbidity strengthens across thresholds, from borderline significance at threshold 1 (Estimate = –0.149, $p = 0.0640$) to significant negative effects at threshold 2 (Estimate = –0.228, $p = 0.0061$) and threshold 3 (Estimate = –0.437, $p = 0.0005$). BMI has consistent positive associations at thresholds 1 and 2 (Estimates = 0.079 and 0.089, both $p < 0.0001$), but the effect sharply increases and becomes statistically nonsignificant at threshold 3 (Estimate = 0.282, $p = 0.1185$), raising concerns of nonlinearity or disproportionate influence. Depression similarly shows strong associations at lower thresholds (threshold 1: Estimate = 0.419, $p = 0.0003$; threshold 2: Estimate = 0.363, $p = 0.0031$), but the effect diminishes substantially at threshold 3 (Estimate = 0.126, $p = 0.4594$), indicating a potential violation of PO.

Additional patterns are observed for education and smoking. For education, the college-or-above group (Education = 3) shows an increasingly protective effect with increasing thresholds (from –0.133 to –0.345), with the middle threshold reaching statistical significance ($p = 0.0427$) and the highest approaching significance ($p = 0.0514$). This trend suggests that education level may differentially impact comorbidity across severity levels. Smoking status, while not significant at any threshold, presents a small positive effect only at the highest threshold (Estimate = 0.282, $p = 0.1185$), hinting at a possible delayed or threshold-specific risk. Collectively, these findings underscore the possibility of PO violations for selected covariates and motivate future sensitivity analyses or partial proportional odds models. Nonetheless, for the purposes of model comparison and summary in the following sections, we proceed under the full proportional odds assumption.

Building upon the diagnostic results assessing whether the covariates in the study satisfy the proportional odds assumption within a GLMM framework, we further evaluated model specifications by comparing the fit of a mixed-effects cumulative logit model with a subject-specific random intercept to that of a generalized linear model including only fixed effects. As shown in Table 16, the mixed-effects model demonstrated substantially better fit, with a lower Akaike Information Criterion (AIC = 13,661 vs. 14,339) and Bayesian Information Criterion (BIC = 13,750 vs. 14,439), confirming that the inclusion of a random intercept improves model adequacy by accounting for subject-level heterogeneity. To verify the plausibility of the random intercept assumption, we examined the distribution of the empirical Bayes estimates. Figure 8 presents a histogram and Q–Q plot of these estimates. The histogram suggests approximate normality, and the Q–Q plot shows that the empirical quantiles closely follow the theoretical quantiles along the diagonal line, with minimal departure from normality at the tails. Together, these diagnostics support both the need for and appropriateness of a normally distributed random intercept. Based on the investigation, we interpret this as strong justification for retaining the random intercept structure in all subsequent mixed-effects models and for attributing between-subject variability to differences not captured by the fixed effects alone.

Building on model validation and the assumption of proportional odds, we next examined fixed effects from the baseline cumulative logit mixed model, which excluded interaction terms between visit and other covariates (Table 18). This modeling approach offers a parsimonious yet interpretable framework for characterizing longitudinal progression in cardiometabolic burden across ordered outcome categories. By adjusting for relevant sociodemographic and behavioral factors while accounting for within-subject correlation via a random intercept, the model enables estimation of population-average effects under the constraint that each covariate exerts a consistent influence across cumulative thresholds of disease burden.

In this context, Visit 2 was associated with significantly lower odds of higher cardiometabolic burden relative to baseline (OR = 0.811, $p = 0.0020$), possibly indicating a modest early reduction or stabilization in risk. Visit 3, however, did not differ significantly from baseline (OR = 0.971, $p = 0.7010$), which may reflect heterogeneity in temporal dynamics or incomplete specification of time-related effects. Female sex was linked to increased odds of higher burden (OR = 1.765, $p < 0.0001$), suggesting underlying differences in disease accumulation, healthcare access, or health behavior reporting. Higher educational attainment (college or above) was associated with significantly reduced odds of greater burden (OR = 1.254, $p = 0.0239$), while no such effect was observed for intermediate educational levels, pointing to a potential threshold effect in the protective role of education.

Several behavioral and psychological factors also showed notable associations. Alcohol use corresponded to increased cardiometabolic burden (OR = 1.234, $p = 0.0007$), consistent with known metabolic effects. Depression showed an inverse association (OR = 0.729, $p = 0.0014$), which may appear counterintuitive but could reflect unmeasured behavioral changes or healthcare utilization patterns among individuals with depressive symptoms. Age and BMI, both modeled as continuous covariates, exhibited strong negative associations with burden (ORs = 0.924 and 0.918, respectively; both $p < 0.0001$), a pattern that diverges from clinical expectations and raises the possibility of nonlinearity, reverse causation, or survival bias in older or leaner individuals.

Smoking status, the Healthy Eating Index (HEI2010), and adherence to physical activity guidelines (PAG2008) did not show statistically significant associations in this specification. These findings align with the results from the Type III tests of fixed effects (Table 19), which provide joint significance assessments for each predictor while controlling for all others. Age ($F(1, 3990) = 398.52$, $p < .0001$), BMI ($F(1, 3990) = 141.60$, $p < .0001$), and sex ($F(1, 3990) = 38.40$, $p < .0001$) emerged as the strongest contributors to the model, reflecting their substantial influence on cardiometabolic risk classification over time. Visit also showed significant global association ($F(2, 3990) = 6.28$, $p = 0.0019$), reinforcing the presence of longitudinal variation. Depression ($F(1, 3990) = 10.18$, $p = 0.0014$), alcohol use ($F(1, 3990) = 11.53$, $p = 0.0007$), and education level ($F(2, 3990) = 3.19$, $p = 0.0411$) had moderate but statistically significant effects. In contrast, HEI2010, PAG2008, and smoking status were not statistically significant (all $p > 0.20$), further supporting their limited explanatory power in this specification.

Following the model without interaction terms between visit and the covariates (Table 18), we extended the analysis by incorporating interaction effects to explicitly model how the relationships between predictors and cardiometabolic burden evolve over time. In longitudinal studies, exposures and risk factors may exert differential impacts at distinct follow-up periods due to changes in physiology, behavior, treatment, or other time-dependent processes. Table 20 presents the fixed effects from a cumulative logit mixed-effects model that includes interactions between visit and key covariates, thereby enabling the detection of these time-varying effects. This modeling strategy acknowledges the non-static nature of risk factor associations in longitudinal data and provides a more nuanced understanding of cardiometabolic risk trajectories than models assuming constant effects across visits.

Including interaction terms between visit and covariates revealed important temporal changes in how demographic and behavioral factors influence cardiometabolic burden over the study period. At baseline, female participants had significantly higher odds of elevated cardiometabolic burden compared to males (estimate = 0.901, odds ratio = 2.46, $p < 0.0001$). However, the interaction between female sex and visit 3 was negative and significant (estimate = -0.683, odds ratio = 0.51, $p < 0.0001$), indicating that the

increased risk observed among females diminished by nearly half at the final visit. This suggests that sex differences in cardiometabolic burden are not constant but decrease over time, possibly reflecting changing physiological or behavioral factors. The interaction between female sex and visit 2 was not statistically significant, indicating stability of sex differences between the first and second visits.

Education showed a more complex pattern over time. While higher education (level 3) was associated with an increased cardiometabolic burden at baseline (estimate = 0.451, odds ratio = 1.57, $p = 0.0012$), its interaction with visit 2 was negative and statistically significant for education level 2 (estimate = -0.389, odds ratio = 0.68, $p = 0.0211$), indicating a decrease in risk at the second visit. Other education interactions with visits 2 and 3 were not significant, suggesting that the effects of education remained relatively stable after the second visit. Smoking status showed a consistent pattern of risk reduction over time, with baseline smoking increasing the odds of cardiometabolic burden (estimate = 0.314, odds ratio = 1.37, $p = 0.0342$), but interactions with visits 2 and 3 were significantly negative (estimates = -0.432, odds ratio = 0.65, $p = 0.0280$ and -0.596, odds ratio = 0.55, $p = 0.0042$, respectively), reflecting a decline in smoking-related risk during follow-up.

Age and body mass index showed nonlinear, time-varying effects. At baseline, both were negatively associated with cardiometabolic burden (age estimate = -0.098, odds ratio = 0.91, $p < 0.0001$; BMI estimate = -0.087, odds ratio = 0.92, $p < 0.0001$), and these negative associations became stronger at visit 2 (age interaction estimate = -0.020, odds ratio = 0.98, $p = 0.0012$; body mass index interaction estimate = -0.040, odds ratio = 0.96, $p = 0.0008$). However, at visit 3, these interactions reversed direction and were positive and highly significant (age interaction estimate = 0.066, odds ratio = 1.07, $p < 0.0001$; BMI estimate = 0.045, odds ratio = 1.05, $p < 0.0001$), indicating that at later follow-up, older age and higher body mass index were associated with increased odds of cardiometabolic burden. This complex pattern highlights the importance of modeling time-varying effects to accurately capture changing risk profiles that would be missed by models assuming constant effects.

Other covariates, including physical activity, depression, alcohol use, and the Healthy Eating Index, did not show statistically significant interactions with visit, indicating that their effects on cardiometabolic burden were stable over time. Their baseline associations were also generally weak or not significant. These results suggest that while some risk factors maintain consistent associations with cardiometabolic burden throughout follow-up, others, particularly sex, education, smoking, age, and body mass index, exhibit dynamic relationships that evolve across visits. Incorporating interactions between visit and covariates is therefore essential to fully characterize the longitudinal patterns of cardiometabolic risk.

Consistent with this, Table 21 presents Type III tests of fixed effects from a multinomial generalized linear mixed model with a cumulative logit link, incorporating interaction terms between visit and covariates. Compared to Table 19, which excludes interaction terms, the significance of main effects for sex, education, age, BMI, depression, and alcohol use remains similar, confirming these covariates as important predictors of cardiometabolic burden. The overall effect of visit is markedly stronger in Table 21 ($F(2, 3970) = 77.81$, $p < 0.0001$) than in Table 19 ($F(2, 3990) = 6.28$, $p = 0.0019$), reflecting additional variation explained by modeling interactions with time.

However, critical insight arises from the interaction tests, which formally assess whether the strength or direction of these effects changes over time. Specifically, the interactions of visit with sex ($F(2, 3970) = 13.42$, $p < 0.0001$), age ($F(2, 3970) = 112.76$, $p < 0.0001$), BMI ($F(2, 3970) = 28.51$, $p < 0.0001$), and smoking status ($F(2, 3970) = 4.68$, $p = 0.0094$) were statistically significant, indicating that the impact of these covariates on cardiometabolic burden varies meaningfully over time. The interaction between visit and education showed a trend toward significance ($F(4, 3970) = 3.18$, $p = 0.0691$), suggesting potential temporal modulation of education effects. In contrast, the non-significant interactions for Healthy Eating Index, physical activity, depression, and alcohol use reinforce the relative temporal stability of their associations.

Table 17 complements the fixed-effects findings by presenting the estimated variance components for the random intercept from multinomial generalized linear mixed models (GLMMs) both without and with

interaction terms between visit and covariates. The model without interactions shows a substantial random intercept variance estimate of 2.194 (SE = 0.15), indicating notable between-subject variability in baseline cardiometabolic burden when covariate effects are assumed constant over time. When interactions between visit and covariates are incorporated, this variance increases modestly to 2.478 (SE = 0.17), reflecting additional heterogeneity in individual trajectories that arises from allowing covariate effects to vary longitudinally. This increase suggests that the inclusion of time-varying effects captures more complex individual differences beyond those explained by static covariate effects. Both models therefore provide complementary insights: the model without interactions summarizes overall subject-level variability assuming stable covariate influences, while the interaction model reveals that allowing effects to change over time accounts for additional, nuanced variation in cardiometabolic burden trajectories. Together, these variance component estimates contextualize the fixed-effect results by quantifying the extent of heterogeneity captured under each modeling approach, highlighting the multifaceted nature of longitudinal cardiometabolic risk.

To summarize for the above analysis, we first estimated a baseline cumulative logit mixed-effects model assuming proportional odds and no interaction terms, which provided a parsimonious and interpretable framework for evaluating the overall influence of sociodemographic, behavioral, and psychological factors on cardiometabolic burden across time. In this specification, female sex, alcohol use, and higher education were associated with increased odds of greater burden, while depression, age, and BMI were linked to reduced odds; smoking status, physical activity, and dietary quality showed no significant associations. However, this model assumed static covariate effects and could not account for longitudinal effect modification.

To address this limitation, we subsequently estimated an extended cumulative logit mixed-effects model incorporating interactions between visit and key covariates. This approach explicitly modeled time-varying associations while retaining the proportional odds assumption within each visit. The interaction model revealed significant temporal variation in the effects of sex, age, BMI, smoking, and education.

Most notably, the elevated cardiometabolic burden associated with female sex and smoking diminished markedly by the final visit, while the protective effects of older age and higher BMI reversed direction over time. These findings were further supported by significant interaction terms in Type III tests and modest increases in the estimated random intercept variance, indicating increased heterogeneity captured through time-varying effects.

However, diagnostics presented in Table 15 revealed strong violations of the proportional odds (PO) assumption for key covariates, most notably *visit*, *sex*, and *BMI*, all of which showed substantial variation in effect size and direction across outcome thresholds. Additional violations were evident for *depression*, *alcohol use*, *education*, and *smoking*, each demonstrating threshold-specific patterns inconsistent with a constant-effects assumption. These findings indicate that a standard cumulative logit model assuming full proportionality may mask important heterogeneity in the relationships between predictors and cardiometabolic burden severity.

To address this, the next section implements a Bayesian cumulative logit mixed-effects model with partial proportional odds, which selectively relaxes the PO constraint for covariates that violate the assumption while retaining it for those that do not. This modeling strategy allows for more accurate estimation of covariate effects, preserves the ordinal outcome structure, and appropriately accounts for within-subject correlation through random intercepts. By incorporating partial proportional odds within a Bayesian framework, we enhance flexibility in modeling complex longitudinal risk profiles without sacrificing interpretability.

We now turn to the results from this model to examine how relaxing proportionality assumption impacts inference on cardiometabolic risk trajectories.

Bayesian Cumulative Logit Mixed-Effects Model with Partial Proportional Odds

To address the violations of the proportional odds (PO) assumption identified in threshold-specific diagnostics (Table 15), we employed a Bayesian cumulative logit mixed-effects model with partial

proportional odds using the brms package in R. This approach allows selected covariates to have threshold-specific (non-proportional) effects while maintaining proportional odds constraints for others, thereby providing a more flexible and accurate representation of covariate influences on the ordinal cardiometabolic burden outcome.

The Bayesian framework facilitates full posterior inference in a complex hierarchical setting, accommodating varying covariate effects and mitigating small-sample biases in higher-order parameters. Modeling the outcome with a cumulative logit link respects the ordinal nature of cardiometabolic burden categories, while subject-level normally distributed random intercepts capture within-subject correlation over repeated measures. Covariates with evidence of PO assumption violations, namely visit, sex, BMI, education, alcohol use, depression, and smoking, were modeled with threshold-specific effects via the cs() function in brms. In contrast, covariates without such evidence of age, adherence to physical activity guidelines (PAG2008), and dietary quality (HEI2010), were modeled under the proportional odds assumption.

The model was fit using 2 Markov chains with 2000 iterations each (1000 warm-up), employing 4 computational cores and an increased target acceptance rate of 0.95 to ensure stable convergence. Posterior summaries of fixed effects and subject-level random intercept variance are presented in Table 22. All Gelman-Rubin diagnostics ($\hat{R}$) equaled 1.00, confirming chain convergence, and effective sample sizes (ESS) for both the bulk and tails of posterior distributions exceeded 700 for all parameters, supporting reliable estimation of both central tendency and posterior tail behavior.

The subject-level random intercept standard deviation was estimated at 1.51 (95% credible interval [CI]: 1.42, 1.62), reflecting substantial between-subject heterogeneity in baseline cardiometabolic burden beyond fixed effects. This underscores the appropriateness of including random intercepts to account for repeated measures correlation.

Among covariates constrained by proportional odds, age exhibited a positive association with cardiometabolic burden (posterior OR = 1.08), contrasting with the negative effect observed in earlier fully proportional models. The 95% CI narrowly included one, indicating borderline statistical significance but improved alignment with clinical expectations that risk increases with age. PAG2008 demonstrated a modest, non-significant protective effect (OR = 0.89, 95% CI includes 1), consistent with the hypothesized benefit of physical activity, while HEI2010 showed a negligible effect (OR = 0.996), suggesting limited influence of dietary quality on burden within this cohort.

Covariates with non-proportional effects revealed more complex, threshold-dependent patterns. BMI displayed a monotonic increase in cardiometabolic burden risk across thresholds, with ORs rising from 1.08 at threshold 1 to 1.11 at threshold 3; all 95% CIs excluded 1, indicating robust associations and correcting the misleading negative effect detected previously under proportional odds constraints. Visit effects were also threshold-specific: Visit 2 effects increased with severity (ORs: 1.01, 1.27, 1.73), whereas Visit 3 effects showed elevated odds at threshold 1 (OR = 1.51), neutrality at threshold 2 (OR = 0.98), and a protective association at threshold 3 (OR = 0.51, 95% CI: 0.39, 0.68), suggesting temporal dynamics in risk profiles.

Female sex demonstrated significant, diminishing protective effects with increasing burden severity (ORs: 0.46, 0.64, 0.68 across thresholds), indicating a reduced influence at higher comorbidity levels. Alcohol use was protective at higher thresholds, with statistically significant ORs of 0.82 and 0.70 at thresholds 2 and 3, respectively. Education showed subtle but meaningful non-proportional effects; college graduates (Education = 3) experienced reduced odds of greater burden at higher thresholds (OR = 0.76 and 0.70), while no significant effects were observed for Education = 2. Depression was positively associated with burden at lower thresholds (OR = 1.58 and 1.35), attenuating to near null at the highest threshold (OR = 1.08). Smoking status exhibited no consistent associations, with ORs near 1 and credible intervals including the null at all thresholds.

Collectively, the Bayesian partial proportional odds generalized linear mixed model (GLMM) employed here offers a flexible and interpretable framework to analyze complex longitudinal ordinal data with threshold-specific covariate effects. This approach addresses limitations of fully proportional odds models by allowing certain covariates to vary across thresholds, thereby more accurately capturing the evolving cardiometabolic burden over time and severity. Convergence and sampling diagnostics support the model's reliability, with all Gelman–Rubin statistics ($\hat{R}$) equal to 1.00 and bulk and tail effective sample sizes exceeding 700, indicating stable posterior estimation.

To evaluate the model's predictive adequacy and overall fit, both numerical and graphical diagnostics were considered. Table 23 summarizes approximate Leave-One-Out Cross-Validation (LOO) and Widely Applicable Information Criterion (WAIC) results, computed via the brms package. These criteria complement each other by quantifying out-of-sample predictive accuracy while penalizing model complexity, a crucial balance in hierarchical Bayesian models incorporating random effects and partial proportional odds structures.

The LOO analysis yielded an expected log predictive density (ELPD) of −6497.1 (SE = 46.9), reflecting strong predictive alignment between the model and observed data. The estimated effective number of parameters, p_loo = 1246.9 (SE = 14.7), underscores the model's flexibility, consistent with its threshold-specific covariate specification and inclusion of random intercepts. Correspondingly, the LOO Information Criterion (LOOIC), computed as $-2 \times$ ELPD, was 12,994.2 (SE = 93.9), serving as a metric of predictive deviance for model comparison. Stability of these estimates was affirmed by Pareto-smoothed importance sampling diagnostics, with 99.97% of observations exhibiting Pareto-k values $\leq$ 0.7. Only two observations fell in the moderate range ($0.7 < k \leq 1$), and none exceeded the critical cutoff of 1. The minimum effective sample size for the most extreme Pareto-k was 302, reinforcing stable posterior inference even in distribution tails.

WAIC results were broadly consistent with LOO findings. The WAIC-based ELPD was −6466.8 (SE = 46.9), with an effective number of parameters of 1216.6 (SE = 14.7), yielding a WAIC of 12,933.5 (SE =

93.8). A diagnostic warning, however, indicated that 14.5% of the pointwise effective number of parameters exceeded the recommended 0.4 threshold, suggesting that WAIC may slightly underestimate uncertainty for a subset of observations. Given this, LOO estimates provide a more robust assessment of predictive fit for this model.

Graphical diagnostics further support the model's adequacy. Figures 9 and 10 display posterior distributions alongside trace plots for all parameters. Intercepts corresponding to the ordinal outcome thresholds (Figure 9, first three rows) exhibit smooth, unimodal posteriors and well-mixed MCMC traces, reflecting precise and convergent threshold estimation. Covariates under the proportional odds assumption such as physical activity adherence (PAG2008), age, and dietary quality (HEI2010), show narrow posterior distributions and stable trace behavior, indicating reliable estimation consistent across thresholds.

Covariates modeled under the partial proportional odds framework, including visit number and sex, demonstrate distinct threshold-specific posterior distributions with well-behaved sampling, confirming the model's capacity to capture non-constant effects across outcome levels. Additional threshold-specific covariates of alcohol use, BMI, education, depression status, and smoking, also display well-defined posteriors and stable MCMC mixing (Figure 10). The posterior distribution for the subject-level random intercept is narrowly concentrated and well-converged, reflecting consistent estimation of between-subject heterogeneity.

Finally, Figure 11 presents a posterior predictive density overlay comparing observed ordinal outcomes with replicated data simulated from the model. The close alignment across outcome categories indicates that the model effectively captures the marginal distribution of the data, corroborating the numerical diagnostics. Collectively, these results demonstrate that the Bayesian partial proportional odds GLMM provides reliable, nuanced estimates of threshold-dependent covariate effects and between-subject variability in longitudinal cardiometabolic burden.

In summary, the model exhibits strong predictive performance, robust convergence, and the flexibility to accommodate both proportional and non-proportional effects within a longitudinal ordinal data context. This methodological framework is well-suited to disentangle complex, evolving patterns of cardiometabolic risk, providing valuable insight into how covariates influence disease burden across time and severity thresholds.

Generalized Estimating Equations (GEE) for Marginal Analysis

While the Bayesian cumulative logit mixed-effects model provides valuable subject-specific insights by modeling individual-level variability through random effects, it is equally important to examine population-averaged associations. In many public health and clinical contexts, the primary interest is in understanding how predictors relate to average outcomes across the population rather than within individuals. Generalized estimating equations (GEE) offer a robust framework for such marginal analyses by estimating average effects while accounting for within-subject correlation via a working correlation structure. Unlike mixed-effects models, GEE does not model random effects explicitly but focuses on population-level interpretations, which are particularly relevant for informing policy or broad clinical strategies. By incorporating GEE alongside the mixed model, we obtain a more complete picture of how covariates influence cardiometabolic burden both at the individual and population levels over time.

A fundamental assumption in cumulative logit models, including those fit by GEE, is the proportional odds (PO) assumption, which requires each predictor to have a constant effect across all thresholds of the ordinal outcome. This assumption implies that the odds ratio for a covariate is the same whether comparing lower versus higher severity categories. Violations of the PO assumption can mask important differences in how predictors relate to mild versus severe comorbidities, leading to oversimplified interpretations that may hinder effective public health decision-making. To assess this, we fitted threshold-specific GEE models (Table 24) to test whether covariate effects vary across outcome levels.

The results clearly indicate that some covariates violate the PO assumption. For instance, the effect of Visit = 2 increases progressively across thresholds, from a negligible and non-significant estimate for Outcome ≥ 1 (0.020, $p = 0.72$) to larger and significant effects at Outcome ≥ 2 (0.155, $p = 0.0033$) and Outcome = 3 (0.366, $p < 0.0001$). Visit = 3 shows an even more striking pattern, with a strong positive effect at the first threshold (0.424, $p < 0.0001$), a small non-significant negative effect at the second (-0.074, $p = 0.35$), and a strong negative effect at the third (-0.833, $p < 0.0001$). Similar threshold-dependent patterns appear for Female sex, whose protective effect weakens at higher thresholds, and Alcohol Use, which shows increasing protection at higher comorbidity levels. These patterns confirm that the impact of certain variables depends on the severity of comorbidity, violating the PO assumption.

Other covariates show more moderate evidence of non-proportionality. Education level 3 is significant only at the middle threshold (-0.193, $p = 0.0292$), and Depression's effect declines as severity increases. Both Age and BMI remain significant across thresholds but increase in effect size with severity (Age from 0.056 to 0.078; BMI from 0.062 to 0.083), suggesting a dose-response relationship rather than a uniform effect. Conversely, Education level 2, physical activity adherence (PAG2008), diet quality (HEI2010), and Smoking demonstrate consistent effects with no violation of the PO assumption. Taken together, these findings argue for a partial proportional odds model that permits some covariates to vary by outcome level, while others maintain constant effects.

Given that GEE methods typically do not support partial proportional odds modeling within a single framework, we proceed by presenting results under the standard PO assumption, acknowledging that some covariate effects may vary by threshold. Future work should explore alternative modeling strategies, such as adjacent-category or continuation-ratio models, or develop GEE extensions allowing for threshold-varying effects. These approaches could provide a more nuanced understanding of how predictors influence different severity levels while preserving the population-averaged interpretation critical for public health.

Table 25 displays parameter estimates from the GEE cumulative logit model assuming PO. Threshold-specific intercepts capture baseline cumulative log-odds of having at least one, two, or three comorbidities when covariates are at reference values: -7.191 (SE = 0.28, $p < 0.0001$), -5.502 (SE = 0.27, $p < 0.0001$), and -3.709 (SE = 0.27, $p < 0.0001$), respectively. These values reflect expected baseline probabilities of increasing comorbidity burden.

Longitudinally, Visit 2 shows a significant increase in odds of higher comorbidity burden (estimate = 0.167, SE = 0.04, OR = 1.182, 95% CI [1.09, 1.28], $p < 0.0001$), whereas Visit 3 does not differ significantly from baseline (estimate = 0.056, SE = 0.06, $p = 0.3639$), suggesting a plateau in progression by the last follow-up. Females consistently have lower odds of higher comorbidities than males (estimate = -0.433, SE = 0.07, OR = 0.649, 95% CI [0.57, 0.74], $p < 0.0001$). Education effects are mixed: the highest education category is modestly protective (estimate = -0.155, SE = 0.08, OR = 0.856, 95% CI [0.74, 0.99], $p = 0.0393$), while the middle category shows no effect. Depression increases comorbidity odds (estimate = 0.249, SE = 0.08, OR = 1.283, 95% CI [1.11, 1.48], $p = 0.0009$), and alcohol use appears protective (estimate = -0.185, SE = 0.05, OR = 0.831, 95% CI [0.75, 0.92], $p = 0.0004$). Smoking, physical activity, and diet quality do not significantly affect comorbidity burden. Age and BMI have strong positive associations: each additional year of age increases odds by 6% (estimate = 0.058, SE = 0.003, OR = 1.060, 95% CI [1.05, 1.06], $p < 0.0001$), and each unit increase in BMI increases odds by 6.3% (estimate = 0.061, SE = 0.005, OR = 1.063, 95% CI [1.05, 1.07], $p < 0.0001$), highlighting their central roles in cardiometabolic risk.

To better capture changes in covariate effects over time, interaction terms between visit and key predictors were incorporated (Table 26). This is critical because models without interactions assume constant covariate effects across visits, potentially missing evolving influences during disease progression. The interaction of Visit 3 with female sex is highly significant (estimate = 0.508, SE = 0.11, OR = 1.662, 95% CI [1.34, 2.07], $p < 0.0001$), indicating the protective effect of female sex diminishes by the third visit, female odds of higher comorbidity categories increase approximately 66% relative to males

compared to baseline. No significant interaction is seen at Visit 2 for females ($p = 0.2233$), suggesting this attenuation occurs later in follow-up.

Age exhibits a complex time-dependent pattern: the Visit 2 × Age interaction is positive and significant (estimate = 0.010, SE = 0.004, OR = 1.010, 95% CI [1.003, 1.018], $p = 0.0079$), indicating a stronger age effect at Visit 2, while Visit 3 × Age interaction is negative and highly significant (estimate = -0.050, SE = 0.005, OR = 0.951, 95% CI [0.940, 0.960], $p < 0.0001$), reflecting a marked reduction of age's impact at the final visit. BMI follows a similar trend with a positive baseline effect (estimate = 0.079, SE = 0.008, OR = 1.082, 95% CI [1.065, 1.100], $p < 0.0001$) and a negative interaction at Visit 3 (estimate = -0.048, SE = 0.01, OR = 0.953, 95% CI [0.933, 0.971], $p < 0.0001$), signaling diminished influence over time. Other interaction terms for education, depression, alcohol use, physical activity adherence, diet quality, and smoking do not reach significance, indicating relatively stable effects. Smoking interactions at Visit 2 (estimate = 0.272, $p = 0.0533$) and Visit 3 (estimate = 0.295, $p = 0.0755$) suggest a possible increasing role deserving further exploration.

To validate these findings, Tables 27 and 28 report Type 3 score tests assessing whether each covariate and interaction term's effects differ from zero at the population level. In the model without interactions (Table 27), significant main effects include visit ($\chi^2(2) = 16.86$, $p = 0.0002$), sex ($\chi^2(1) = 41.67$, $p < 0.0001$), education ($\chi^2(2) = 6.13$, $p = 0.0466$), depression ($\chi^2(1) = 10.99$, $p = 0.0009$), age ($\chi^2(1) = 354.65$, $p < 0.0001$), alcohol use ($\chi^2(1) = 12.35$, $p = 0.0004$), and BMI ($\chi^2(1) = 112.17$, $p < 0.0001$), confirming their significant associations with higher comorbidity categories. Physical activity adherence, diet quality, and smoking do not show significant effects here.

Upon adding interaction terms (Table 28), the significance of visit sharply increases ($\chi^2(2) = 110.92$, $p < 0.0001$), reflecting improved model fit by capturing time-varying covariate effects. Main effects of sex ($\chi^2(1) = 43.55$, $p < 0.0001$), education ($\chi^2(2) = 6.21$, $p = 0.0448$), depression ($\chi^2(1) = 11.00$, $p = 0.0009$), age ($\chi^2(1) = 354.58$, $p < 0.0001$), alcohol use ($\chi^2(1) = 10.37$, $p = 0.0013$), and BMI ($\chi^2(1) = 113.36$, $p < 0.0001$) remain significant. Importantly, interactions of visit with sex ($\chi^2(2) = 23.02$, $p < 0.0001$), age

($\chi^2(2) = 143.18$, $p < 0.0001$), and BMI ($\chi^2(2) = 42.87$, $p < 0.0001$) are significant, confirming that these covariate effects vary significantly across visits. Smoking's interaction term approaches significance ($\chi^2(2) = 5.11$, $p = 0.0775$), suggesting possible temporal variation.

In summary, these results underscore the dynamic nature of sex, age, and BMI in influencing cardiometabolic comorbidity progression. Incorporating visit-by-covariate interactions in GEE models enhances our ability to detect evolving risk factor effects over time, providing richer population-level insights that are critical for tailoring effective public health interventions.

Discussion

Summary of Results

This study provides a comprehensive longitudinal assessment of ordinal cardiometabolic multimorbidity, defined as the co-occurrence of diabetes, high cholesterol, and hypertension, using an integrated series of descriptive, exploratory, and model-based statistical approaches. At baseline, the analytic sample of 2,000 adults was 63.9% female, with educational attainment evenly split across categories, and 25.6% screened positive for depressive symptoms. Behaviorally, 62.1% met physical activity guidelines, 46.5% reported alcohol use, and 17.1% were current smokers. Over three visits, marked temporal shifts emerged in both comorbidities and their underlying risk factors. The prevalence of diabetes nearly doubled from 19.9% to 36.7%, hypertension rose from 29.9% to 40.8%, and high cholesterol increased more modestly from 43.6% to 47.9%. The proportion of participants reporting all three conditions rose from 7.2% at Visit 1 to 13.9% at Visit 2 before declining slightly to 7.5% at Visit 3, while those with no conditions declined steadily from 39.5% to 19.9%. These descriptive trends reflect the accumulation and partial remission of disease over time, highlighting the need for flexible modeling approaches that accommodate both progression and recovery.

Exploratory analyses further underscored the cohort's complex and evolving risk landscape. Continuous predictors such as age, BMI, and dietary quality (HEI-2010) showed wide variability. Age followed a

slightly right-skewed distribution centered around midlife (mean = 53.5 years), while HEI-2010 was approximately normal (mean = 60.6), with a left tail indicating a subset with poor diet quality. BMI displayed a right-skewed distribution with substantial outlier influence; the modal value hovered near 30 kg/m^2, and nearly half the cohort qualified as obese by Visit 3. A single extreme BMI value exceeding 70 kg/m^2 was excluded from multivariable analyses to preserve model stability. Longitudinal dynamics showed strong between-subject stability in BMI (ICC = 0.92), but meaningful within-subject variation (within-person SD = 2.73 kg/m^2), supporting its treatment as a time-varying covariate. Categorical transitions in comorbidities revealed net increases in diabetes and obesity, but bidirectional shifts in cholesterol and alcohol use, likely reflecting treatment responses or behavioral modification. Smoking showed a gradual decline over time, while obesity incidence increased steadily. Composite comorbidity status exhibited substantial within-subject turnover across visits: nearly 30% worsened between Visits 1 and 2, with another 35.9% worsening from Visit 2 to 3; meanwhile, roughly 12% and 29% of participants improved during those intervals, respectively.

Stratified visualizations confirmed that older age, higher BMI, depressive symptoms, and lower educational attainment were associated with higher comorbidity burden. Physical activity and alcohol use were generally associated with lower burden, although selection and reverse causation cannot be ruled out. Sex differences were evident at baseline, with females showing lower comorbidity levels; however, these differences diminished over time, consistent with converging risk profiles. The observed heterogeneity and temporal variation in risk exposure and disease burden underscore the importance of longitudinal modeling strategies that can accommodate time-varying effects, threshold-specific risk contributions, and both fixed and dynamic covariates in shaping the trajectory of cardiometabolic multimorbidity.

Building on these patterns on EDA, collinearity diagnostics revealed that while most predictor associations were weak to moderate, a few key relationships warrant modeling consideration. Cross-sectionally, BMI and obesity status showed a very strong correlation at all visits ($r = 0.75–0.77$, $p <$

0.0001), and this remained high longitudinally ($r = 0.90$), confirming their definitional overlap and supporting the choice to retain BMI alone in multivariable models for its granularity. HEI-2010 was weakly and inversely correlated with BMI at later visits ($r \approx –0.085$, $p = 0.0001$), and age was modestly positively associated with diet quality ($\rho \approx 0.15$, $p < 0.0001$). Behavioral risk factors such as smoking and alcohol use were initially associated ($\chi^2 = 46.0$ and 25.5, $p < 0.0001$), but this link weakened over time. Education level was weakly related to diet quality and modestly linked to physical activity, reflecting socioeconomic patterning. Longitudinally, BMI remained highly stable between individuals (ICC = 0.921) but showed meaningful within-subject variability, validating its role as a dynamic predictor. Repeated-measures models confirmed that higher BMI was associated with poorer diet ($\beta = –0.0384$, $p < 0.0001$), and both age and BMI increased over time ($\beta = 0.3426$, $p < 0.0001$). GEE models of binary predictors showed interpretable correlations between smoking and alcohol use (0.16–0.30), sex and health behaviors (0.15–0.29), and depression with both smoking and alcohol (≈0.27–0.30), all significant at $p < 0.0001$. These results emphasize the importance of accounting for evolving predictor relationships over time, modeling exposures dynamically where appropriate, and avoiding collinear pairs in multivariable frameworks.

Cross-sectional analyses of cumulative cardiometabolic multimorbidity across three visits were conducted using partial proportional odds ordinal logistic regression models, carefully addressing violations of the proportional odds (PO) assumption identified through graphical diagnostics and formal Brant testing. Early visits (Visit 1 and 2) exhibited significant non-proportionality for key predictors including age, sex, and education level, as well as suggestive departures for behavioral factors such as alcohol use. Likelihood ratio tests further substantiated the necessity of relaxing the PO assumption at these time points, while Visit 3 displayed largely proportional effects, indicating a shift toward model stability over time. Continuous covariates of age, BMI, and HEI2010 diet quality, demonstrated adequate logit-linearity, supporting their modeling as continuous predictors. These diagnostic results informed a robust modeling

strategy that balances interpretability with the flexibility required to accurately capture the complex and evolving relationships governing cardiometabolic burden.

Model-based inferences revealed a nuanced and temporally dynamic landscape of risk factor associations with ordinal comorbidity burden. Male sex and younger age emerged as consistent predictors of higher comorbidity early in follow-up, though the sex effect attenuated by Visit 3, reflecting potential convergence of risk profiles. Alcohol use was associated with increased odds of moderate and severe comorbidity at earlier visits but showed no significant impact later, underscoring shifting behavioral influences. Education's effects varied both by visit and outcome severity, while depression exhibited a protective or inverse association that strengthened over time. Notably, physical activity became positively associated with higher comorbidity only at Visit 3, suggesting possible reverse causation or changes in behavior due to disease progression. BMI demonstrated an inverse relationship at lower comorbidity thresholds but revealed a U-shaped risk pattern at the highest burden level by Visit 3, emphasizing complex nonlinear effects. Smoking status was inversely related to severe comorbidity only at Visit 3, and diet quality showed no significant associations throughout. Collectively, these findings highlight the imperative for longitudinal modeling approaches that accommodate threshold-specific effects, time-varying covariates, and within-subject correlation structures to fully characterize the dynamic trajectories of cardiometabolic multimorbidity.

In this longitudinal analysis of cardiometabolic comorbidity burden, cumulative logit generalized linear mixed-effects models (GLMMs) with subject-specific random intercepts were utilized to account for within-subject correlation and individual heterogeneity. The baseline proportional odds (PO) model without interaction terms indicated that female sex (OR = 1.77, $p < 0.0001$) and alcohol use (OR = 1.23, $p = 0.0007$) were associated with increased odds of higher cardiometabolic burden, while higher education (college or above) conferred a modest protective effect (OR = 0.80, $p = 0.0239$). Depression (OR = 0.73, $p = 0.0014$), age (OR = 0.92 per year, $p < 0.0001$), and BMI (OR = 0.92 per unit, $p < 0.0001$) showed inverse associations. Smoking, physical activity, and diet quality were not statistically significant

predictors. Model diagnostics confirmed that incorporating a random intercept substantially improved model fit (AIC reduced from 14,339 to 13,661), and detailed evaluation of the random intercept estimates demonstrated approximate normality, validating the assumption of normally distributed subject-level effects. This affirmed the appropriateness of mixed-effects modeling to capture unobserved heterogeneity beyond fixed covariates. Threshold-specific analyses from Cumulative Binary Logistic Mixed Models revealed notable violations of the PO assumption for key variables including visit, sex, and BMI, suggesting that the proportional odds constraint may oversimplify important threshold-dependent associations.

When interactions between visit and covariates were incorporated, substantial temporal dynamics in risk factor effects emerged. Female sex was linked to markedly higher odds of elevated cardiometabolic burden at baseline (OR = 2.46, $p < 0.0001$), but this risk significantly attenuated by Visit 3 (interaction OR = 0.51, $p < 0.0001$), indicating a convergence of sex differences over time. Smoking's initial risk elevation (OR = 1.37, $p = 0.034$) diminished substantially in later visits (interaction ORs 0.65 and 0.55 at Visits 2 and 3, respectively; both $p < 0.05$). Age and BMI displayed complex, nonlinear time-varying effects, with protective associations at baseline (age OR = 0.91, BMI OR = 0.92; both $p < 0.0001$) strengthening at Visit 2 and reversing at Visit 3 (age interaction OR = 1.07, BMI interaction OR = 1.05; both $p < 0.0001$), reflecting increasing risk with advancing age and BMI in later follow-up. These time-varying effects increased the random intercept variance from 2.19 to 2.48, highlighting greater captured heterogeneity. Other covariates such as physical activity, depression, alcohol use, and diet quality exhibited stable, generally nonsignificant associations across visits. Collectively, these results underscore the critical need for flexible modeling strategies that incorporate both longitudinal effect modification and partial relaxation of proportional odds assumptions to accurately characterize complex cardiometabolic risk trajectories.

To address identified violations of the proportional odds assumption, we implemented a Bayesian cumulative logit mixed-effects model with partial proportional odds, enabling threshold-specific effects

for covariates exhibiting non-proportionality while maintaining proportionality for others. This flexible modeling approach, which included subject-level normally distributed random intercepts with an estimated standard deviation of 1.51 (95% credible interval: 1.42, 1.62), effectively captured substantial individual heterogeneity beyond fixed effects. Convergence diagnostics were robust, with all Gelman-Rubin statistics equal to 1.00 and effective sample sizes exceeding 700, ensuring reliable posterior inference. Among covariates constrained under proportional odds, age demonstrated a borderline significant positive association with cardiometabolic burden (posterior OR ~1.08; 95% credible interval narrowly including 1), reflecting a clinically plausible increase in risk with advancing age, unlike previous models that suggested inverse effects. Adherence to physical activity guidelines showed a modest, nonsignificant protective trend (OR ~0.89), while dietary quality had negligible influence (OR ~0.996).

Covariates modeled with threshold-specific effects revealed complex and interpretable longitudinal patterns. Body mass index was positively and significantly associated with increasing cardiometabolic burden across all thresholds, with odds ratios rising monotonically from 1.08 at the lowest threshold to 1.11 at the highest, each with 95% credible intervals excluding 1, correcting prior misleading inverse associations. The effects of study visit also varied meaningfully by threshold: Visit 2 showed increasing odds of higher burden severity across thresholds (ORs of approximately 1.01, 1.27, and 1.73), while Visit 3 exhibited elevated risk at the lowest threshold (OR = 1.51), no effect at the middle threshold (OR = 0.98), and a protective effect at the highest severity threshold (OR = 0.51, 95% CI: 0.39–0.68), indicating dynamic temporal changes in cardiometabolic risk profiles. Female sex conferred significant protective effects that diminished with burden severity (ORs of 0.46, 0.64, and 0.68, all statistically significant), highlighting a reduced influence of sex on more severe comorbid states. Alcohol use was similarly protective at higher thresholds, with statistically significant odds ratios of 0.82 and 0.70 at the middle and highest severity levels. Higher education (college or above) was associated with significantly reduced odds of greater burden at higher thresholds (ORs 0.76 and 0.70), whereas depression's positive

association with cardiometabolic burden was strongest at lower severity levels (OR = 1.58 and 1.35) but attenuated near null (OR = 1.08) at the highest threshold. Smoking status showed no consistent significant association across thresholds, with odds ratios near unity and credible intervals spanning the null.

Model fit and predictive performance were excellent, as evidenced by Leave-One-Out Cross-Validation yielding an expected log predictive density of −6497.1 (SE = 46.9) and a stable effective number of parameters (~1247), supported by Pareto-smoothed importance sampling diagnostics where 99.97% of observations had Pareto-k values ≤ 0.7, indicating no undue influence by extreme observations. WAIC metrics corroborated these findings. Graphical diagnostics demonstrated well-mixed MCMC chains with unimodal posterior distributions for all parameters, including both proportional and threshold-specific covariates. Posterior predictive checks confirmed close alignment between observed and simulated ordinal outcome distributions, affirming the model's capacity to capture the complex distribution of cardiometabolic burden over time and severity. Overall, this Bayesian partial proportional odds mixed-effects framework robustly disentangles evolving, threshold-dependent covariate effects and individual heterogeneity in longitudinal cardiometabolic risk, offering an improved, interpretable tool for analyzing complex ordinal trajectories in clinical epidemiology.

Complementing the subject-specific Bayesian GLMM analysis, population-averaged associations were estimated using generalized estimating equations (GEE) to provide a broader understanding of how covariates influence cardiometabolic burden at the population level over time. The standard GEE cumulative logit model, assuming proportional odds (PO), identified several significant predictors of higher cardiometabolic burden. Age and BMI emerged as consistently strong risk factors, with each additional year of age associated with a 6% increase in the odds of higher burden (OR = 1.06, 95% CI: 1.05–1.06, $p < 0.0001$) and each unit increase in BMI associated with a 6.3% increase in odds (OR = 1.063, 95% CI: 1.05–1.07, $p < 0.0001$). Female sex remained significantly protective (OR = 0.649, 95% CI: 0.57–0.74, $p < 0.0001$), while depression (OR = 1.283, 95% CI: 1.11–1.48, $p = 0.0009$) and lower education were associated with elevated risk. Alcohol use was linked to significantly lower odds of

greater burden (OR = 0.831, 95% CI: 0.75–0.92, $p = 0.0004$), aligning with findings from the GLMM analysis. Smoking, physical activity adherence, and dietary quality did not exhibit significant marginal associations in this framework.

To assess whether covariate effects varied over time, a GEE model including visit-by-covariate interaction terms was estimated. This revealed substantial temporal heterogeneity. The interaction between Visit 3 and female sex was highly significant (OR = 1.662, 95% CI: 1.34–2.07, $p < 0.0001$), indicating that the protective effect of female sex diminished by the final visit. Age's influence also shifted across visits: while the effect increased at Visit 2 ($p = 0.0079$), it significantly declined at Visit 3 (OR = 0.951, 95% CI: 0.94–0.96, $p < 0.0001$), suggesting a waning age-related risk in later follow-up. A similar attenuation was observed for BMI (interaction OR = 0.953, 95% CI: 0.93–0.97, $p < 0.0001$). Type 3 score tests confirmed significant interaction effects for sex, age, and BMI ($p < 0.0001$ for all), emphasizing evolving roles of these covariates over time. Smoking's time-varying association approached significance ($p = 0.0775$), warranting further investigation. These GEE results reinforce the need for dynamic modeling approaches in public health contexts and confirm key findings from the Bayesian GLMM: while age, BMI, sex, depression, and alcohol use influence cardiometabolic burden, their effects can vary significantly over time and severity, highlighting both consistent and evolving population-level risk factors.

Clinical Interpretation

This longitudinal analysis modeled cardiometabolic multimorbidity which defined by ordinal progression from zero to three conditions (type 2 diabetes, hypertension, high cholesterol), to evaluate how individual- and population-level risk factors influence cardiometabolic burden and, by implication, downstream cardiovascular disease (CVD) risk. Each level on the ordinal scale represents a cumulative increase in cardiometabolic dysfunction and is strongly associated with incident major cardiovascular events, all-cause mortality, and healthcare utilization (Nkomo et al., 2008; Chong et al., 2024).

Subject-Specific Effects from GLMM

Using a Bayesian cumulative logit generalized linear mixed-effects model (GLMM), we examined subject-specific associations while accounting for within-person variability via random intercepts. The results revealed clear threshold-specific patterns for multiple covariates.

Body mass index (BMI) emerged as a robust, progressive risk factor. Each unit increase in BMI was associated with increased odds of transitioning to more severe comorbidity levels, with posterior odds ratios rising from 1.08 to 1.11 across thresholds. These findings mirror well-documented links between excess adiposity and progression from metabolic syndrome to cardiovascular morbidity (Després et al., 2008; Lavie et al., 2009). The fact that BMI's effect intensified with severity underscores its importance in both primary prevention and risk stratification.

Sex differences were evident: females had significantly lower odds of being in higher comorbidity categories, particularly at early thresholds (OR = 0.46 for ≥1 condition), but this protective effect weakened with increasing severity (OR = 0.68 for all three conditions). This aligns with prior evidence that women initially present with lower cardiometabolic burden but experience risk convergence over time due to menopause-related hormonal changes and delayed diagnosis (Mosca et al., 2011).

Age was associated with a small but consistent increase in risk under the partial proportional odds model, with posterior ORs approaching 1.08, in contrast to inverse associations in simpler models. This result accords with clinical evidence linking aging to endothelial dysfunction and cumulative vascular risk (Lakatta & Levy, 2003).

Depressive symptoms were positively associated with early-stage multimorbidity (OR = 1.58 at threshold 1), but this effect attenuated at higher thresholds (OR = 1.08), potentially reflecting differential pathways in early versus advanced disease accumulation. This supports growing recognition of depression as a contributor to cardiometabolic syndrome through behavioral and neuroendocrine mechanisms (Whooley & Wong, 2013).

Alcohol use showed a threshold-dependent protective effect (OR = 0.82 and 0.70 at thresholds 2 and 3, respectively), potentially reflecting moderate drinking patterns associated with lower CVD incidence (Ronksley et al., 2011), though causality remains debated. The protective effects of higher education (OR = 0.70 at highest severity) further emphasize the role of socioeconomic factors in chronic disease risk (Stringhini et al., 2010).

Notably, diet quality (HEI-2010), smoking status, and physical activity were not statistically significant in the subject-specific model, despite being critical lifestyle exposures. This may be due to confounding, measurement limitations, or nonlinear effects not captured by the model.

Population-Level Effects from GEE

To complement subject-level inferences, we implemented a cumulative logit Generalized Estimating Equation (GEE) model to estimate population-averaged associations. This marginal framework offers insights more directly relevant to public health and policy decision-making, where average effects across the population guide interventions (Zeger & Liang, 1986; Pepe & Anderson, 1994).

In the standard GEE model assuming proportional odds, age and BMI remained dominant risk factors. Each additional year of age increased the odds of greater multimorbidity by 6% (OR = 1.060, 95% CI: 1.05–1.06, $p < 0.0001$), and each unit increase in BMI raised the odds by 6.3% (OR = 1.063, 95% CI: 1.05–1.07, $p < 0.0001$). These results underscore the population-level burden of aging and obesity in driving cardiometabolic diseases, consistent with large cohort studies such as NHANES and Framingham (Wilson et al., 1998; Ford et al., 2004).

Female sex remained significantly protective in the marginal model (OR = 0.649, 95% CI: 0.57–0.74, $p < 0.0001$), although interaction analysis revealed that this advantage eroded by Visit 3 (interaction OR = 1.662, $p < 0.0001$), consistent with longitudinal epidemiologic trends showing diminishing sex disparities in CVD risk post-menopause (Maas & Appelman, 2010).

Depression maintained a positive association with higher comorbidity (OR = 1.283, 95% CI: 1.11–1.48, p = 0.0009), reinforcing its value as a population-level screening target for multimorbidity prevention (Read et al., 2017). Alcohol use again conferred a protective effect (OR = 0.831, 95% CI: 0.75–0.92, p = 0.0004), although the clinical interpretation requires caution due to potential confounding and selection bias (Brien et al., 2011). In contrast, physical activity, diet quality, and smoking did not show statistically significant marginal effects. This may reflect overlapping variance with BMI and age, or limitations in capturing behavioral intensity or changes over time.

Critically, visit-by-covariate interactions demonstrated that the effects of age, sex, and BMI varied significantly over time ($p < 0.0001$ for all Type 3 tests). For example, BMI's impact declined by Visit 3 (interaction OR = 0.953, $p < 0.0001$), suggesting potential effects of treatment or disease adaptation. Age showed a similar pattern (Visit 3 interaction OR = 0.951, $p < 0.0001$), hinting at survivor bias or differential trajectories among older adults. These time-varying associations mirror real-world clinical phenomena, where risk profiles shift due to intervention, comorbidity saturation, or care engagement (Barnett et al., 2012; Tinetti et al., 2012).

Together, the marginal GEE and subject-specific GLMM models provide a coherent picture: BMI, age, sex, depression, and education consistently influence cardiometabolic burden, though their magnitude and clinical relevance may depend on the stage of multimorbidity and follow-up interval. These results support both personalized care and population-level prevention strategies and underscore the importance of models that allow for dynamic, threshold-specific, and time-varying effects in chronic disease surveillance and management.

Limitations

Outcome Definition and Scope

This study defines cardiometabolic multimorbidity as the count of three conditions: type 2 diabetes, hypertension, and high cholesterol. While this operationalization enables straightforward ordinal

modeling, it omits other common and clinically significant conditions such as coronary artery disease, stroke, peripheral artery disease, and metabolic syndrome. These exclusions constrain the scope of inference and likely underestimate the total burden of cardiometabolic risk. Patients with controlled hypertension and hyperlipidemia are treated equivalently to those with prior myocardial infarction, ignoring vast differences in prognosis and care needs. The literature emphasizes that multimorbidity is a heterogeneous and multidimensional construct (Violan et al., 2014; Fortin et al., 2005); limiting it to three upstream conditions may reduce the utility of the outcome for informing disease progression, care planning, or prevention strategies.

Moreover, the outcome treats each increment from 0 to 3 conditions as ordinally equivalent, assuming equal clinical weight and spacing. This simplification ignores key differences in severity or chronicity, such as controlled versus uncontrolled diabetes. It also overlooks the influence of condition duration and treatment status, which are important prognostic indicators. These modeling choices introduce outcome misclassification and may attenuate associations with key predictors. While ordinal models are appealing for progression analysis, their validity depends on careful outcome construction that aligns with clinical relevance and statistical assumptions (McCullagh, 1980; Greenland, 2005).

Lack of Clinically Anchored Hypothesis and Objective-Specific Modeling

Despite using sophisticated modeling approaches, including cumulative logit models with time-varying effects and partial proportional odds structures, the study lacks a clearly articulated clinical hypothesis. Advanced models are most informative when addressing actionable questions, such as whether treating depression reduces multimorbidity progression, or whether increasing physical activity modifies BMI trajectories in relation to disease accumulation. Without such goals, the analysis becomes descriptive, offering covariate associations without clear guidance for clinical or public health interventions (Greenland, 2005; Rubin, 2008).

This limitation is particularly salient given the growing emphasis on target trial emulation in observational data (Hernán & Robins, 2016; Petersen & van der Laan, 2014). These frameworks tie modeling to well-defined intervention scenarios and clinical contrasts. For example, the finding that alcohol use is inversely associated with disease burden lacks interpretability without clarifying whether this is causal, confounded, or due to reverse causation. Similarly, findings related to depression, education, and sex differences remain speculative in the absence of predefined contrasts or mechanistic hypotheses. Without grounding in such frameworks, the analysis may overinterpret complex, socially patterned exposures without adequate contextualization (Westreich & Greenland, 2013).

Measurement Bias, Misclassification, and Residual Confounding

Key behavioral predictors, including smoking, alcohol use, physical activity, and dietary quality were self-reported, introducing exposure misclassification through recall and social desirability bias. These sources of error can attenuate associations and generate subgroup-specific misclassification (Althubaiti, 2016; Kipnis et al., 2003). Depression was modeled as a binary indicator from a screening instrument, ignoring the severity, chronicity, or recurrence of symptoms. This simplification can mask dose-response relationships and conflate mild with major depressive episodes, potentially biasing estimates (de la Torre-Luque et al., 2019; Streiner, 2002).

In addition, the study includes education as a socioeconomic proxy but omits key social determinants like income, employment, healthcare access, and neighborhood context. These unmeasured confounders are likely associated with both exposures and outcomes, raising concerns about residual confounding (Stringhini et al., 2010; Braveman et al., 2011). The apparent attenuation or reversal of age and BMI effects at later visits may reflect selective attrition or healthy survivor bias, whereby sicker individuals are underrepresented at follow-up (Ferrucci et al., 2004). Such biases complicate interpretation of time-varying associations and require sensitivity analyses, such as inverse probability weighting, to adjust for dropout.

Modeling Assumptions and Flexibility

Cumulative logit models assume proportional odds (PO), where covariate effects are constant across outcome thresholds. This assumption, commonly violated in health research, was not met for key predictors such as sex, age, and BMI (Peterson & Harrell, 1990; Williams, 2006). The GEE model relies on the PO assumption and therefore may obscure meaningful threshold-specific effects. For example, BMI may influence initial disease onset differently than later-stage progression. Ignoring this violates model assumptions, biases population-level estimates, and hampers risk stratification.

The GLMM used here allows partial relaxation of the PO assumption but did not include random slopes, restricting the model's ability to account for individual-specific changes over time in covariates like BMI and physical activity. This limitation reduces the capacity to capture dynamic within-person changes and may underrepresent true heterogeneity in disease progression (Verbeke & Molenberghs, 2000).

Partial proportional odds (PPO) models represent a major advancement by enabling threshold-specific covariate effects, uncovering clinically relevant variation missed by PO models (Peterson & Harrell, 1990; Williams, 2006). However, PPO models introduce additional parameters, increasing the risk of overfitting, especially with sparse data. They also require more extensive diagnostics, including model fit and sensitivity to threshold specification (Agresti, 2010; Gelman et al., 2013). Without these, PPO models may yield unstable estimates.

Furthermore, there is currently no extension of PPO models within the GEE framework, meaning all population-averaged estimates rely on the unrealistic PO assumption (Pan, 2001; Prentice & Zhao, 1991). This limitation narrows the interpretability of results for public health use. PPO models also complicate communication: multiple threshold-specific odds ratios per covariate challenge standard interpretation, especially when effects change direction across levels. Without careful visualization and explanation, the models risk being inaccessible to practitioners (Agresti, 2010).

Finally, model complexity raises concerns about overfitting and generalizability. While diagnostics suggest convergence and reasonable fit, the low prevalence of advanced multimorbidity limits power to estimate effects at extreme thresholds. Bayesian models are robust but must be validated to avoid overfitting noise (Gelman et al., 2013). Moderate follow-up duration also limits inference on long-term trajectories. These modeling trade-offs underscore the need for methodological transparency and alignment of model choices with clinical or policy objectives.

Future Work

Model Diagnostics and Validation

Ensuring the robustness of ordinal regression models in longitudinal data requires tailored diagnostic strategies distinct for subject-specific models (GLMMs) and population-averaged models (GEEs). For GLMMs, diagnostic methods must address the ordinal nature and hierarchical data structure. Conditional residuals like randomized quantile residuals (RQRs) enable simulation-based checks of local fit and assumption violations (Dunn & Smyth, 1996; Li & Shepherd, 2012). Complementary Pearson residuals, although limited by ordinal constraints, can be examined over time or across covariate strata to detect heteroscedasticity or nonlinear threshold effects. Predictive calibration is best evaluated using calibration plots comparing observed versus predicted category probabilities, while posterior predictive checks in Bayesian GLMMs provide a rigorous assessment of model plausibility (Gelman et al., 2013). Discrimination metrics such as ordinal c-statistics and Harrell's C-index quantify predictive accuracy across multiple disease burden levels (Van Calster et al., 2012; Harrell, 2015). Influence diagnostics—Cook's Distance and DFBETAS adapted for mixed models to help identify subjects or clusters disproportionately impacting fixed or random effects, particularly important in sparse outcome categories with potential overfitting risks (Lesaffre & Verbeke, 1998).

In contrast, GEEs, which yield marginal population-averaged estimates, require diagnostics focused on marginal residuals (Pearson or deviance residuals) despite complications from within-subject correlation

and misspecified working correlation structures (Zeger & Liang, 1986; Hardin & Hilbe, 2003). Residual plots stratified by fitted values, time, or covariates (e.g., sex, BMI) can reveal misfit or omitted interactions. Testing alternative working correlation matrices (exchangeable, autoregressive, unstructured) and assessing differences in standard errors or coefficient stability gauge model robustness. Calibration can be evaluated by observed vs. predicted probability plots, Brier scores, or Hosmer–Lemeshow-type tests extended to ordinal outcomes (Lipsitz et al., 1996). Discrimination in GEE models, while challenging, can be assessed via cumulative or category-specific ROC curves and ordinal C-indices based on marginal predictions (Van Calster et al., 2012). Influence diagnostics such as leverage scores and case-deletion analyses are crucial to detect subjects whose exclusion markedly alters parameter estimates, especially in longitudinal studies prone to attrition or informative missingness (Preisser & Qaqish, 1996). Combining these diagnostics with sensitivity analyses under different correlation structures or cross-comparison with GLMM results enhances confidence in marginal inferences.

Sensitivity Analyses and Robustness Checks

Robust inference demands explicit testing of assumptions, notably the functional form of continuous predictors like BMI, physical activity, and diet. The linearity assumption on the log-odds scale inherent in cumulative logit models can be relaxed using flexible approaches such as restricted cubic splines or penalized splines (Harrell, 2015; Royston & Sauerbrei, 2008). Model comparisons via likelihood ratio tests, AIC, or BIC, supplemented with component-plus-residual plots, allow detection and visualization of nonlinear effects that might otherwise bias inference. Alternative outcome definitions warrant exploration to test stability; expanding cardiometabolic multimorbidity definitions to include conditions like coronary artery disease or using severity-weighted composite scores better captures clinical complexity (Charlson et al., 1987; Fortin et al., 2005). This also facilitates examination of whether associations, e.g., with depression or physical activity, hold consistently across outcome operationalizations. Investigating dynamic outcomes, such as time-to-first event or progression-free survival, complements count-based models and integrates temporal dimensions.

Ordinal outcomes are sensitive to threshold specification. Future work should re-estimate models using alternative threshold schemes, binary or multi-level splits, category collapsing to test whether covariate effects remain stable and directionally consistent. This is vital when outcome categories are sparse or unevenly distributed, where threshold choices can substantially influence estimates (Agresti, 2010). Graphical displays of threshold-specific log-odds from partial proportional odds models can help judge whether category collapsing is justified without losing clinically relevant distinctions. Finally, alternative operationalizations of behavioral variables, continuous depressive symptom scores instead of dichotomies, finer alcohol consumption categories, should be tested for robustness, recognizing known nonlinear effects such as the U-shaped alcohol-cardiometabolic risk relationship (Mukamal et al., 2005). Random slopes for key time-varying covariates (BMI, physical activity, depression) can model individual heterogeneity in trajectories, addressing limitations of random intercept-only models (Verbeke & Molenberghs, 2000). Although computationally demanding, especially in Bayesian GLMMs, this approach refines personalized risk prediction and inference.

Model Extensions for Improved Inference

To improve inference rigor, future studies should explore partial proportional odds (PPO) extensions in GEE frameworks. Since standard GEE imposes the proportional odds assumption, semi-parametric or stratified GEE models, fitting separate cumulative logit models per threshold, could approximate threshold-specific effects marginally, relaxing PO constraints (Liang & Zeger, 1986; Heagerty & Zeger, 1996). Further, generalized additive model extensions of GEE (GAM-GEE) might flexibly capture nonlinear covariate effects across outcome levels.

Enhancements to Bayesian hierarchical models should include shrinkage priors like horseshoe or spike-and-slab to stabilize parameter estimates under sparse data, particularly important for threshold-specific coefficients in PPO models (Carvalho et al., 2010; Piironen & Vehtari, 2017). Bayesian model validation through leave-one-out cross-validation (LOO-CV) or the widely applicable information criterion (WAIC)

can guide model choice and prevent overfitting (Gelman et al., 2013). Hierarchical priors borrowing strength across thresholds can reduce estimation variance while preserving flexibility.

Finally, time-varying effect models deserve systematic application to capture how covariate effects evolve over follow-up. Varying-coefficient models (Hastie & Tibshirani, 1993) or functional regression approaches (Fan & Zhang, 2008) permit smooth temporal variation of regression coefficients, enhancing understanding of dynamic risk factor influences like depression or BMI. Bayesian implementations using Gaussian process priors or B-splines allow modeling of smooth time-dependent effects, helping to identify critical windows for intervention and providing nuanced, personalized risk profiles.

References

Althubaiti, A. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of multidisciplinary healthcare*, 211-217.

Aggarwal, R., Chiu, N., Wadhera, R. K., Moran, A. E., Raber, I., Shen, C., ... & Kazi, D. S. (2021). Racial/ethnic disparities in hypertension prevalence, awareness, treatment, and control in the United States, 2013 to 2018. *Hypertension*, *78*(6), 1719-1726.

Agresti, A. (2010). *Analysis of ordinal categorical data*. John Wiley & Sons.

Alberti, K. G., Eckel, R. H., Grundy, S. M., Zimmet, P. Z., Cleeman, J. I., Donato, K. A., ... & Smith Jr, S. C. (2009). Harmonizing the metabolic syndrome: a joint interim statement of the international diabetes federation task force on epidemiology and prevention; national heart, lung, and blood institute; American heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. *Circulation*, *120*(16), 1640-1645.

Allen, N. B., Siddique, J., Wilkins, J. T., Shay, C., Lewis, C. E., Goff, D. C., ... & Lloyd-Jones, D. (2014). Blood pressure trajectories in early adulthood and subclinical atherosclerosis in middle age. *Jama*, *311*(5), 490-497.

Ambrose, J. A., & Barua, R. S. (2004). The pathophysiology of cigarette smoking and cardiovascular disease: an update. *Journal of the American college of cardiology*, *43*(10), 1731-1737.

Barnett, K., Mercer, S. W., Norbury, M., Watt, G., Wyke, S., & Guthrie, B. (2012). Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet*, *380*(9836), 37-43.

Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., ... & Muntner, P. (2017). Heart disease and stroke statistics—2017 update: a report from the American Heart Association. *circulation*, *135*(10), e146-e603.

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, 1171-1178.

Braveman, P. A., Kumanyika, S., Fielding, J., LaVeist, T., Borrell, L. N., Manderscheid, R., & Troutman, A. (2011). Health disparities and health equity: the issue is justice. *American journal of public health*, *101*(S1), S149-S155.

Braveman, P., Egerter, S., & Williams, D. R. (2011). The social determinants of health: coming of age. *Annual review of public health*, *32*(1), 381-398.

Brien, S. E., Ronksley, P. E., Turner, B. J., Mukamal, K. J., & Ghali, W. A. (2011). Effect of alcohol consumption on biological markers associated with risk of coronary heart disease: systematic review and meta-analysis of interventional studies. *Bmj*, *342*.

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, *80*, 1-28.

Bürkner, P. C. (2017). Advanced Bayesian multilevel modeling with the R package brms. *arXiv preprint arXiv:1705.11123*.

Carnethon, M. R., Gulati, M., & Greenland, P. (2005). Prevalence and cardiovascular disease correlates of low cardiorespiratory fitness in adolescents and adults. *Jama*, *294*(23), 2981-2988.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, *76*, 1-32.

Carvalho, C. M., Polson, N. G., & Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, 465-480.

Centers for Disease Control and Prevention. (2014). National diabetes statistics report: estimates of diabetes and its burden in the United States, 2014. *Atlanta, GA: US Department of Health and Human Services*, *2014*.

Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases*, *40*(5), 373-383.

Chong, B., Jayabaskaran, J., Jauhari, S. M., Chan, S. P., Goh, R., Kueh, M. T. W., ... & Chan, M. Y. (2024). Global burden of cardiovascular diseases: projections from 2025 to 2050. *European Journal of Preventive Cardiology*, zwae281.

Cutler, D. M., & Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of health economics*, *29*(1), 1-28.

de la Torre-Luque, A., de la Fuente, J., Prina, M., Sanchez-Niubo, A., Haro, J. M., & Ayuso-Mateos, J. L. (2019). Long-term trajectories of depressive symptoms in old age: relationships with sociodemographic and health-related factors. *Journal of affective disorders*, *246*, 329-337.

Després, J. P., & Lemieux, I. (2006). Abdominal obesity and metabolic syndrome. *Nature*, *444*(7121), 881-887.

Diez Roux, A. V. (2012). Conceptual approaches to the study of health disparities. *Annual review of public health*, *33*(1), 41-58.

Dimitriadis, K., Theofilis, P., Iliakis, P., Pyrpyris, N., Dri, E., Sakalidis, A., ... & Tsioufis, K. (2024). Management of dyslipidemia in coronary artery disease: The present and the future. *Coronary Artery Disease*, *35*(6), 516-524.

Dunn, P. K., & Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and graphical statistics*, *5*(3), 236-244.

Fan, J., & Zhang, W. (2008). Statistical methods with varying coefficient models. *Statistics and its Interface*, *1*(1), 179.

Ferrucci, L., Guralnik, J. M., Studenski, S., Fried, L. P., Cutler Jr, G. B., Walston, J. D., & Interventions on Frailty Working Group. (2004). Designing randomized, controlled trials aimed at preventing or delaying functional decline and disability in frail, older persons: a consensus report. *Journal of the American Geriatrics Society*, *52*(4), 625-634.

Field, A. E., Coakley, E. H., Must, A., Spadano, J. L., Laird, N., Dietz, W. H., ... & Colditz, G. A. (2001). Impact of overweight on the risk of developing common chronic diseases during a 10-year period. *Archives of internal medicine*, *161*(13), 1581-1586.

Ford, E. S., Ajani, U. A., Croft, J. B., Critchley, J. A., Labarthe, D. R., Kottke, T. E., ... & Capewell, S. (2007). Explaining the decrease in US deaths from coronary disease, 1980–2000. *New England Journal of Medicine*, *356*(23), 2388-2398.

Ford, E. S., Giles, W. H., & Mokdad, A. H. (2004). Increasing prevalence of the metabolic syndrome among US adults. *Diabetes care*, *27*(10), 2444-2449.

Fortin, M., Bravo, G., Hudon, C., Vanasse, A., & Lapointe, L. (2005). Prevalence of multimorbidity among adults seen in family practice. *The Annals of Family Medicine*, *3*(3), 223-228.

Fortin, M., Hudon, C., Haggerty, J., van den Akker, M., & Almirall, J. (2010). Prevalence estimates of multimorbidity: a comparative study of two sources. *BMC health services research*, *10*, 1-6.

Friedman, G. D., Cutter, G. R., Donahue, R. P., Hughes, G. H., Hulley, S. B., Jacobs Jr, D. R., ... & Savage, P. J. (1988). CARDIA: study design, recruitment, and some characteristics of the examined subjects. *Journal of clinical epidemiology*, *41*(11), 1105-1116.

Gao, Y., Shah, L. M., Ding, J., & Martin, S. S. (2023). US trends in cholesterol screening, lipid levels, and lipid-lowering medication use in US adults, 1999 to 2018. *Journal of the American Heart Association*, *12*(3), e028205.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and computing*, *24*, 997-1016.

Goldstein, H. (2011). *Multilevel statistical models*. John Wiley & Sons.

Greenland, S. (2005). Multiple-bias modelling for analysis of observational data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *168*(2), 267-306.

Gregg, E. W., Zhuo, X., Cheng, Y. J., Albright, A. L., Narayan, K. V., & Thompson, T. J. (2014). Trends in lifetime risk and years of life lost due to diabetes in the USA, 1985–2011: a modelling study. *The lancet Diabetes & endocrinology*, *2*(11), 867-874.

Grundy, S. M., Cleeman, J. I., Daniels, S. R., Donato, K. A., Eckel, R. H., ... & American Heart Association. (2005). Diagnosis and management of the metabolic syndrome. An American Heart Association. *Cardiology in review*, *13*(6), 322-327.

Harrell, F. E. (2001). *Regression modeling strategies* (Vol. 54). New York: Springer-Verlag.

Hastie, T., & Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *55*(4), 757-779.

Heagerty, P. J., & Zeger, S. L. (1996). Marginal regression models for clustered ordinal measurements. *Journal of the American Statistical Association*, *91*(435), 1024-1036.

Heagerty, P. J., & Zeger, S. L. (2000). Marginalized multilevel models and likelihood inference (with comments and a rejoinder by the authors). *Statistical Science*, *15*(1), 1-26.

Hernán, M. A., Hsu, J., & Healy, B. (2019). A second chance to get causal inference right: a classification of data science tasks. *Chance*, *32*(1), 42-49.

Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American journal of epidemiology*, *183*(8), 758-764.

Kipnis, V., Midthune, D., Freedman, L., Bingham, S., Day, N. E., Riboli, E., ... & Carroll, R. J. (2002). Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public health nutrition*, *5*(6a), 915-923.

Kopelman, P. G. (2000). Obesity as a medical problem. *Nature*, *404*(6778), 635-643.

Lakatta, E. G., & Levy, D. (2003). Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: Part I: aging arteries: a “set up” for vascular disease. *Circulation*, *107*(1), 139-146.

Lavie, C. J., Milani, R. V., & Ventura, H. O. (2009). Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss. *Journal of the American college of cardiology*, *53*(21), 1925-1932.

Li, C., & Shepherd, B. E. (2012). A new residual for ordinal outcomes. *Biometrika*, *99*(2), 473-480.

Liao, D., Cai, J., Barnes, R. W., Tyroler, H. A., Rautaharju, P., Holme, I., & Heiss, G. (1996). Association of cardiac automatic function and the development of hypertension: The ARIC study. *American journal of hypertension*, *9*(12), 1147-1156.

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13-22.

Libby, P., Ridker, P. M., & Maseri, A. (2002). Inflammation and atherosclerosis. *Circulation*, *105*(9), 1135-1143.

Lipsitz, S. R., Fitzmaurice, G. M., Orav, E. J., & Laird, N. M. (1994). Performance of generalized estimating equations in practical situations. *Biometrics*, 270-278.

Lesaffre, E., & Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, 570-582.

Maas, A. H., & Appelman, Y. E. (2010). Gender differences in coronary heart disease. *Netherlands Heart Journal*, *18*, 598-603.

Mayer-Davis, E. J., Lawrence, J. M., Dabelea, D., Divers, J., Isom, S., Dolan, L., ... & Wagenknecht, L. (2017). Incidence trends of type 1 and type 2 diabetes among youths, 2002–2012. *New England Journal of Medicine*, *376*(15), 1419-1429.

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (Methodological)*, *42*(2), 109-127.

Menke, A., Casagrande, S., Geiss, L., & Cowie, C. C. (2015). Prevalence of and trends in diabetes among adults in the United States, 1988-2012. *Jama*, *314*(10), 1021-1029.

Mosca, L., Barrett-Connor, E., & Kass Wenger, N. (2011). Sex/gender differences in cardiovascular disease prevention: what a difference a decade makes. *Circulation*, *124*(19), 2145-2154.

Mukamal, K. J., Conigrave, K. M., Mittleman, M. A., Camargo Jr, C. A., Stampfer, M. J., Willett, W. C., & Rimm, E. B. (2003). Roles of drinking pattern and type of alcohol consumed in coronary heart disease in men. *New England Journal of Medicine*, *348*(2), 109-118.

Muntner, P., Carey, R. M., Gidding, S., Jones, D. W., Taler, S. J., Wright, J. T., & Whelton, P. K. (2018). Potential US population impact of the 2017 ACC/AHA high blood pressure guideline. *Journal of the American College of Cardiology*, *71*(2), 109-118.

Muntner, P., Hardy, S. T., Fine, L. J., Jaeger, B. C., Wozniak, G., Levitan, E. B., & Colantonio, L. D. (2020). Trends in blood pressure control among US adults with hypertension, 1999-2000 to 2017-2018. *Jama*, *324*(12), 1190-1200.

Nkomo, V. T., Gardin, J. M., Skelton, T. N., Gottdiener, J. S., Scott, C. G., & Enriquez-Sarano, M. (2006). Burden of valvular heart diseases: a population-based study. *The lancet*, *368*(9540), 1005-1011.

Norby, F. L., Soliman, E. Z., Chen, L. Y., Bengtson, L. G., Loehr, L. R., Agarwal, S. K., & Alonso, A. (2016). Trajectories of cardiovascular risk factors and incidence of atrial fibrillation over a 25-year follow-up: the ARIC Study (Atherosclerosis Risk in Communities). *Circulation*, *134*(8), 599-610.

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics*, *57*(1), 120-125.

Pedamallu, H., Zmora, R., Perak, A. M., & Allen, N. B. (2023). Life course cardiovascular health: risk factors, outcomes, and interventions. *Circulation research*, *132*(12), 1570-1583.

Penninx, B. W., Beekman, A. T., Honig, A., Deeg, D. J., Schoevers, R. A., Van Eijk, J. T., & Van Tilburg, W. (2001). Depression and cardiac mortality: results from a community-based longitudinal study. *Archives of general psychiatry*, *58*(3), 221-227.

Peterson, B., & Harrell Jr, F. E. (1990). Partial proportional odds models for ordinal response variables. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *39*(2), 205-217.

Petersen, M. L., & van der Laan, M. J. (2014). Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology*, *25*(3), 418-426.

Piironen, J., & Vehtari, A. (2017). Sparsity information and regularization in the horseshoe and other shrinkage priors.

Pletcher, M. J., Vittinghoff, E., Thanataveerat, A., Bibbins-Domingo, K., & Moran, A. E. (2016). Young adult exposure to cardiovascular risk factors and risk of events later in life: the Framingham Offspring Study. *PloS one*, *11*(5), e0154288.

Powell, K. E., Paluch, A. E., & Blair, S. N. (2011). Physical activity for health: What kind? How much? How intense? On top of what?. *Annual review of public health*, *32*(1), 349-365.

Preisser, J. S., & Qaqish, B. F. (1996). Deletion diagnostics for generalised estimating equations. *Biometrika*, *83*(3), 551-562.

Prentice, R. L., & Zhao, L. P. (1991). Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. *Biometrics*, 825-839.

Read, J. R., Sharpe, L., Modini, M., & Dear, B. F. (2017). Multimorbidity and depression: a systematic review and meta-analysis. *Journal of affective disorders*, *221*, 36-46.

Reedy, J., Krebs-Smith, S. M., Miller, P. E., Liese, A. D., Kahle, L. L., Park, Y., & Subar, A. F. (2014). Higher diet quality is associated with decreased risk of all-cause, cardiovascular disease, and cancer mortality among older adults. *The Journal of nutrition*, *144*(6), 881-889.

Regitz-Zagrosek, V., Oertelt-Prigione, S., Prescott, E., Franconi, F., ... & Stangl, V. (2016). Gender in cardiovascular diseases: impact on clinical manifestations, management, and outcomes. *European heart journal*, *37*(1), 24-34.

Ronksley, P. E., Brien, S. E., Turner, B. J., Mukamal, K. J., & Ghali, W. A. (2011). Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. *Bmj*, *342*, d671.

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology* (Vol. 3). Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Royston, P., & Sauerbrei, W. (2008). *Multivariable model-building: a pragmatic approach to regression anaylsis based on fractional polynomials for modelling continuous variables*. John Wiley & Sons.

Rubin, D. B. (2008). For objective causal inference, design trumps analysis.

Streiner, D. L. (2002). Breaking up is hard to do: the heartbreak of dichotomizing continuous data. *The Canadian Journal of Psychiatry*, *47*(3), 262-266.

Stringhini, S., Sabia, S., Shipley, M., Brunner, E., Nabi, H., Kivimaki, M., & Singh-Manoux, A. (2010). Association of socioeconomic position with health behaviors and mortality. *Jama*, *303*(12), 1159-1166.

Sullivan Pepe, M., & Anderson, G. L. (1994). A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. *Communications in statistics-simulation and computation*, *23*(4), 939-951.

Tian, Y., Li, D., Cui, H., Zhang, X., Fan, X., & Lu, F. (2024). Epidemiology of multimorbidity associated with atherosclerotic cardiovascular disease in the United States, 1999–2018. *BMC public health*, *24*(1), 267.

Tinetti, M. E., Fried, T. R., & Boyd, C. M. (2012). Designing health care for the most common chronic condition—multimorbidity. *Jama*, *307*(23), 2493-2494.

THORPE, J., Roland, J., RICHARD, P., BOWIE, J. V., LAVEIST, T. A., & GASKIN, D. J. (2013). Economic burden of men's health disparities in the United States. *International Journal of Men's Health*, *12*(3).

Twisk, J. W. (2013). *Applied longitudinal data analysis for epidemiology: a practical guide*. cambridge university press.

Van Calster, B., Nieboer, D., Vergouwe, Y., De Cock, B., Pencina, M. J., & Steyerberg, E. W. (2016). A calibration hierarchy for risk models was defined: from utopia to empirical data. *Journal of clinical epidemiology*, *74*, 167-176.

Verbeke, G., & Molenberghs, G.,(1997). *Linear mixed models for longitudinal data* (pp. 63-153). Springer New York.

Violan, C., Foguet-Boreu, Q., Flores-Mateo, G., Salisbury, C., Blom, J., Freitag, M., ... & Valderas, J. M. (2014). Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. *PloS one*, *9*(7), e102149.

Westreich, D., & Greenland, S. (2013). The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *American journal of epidemiology*, *177*(4), 292-298.

Whooley, M. A., & Wong, J. M. (2013). Depression and cardiovascular disorders. *Annual review of clinical psychology*, *9*(1), 327-354.

Willi, C., Bodenmann, P., Ghali, W. A., Faris, P. D., & Cornuz, J. (2007). Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. *Jama*, *298*(22), 2654-2664.

Williams, D. R., Mohammed, S. A., Leavell, J., & Collins, C. (2010). Race, socioeconomic status, and health: complexities, ongoing challenges, and research opportunities. *Annals of the new York Academy of Sciences*, *1186*(1), 69-101.

Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *The stata journal*, *6*(1), 58-82.

Wilkins, J. T., Ning, H., Berry, J., Zhao, L., Dyer, A. R., & Lloyd-Jones, D. M. (2012). Lifetime risk and years lived free of total cardiovascular disease. *Jama*, *308*(17), 1795-1801.

Wilson, P. W., D'Agostino, R. B., Parise, H., Sullivan, L., & Meigs, J. B. (2005). Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation*, *112*(20), 3066-3072.

Winkleby, M. A., Jatulis, D. E., Frank, E., & Fortmann, S. P. (1992). Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *American journal of public health*, *82*(6), 816-820.

Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 121-130.

Table 1. Baseline and Longitudinal Characteristics of Study Participants Across Visits (N = 2000)

Variable	Visit(s)	Category/Value	Count (%), N=2000
Sex	All Visits	Male	722 (36.10)
		Female	1278 (63.90)
Education Level	All Visits	Less than High School	734 (36.70)
		Completed High School	537 (26.85)
		College or Higher	729 (36.45)
CES-D-based Depression Status	All Visits	No Depression	1488 (74.40)
		Depressed	512 (25.60)
Physical Activity (2008 Guidelines)	All Visits	Did Not Meet	758 (37.90)
		Met	1242 (62.10)
Diabetes Mellitus	Visit 1	No Diabetes	1603 (80.15)
		Has Diabetes (any type)	397 (19.85)
	Visit 2	No Diabetes	1430 (71.50)
		Has Diabetes (any type)	570 (28.50)
	Visit 3	No Diabetes	1267 (63.35)
		Has Diabetes (any type)	633 (36.65)
High Cholesterol	Visit 1	No	1129 (56.45)
		Yes	871 (43.55)
	Visit 2	No	1050 (52.50)
		Yes	950 (47.50)
	Visit 3	No	1042 (52.10)
		Yes	958 (47.90)
Hypertension (High Blood Pressure)	Visit 1	No	1403 (70.15)
		Yes	597 (29.85)
	Visit 2	No	1181 (59.05)
		Yes	819 (40.95)
	Visit 3	No	1184 (59.20)
		Yes	816 (40.80)
Composite Comorbidity Outcome*	Visit 1	0 Conditions	789 (39.45)
		1 Condition	701 (35.05)
		2 Conditions	366 (18.30)
		3 Conditions	144 (7.20)
	Visit 2	0 Conditions	642 (32.10)
		1 Condition	654 (32.70)
		2 Conditions	427 (21.35)
		3 Conditions	277 (13.85)
	Visit 3	0 Conditions	398 (19.90)
		1 Condition	847 (42.35)
		2 Conditions	605 (30.25)
		3 Conditions	150 (7.50)

Alcohol Use	**Visit 1**	**No**	1070 (53.50)
		Yes	930 (46.50)
	Visit 2	**No**	881 (44.05)
		Yes	1119 (55.95)
	Visit 3	**No**	1029 (51.45)
		Yes	971 (48.55)
Smoking Status	**Visit 1**	**No**	1658 (82.90)
		Yes	342 (17.10)
	Visit 2	**No**	1729 (86.45)
		Yes	271 (13.55)
	Visit 3	**No**	1744 (87.20)
		Yes	256 (12.80)
Obesity (BMI-Based)	**Visit 1**	**Not Obese**	1152 (57.60)
		Obese	848 (42.40)
	Visit 2	**Not Obese**	1129 (56.45)
		Obese	871 (43.55)
	Visit 3	**Not Obese**	1019 (50.95)
		Obese	981 (49.05)
Age at Visit, Mean ± SD	**Visit 1**	47.78 ± 11.8	
	Visit 2	53.77 ± 11.8	
	Visit 3	58.76 ± 11.8	
Body Mass Index (BMI), Mean ± SD	**Visit 1**	29.84 ± 5.7	
	Visit 2	30.16 ± 5.9	
	Visit 3	30.69 ± 6.0	
Healthy Eating Index–2010 (HEI-2010), Mean ± SD	**All Visits**	60.63 ± 13.8	

** Composite Outcome is the total number of conditions present among Diabetes Mellitus, High Cholesterol, and Hypertension.*

Table 2. Top Three Participants with Highest Body Mass Index (BMI) Values Across Visits

Visit(s)	Top 1 (ID)	BMI Value	Top 2 (ID)	BMI Value	Top 3 (ID)	BMI Value
Visit 1	1695	70.3	215	62.2	1875	59.5
Visit 2	1695	70.9	1875	60.6	963	60.5
Visit 3	1695	70.4	1875	62.3	963	60.7

Table 3. Temporal Changes in Categorical Health and Behavioral Statuses Between Consecutive Visits

Variable	Interval	Status	Count (%), N=2000
Diabetes Mellitus	**Visit 1 to Visit 2**	**Decrease**	38 (1.90)
		Increase	211 (10.55)
	Visit 2 to Visit 3	**Decrease**	0 (0.00)
		Increase	163 (8.15)
High Cholesterol	**Visit 1 to Visit 2**	**Decrease**	208 (10.40)
		Increase	287 (14.35)
	Visit 2 to Visit 3	**Decrease**	480 (24.00)
		Increase	488 (24.40)
Hypertension (High Blood Pressure)	**Visit 1 to Visit 2**	**Decrease**	66 (3.30)
		Increase	288 (14.40)
	Visit 2 to Visit 3	**Decrease**	446 (22.30)
		Increase	443 (22.15)
Alcohol Use	**Visit 1 to Visit 2**	**Decrease**	171 (8.55)
		Increase	360 (18.00)
	Visit 2 to Visit 3	**Decrease**	566 (28.30)
		Increase	418 (20.90)
Smoking Status	**Visit 1 to Visit 2**	**Decrease**	119 (5.95)
		Increase	48 (2.40)
	Visit 2 to Visit 3	**Decrease**	214 (10.70)
		Increase	199 (9.95)
Obesity (BMI-Based)	**Visit 1 to Visit 2**	**Decrease**	120 (6.00)
		Increase	143 (7.15)
	Visit 2 to Visit 3	**Decrease**	37 (1.85)
		Increase	147 (7.35)
Composite Comorbidity Outcome	**Visit 1 to Visit 2**	**Decrease**	231 (11.55)
		Increase	599 (29.95)
	Visit 2 to Visit 3	**Decrease**	573 (28.65)
		Increase	717 (35.85)

Table 4. Variability and Reliability Metrics for Body Mass Index (BMI)

Metrics	Estimate
Within-subject Variance	2.73
Intraclass Correlation Coefficient (ICC)	0.92

Table 5. Cross-Sectional Associations Between Key Predictors Across Visits: Baseline Diagnostics for Potential Collinearity

Variable 1	Variable 2	Comparison Method	Visit(s)	Statistic	P-value
BMI	**Obesity**	**Point-Biserial Correlation**	**Visit 1**	0.766	<.0001
			Visit 2	0.757	<.0001
			Visit 3	0.750	<.0001
HEI2010	**BMI**	**Spearman Correlation**	**Visit 1**	-0.0442	0.0485
			Visit 2	-0.0854	0.0001
			Visit 3	-0.0860	0.0001
Alcohol Use	**Smoking Status**	**Chi-squared Statistics**	**Visit 1**	46.0156	<.0001
			Visit 2	25.5051	<.0001
			Visit 3	0.5851	0.4443
Age at Visit	**BMI**	**Spearman Correlation**	**Visit 1**	0.0437	0.0506
			Visit 2	-0.0380	0.0894
			Visit 3	-0.0310	0.1665
Age at Visit	**HEI2010**	**Spearman Correlation**	**Visit 1**	0.1566	<.0001
			Visit 2	0.1561	<.0001
			Visit 3	0.1544	<.0001
Sex	**Smoking Status**	**Chi-squared Statistics**	**Visit 1**	27.6672	<.0001
			Visit 2	22.8864	<.0001
			Visit 3	1.4309	0.2316
Sex	**Alcohol Use**	**Chi-squared Statistics**	**Visit 1**	92.9188	<.0001
			Visit 2	32.7687	<.0001
			Visit 3	0.1781	0.6730
Depression	**Smoking Status**	**Chi-squared Statistics**	**Visit 1**	7.7383	0.0054
			Visit 2	2.2284	0.1355
			Visit 3	1.6222	0.2028
Depression	**Alcohol Use**	**Chi-squared Statistics**	**Visit 1**	28.8175	<.0001
			Visit 2	31.7236	<.0001
			Visit 3	8.9104	0.0028
Education Level	**HEI2010**	**Spearman Correlation**	**All Visits**	-0.0639	0.0043
Education Level	**PAG2008**	**Chi-squared Statistics**	**All Visits**	6.5753	0.0373

Table 6. Repeated Measures Point-Biserial Correlation Between BMI and Obesity Status

Response Variable	Predictor Variable	Estimated Mean Difference	Between-subject Variance	Within-subject Variance	Approximate Correlation $r_{pb}^{approx} = \frac{\mu_1 - \mu_0}{\sqrt{\sigma^2_{total}}}$
BMI	**Obesity**	4.07	18.31	2.28	0.90

Table 7. Longitudinal Association Between Continuous Variables with Subject-Level Variance, Residual Variance, and Intraclass Correlation Coefficient (ICC) from Mixed Effects Models

Response Variable	Predictor Variable	Estimate (SE)	t-value (DF)	P-value	Subject-Level Variance	Residual Variance	Intraclass Correlation Coefficient (ICC)
BMI	**HEI2010**	-0.0384 (0.0092)	-4.17 (1998)	<.0001	31.55	2.72	0.921
Age	**HEI2010**	0.1418 (0.0188)	7.55 (1998)	<.0001	124.25	30.94	0.801
Age	**BMI**	0.3426 (0.0349)	9.81 (1998)	<.0001	134.06	29.66	0.819
BMI	**Age**	0.0622 (0.0042)	14.83 (1998)	<.0001	32.72	2.54	0.928

Table 8. Working Correlations and GEE Z-Statistics for Pairwise Associations Between Binary Variables Across Repeated Measures

Response Variable	Predictor Variable	Working Correlation	Z-value	P-value
Smoking Status	**Alcohol Use**	0.2974	-6.33	<.0001
Alcohol Use	**Smoking Status**	0.1626	-6.49	<.0001
Sex	**Smoking Status**	0.9999	4.84	<.0001
Smoking Status	**Sex**	0.2910	5.19	<.0001
Sex	**Alcohol Use**	0.9999	7.39	<.0001
Alcohol Use	**Sex**	0.1543	7.45	<.0001
Depression	**Smoking Status**	0.9999	-6.11	<.0001
Smoking Status	**Depression**	0.2709	36.11	<.0001
Depression	**Alcohol Use**	0.9999	1.49	0.1356
Alcohol Use	**Depression**	0.2974	-6.33	<.0001

Table 9. Brant Test for Proportional Odds Assumption at Each Visit

Predictor Variables	χ^2 (DF), visit 1	P-value	χ^2 (DF), visit 2	P-value	χ^2 (DF), visit 3	P-value
Omnibus	**53.79 (20)**	**<0.01**	**35.44 (20)**	**0.02**	18.02	0.59
Sex	**15.47 (2)**	**<0.01**	1.42 (2)	0.49	1.46	0.48
Education=2	0.7 (2)	0.70	1.69 (2)	0.43	0.13	0.94
Education=3	2.9 (2)	0.23	1.39 (2)	0.5	2.25	0.33
HEI2010	1.75 (2)	0.42	2.47 (2)	0.29	1.28	0.53
PAG2008	1.34 (2)	0.51	0.81 (2)	0.67	1.57	0.46
Depression	4.03 (2)	0.13	0.33 (2)	0.85	0.58	0.75
Age	**16.42 (2)**	**<0.01**	**14.52 (2)**	**<0.01**	1.79	0.41
Alcohol Use	2.99 (2)	0.22	2.27 (2)	0.32	0.27	0.87
BMI	0.88 (2)	0.65	4.02 (2)	0.13	3.01	0.22
Smoking Status	1.79 (2)	0.41	0.19 (2)	0.91	**5.47**	**0.06**

Note. *The Brant test assesses the proportional odds (PO) assumption for each predictor in an ordinal logistic regression model. The χ^2 statistic and corresponding degrees of freedom (df) are reported for each variable at Visits 1, 2, and 3. The p-value indicates whether the assumption is violated ($p < 0.05$). Bolded p-values indicate either statistically significant violations ($p < 0.05$) or borderline evidence ($0.05 \leq p < 0.10$) of non-proportionality. A significant omnibus test suggests that at least one variable violates the PO assumption. Significant results for individual predictors imply that the log-odds of the outcome are not consistent across cumulative thresholds, indicating potential non-proportionality. Results vary by visit, highlighting the importance of testing PO assumptions longitudinally.*

Table 10. Likelihood Ratio Test comparing proportional odds and non-proportional odds models across visits. The test compares the ordinal logistic model (with cumulative logit link) to the nominal logistic model (generalized logit), with the null hypothesis that the proportional odds assumption holds, ($\alpha = 0.05$).

Visit	Ordinal Logistic Model -2log L (df=10)	Nominal Logistic Model -2log L (df=30)	Likelihood Ratio Test (df=20)	P-value
Visit 1	4416.338	4359.822	56.516	<.0001
Visit 2	4783.098	4738.057	45.041	0.0011
Visit 3	4843.540	4825.248	18.292	0.5974

Table 11. Parameter Estimates from the Partial Proportional Odds Logistic Model (Visit 1)

Coefficients	Threshold	Estimate	Standard Error	Z-value	P-value
Intercept	–	5.357	0.42	12.791	<.0001
Sex	**≥ 1**	0.733	0.11	6.411	<.0001
	≥ 2	0.614	0.12	5.330	<.0001
	≥ 3	0.515	0.17	2.985	0.0028
Education Level 2	**≥ 1**	-0.227	0.13	-1.737	0.0823
	≥ 2	0.225	0.13	1.685	0.0920
	≥ 3	0.279	0.20	1.391	0.1643
Education Level 3	**≥ 1**	-0.039	0.12	-0.325	0.7452
	≥ 2	0.479	0.13	3.781	0.0001
	≥ 3	0.534	0.20	2.682	0.0073
HEI2010	**Proportional**	0.001	0.003	0.255	0.7989
PAG2008	**Proportional**	0.118	0.09	1.297	0.1947
Depression	**≥ 1**	-0.358	0.12	-2.924	0.0035
	≥ 2	-0.081	0.12	-0.659	0.5100
	≥ 3	0.099	0.19	0.515	0.6066
Age	**≥ 1**	-0.085	0.005	-18.492	<0.0001
	≥ 2	-0.059	0.005	-13.062	<0.0001
	≥ 3	-0.032	0.005	-6.431	<0.0001
Alcohol Use	**≥ 1**	0.007	0.11	0.063	0.9495
	≥ 2	0.391	0.11	3.466	0.0005
	≥ 3	0.836	0.19	4.474	<0.0001
BMI	**Proportional**	-0.076	0.01	-9.516	<0.0001
Smoking Status	**Proportional**	0.098	0.12	0.812	0.4170

Table 12. Parameter Estimates from the Partial Proportional Odds Logistic Model (Visit 2)

Coefficients	Threshold	Estimate	Standard Error	Z-value	P-value
Intercept	–	6.356	0.43	14.837	<.0001
Sex	**Proportional**	0.510	0.09	5.536	<.0001
Education Level 2	**≥ 1**	-0.343	0.14	-2.489	0.0128
	≥ 2	-0.015	0.13	-0.115	0.9084
	≥ 3	-0.074	0.16	-0.459	0.6463
Education Level 3	**≥ 1**	-0.111	0.13	-0.890	0.3732
	≥ 2	0.195	0.12	1.647	0.0996
	≥ 3	0.304	0.16	1.904	0.0569
HEI2010	**Proportional**	0.003	0.003	0.866	0.3863
PAG2008	**Proportional**	0.050	0.09	0.564	0.5729
Depression	**Proportional**	-0.176	0.10	-1.783	0.0747
Age	**≥ 1**	-0.083	0.005	-17.846	<.0001
	≥ 2	-0.078	0.005	-16.622	<.0001
	≥ 3	-0.061	0.005	-11.353	<.0001
Alcohol Use	**≥ 1**	0.073	0.11	0.670	0.5029
	≥ 2	0.288	0.10	2.811	0.0050
	≥ 3	0.482	0.13	3.584	0.0003
BMI	**≥ 1**	-0.107	0.01	-11.648	<.0001
	≥ 2	-0.071	0.01	-8.737	<.0001
	≥ 3	-0.062	0.01	-6.576	<.0001
Smoking Status	**Proportional**	-0.170	0.13	-1.336	0.1816

Table 13. Parameter Estimates from the Partial Proportional Odds Logistic Model (Visit 3)

Coefficients	Threshold	Estimate	Standard Error	Z-value	P-value
Intercept	–	2.634	0.39	6.807	<.0001
Sex	**≥ 1**	0.165	0.12	1.369	0.1708
	≥ 2	0.143	0.10	1.418	0.1562
	≥ 3	0.236	0.17	1.380	0.1677
Education Level 2	**Proportional**	0.088	0.11	0.825	0.4094
Education Level 3	**Proportional**	0.161	0.10	1.635	0.1020
HEI2010	**Proportional**	-0.001	0.003	-0.397	0.6911
PAG2008	**≥ 1**	-0.098	0.117	-0.835	0.4037
	≥ 2	0.227	0.097	2.345	0.0190
	≥ 3	0.367	0.163	2.249	0.0245
Depression	**Proportional**	-0.382	0.098	-3.897	<.0001
Age	**Proportional**	-0.026	0.004	-6.994	<.0001
Alcohol Use	**Proportional**	0.065	0.083	0.789	0.4299
BMI	**≥ 1**	-0.083	0.008	-10.327	<.0001
	≥ 2	-0.025	0.007	-3.389	0.0007
	≥ 3	0.038	0.009	4.349	<.0001
Smoking Status	**≥ 1**	-0.280	0.183	-1.534	0.1249
	≥ 2	-0.147	0.139	-1.060	0.2891
	≥ 3	-0.480	0.215	-2.237	0.0253

Table 14. Evaluation of Partial Proportional Odds Model Fit Across Study Visits

Visit	Residual Deviance (DF)	Log-likelihood	AIC	BIC	Number of Outliers (|Pearson Residual| > 4)	Max. Absolute Pearson Residual (Observation ID, Threshold Category)
Visit 1	4513.153 (5974)	-2256.58	4551.153	4657.56	29	6.38 (ID: 1841, Category: 3)
Visit 2	4813.354 (5976)	-2406.68	4855.354	4972.96	11	5.23 (ID: 1256, Category: 1)
Visit 3	4922.041 (5978)	-2461.02	4960.041	5066.45	40	5.99 (ID: 1052, Category: 3)

Table 15. Threshold-Specific Fixed Effects Estimates from Cumulative Binary Logistic Mixed Models for Assessing the Proportional Odds Assumption

	Outcome ≥ 1				Outcome ≥ 2				Outcome = 3			
Variables	**Estimate**	**SE**	**t-value (df)**	**P-value**	**Estimate**	**SE**	**t-value (df)**	**P-value**	**Estimate**	**SE**	**t-value (df)**	**P-value**
Intercept	-4.357	0.42	-10.28 (1992)	<.0001	-8.206	0.51	-16.02 (1992)	<.0001	-11.598	0.81	-14.27 (1992)	<.0001
Visit=2	0.032	0.09	0.37 (3992)	0.7110	0.234	0.09	2.51 (3992)	0.0120	0.541	0.14	3.90 (3992)	<.0001
Visit=3	0.550	0.10	5.64 (3992)	<.0001	-0.082	0.10	-0.80 (3992)	0.4234	-1.023	0.17	-6.04 (3992)	<.0001
Female=1	-0.763	0.11	-7.18 (3992)	<.0001	-0.453	0.11	-3.95 (3992)	<.0001	-0.433	0.16	-2.68 (3992)	0.0074
Education=2	0.020	0.12	0.16 (3992)	0.8735	-0.040	0.14	-0.29 (3992)	0.7681	0.018	0.19	0.10 (3992)	0.9227
Education=3	-0.133	0.11	-1.18 (3992)	0.2400	-0.253	0.12	-2.03 (3992)	0.0427	-0.345	0.18	-1.95 (3992)	0.0514
PAG2008=1	-0.129	0.10	-1.28 (3992)	0.2008	-0.203	0.11	-1.84 (3992)	0.0660	-0.196	0.15	-1.28 (3992)	0.2004
Depression=1	0.419	0.12	3.63 (3992)	0.0003	0.363	0.12	2.96 (3992)	0.0031	0.126	0.17	0.74 (3992)	0.4594
Alcohol Use=1	-0.149	0.08	-1.85 (3992)	0.0640	-0.228	0.08	-2.74 (3992)	0.0061	-0.437	0.13	-3.49 (3992)	0.0005
Smoking Status=1	0.031	0.12	0.27 (3992)	0.7896	-0.043	0.12	-0.35 (3992)	0.7268	0.282	0.18	1.56 (3992)	0.1185
HEI2010	-0.004	0.004	-1.04 (3992)	0.2974	0.002	0.004	0.53 (3992)	0.5995	0.005	0.006	0.83 (3992)	0.4063
Age	0.073	0.005	16.14 (3992)	<.0001	0.087	0.005	16.48 (3992)	<.0001	0.096	0.008	11.76 (3992)	<.0001
BMI	0.079	0.009	9.07 (3992)	<.0001	0.089	0.009	9.75 (3992)	<.0001	0.282	0.18	1.56 (3992)	0.1185

Table 16. Model fit statistics for evaluating the inclusion of a random intercept.

Model	AIC	BIC
Mixed-effects model (with random intercept)	13661	13750
Generalized linear model (fixed effects only)	14339	14439

Table 17. Estimated Variance Component for the Random Intercept from Multinomial GLMMs With and Without Interaction Terms. Estimates and standard errors (SE) are reported.

Models	Estimate	Standard Error
Without covariates interact with visit	2.194	0.15
With covariates interact with visit	2.478	0.17

Table 18. Fixed Effects Estimates from a Multinomial Generalized Linear Mixed Model (GLMM) with Cumulative Logit Link. The model excludes interaction terms between visit and other predictors. Standard errors (SE), odds ratio (OR) estimate, *t*-statistics with degrees of freedom (df), and p-values are reported.

Predictor	Estimate	SE	OR Estimate	t-value (df)	P-value
Intercept, outcome=0	4.936	0.37	—	13.24 (1992)	<.0001
Intercept, outcome=1	7.388	0.38	—	19.27 (1992)	<.0001
Intercept, outcome=2	9.664	0.40	—	24.38 (1992)	<.0001
Visit=2	-0.209	0.07	0.811	-3.09 (3990)	0.0020
Visit=3	-0.029	0.08	0.971	-0.38 (3990)	0.7010
Female=1	0.568	0.09	1.765	6.20 (3990)	<.0001
Education=2	0.005	0.11	1.005	0.05 (3990)	0.9636
Education=3	0.226	0.10	1.254	2.26 (3990)	0.0239
PAG2008=1	0.0015	0.09	1.002	1.27 (3990)	0.2032
Depression=1	-0.316	0.10	0.729	-3.19 (3990)	0.0014
Alcohol Use=1	0.210	0.06	1.234	3.40 (3990)	0.0007
Smoking Status=1	0.007	0.09	1.007	0.07 (3990)	0.9412
HEI2010	0.0015	0.003	1.002	0.48 (3990)	0.6332
Age	-0.079	0.004	0.924	-19.96 (3990)	<.0001
BMI	-0.086	0.007	0.918	-11.90 (3990)	<.0001

Table 19. Type III Tests of Fixed Effects from the Multinomial Generalized Linear Mixed Model (GLMM). The model excludes interaction terms involving visit. F-statistics are reported with numerator and denominator degrees of freedom (DF), along with corresponding *p*-values.

Effect	Numerator DF	Denominator DF	F-value	p-value
Visit	2	3990	6.28	0.0019
Sex	1	3990	38.40	<.0001
Education	2	3990	3.19	0.0411
HEI2010	1	3990	0.23	0.6332
PAG2008	1	3990	1.62	0.2032
Depression	1	3990	10.18	0.0014
Age	1	3990	398.52	<.0001
Alcohol Use	1	3990	11.53	0.0007
BMI	1	3990	141.60	<.0001
Smoking Status	1	3990	0.01	0.9412

Table 20. Fixed Effects Estimates from a Multinomial Generalized Linear Mixed Model (GLMM) with Cumulative Logit Link, Including Interaction Terms Between Visit and Covariates. The model accounts for subject-level random intercepts and was estimated using adaptive quadrature. Reported are coefficient estimates, standard errors (SE), odds ratio (OR) estimate, *t*-statistics with corresponding denominator degrees of freedom (df), and *p*-values.

Predictor	Estimate	SE	OR Estimate	t-value (df)	P-value
Intercept, outcome=0	5.138	0.53	—	9.67 (1992)	<.0001
Intercept, outcome=1	7.701	0.54	—	14.29 (1992)	<.0001
Intercept, outcome=2	10.069	0.55	—	18.39 (1992)	<.0001
Visit=2	2.639	0.63	13.999	4.20 (3970)	<.0001
Visit=3	-4.087	0.61	0.017	-6.76 (3970)	<.0001
Female=1	0.901	0.13	2.462	7.03 (3970)	<.0001
Education=2	0.202	0.15	1.224	1.34 (3970)	0.1810
Education=3	0.451	0.14	1.570	3.24 (3970)	0.0012
PAG2008=1	0.179	0.12	1.196	1.45 (3970)	0.1459
Depression=1	-0.098	0.006	0.907	-1.58 (3970)	0.1133
Alcohol Use=1	0.202	0.11	1.224	1.83 (3970)	0.0679
Smoking Status=1	0.314	0.15	1.369	2.12 (3970)	0.0342
HEI2010	0.0006	0.004	1.001	1.27 (3970)	0.2051
Age	-0.098	0.006	0.907	-17.49 (3970)	<.0001
BMI	-0.087	0.010	0.917	-8.36 (3970)	<.0001
Visit × Female (Visit = 2, Female = 1)	-0.151	0.14	0.860	-1.06 (3970)	0.2895
Visit × Female (Visit = 3, Female = 1)	-0.683	0.14	0.505	-4.89 (3970)	<.0001
Visit × Education (Visit = 2, Education = 2)	-0.389	0.17	0.678	-2.31 (3970)	0.0211
Visit × Education (Visit = 2, Education = 3)	-0.256	0.16	0.774	-1.65 (3970)	0.0982
Visit × Education (Visit = 3, Education = 2)	0.009	0.17	1.009	0.05 (3970)	0.9586
Visit × Education (Visit = 3, Education = 3)	-0.149	0.15	0.862	-0.98 (3970)	0.3266
Visit × PAG2008 (Visit = 2, PAG2008 = 1)	-0.110	0.14	0.896	-0.81 (3970)	0.4196
Visit × PAG2008 (Visit = 3, PAG2008 = 1)	-0.050	0.13	0.951	-0.37 (3970)	0.7129
Visit × Depression (Visit = 2, Depression = 1)	-0.034	0.15	0.967	-0.22 (3970)	0.8242
Visit × Depression (Visit = 3, Depression = 1)	-0.192	0.15	0.825	-1.28 (3970)	0.2003
Visit × Alcohol Use (Visit = 2, Alcohol Use = 1)	0.031	0.14	1.031	0.22 (3970)	0.8293
Visit × Alcohol Use (Visit = 3, Alcohol Use = 1)	-0.081	0.15	0.922	-0.54 (3970)	0.5858
Visit × Smoking (Visit = 2, Smoking = 1)	-0.432	0.20	0.649	-2.20 (3970)	0.0280
Visit × Smoking (Visit = 3, Smoking = 1)	-0.596	0.21	0.551	-2.87 (3970)	0.0042
Visit × HEI2010 (Visit = 2)	-0.001	0.005	0.999	-0.16 (3970)	0.8725
Visit × HEI2010 (Visit = 3)	-0.006	0.005	0.994	-1.17 (3970)	0.2421
Visit × Age (Visit = 2)	-0.020	0.006	0.980	-3.24 (3970)	0.0012
Visit × Age (Visit = 3)	0.066	0.006	1.068	10.93 (3970)	<.0001
Visit × BMI (Visit = 2)	-0.040	0.012	0.961	-3.36 (3970)	0.0008
Visit × BMI (Visit = 3)	0.045	0.012	1.046	3.91 (3970)	<.0001

Table 21. Type III Tests of Fixed Effects from a Multinomial Generalized Linear Mixed Model (GLMM) with Cumulative Logit Link. The model includes interaction terms between visit and other covariates. F-statistics are reported with numerator and denominator degrees of freedom (DF), along with corresponding *p*-values.

Effect	Numerator DF	Denominator DF	F-value	p-value
Visit	2	3970	77.81	<.0001
Sex	1	3970	42.03	<.0001
Education	2	3970	4.87	0.0077
HEI2010	1	3970	1.08	0.2994
PAG2008	1	3970	1.83	0.1767
Depression	1	3970	7.96	0.0048
Age	1	3970	398.67	<.0001
Alcohol Use	1	3970	8.33	0.0039
BMI	1	3970	127.67	<.0001
Smoking Status	1	3970	0.09	0.7677
Visit × Sex	2	3970	13.42	<.0001
Visit × Education	4	3970	3.18	0.0691
Visit × HEI2010	2	3970	0.82	0.4426
Visit × PAG2008	2	3970	0.33	0.7209
Visit × Depression	2	3970	0.95	0.3873
Visit × Age	2	3970	112.76	<.0001
Visit × Alcohol Use	2	3970	0.31	0.7347
Visit × BMI	2	3970	28.51	<.0001
Visit × Smoking	2	3970	4.68	0.0094

Table 22. Posterior Summaries of Fixed Effects and Random Intercepts from a Bayesian Cumulative Logit Mixed-Effects Model with Partial Proportional Odds Specification

Parameter	Estimate	Standard Error	OR Estimates	95% CI (Lower)	95% CI (Upper)	$\hat{R}$	Bulk ESS	Tail ESS
Subject-Level Random Effects								
SD(Intercept) of ID Level	1.51	0.05	—	1.42	1.62	1.00	622	817
Threshold-Specific Intercepts								
Cumulative Logit Intercept: Threshold 1	4.78	0.40	—	3.99	5.57	1.00	854	1153
Cumulative Logit Intercept: Threshold 2	7.42	0.41	—	6.59	8.23	1.00	821	1053
Cumulative Logit Intercept: Threshold 3	9.88	0.49	—	8.91	10.82	1.00	1055	1101
Proportional Odds Effects								
PAG2008=1	-0.12	0.09	0.887	-0.29	0.05	1.00	1111	1380
Age	0.08	<.01	1.083	-0.07	0.09	1.00	994	1383
HEI2010	-0.004	<.01	0.996	-0.01	0.01	1.00	923	1174
Non-Proportional Effects								
Visit 2, Threshold 1	0.01	0.08	1.010	-0.16	0.17	1.00	2002	1669
Visit 2, Threshold 2	0.24	0.08	1.271	0.07	0.40	1.00	1390	1492
Visit 2, Threshold 3	0.55	0.12	1.733	0.32	0.80	1.00	1777	1619
Visit 3, Threshold 1	0.41	0.10	1.507	0.21	0.60	1.00	1681	1821
Visit 3, Threshold 2	-0.02	0.09	0.980	-0.21	0.15	1.00	1237	1627
Visit 3, Threshold 3	-0.68	0.14	0.507	-0.95	-0.39	1.00	1394	1761
Female=1, Threshold 1	-0.78	0.11	0.458	-0.99	-0.56	1.00	873	1074
Female=1, Threshold 2	-0.44	0.10	0.644	-0.64	-0.23	1.00	753	1298
Female=1, Threshold 3	-0.39	0.14	0.677	-0.67	-0.12	1.00	1152	1609
Alcohol Use=1, Threshold 1	-0.12	0.08	0.887	-0.28	0.04	1.00	1593	1511
Alcohol Use=1, Threshold 2	-0.20	0.08	0.819	-0.36	-0.04	1.00	1525	1723
Alcohol Use=1, Threshold 3	-0.36	0.11	0.698	-0.59	-0.14	1.00	1954	1703
BMI, Threshold 1	0.08	0.01	1.083	0.06	0.10	1.00	1036	1415
BMI, Threshold 2	0.09	0.01	1.094	0.07	0.10	1.00	1031	1011
BMI, Threshold 3	0.10	0.01	1.105	0.07	0.12	1.00	1238	1498
Education=2, Threshold 1	0.02	0.13	1.020	-0.25	0.26	1.00	828	1123
Education=2, Threshold 2	-0.07	0.13	0.932	-0.32	0.18	1.00	724	1297
Education=2, Threshold 3	0.03	0.16	1.030	-0.29	0.36	1.00	1026	1563
Education=3, Threshold 1	-0.16	0.12	0.852	-0.40	0.07	1.00	897	1252
Education=3, Threshold 2	-0.27	0.12	0.763	-0.50	-0.04	1.00	831	1231
Education=3, Threshold 3	-0.35	0.15	0.704	-0.66	-0.06	1.00	996	1209
Depression=1, Threshold 1	0.46	0.12	1.584	0.23	0.69	1.00	994	1454
Depression=1, Threshold 2	0.30	0.12	1.350	0.06	0.53	1.00	843	1278
Depression=1, Threshold 3	0.08	0.15	1.083	-0.22	0.36	1.00	995	1236
Smoking=1, Threshold 1	-0.01	0.12	0.990	-0.24	0.23	1.00	2138	1587
Smoking=1, Threshold 2	-0.06	0.12	0.942	-0.29	0.16	1.00	1961	1631
Smoking=1, Threshold 3	0.13	0.16	1.13	0.19	0.44	1.00	1894	1510

***Note**: Reported are posterior means, standard errors (SE), exponentiated odds ratio (OR) estimates for covariate effects, 95% credible intervals, Gelman–Rubin convergence diagnostics ($\hat{R}$), and effective sample sizes for the bulk and tail of the posterior distributions (Bulk ESS, Tail ESS). Threshold-specific intercepts anchor the ordinal outcome scale and are not interpretable as odds ratios. Covariates with threshold-varying effects are grouped under non-proportional odds.*

Table 23. Model Fit Indices for the Bayesian Cumulative Logit Mixed-Effects Model with Partial Proportional Odds Specification, Based on Approximate Leave-One-Out Cross-Validation (LOO) and the Widely Applicable Information Criterion (WAIC).

Metric	Estimate	Standard Error	Notes
Leave-One-Out Cross Validation (LOOCV)			
Expected Log Predictive Density	-6497.1	46.9	Higher values indicate better predictive accuracy
Effective Number of Parameters	1246.9	14.7	Model flexibility; larger values indicate greater complexity
Leave-One-Out Information Criterion (LOOIC)	12994.2	93.9	LOOIC $= -2 \times$ Expected Log Predictive Density
Pareto-k Diagnostics	—	—	5995 (99.97%) observations had reliable $k \leq 0.7$; 2 (0.03%) had $0.7 < k \leq 1$; none >1
Minimum Effective Sample Size	302	—	Effective sample size for most extreme k; indicates stable estimation
Widely Applicable Information Criterion			
Expected Log Predictive Density	-6466.8	46.9	Same interpretation as above
Effective Number of Parameters	1216.6	14.7	Slightly lower than the value of Leave-One-Out
Widely Applicable Information Criterion (WAIC)	12933.5	93.8	WAIC= $-2 \times$ Expected Log Predictive Density
WAIC Reliability Warning	—	—	869 (14.5%) of pointwise Effective Number of Parameters > 0.4; suggests LOOCV is more stable for this model

Table 24. Threshold-Specific GEE Estimates from Binary Logistic Models for Assessing the Proportional Odds Assumption

	Outcome ≥ 1				Outcome ≥ 2				Outcome = 3			
Variables	**Estimate**	**SE**	**Z-value**	**P-value**	**Estimate**	**SE**	**Z-value**	**P-value**	**Estimate**	**SE**	**Z-value**	**P-value**
Intercept	-3.393	0.32	-10.45	<.0001	-5.803	0.34	-16.98	<.0001	-9.110	0.52	-17.46	<.0001
Visit=2	0.020	0.05	0.36	0.7177	0.155	0.05	2.93	0.0033	0.366	0.09	4.06	<.0001
Visit=3	0.424	0.08	5.35	<.0001	-0.074	0.08	-0.94	0.3460	-0.833	0.15	-5.73	<.0001
Female=1	-0.612	0.08	-7.56	<.0001	-0.305	0.08	-3.78	0.0002	-0.263	0.12	-2.20	0.0282
Education=2	0.017	0.10	0.17	0.8637	-0.035	0.09	-0.37	0.7101	0.094	0.14	0.66	0.5073
Education=3	-0.089	0.09	-1.03	0.3022	-0.193	0.09	-2.18	0.0292	-0.266	0.14	-1.88	0.0596
PAG2008=1	-0.072	0.08	-0.93	0.3526	-0.146	0.08	-1.90	0.0573	-0.099	0.12	-0.85	0.3973
Depression=1	0.296	0.09	3.25	0.0011	0.245	0.09	2.83	0.0047	0.149	0.13	1.15	0.2483
Alcohol Use=1	-0.112	0.06	-1.74	0.0823	-0.163	0.06	-2.74	0.0062	-0.319	0.09	-3.36	0.0008
Smoking Status=1	0.022	0.09	0.23	0.8146	-0.025	0.09	-0.28	0.7826	0.250	0.15	1.70	0.0890
HEI2010	-0.003	0.003	-1.18	0.2374	0.001	0.003	0.50	0.6188	0.004	0.004	0.99	0.3235
Age	0.056	0.003	17.76	<.0001	0.062	0.004	17.91	<.0001	0.078	0.006	13.13	<.0001
BMI	0.062	0.007	8.76	<.0001	0.063	0.006	9.94	<.0001	0.083	0.008	10.27	<.0001

Table 25. GEE Parameter Estimates from a Cumulative Logit Model for Ordinal Outcomes (No Visit-by-Covariate Interactions)

Parameter	Estimate	SE	OR Estimate	95% CI (Lower)	95% CI (Upper)	Z-value	P-value
Intercept, Threshold 1	-7.191	0.28	—	-7.732	-6.650	-26.06	<.0001
Intercept, Threshold 2	-5.502	0.27	—	-6.036	-4.969	-20.21	<.0001
Intercept, Threshold 3	-3.709	0.27	—	-4.230	-3.189	-13.96	<.0001
Visit = 2	0.167	0.04	1.182	0.086	0.247	4.03	<.0001
Visit = 3	0.056	0.06	1.058	-0.064	0.175	0.91	0.3639
Female = 1	-0.433	0.07	0.649	-0.564	-0.301	-6.46	<.0001
Education = 2	0.015	0.08	1.015	-0.145	0.175	0.18	0.8548
Education = 3	-0.155	0.08	0.856	-0.302	-0.008	-2.06	0.0393
PAG2008 = 1	-0.097	0.07	0.908	-0.226	0.033	-1.46	0.1448
Depression = 1	0.249	0.08	1.283	0.102	0.396	3.31	0.0009
Alcohol Use = 1	-0.185	0.05	0.831	-0.287	-0.082	-3.53	0.0004
Smoking = 1	0.061	0.08	1.063	-0.095	0.217	0.76	0.4465
HEI2010	-0.0007	0.002	0.999	-0.005	0.004	-0.30	0.7662
Age	0.058	0.003	1.060	0.053	0.064	21.70	<.0001
BMI	0.061	0.005	1.063	0.053	0.074	11.73	<.0001

***Note.** Model estimated using generalized estimating equations (GEE) with a multinomial distribution and cumulative logit link. Estimates include threshold-specific intercepts and covariate effects, with corresponding standard errors (SE), odds ratios (OR), 95% confidence intervals, Z-values, and p-values. An independent working correlation structure was used to account for within-subject correlation across repeated visits.*

Table 26. GEE Parameter Estimates from a Cumulative Logit Model for Ordinal Outcomes with Interaction Terms.

Parameter	Estimate	SE	OR Estimate	95% CI (Lower)	95% CI (Upper)	Z-value	P-value
Intercept, Threshold 1	-8.304	0.42	—	-9.125	-7.482	-19.81	<.0001
Intercept, Threshold 2	-6.593	0.42	—	-7.411	-5.775	-15.80	<.0001
Intercept, Threshold 3	-4.763	0.41	—	-5.571	-3.956	-11.56	<.0001
Visit = 2	-0.844	0.42	0.430	-1.667	-0.021	-2.01	0.0444
Visit = 3	3.803	0.51	44.835	2.810	4.797	7.50	<.0001
Female = 1	-0.665	0.09	0.514	-0.850	-0.480	-7.04	<.0001
Education = 2	-0.0003	0.11	0.999	-0.223	0.222	-0.0003	0.9976
Education = 3	-0.206	0.10	0.814	-0.409	-0.004	-2.00	0.0456
PAG2008 = 1	-0.107	0.09	0.899	-0.284	0.071	-1.18	0.2383
Depression = 1	0.220	0.10	1.246	0.023	0.417	2.19	0.0286
Alcohol Use = 1	-0.199	0.09	0.820	-0.377	-0.021	-2.20	0.0281
Smoking = 1	-0.085	0.12	0.919	-0.318	0.148	-0.71	0.4753
HEI2010	-0.0002	0.003	0.999	-0.007	0.006	-0.06	0.9544
Age	0.075	0.004	1.078	0.066	0.083	17.57	<.0001
BMI	0.079	0.008	1.082	0.063	0.095	9.84	<.0001
Visit × Female (Visit = 2, Female = 1)	0.108	0.09	1.114	-0.066	0.284	1.22	0.2233
Visit × Female (Visit = 3, Female = 1)	0.508	0.11	1.662	0.291	0.725	4.59	<.0001
Visit × Education (Visit = 2, Education = 2)	0.167	0.10	1.182	-0.034	0.368	1.63	0.1032
Visit × Education (Visit = 2, Education = 3)	0.108	0.09	1.114	-0.074	0.291	1.16	0.2457
Visit × Education (Visit = 3, Education = 2)	-0.086	0.13	0.918	-0.349	0.178	-0.64	0.5238
Visit × Education (Visit = 3, Education = 3)	0.042	0.12	1.043	-0.194	0.278	0.35	0.7261
Visit × PAG2008 (Visit = 2, PAG2008 = 1)	0.047	0.08	1.048	-0.113	0.208	0.58	0.5634
Visit × PAG2008 (Visit = 3, PAG2008 = 1)	-0.015	0.11	0.985	-0.225	0.195	-0.14	0.8887
Visit × Depression (Visit = 2, Depression = 1)	-0.013	0.09	0.987	-0.184	0.158	-0.15	0.8795
Visit × Depression (Visit = 3, Depression = 1)	0.124	0.11	1.132	-0.098	0.347	1.10	0.2726
Visit × Alcohol Use (Visit = 2, Alcohol Use = 1)	-0.059	0.11	0.943	-0.275	0.157	-0.53	0.5938
Visit × Alcohol Use (Visit = 3, Alcohol Use = 1)	0.130	0.12	1.139	-0.110	0.370	1.06	0.2869
Visit × Smoking (Visit = 2, Smoking = 1)	0.272	0.14	1.313	-0.004	0.549	1.93	0.0533
Visit × Smoking (Visit = 3, Smoking = 1)	0.295	0.17	1.343	-0.030	0.620	1.78	0.0755
Visit × HEI2010 (Visit = 2)	-0.003	0.003	0.997	-0.009	0.003	-0.84	0.4036
Visit × HEI2010 (Visit = 3)	0.002	0.004	1.002	-0.006	0.009	0.40	0.6928
Visit × Age (Visit = 2)	0.010	0.004	1.010	0.003	0.018	2.66	0.0079
Visit × Age (Visit = 3)	-0.050	0.005	0.951	-0.060	-0.040	-9.98	<.0001
Visit × BMI (Visit = 2)	0.011	0.008	1.011	-0.005	0.027	1.39	0.1637
Visit × BMI (Visit = 3)	-0.048	0.010	0.953	-0.067	-0.029	-5.05	<.0001

Note. *Model estimated using generalized estimating equations (GEE) with a multinomial distribution and cumulative logit link. Estimates include threshold-specific intercepts and covariate effects, with corresponding standard errors (SE), odds ratios (OR), 95% confidence intervals, Z-values, and p-values. An independent working correlation structure was used to account for within-subject correlation across repeated visits.*

Table 27. Type 3 Score Statistics from the Generalized Estimating Equations Model Without Interaction Terms (Multinomial Distribution, Cumulative Logit Link)

Source	DF	Chi-squared Statistic	P-value
Visit	2	16.86	0.0002
Sex	1	41.67	<.0001
Education	2	6.13	0.0466
HEI2010	1	0.09	0.7670
PAG2008	1	2.11	0.1466
Depression	1	10.99	0.0009
Age	1	354.65	<.0001
Alcohol Use	1	12.35	0.0004
BMI	1	112.17	<.0001
Smoking	1	0.58	0.4463

Table 28. Type 3 Score Statistics from the Generalized Estimating Equations Model with Interaction Terms (Multinomial Distribution, Cumulative Logit Link)

Source	DF	Chi-squared Statistic	P-value
Visit	2	110.92	<.0001
Sex	1	43.55	<.0001
Education	2	6.21	0.0448
HEI2010	1	0.05	0.8312
PAG2008	1	1.96	0.1615
Depression	1	11.00	0.0009
Age	1	354.58	<.0001
Alcohol Use	1	10.37	0.0013
BMI	1	113.36	<.0001
Smoking	1	1.60	0.2061
Visit × Sex	2	23.02	<.0001
Visit × Education	4	5.04	0.2830
Visit × HEI2010	2	1.28	0.5270
Visit × PAG2008	2	0.48	0.7858
Visit × Depression	2	1.70	0.4272
Visit × Age	2	143.18	<.0001
Visit × Alcohol Use	2	2.48	0.2897
Visit × BMI	2	42.87	<.0001
Visit × Smoking	2	5.11	0.0775

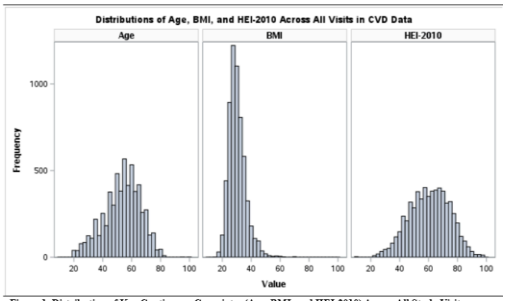


Figure 1. Distribution of Key Continuous Covariates (Age, BMI, and HEI-2010) Across All Study Visits

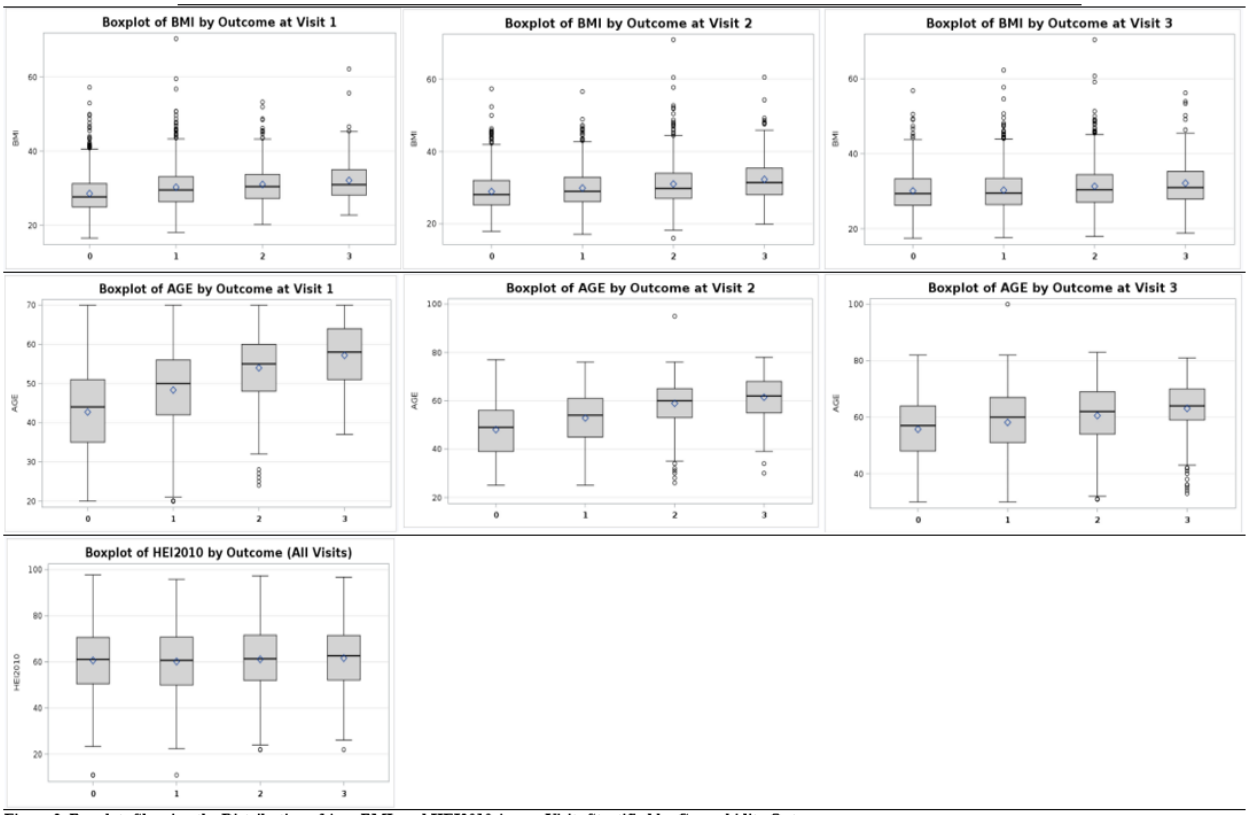


Figure 2. Boxplots Showing the Distribution of Age, BMI, and HEI2010 Across Visits Stratified by Comorbidity Outcome

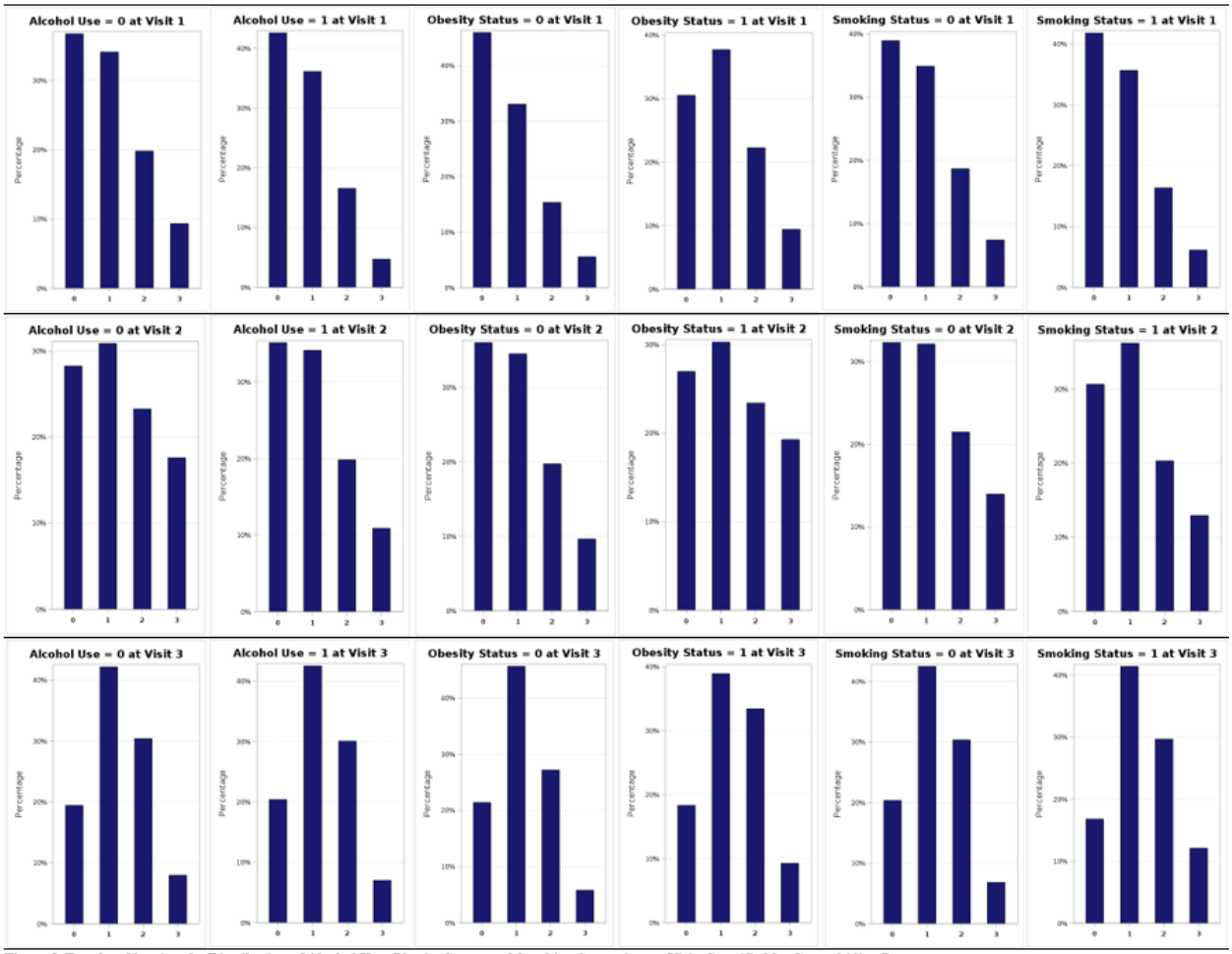


Figure 3. Barplots Showing the Distribution of Alcohol Use, Obesity Status and Smoking Status Across Visits Stratified by Comorbidity Outcome

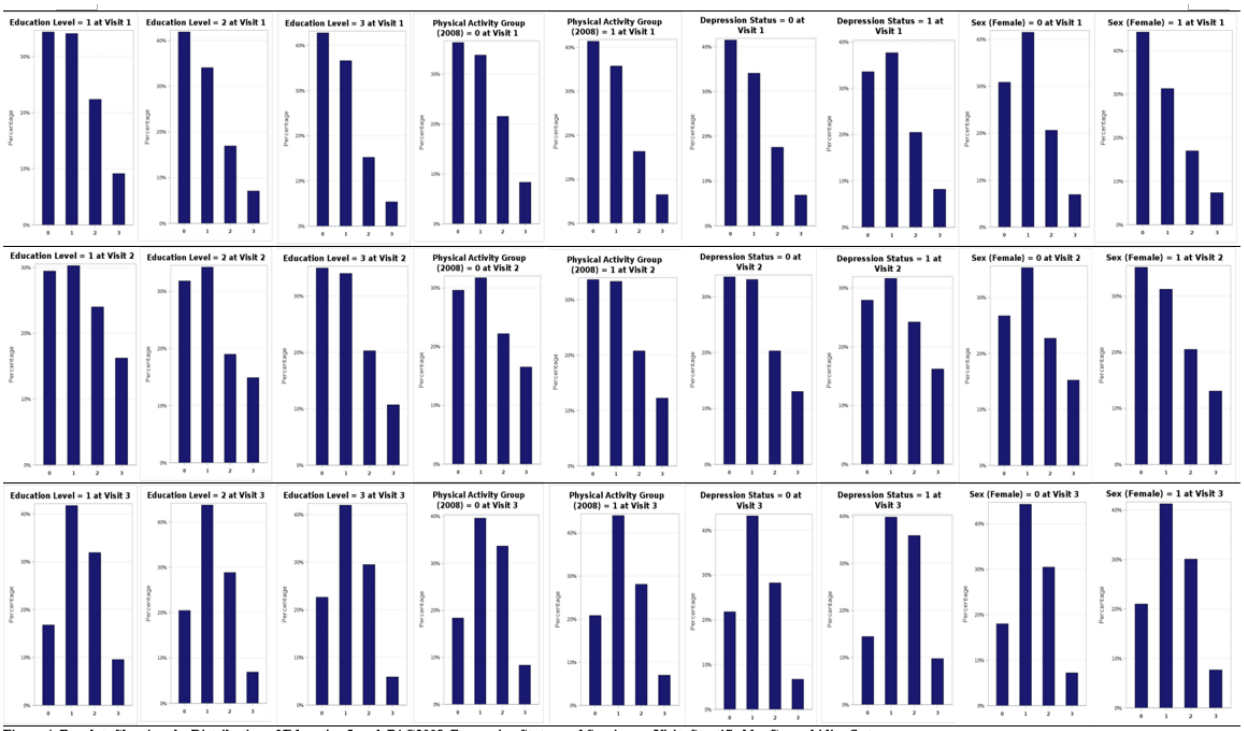


Figure 4, Barplots Showing the Distribution of Education Level, PAG2008, Depression Status and Sex Across Visits Stratified by Comorbidity Outcomes

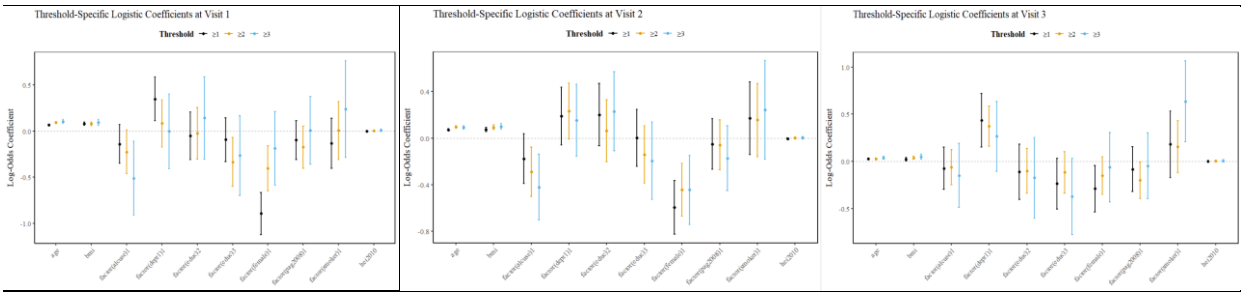
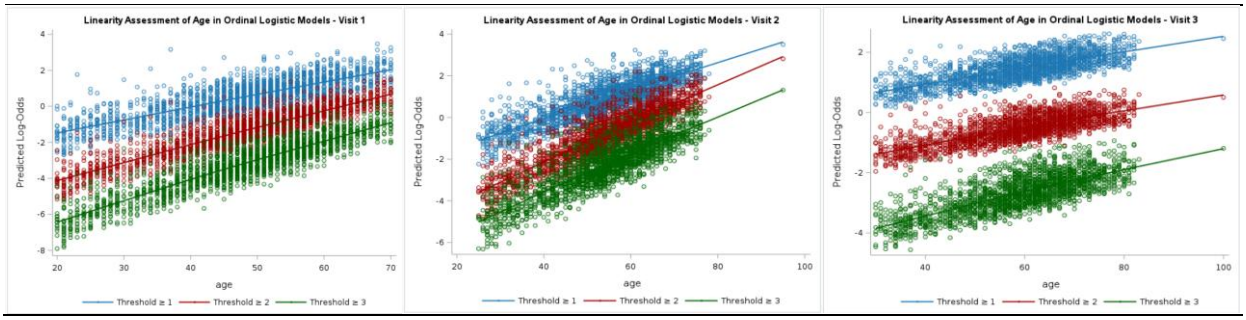


Figure 5. Log-odds coefficients and 95% confidence intervals for predictors of the ordinal outcome across three cumulative thresholds across three visits. Each panel (or color-coded group) represents a separate binary logistic regression comparing different levels of the outcome (≥1, ≥2, =3). Deviations across thresholds suggest possible violations of the proportional odds assumption for certain predictors.



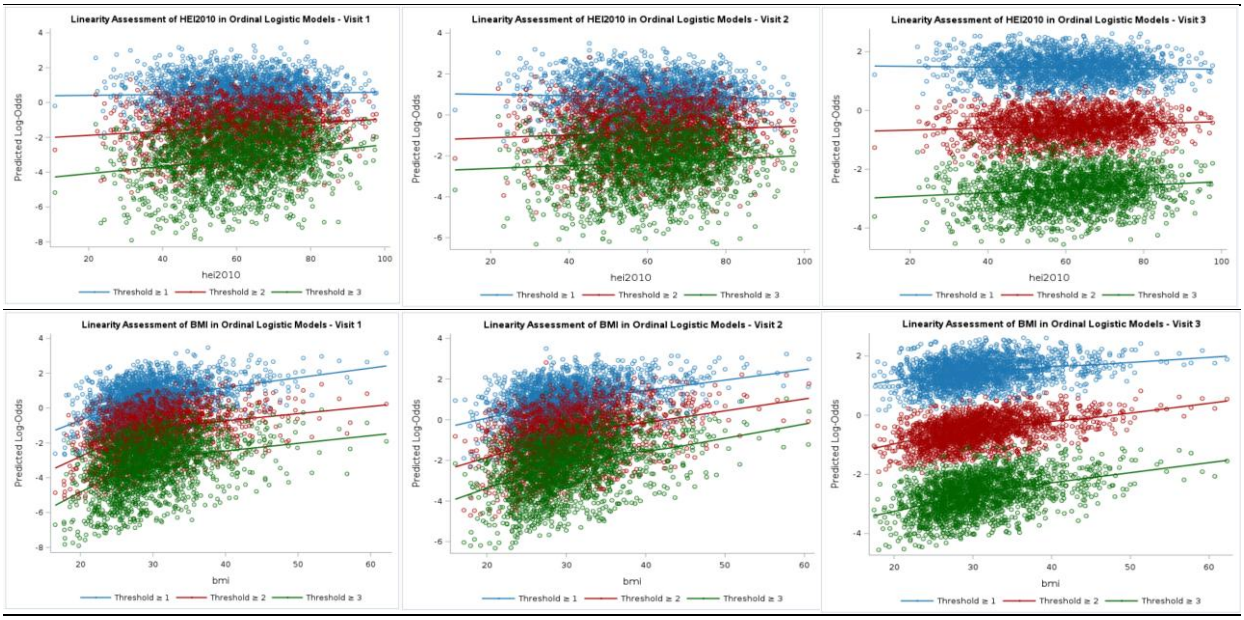
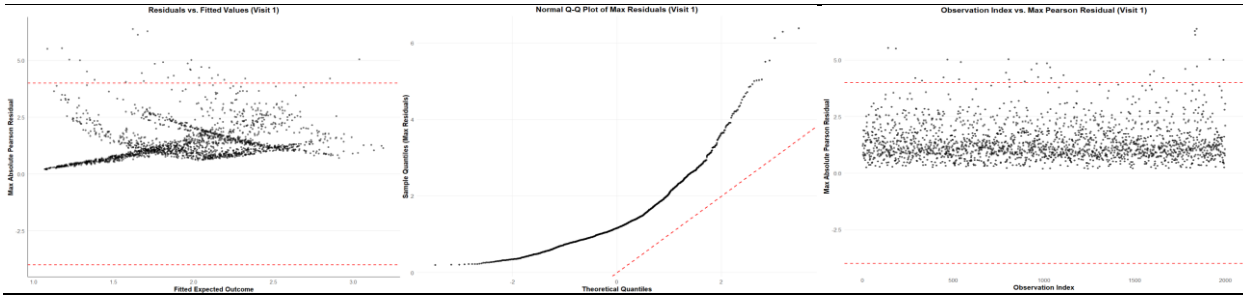


Figure 6. Linearity Assessment of Age, HEI-2010, and BMI in Ordinal Logistic Models Across Visits. Predicted log-odds from cumulative logit models are plotted against each continuous covariate to assess linearity on the logit scale. Separate binary logistic regressions were fit using threshold indicators (i.e., outcome ≥ 1, ≥ 2, ≥ 3), with LOESS curves overlaid to visualize departures from linearity. Each line corresponds to a different threshold, and scatter points represent individual-level predictions. Approximate linearity of the smoothed trends supports the logit-linear assumption for each predictor.



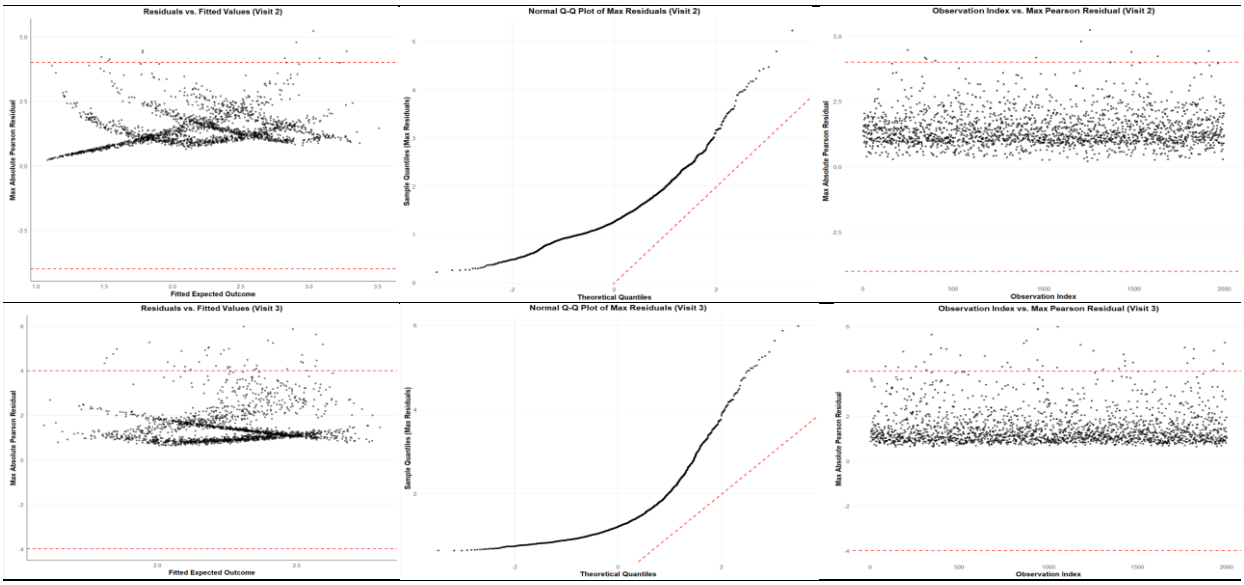


Figure 7. Diagnostic Plots of Pearson Residuals Across Study Visits.

Each panel presents diagnostic assessments from the partial proportional odds model for each visit, including (1) Pearson residuals versus fitted expected outcomes, (2) normal Q-Q plots evaluating residual distribution, and (3) Pearson residuals by observation index. Horizontal dashed lines indicate the threshold of |Pearson residual| > 4, used to flag unusually large residuals for detecting outliers and assessing model fit.

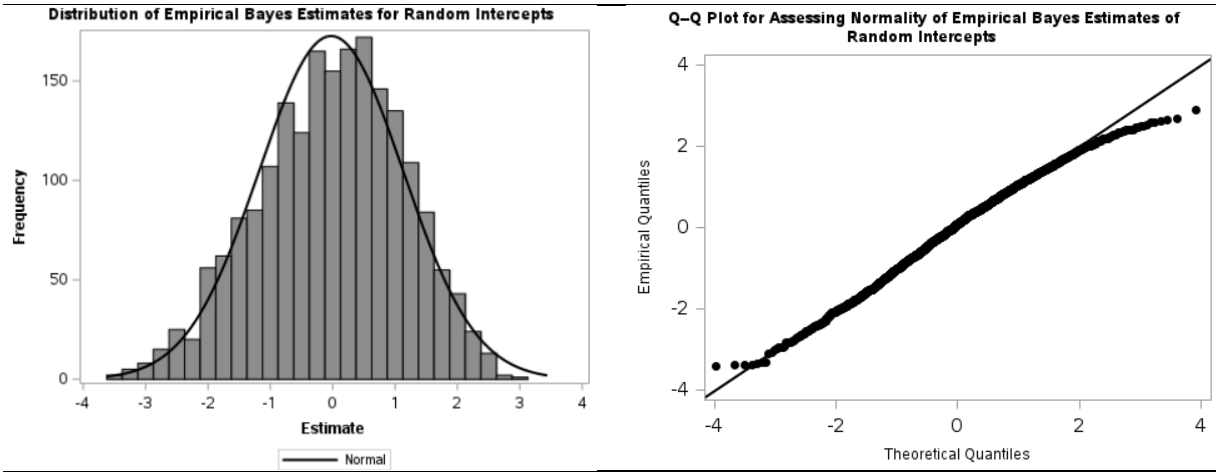


Figure 8. Assessment of the Random Intercept Normality Assumption Using Histogram and Q–Q Plot.

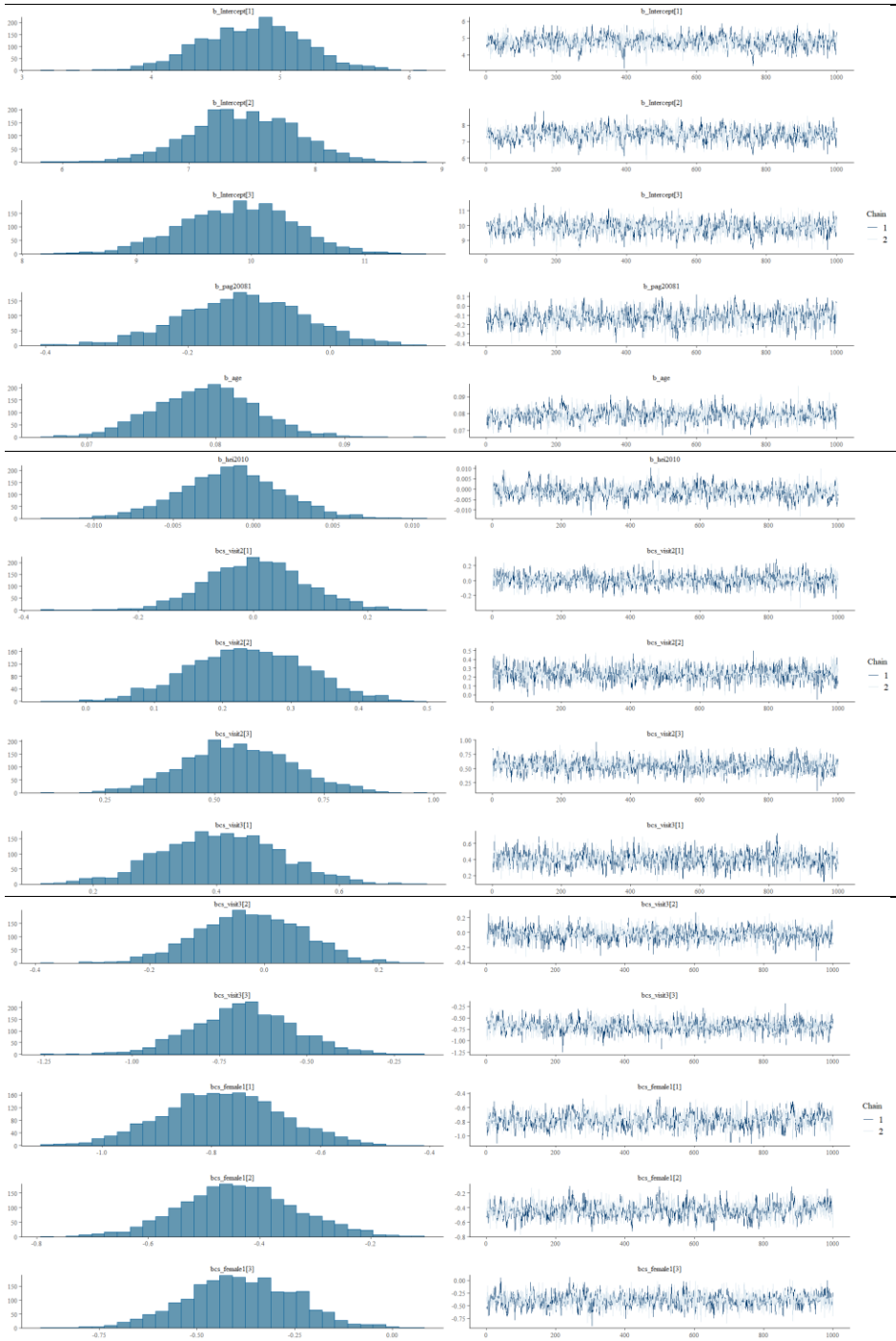


Figure 9. Posterior densities (left) and MCMC trace plots (right) for selected parameters from the Bayesian cumulative logistic mixed-effects model. Rows 1–3 show threshold-specific intercepts defining ordinal cut-points. Rows 4–6 display proportional effects for PAG2008, age, and HEI2010, each modeled with a constant slope across thresholds. Rows 7–15 depict non-proportional effects for visit (levels 2 and 3) and female (level 1), with threshold-specific coefficients estimated via the cs() function.

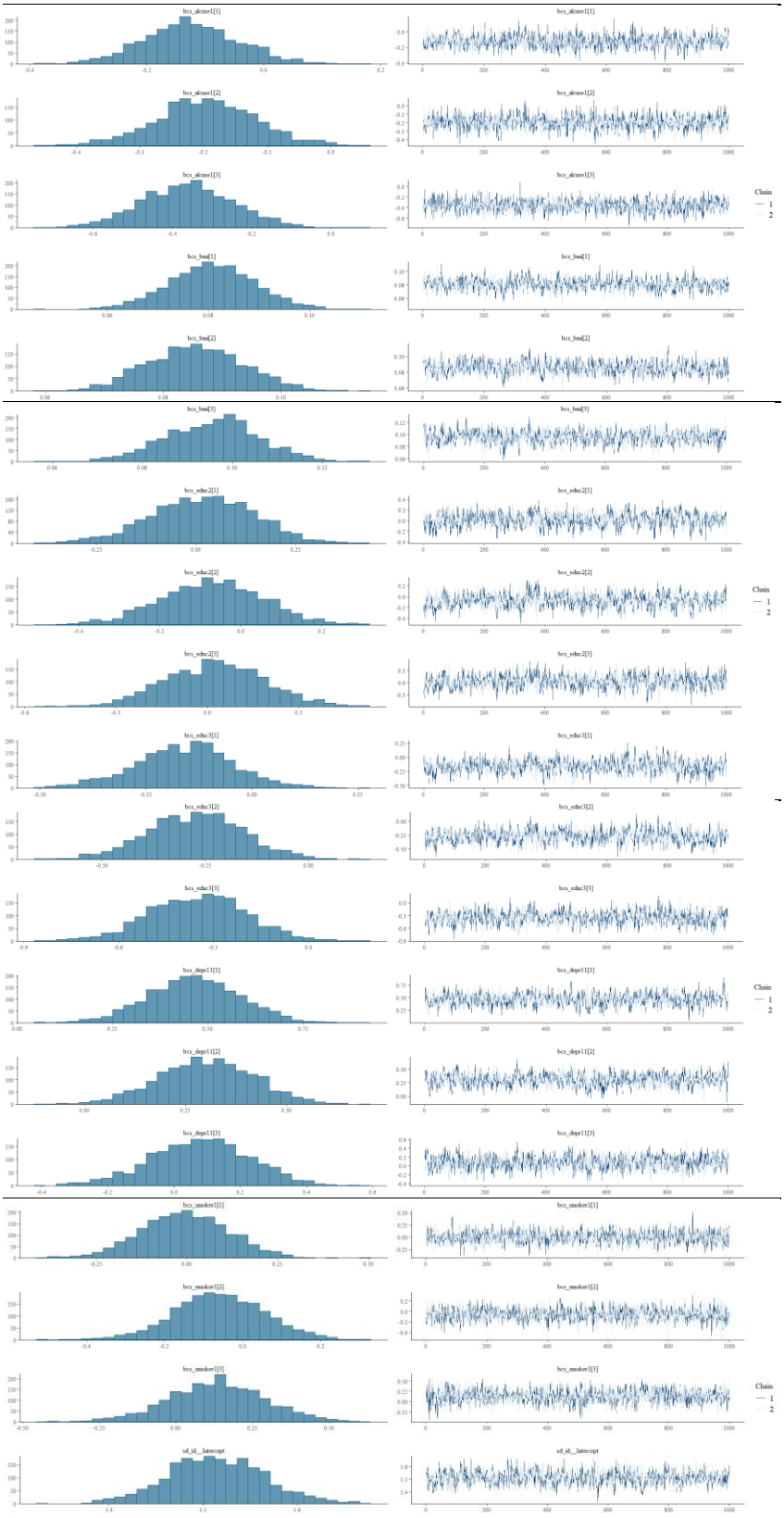


Figure 10. Posterior densities (left) and MCMC trace plots (right) for additional parameters from the Bayesian cumulative logistic mixed-effects model. Covariates shown were modeled with non-proportional effects using the cs() function, allowing threshold-specific estimates. These include alcohol use of level 1, BMI of continuous, education levels 2 and 3, depression of level 1, and smoking status of level 1, each with three threshold-specific coefficients. The final row displays the posterior for the subject-level random intercept standard deviation.

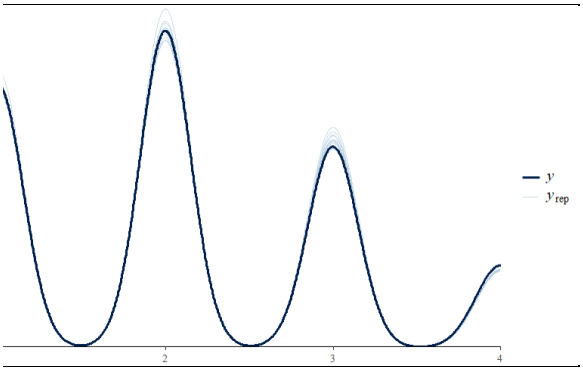


Figure 11. Figure X. Posterior Predictive Density Overlay Plot Comparing Observed and Simulated Outcomes.
The plot compares the empirical distribution of the observed outcome (y) to the distributions of posterior predictive replicates (y_{rep}) simulated from the fitted cumulative logit mixed-effects model. Close alignment between y and y_{rep} densities indicate adequate model fit in terms of marginal outcome distribution.