Contents

S1 Visualization of the Joint Model

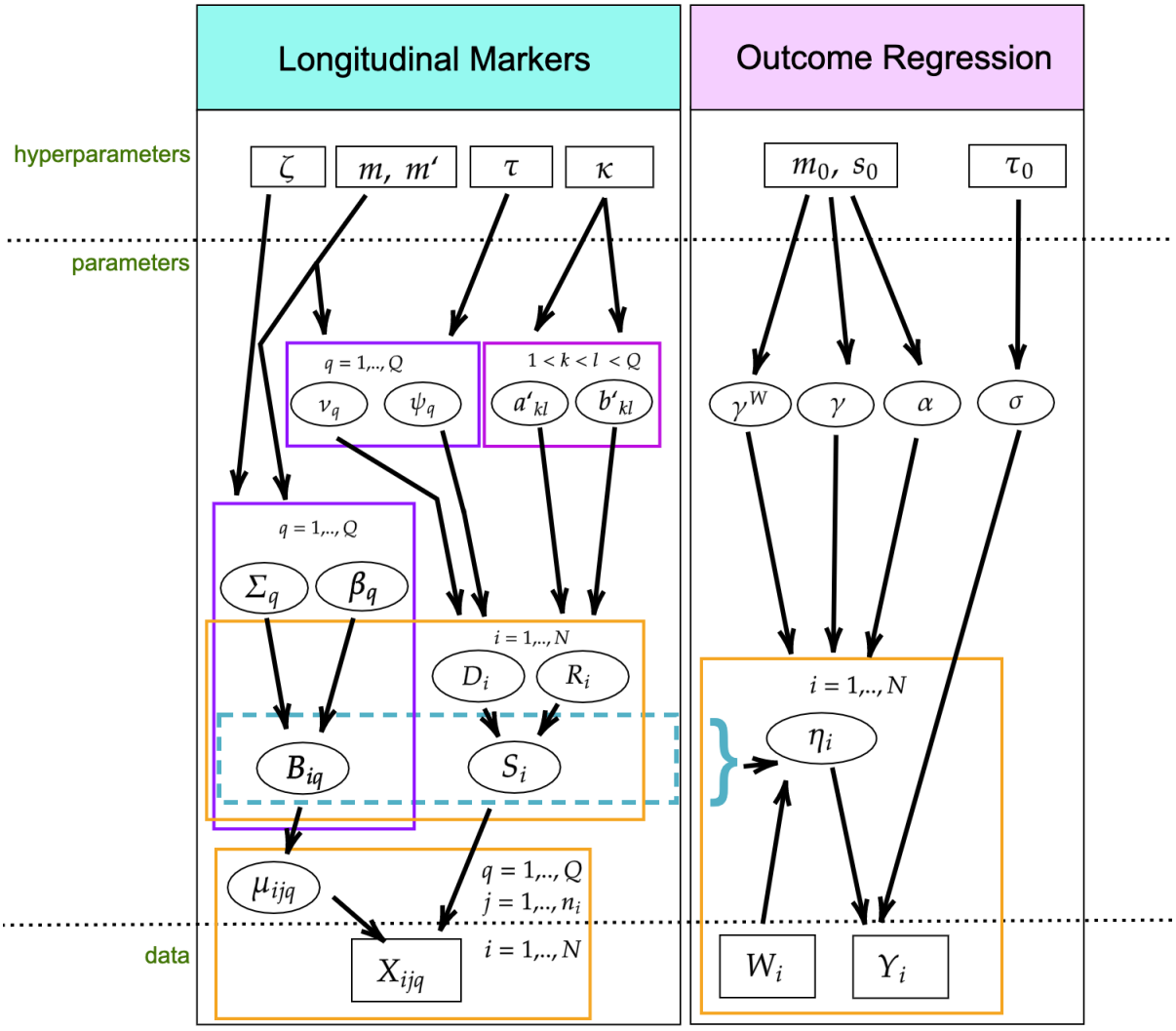

Figure S1: A visualization of the relationships between the model parameters and data. This directed graph shows the hierarchical form of our model framework. The quantities in squares are either data or hyperparameters; the unknown quantities are displayed in circles. The arrows connecting variables indicate that the parent parameterizes the distribution of the child node. The rectangular "plates" that enclose variables indicate that a similar graphical structure is repeated over the index. The index in a plate indicates nodes, hyperparameter levels and subjects.

S2 Posterior Predictive Model Checking

To assess our model's validity on the SWAN data, we conduct posterior predictive checks for both the trajectories submodel and the outcome submodel.

For the outcome submodel, we generated simulated data from the posterior predictive distribution. The posterior predictive distribution for the predicted outcome, $\widetilde{Y}$ can be written as:

$$p(\widetilde{Y}|Y) = \int p(\widetilde{Y}|\theta, X)p(\theta, X|Y)d\theta dX$$

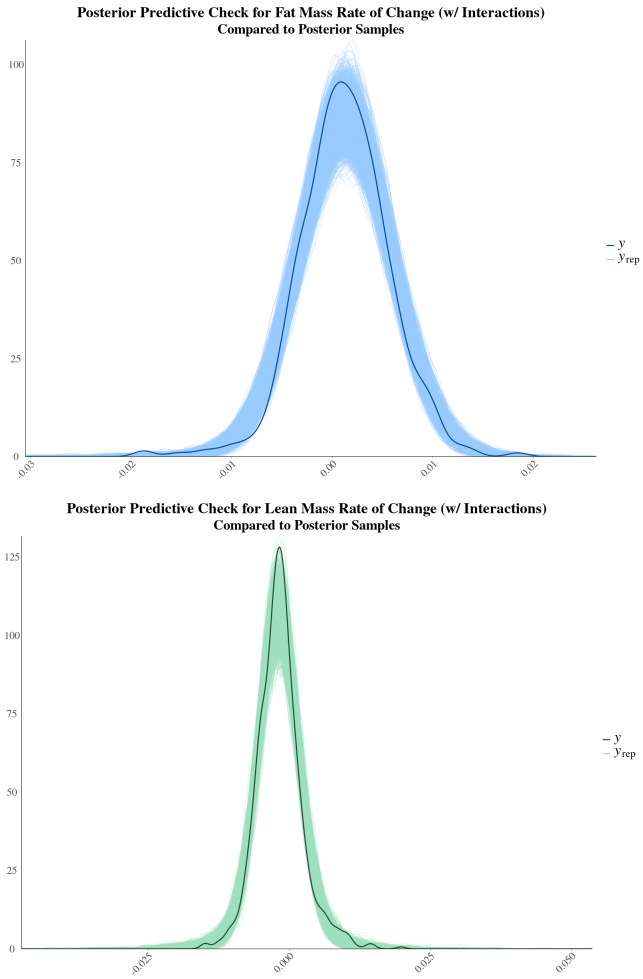

Figure S2: Visualizations of the posterior predictive checks performed for the fat mass rate of change with variance interactions (top) and lean mass rate of change with variance interactions (bottom). The observed outcomes (y) are represented by the solid lines and the model-generated outcomes (yrep) are represented by the thin semi-opaque lines. We see that the model-generated outcomes cover the observed outcomes for both models, indicating that our model is generating reasonable estimates of the outcomes.

where θ are the unknown model parameters and X are the predictor variables used in the outcome regression. For each draw of the model parameters from the posterior distribution, $p(\theta|Y, X)$, we can draw a vector $\widetilde{Y}$ from the posterior predictive distribution by conditioning on the draw of the model parameters and then simulating from the data model (Gabry et al., 2019).

We then plotted 1,000 draws of this model-generated data against the observed outcome, which is shown in Figure S2. For both the fat mass rate of change and the lean mass rate of change models, we see that the simulated replicated data from the model overlap the observed data, indicating that our model is producing reasonable predictions.

For the trajectories submodel, we define the following statistic: $T(x_{itq}; b_{iqp}, t)_q = \sum_t (x_{itq} - \mu(b_{iqp}, t))^2/(\sigma^2_{iq}))$ where $\mu(b_{iqp}, t))$ is the estimated individual i's mean trajectory for hormone q and σ^2_{iq} is the estimated variance of individual i's trajectory for hormone q. By doing this, we can compare $T_i(x^{obs}_{itq}; b_{iqp}, t)_q$ (which is a function

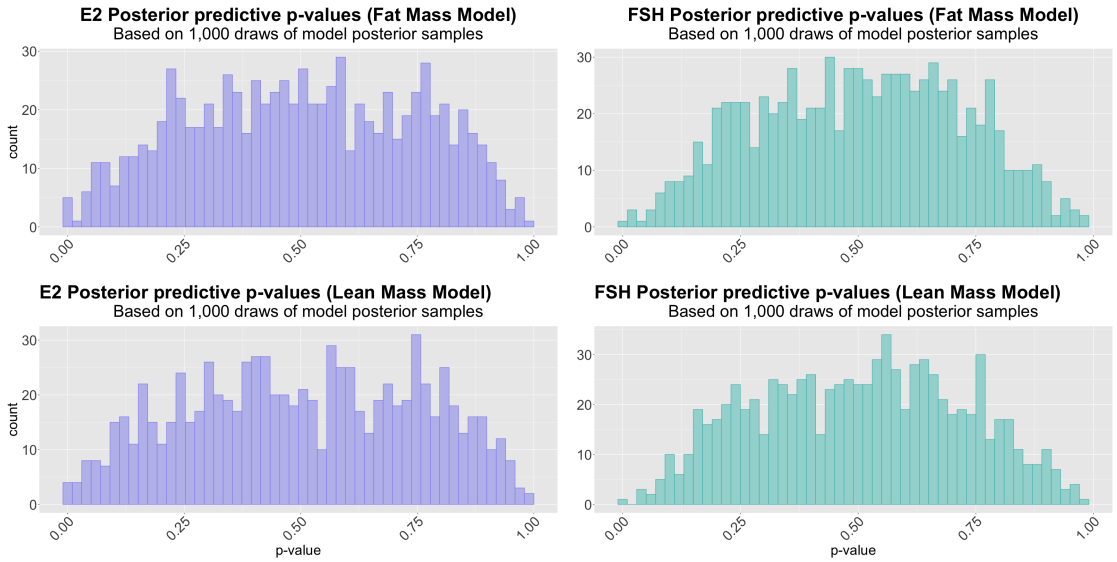

Figure S3: Posterior predictive check of E2, FSH trajectories across all individuals for both the fat mass and lean mass models. The median p-value for each 1,000 draws of posterior samples was 0.5.

of the observed data and the estimated parameters) with $T_i(x_{itq}^{sim}; b_{ipq}, t)_q$ (which is a function of the model generated data using the model estimated parameters). If there are large discrepancies between $T_i(x_{itq}^{obs}; b_{iqp}, t, \sigma_{iq}^2)_q$ and $T_i(x_{itq}^{sim}; b_{iqp}, t, \sigma_{iq}^2)_q$, this could indicate poor model fit (Gelman et al., 2013).

One way to compare these two $T(x_{itq})$ statistics is to compute the 'posterior predictive p-value', which is $P(T_i(x_{itq}^{obs}; b_{iqp}, t, \sigma_{iq}^2)_q < T_i(x_{itq}^{sim}; b_{iqp}, t, \sigma_{iq}^2)_q | (x_{itq}^{obs}))$. For E2 and FSH, we keep (x_{it}^{obs}) fixed at the observed values and compute 1,000 values of $T_i(x_{itq}^{sim}; b_{iqp}, t)$ from the posterior of $b_{iqp}, t, \sigma_{iq}^2$. We then compare these values with 1,000 draws from $T_i(x_{itq}^{sim}; b_{iqp}, t, \sigma_{iq}^2)_q$. Figure S3 displays the histograms of the resulting p-values for each individual's hormone trajectory for the two models. Across all of the hormones, most of the the computed p-values were between 0.25 and 0.75. Further analysis of the p-values across the quantiles of the distribution shows that the generated data from the model reasonably captures the individual trends. This provides justification that both the trajectories submodel and the outcome submodel are good fits for the data.

Parameter	Post. Mean	2.5% CrI	97.5% CrI
β_{11}	-37.37	-59.01	-15.79
β_{12}	3.78	0.78	6.84
β_{21}	26.89	4.06	48.02
β_{22}	-6.48	-9.00	-4.08
$\Sigma_1[1,1]$	48.07	39.36	57.96
$\Sigma_1[1,2]$	-1.50	-2.43	0.55
$\Sigma_1[2,2]$	0.54	0.36	0.72
$\Sigma_2[1,1]$	80.40	70.93	91.43
$\Sigma_2[1,2]$	0.13	-0.63	0.88
$\Sigma_2[2,2]$	0.56	0.45	0.67
ν_1	-293.86	-312.52	-274.63
ν_2	-701.88	-725.78	-677.63
ξ_1	151.41	131.17	172.56
ξ_2	247.34	225.05	269.14
α_1*	9.00	6.72	12.02
β_1*	33.87	24.83	45.69
σ^2	3.61	3.31	3.90

Table S1: Evaluation of the posterior means and 95% CrI estimates for the other parameters in the fat mass rate of change model. All values except for α_1, β_1 (indicated with asterisk) have been multiplied by 10^3. α_1, β_1 have been presented in their original values.

Parameter	Post. Mean	2.5% CrI	97.5% CrI
β_{11}	-37.42	-58.27	-16.74
β_{12}	3.67	0.47	6.70
β_{21}	27.09	4.91	49.10
β_{22}	-6.47	-8.92	-3.98
$\Sigma_1[1,1]$	47.92	39.14	57.88
$\Sigma_1[1,2]$	-1.54	-2.44	-0.66
$\Sigma_1[2,2]$	0.52	0.33	0.71
$\Sigma_2[1,1]$	80.55	70.92	91.29
$\Sigma_2[1,2]$	0.08	-0.69	0.84
$\Sigma_2[2,2]$	0.56	0.44	0.68
ν_1	-293.17	-312.10	-273.82
ν_2	-702.11	-725.35	-679.66
ξ_1	141.88	119.81	161.74
ξ_2	241.88	220.67	262.82
α_1*	9.15	6.79	12.06
β_1*	34.41	24.90	46.00
σ_1^2	2.74	2.28	3.10
σ_2^2	6.71	4.77	10.70
Π_1*	0.86	0.66	0.97

Table S2: Evaluation of of the posterior means and 95% CrI estimates for the other model parameters in the lean mass rate of change model. All values except for α_1, β_1, Π_1 (indicated with asterisk) have been multiplied by 10^3. α_1, β_1, Π_1 have been presented in their original values.

S3 Data Analysis: Posterior Means and 95% CrIs for Additional Model Parameters

In this section, we present the estimated posterior means and 95% credible intervals for the other parameters from the data application (Section 5 in the main text). Table S1 contains the estimates for the fat mass rate of change model and Table S2 contains the estimates for the lean mass rate of change model.

Parameter	Truth	Average Post. Mean	Bias	Coverage %	Average Interval Length
β_{11}	0	0.00	0.00	95.5	0.14
β_{12}	2	2.00	0.00	95.0	0.13
β_{21}	2	2.00	0.01	95.5	0.15
β_{22}	1	1.00	0.00	95.5	0.09
$\Sigma_1[1,1]$	1	1.00	0.00	95.0	0.22
$\Sigma_1[1,2]$	-0.05	-0.05	0.00	96.0	0.14
$\Sigma_1[2,2]$	1.01	1.00	0.01	94.0	0.19
$\Sigma_2[1,1]$	1	1.00	0.00	96.0	0.24
$\Sigma_2[1,2]$	-0.1	-0.10	0.00	97.5	0.10
$\Sigma_2[2,2]$	0.5	0.50	0.00	95.0	0.10
ν_1	0	0.00	0.00	95.5	0.06
ν_2	0.25	0.25	0.00	97.0	0.04
ξ_1	0.38	0.37	0.00	95.0	0.05
ξ_2	0.25	0.25	0.00	97.0	0.04
a_1	1	1.01	0.01	96.0	0.01
b_1	5	5.06	0.06	96.0	1.32

Table S3: Two Trajectory Simulation Setting: Evaluation of bias, coverage, and 95% credible interval length across 200 simulation replicates for the $\mathbf{B}_i$ and $\mathbf{S}_i$ parameters for the JMIV model. Our model achieves $> 90\%$ coverage across all parameters and maintains low bias.

S4 Two Biomarker Simulation Study and Three Biomarker Simulation Study

In this section, we present the bias, coverage and average interval length for the other JMIV model parameters from running 200 simulation replicates. For each simulation replicate, we ran two chains with 2,000 steps and 1,000 burn in. The data generation parameters are detailed in Section 4.1 and 4.2 of the main text. Table S3 contains the estimates for the two biomarker simulation study and Table S4 contains the estimates for the three biomarker simulation study.

S5 Simulation 3: Linear Approximation of Nonlinearity

In this simulation study, we study how well our model performs when the true relationship between some of the longitudinal means and variances terms and the cross-sectional outcome is nonlinear, but we approximate this relationship with a linear form.

Step 1: Estimate Linear Approximation Coefficients

We use the same data generation parameters for the longitudinal markers as in Section 4.1 of the main text. For the outcome model, we generate the mean $\eta(\boldsymbol{B}_i, \boldsymbol{S}_i)$ as:

$$\eta(\mathbf{B}_i, \mathbf{S}_i) = 2b_{i11} + b_{i12} - b_{i21} + 0.5b_{i22} + 2s_{i11} - s_{i21} + 2s_{i22} + 0.5b_{i21}^2 + 0.75s_{i11}^2$$

so that the individual slope of the first biomarker (b_{i12}) and the variance of the first biomarker (s_{i11}) are quadratically related to the outcome.

Parameter	Truth	Average Post. Mean	Bias	Coverage %	Average Interval Length
β_{11}	0	0.00	0.00	95.0	0.14
β_{12}	2	1.99	0.00	93.5	0.35
β_{21}	2	1.99	0.00	96.0	0.14
β_{22}	1	0.99	0.00	95.5	0.09
β_{31}	1	1.99	0.00	93.0	0.14
β_{32}	1	1.00	0.00	95.5	0.13
$\Sigma_1[1,1]$	1	1.00	0.00	94.0	0.21
$\Sigma_1[1,2]$	-0.05	-0.05	0.00	97.0	0.14
$\Sigma_1[2,2]$	1.00	0.99	0.01	96.0	0.18
$\Sigma_2[1,1]$	1	1.00	0.00	96.5	0.22
$\Sigma_2[1,2]$	-0.1	-0.10	0.00	93.5	0.10
$\Sigma_2[2,2]$	0.5	0.50	0.00	97.5	0.09
$\Sigma_3[1,1]$	1	1.00	0.00	96.0	0.21
$\Sigma_3[1,2]$	-0.25	-0.25	0.00	94.5	0.14
$\Sigma_3[2,2]$	1	1.00	0.00	92.0	0.18
ν_1	0.00	0.00	0.00	94.0	0.06
ν_2	0.25	0.25	0.00	95.5	0.04
ν_3	0.25	0.25	0.00	94.5	0.07
ξ_1	0.375	0.38	0.00	97.0	0.06
ξ_2	0.25	0.25	0.00	94.5	0.04
ξ_3	0.25	0.25	0.00	94.5	0.07
a_{12}	1	1.01	0.01	96.0	0.19
a_{13}	1	1.01	0.01	94.5	0.19
a_{23}	2	2.04	0.04	93.5	0.50
b_{12}	5	5.09	0.09	96.5	1.29
b_{13}	5	5.07	0.07	92.5	1.29
b_{23}	2	2.05	0.05	94.0	0.50

Table S4: Three Trajectory Simulation Setting: Evaluation of bias, coverage, and 95% credible interval length across 200 simulation replicates for the $\mathbf{B}_i$ and $\mathbf{S}_i$ parameters for the JMIV model. Our model achieves $> 90\%$ coverage across all parameters and maintains low bias.

To estimate the "linear approximation" coefficients, we simulate data for 1 million individuals and generate the outcome data as:

$$Y_i \sim \mathcal{N}(\eta(\mathbf{B}_i, \mathbf{S}_i), 0.01)$$

We then fit a linear approximation using the lm() function in R and the following model:

$$Y_i = \alpha_{11} * b_{i11} + \alpha_{12} * b_{i12} + \alpha_{21} * b_{i21} + \alpha_{22} * b_{i22} + \gamma_{11} * s_{i11} + \gamma_{21} * s_{i21} + \gamma_{22} * s_{i22}$$

and collect the estimated coefficients. These coefficients are the "target" coefficients that we want to approximate. These targets are shown in Table S5 as the "truth" values.

Step 2: Joint Model Simulation Replicates

We follow the same data generation in Step 1 for the longitudinal markers. For the outcome model, we generate the outcome data as:

$$Y_i \sim \mathcal{N}(\eta(\mathbf{B}_i, \mathbf{S}_i), 0.5)$$

After generating this data, we apply our joint model, but modeled with the same linear mean function, $\eta(\mathbf{B}_i, \mathbf{S}_i)$, as in Section 4 of main text and collect the estimated coefficients. We do this for 200 independent replicates and evaluate model performance using the same criteria (bias, coverage, and average 95% CrI length) as in the previous simulation studies. Table S5 displays the results for the outcome mean and variance parameters. We find that the model maintains low bias and high coverage of the truth (¿ 90% coverage). This indicates that our model can recover the estimated parameters from a linear approximation of the model, when the true form of the outcome mean may be nonlinear.

Truth	Bias	Coverage (%)	Average Interval Length
α_{11} = 1.99	0.02	93.50	0.41
α_{12} = 1.44	0.03	93.00	0.34
α_{21} = -1.06	0.00	94.58	0.38
α_{22} = 0.56	-0.03	94.50	0.51
γ_{11} = 1.97	-0.01	96.00	0.69
γ_{12} = -0.99	0.03	94.50	1.17
γ_{22} = -2.06	-0.03	92.50	0.58

Table S5: Simulation III: bias, coverage, and 95% credible interval (or confidence interval) length across 200 simulation replicates. With the linear approximation, our model maintains low bias and high coverage of the true (linear approximating) parameters.

References

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., and Gelman, A. (2019). Visualization in bayesian workflow. *Journal of Royal Statistical Society: Series A*, 182:389–402.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. CRC Press.