

ARTICLE TYPE

Functional Mixed-Effect Modeling of Multivariate Jaw Kinematics

Palmer Swanson¹ | Author Two^{2,3} | Author Three³¹Department of Statistics, Florida State University, Florida, United States of America²Department Name, Institution Name, State Name, Country Name³Department Name, Institution Name, State Name, Country Name**Correspondence**

Corresponding author Palmer Swanson, Department of Statistics, Florida State University 214 Rogers Building (OSB), 117 N. Woodward Ave. P.O. Box 3064330, Tallahassee, FL 32306-4330

Email: pjswanson@fsu.edu

Present address

This is sample for present address text this is sample for present address text.

Abstract

Temporomandibular joint (TMJ) disorders affect up to one-third of adults and are associated with pain, joint clicking, and compromised jaw function, significantly impacting quality of life. Understanding how individual-level characteristics influence jaw motion is therefore of clinical interest. In this work, we model three-dimensional jaw kinematic data using a Bayesian function-on-scalar regression framework. The data consist of tri-variate functional outcomes for 40 participants across multiple tasks, with motion tracked at key anatomical points on the jaw.

We extend prior work by Goldsmith to develop a hierarchical spline-based model for tri-variate functional data that accounts for subject-specific random effects and fixed covariates. To address computational challenges associated with high-dimensional covariance estimation in multivariate models, we propose a modified version of the model that avoids large inverse-Wishart priors at the level-1 error, instead shifting covariance structure into the random effects. We compare performance using a Gibbs sampler and a Hamiltonian Monte Carlo (HMC) implementation in STAN.

Simulation studies evaluate both methods across varying numbers of covariates, knots, and noise levels. Results suggest that while the Gibbs sampler generally achieves higher posterior coverage and lower IMSE, the STAN implementation achieves substantially faster computation times. We apply the model to real jaw tracking data, finding that demographic and clinical covariates significantly influence jaw motion.

KEY WORDS

Bayesian Regression, Trivariate Data, Functional Response, Kinematic Data, Gibbs Sampler, Hamiltonian Monte Carlo

1 | INTRODUCTION

For humans, the face plays a crucial role in identification, communication, and proper functioning of certain senses. The jaw plays a jointly important role in these activities, as well as mastication and communication. Unfortunately, many people - up to a third of adults - suffer from dysfunction of the temporomandibular joint, also known as TMJ disorders¹. Symptoms of TMJ disorders commonly include pain, clicking of the jaw joint, dislocation of the jaw, headaches, and dizziness. Any one of these can greatly reduce the quality of life of an individual. For that reason, it is important to better understand the impact of various covariates on human jaw movement.

This work focuses on modeling the three-dimensional kinematic motion of various regions of the jaw, with particular attention to how this motion is influenced by scalar covariates. The motivating data set for this study involved three-dimensional kinematic data of subjects' jaws at various positions and for various oral tasks. Each subject was fitted with small sensors on their central incisor, left condylar apex, and right condylar apex. Subjects completed repeated open-close, lateral left, lateral right, and anterior/posterior shift tasks. Our data consisted of 776 three-dimensional function observations for 40 subjects (25 males and 25 females), across the four tasks.

Kinematic data is obtained for each repeated task. In other words, we have observed X, Y, and Z motion data for each subject, repetition, and task. In other words, we are provided tri-variate functional outcomes $P_{ijk}^X(t)$, $P_{ijk}^Y(t)$, $P_{ijk}^Z(t)$ for subject i , repetition j , and task k . Fig. 1 provides an example of the observed data. The data is facetted into each Cartesian coordinate, and trajectories are grouped together by subject.

The goal of this analysis is to understand the impact of multiple different covariates such as age, sex, and other demographic/clinical variables on three-dimensional kinematic motion, while incorporating subject specific random effects.

1.1 | Description of 3D Kinematic Data

Tempromandibular disorders (TMD) are common with patients experiencing limited tempromandibular joint (TMJ) function. This limited function often affects several critical components of daily living. Symptoms such as joint sounds and mild pain are often found in both patients with TMJ disorders and those with dentofacial deformities. Despite this, the effect of skeletal malocclusion classification (Class II and III) and sex on TMJ function is an area of minimal understanding. Additionally, orthognathic surgery can alleviate some or all of these symptoms. Mandibular kinematic data reflects TMJ functional motion and has vast potential for clinical use.

Patients enrolled in an IRB-approved prospective natural history study on morphological differences across dentofacial deformities at the NIH between 2018 and 2022. Forty patients with Class I, II, and III (asymptomatic=24, symptomatic=16) completed a biomechanical assessment of TMJ function. The average age was 32.73 years, with a standard deviation of 14.81 years. Twenty-five of the subjects were female, while fifteen were male. TMJ symptoms include popping or clicking sounds, discomfort, and mild pain. 3D mandibular kinematics were captured using a custom jaw tracking system during maximum open-close, lateral (left and right), and anterior oral tasks.

Each subject is fitted with several markers on several parts of the jaw: the lower central incisor as well as the left and right apex of the mandible. Data combined with CBCT images resulted in kinematic trajectories of the mandibular ventral incisor (CI), and movement of the bilateral condylar apices (CA). For each subject and task, numerous replications are run and three dimensional kinematic data in cartesian coordinates relative to the starting position are recorded; for each task, the X, Y, and Z coordinates of jaw movement are observed. In other words, we are provided tri-variate functional outcomes $P_{ijk}^X(t)$, $P_{ijk}^Y(t)$, $P_{ijk}^Z(t)$ for subject i , repetition j , and task k . An example set of data is provided below:

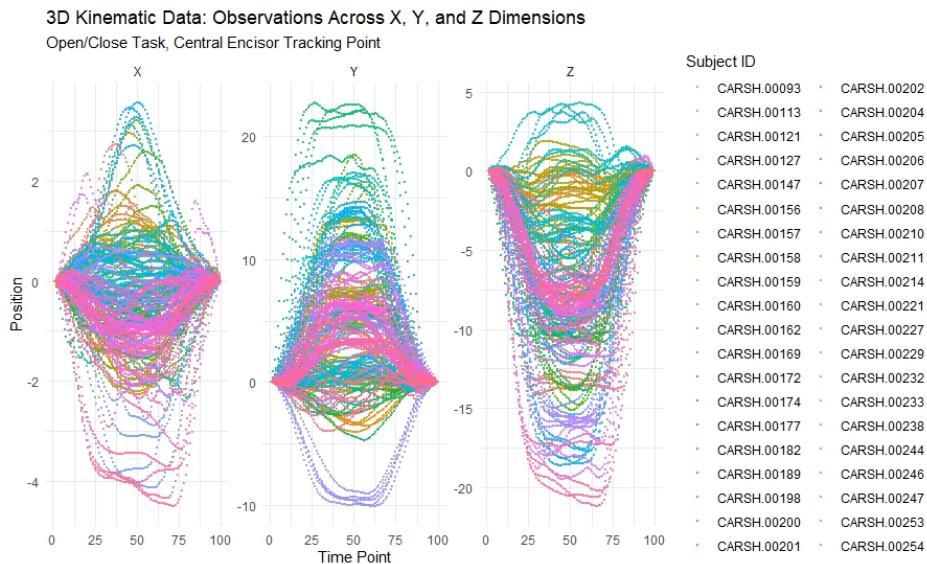


FIGURE 1 Observed Kinematic Data for the Central Encisor Tracking Point, Open/Close Task. Subjects are grouped by color.

1.2 | Description of Covariates

In addition to kinematic data on each subjects tasks and replications, we also have data on baseline scalar covariates. Baseline scalar covariates consisted of sex, age, ethnicity, skeletal classification, and the presence of TMD symptoms. Age was modeled

as a continuous predictor, while the remaining covariates were categorical. A summary of baseline covariates is provided in Table ???. The "average" subject was a 32 year old caucasian female with a class 3 skeletal type and no TMD symptoms. When creating the fixed effect design matrix, the categorical variables are encoded via dummy variables.

2 | METHODS AND BACKGROUND

It is helpful to consider this problem from the point of view that we have functional data for a group of subjects and replications.

In traditional statistical analysis, data often consists of scalar values, meaning individual numerical measurements. Classical statistical methods, such as regression models, hypothesis testing, and probability distributions, are well-suited for analyzing this type of data. However, in many real-world applications, data is not limited to scalar observations. Instead, it can take the form of curves or functions observed over a continuum, such as time or space. Functional data of this nature appears in numerous disciplines. For instance, growth curves are frequently analyzed in biological and medical research to study changes in body measurements over time¹. Similarly, hormone levels fluctuate within an individual's body throughout the day, and their trajectories provide insights into biological processes². In engineering and environmental sciences, sensor data often consists of time series that capture continuous readings of temperature, pressure, or other physical phenomena³. Even in finance, stock prices evolve over time as continuous functions, and understanding their dynamics is critical for optimizing investment strategies⁴. In other words, data often presents itself as curves, functions, or other non scalar manifestations.

There are a variety of methods for analysis with functional data depending on if the response, predictors, or both are functional. One of the early attempts at doing so was the 1982 paper by Ramsay⁵. The focus of this work is the scenario where the response is functional and we are attempting to understand the connection between scalar predictors and said functional outcome. This scenario is commonly referred to as function-on-scalar (FOS) regression. Ramsay and Silverman⁶ introduced a basic model for FOS regression that took the form of:

$$\mathbf{y}(t) = \mathbf{Z}\boldsymbol{\beta}(t) + \boldsymbol{\epsilon}(t) \quad (1)$$

Typically, the observed data $\mathbf{y}(t)$ is not functional in observation. Rather, we assume that the observed data is simply a realization of the true response curves as time points t_1, \dots, t_n ⁷. This introduces an inherent issue of dimensionality: while the true response is theoretically infinite-dimensional, we must develop models that approximate it using a finite number of parameters. Fortunately, there are several ways to do this, including Fourier series, wavelets, principal components, and the focus of this work, splines⁸. The use of splines also allows for the incorporation of penalty terms, which can help control overfitting and improve generalisability.

There are extensive pools of literature to pull from with respect to functional data analysis and function on scalar regression in particular. Beginning in the 1950's, principal components were used to model the mean of curves.

2.1 | Goldsmith's Univariate Model

This work is motivated by the model elucidated by Goldsmith. Therefore a brief description of the "base" model is appropriate. We observe functional data, at this stage univariate, for each subject, collected over multiple visits, resulting in a total number of observations n across all subjects. This data consist of curves representing possibly multiple dimensions of positions tracked over time, along with a set of scalar covariates for each subject. Goldsmith's as well as our model includes fixed effects to capture the influence of these covariates, subject-specific random effects to account for individual differences, and residual curves that may be correlated. We and Goldsmith estimate these effects using penalized splines, specifically cubic B-splines with a penalty that controls smoothness

Goldsmith assumes that the data is comprised of functional responses and scalar covariates for subjects $i = 1, \dots, I$ and visits $j = 1, \dots, J_i$, meaning that we have $n = \sum_i J_i$ total observations. He formulates a model with functional responses observed on a uniform grid of length D for all subjects and visits, giving the following model:

$$\mathbf{y}_{ij} = \omega_i \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b} + \boldsymbol{\epsilon}_{ij} \quad (2)$$

where \mathbf{y}_{ij} denotes a $1 \times D$ functional outcome vector for subject i and replication j , ω_i denotes a fixed effect design vector, \mathbf{z}_{ij} denotes a random effects design vector, $\boldsymbol{\beta}$ denotes the fixed effect functional coefficient matrix, \mathbf{b} denotes the random effect coefficient matrix, and $\boldsymbol{\epsilon}_{ij}$ denotes a $1 \times D$ residual vector following a multivariate normal distribution centered at 0⁹

β and b are expanded via splines. A spline matrix Θ is created of size $D \times K$, where K is the number of basis functions used. Correspondingly, matrices of spline coefficients are introduced, meaning that β and b are rewritten as $[\Theta B_w]^T$ and $[\Theta B_Z]^T$ respectively.

The columns of B_w and B_Z are penalized with a penalty matrix P that is selected *a priori*. Variances in B_w are unique to each coefficient, whereas the variances in the random effects are pooled. This gives Goldsmith the following model for the univariate case:

$$\begin{aligned} y_{ij} &= \omega_i B_W^T \Theta^T + z_{ij} B_Z^T \Theta^T + \epsilon_{ij} \\ \epsilon_{ij} &\sim N[0, \Sigma] \\ B_{Z,i} &\sim N[0, \sigma_z^2 P^{-1}] \quad \text{for } i = 1, \dots, I \\ B_{W,k} &\sim N[0, \sigma_{W,k}^2 P^{-1}] \quad \text{for } k = 0, \dots, p \end{aligned} \tag{3}$$

By placing an inverse Wishart prior on ϵ_{ij} , inverse gamma priors on the variance terms, and recentering the model, Goldsmith presents the following:

$$\begin{aligned} Y &= ZB_Z^T \Theta^T + \epsilon \\ \epsilon &\sim N[0, \Sigma \otimes I_n] \\ \Sigma &\sim IW[v, \Psi] \\ B_{Z,i} &\sim N[B_w w_i^T, \sigma_z^2 P^{-1}] \quad \text{for } i = 1, \dots, I \\ \sigma_z^2 &\sim IG[a_z, b_z] \\ B_{W,k} &\sim N[0, \sigma_{W,k}^2 P^{-1}] \\ \sigma_{W,k}^2 &\sim IG[a_{W,k}, b_{W,k}] \quad \text{for } k = 1, \dots, p \end{aligned} \tag{4}$$

In the paper, he provides derivations to obtain conditional distributions for model parameters and uses these to implement a Gibbs sampler.

2.2 | Higher Dimensions

Goldsmith provides a model for bivariate data; his motivating dataset is bivariate in nature (hand movement). He provides a model for two dimensions⁹. He concatenates data and spline coefficient matrices. Specifically, $Y = [Y_1, Y_2, Y_3]$, where $Y_i, i \in (1, 2)$ corresponds to functional outcomes in the X and Y dimensions, $B_W^T = [B_{1,W}^T, B_{2,W}^T]$, corresponding to the fixed effect spline coefficients in two dimensions, and $B_Z^T = [B_{1,Z}^T, B_{2,Z}^T]$ corresponding to the random effect spline coefficients in two dimensions. Taking advantage of Kronecker product structure, Goldsmith writes out the following model:

$$\begin{aligned} Y &= ZB_Z^T (I_2 \otimes \Theta^T) + \epsilon \\ \epsilon &\sim N[0, \Sigma \otimes I_n] \\ \Sigma &\sim IW[v, \Psi] \\ B_{Z,i} &\sim N[B_w w_i^T, \text{diag}(\sigma_{1,Z}^2, \sigma_{2,Z}^2) \otimes P^{-1}] \quad \text{for } i = 1, \dots, I \\ \sigma_{m,Z}^2 &\sim IG[a_{m,Z}, b_{m,Z}] \quad \text{for } m = 1, 2 \\ B_{W,i} &\sim N[0, \text{diag}(\sigma_{1,W_k}^2, \sigma_{2,W_k}^2) \otimes P^{-1}] \\ \sigma_{m,W_k}^2 &\sim IG[a_{m,W_k}, b_{m,W_k}] \quad \text{for } m = 1, 2 \quad \text{and } k = 1, \dots, p \end{aligned} \tag{5}$$

In this model, the fixed and random effects for each dimension are assumed independent and penalized separately. The uni-variate Gibbs sampler is easily adapted to the two dimensional case.

2.3 | Note on Penalty Matrix P

In Goldsmith, the penalty matrix P is constructed by using $\alpha P_0 + (1 - \alpha)P_2$, where P_0 is the identity matrix and P_2 is the second derivative penalty matrix. This weighted sum is implemented to alleviate the noninvertability of P_2 . Goldsmith suggests setting α to be small, less than 0.01 in order to avoid enforcing shrinkage. In general, I concur with this heuristic, though I would caution against setting α to be *too* small. The reason for this is that the magnitude of the smallest eigenvalue of the resulting P matrix move linearly with α ; if α is too small, then the smallest eigenvalue could be numerically/computationally 0, making P almost non-invertible.

2.4 | Working Models

In this section, we describe two potential working models for the analysis of tri-variate (three-dimensional) functional data, extending the bivariate framework of Goldsmith⁹. Our goal is to provide a flexible modeling structure while addressing computational and statistical challenges posed by high-dimensional covariance estimation.

2.4.1 | Goldsmith's Extended Model

As mentioned earlier, Goldsmth laid out a model for bivariate data. For the extension to three dimensions, we continue to follow his lead, giving the following model for three dimensional functional data:

$$\begin{aligned}
 Y &= ZB_Z^T (I_3 \otimes \Theta^T) + \epsilon \\
 \epsilon &\sim N[0, \Sigma \otimes I_n] \\
 \Sigma &\sim IW[v, \Psi] \\
 B_{Z,i} &\sim N[B_w w_i^T, \text{diag}(\sigma_{1,Z}^2, \sigma_{2,Z}^2, \sigma_{3,Z}^2) \otimes P^{-1}] \quad \text{for } i = 1, \dots, I \\
 \sigma_{m,Z}^2 &\sim IG[a_{m,Z}, b_{m,Z}] \quad \text{for } m = 1, 2, 3 \\
 B_{W,i} &\sim N[0, \text{diag}(\sigma_{1,W_k}^2, \sigma_{2,W_k}^2, \sigma_{3,W_k}^2) \otimes P^{-1}] \\
 \sigma_{m,W_k}^2 &\sim IG[a_{m,W_k}, b_{m,W_k}] \quad \text{for } m = 1, 2, 3 \quad \text{and } k = 1, \dots, p
 \end{aligned} \tag{6}$$

2.4.2 | Alternative Model

In his model, Goldsmith places an inverse-Wishart prior on the level 1 covariance structure. There is one cause for concern with this: for observations in three dimensions with each dimension being observed across a grid of D , the level 1 covariance matrix would have to be $3 * D \times 3 * D$. In the case of $D = 100$, that would mean that our level 1 covariance matrix would be 300×300 . As the dimensionality of the matrix involved increases, the scale parameter also must increase to for the distribution to remain proper. A result of this is that the Inverse Wishart $IW(nS + \Psi, n + v)$ posterior distribution tends to concentrate more tightly around its prior expected value: $\Psi/(v - p - 1)^{10}$. When the scale matrix is the identity matrix, this expected value is proportional to the identity matrix itself; using an Inverse Wishart prior with an identity scale matrix can lead to over-shrinkage of the estimated covariance matrix towards the identity matrix.

In addition, we endeavor to use STAN, which encounters computational difficulties for high dimensional inverse wishart distributions. To this end, we propose a slightly modified model that still accounts for correlation will not needing to delve into the uncertainty of higher order inverse Wishart distributions:

$$\begin{aligned}
Y &= ZB_Z^T (I_3 \otimes \Theta^T) + \epsilon \\
\epsilon &\sim N[0, \sigma_0^2 I_n] \\
B_{Z,i} &\sim N[B_w w_i^T, \Omega_Z \otimes P^{-1}] \quad \text{for } i = 1, \dots, I \\
\Omega_Z &\sim IW[v_z, \Psi_z] \\
B_{W,i} &\sim N[0, \Omega_{W,k} \otimes P^{-1}] \\
\Omega_{W,k} &\sim IW[v_{w,k} \Psi_{w,k}] \quad \text{for } k = 1, \dots, p
\end{aligned} \tag{7}$$

In other words, we have moved the dependence from the level 1 error structure into the fixed and random effects. The intuition behind this is that now, rather than a 300×300 covariance matrix, our model only has several 3×3 . The intuition to this approach is that not only is the computational burden reduced, but the mixing in MCMC sampling (especially in STAN) is also improved. Furthermore, this also allows for more interpretable covariance structures at the subject and covariate levels. This setup allows for data borrowing across dimensions without forcing the residual error covariance to absorb all variation.

3 | SIMULATION

In statistics, simulation studies are often used; they can offer empirical evidence of a new method. Additionally the controlled environment under which they are run allows for easy measurement of accuracy and repeatability. This work is no different. We demonstrate the performance of these two competing models by simulating three dimensional kinematic data. Code required to reproduce these data generation and simulation steps may be found at:XXX.

3.1 | Data Generation

In general, the generated data is generated assuming a grid length $D = 25$ in all three dimensions for a total grid length of 75. The data follows a model structured in the following fashion:

$$\underbrace{Y_{gen}}_{n \times 3D} = \underbrace{\underbrace{X}_{n \times p} * \underbrace{\beta}_{p \times 3D}}_{n \times 3D} + \underbrace{\underbrace{Z}_{n \times I} * \underbrace{b}_{I \times 3D}}_{n \times 3D} + \underbrace{\epsilon}_{n \times 3D} \tag{8}$$

We generate smooth functional coefficients using B-splines for three-dimensional data. In doing so, we create a coefficient matrix for each dimension (X, Y, and Z) by applying a B-spline basis and random coefficients. This leaves us with a $p \times 3 * D$ grid of coefficients, where each block corresponds to smooth coefficients functions in each dimension. The subject design matrix X is generated from a normal distribution with a user specified variance.

To obtain the fixed effect components of this data, we generate an $n \times p$ design matrix comprised of baseline covariates for each of the I subjects over J_i replications and multiply this by the matrix containing the grid of coefficients. For subject specific random effects, an $n \times I$ design matrix is generated and multiplied by random coefficients generated from a standard normal distribution. To generate level-1 errors, we draw $n \times 3 * D$ values from a standard normal distribution.

An example of this simulated data is shown below:

3.2 | Results

In our simulation, we varied several key components of the model: the number of knots used in data estimation, the number of covariates, and the variance of both fixed and random effect variances, using both the STAN code and comparing the results to Goldsmith's tri-variate Gibbs sampler. For the STAN model, we ran four parallel chains, each chain containing 2000 iterations (1000) burn in, for a total of 8000 iterations with the first 4000 discarded as burnin. The for the tri-variate Goldsmith sampler, one chain with 8000 iterations and 4000 burn in iterations were run to have analysis be roughly comparable.

Goldsmith's model spent a considerable amount of time discussing various initial values and hyperparameter values. In particular he discusses an initial value for the inverse Wishart Scale parameter. As our model does away with most of those

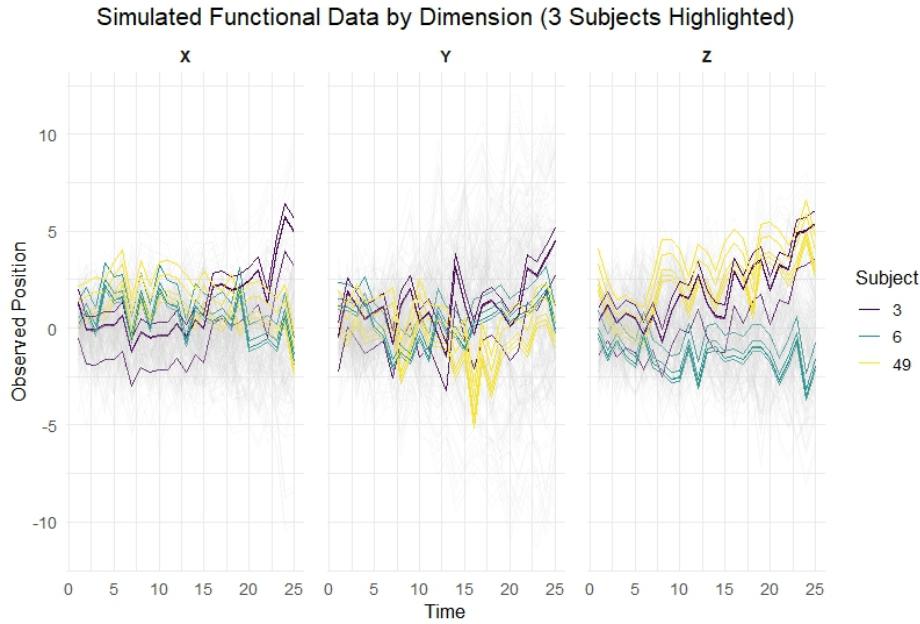


FIGURE 2 Example of Simulated Data. Three random subjects are highlighted.

initial values, we did not set initial values and largely used uninformative diagonal matrices as initial scale parameters for our smaller inverse Wishart distribution.

For each condition, 10 datasets were generated and both the Gibbs and STAN samplers were run, and several measures of performance were recorded including computation time (in seconds), integrated mean squared error (compared to the "true" coefficient curves), and mean coverage across all coefficient curves. The results are shown in Table A2.

In terms of coverage, both the STAN model and the Gibbs sampler model often achieve coverage > 90%. For some cases however, STAN coverage fell to 79.8%. The two models had comparable IMSE values. In terms of computation time, Gibbs sampling was consistently much slower than STAN, occasionally by an order of magnitude. Increasing the number of knots in the fitted model often improved coverage and reduced IMSE.

4 | EMPERICAL RESULTS

Now, we turn our attention to the empirical data described earlier. The focus is on the estimation of the the significance of several scalar covariates on the overall three dimensional kinematic motion of the jaw. The categorical variables are encoded as dummy variables for the purposes of design matrices. We used the modified Gibbs sampler with one chain. Below are some results from this:

A table in the appendix details the significant variables. The covariate age is the most consistently significant variable; it is significant across all tasks and at all tracking points. Additionally, it is significant across a wide range of time points. Sex is another covariate that is very commonly significant, though typically over shorter windows of time than age. Various ethnicities are also commonly significant, though mostly in the Z dimension and like sex, across shorter time spans than age. Skeletal class also seems to mostly significant in the Z dimension, likewise at shorter windows of time than age. The Presence of TMD symptoms is significant only in limited contexts (at one task, one tracking point, and only from time 40 to 62).

5 | CONCLUSIONS

The purpose of this work has been to on the using functional regression to analyze three dimensional kinematic data. Our main model is an extension of an existing method, allowing for the modelling of mean effects, subject level random effects, and

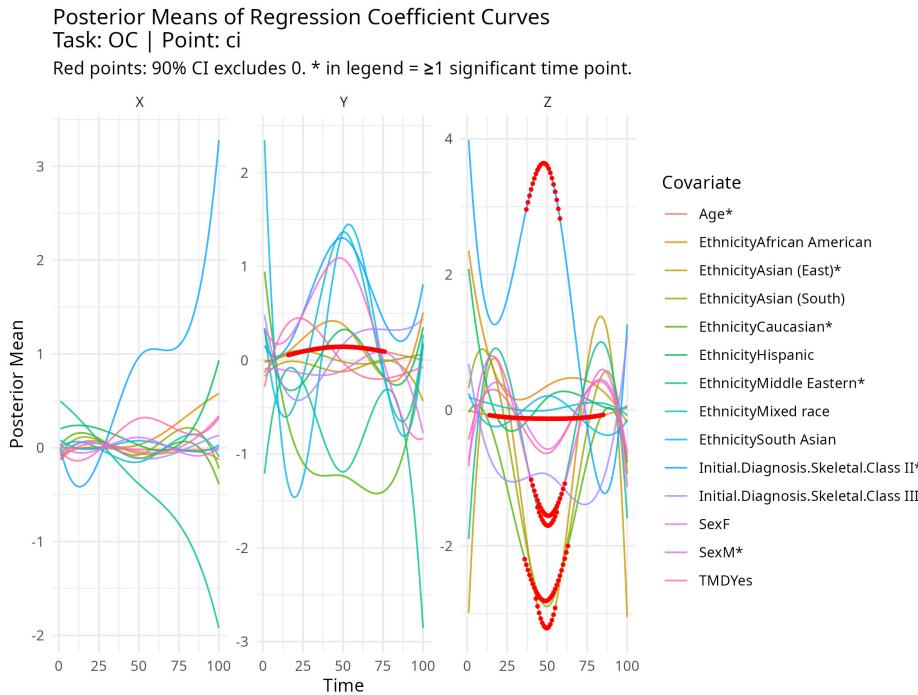


FIGURE 3 Results for Open/Close Task, central encisor tracking point. Covariates that contained at least one significant time point are denoted with an asterisk in the legend. Red points are used to denote time points for the coefficient curves that do not include 0 in their 90% posterior credibility interval.

correlated errors across all three dimensions. Attempting to simplify the model estimation by moving a large covariance structure into the random and fixed effects did not work as well as was hoped.

Despite this, analysis was performed using a modified version of an existing Gibbs sampler, , which provided stable estimation while retaining flexibility in the modeling framework. This approach allowed us to identify covariates that exert consistent and interpretable effects on jaw motion. In particular, age emerged as the most robust and persistent factor, while sex, ethnicity, and skeletal class were significant in more localized contexts, typically in the Z dimension and over shorter time spans. The presence of TMD symptoms was found to have only limited impact, significant in a restricted set of tasks, points, and time intervals.

More generally, this work demonstrates that high-dimensional functional outcomes, when carefully structured, via hierarchical models for example, can yield interpretable insights while maintaining statistical rigor.

There are numerous avenues for future exploration of this work's results. One such idea could be the inclusion of the selection of the roughness penalty α into the model instead of selecting it *a priori*. This would make the model truly Bayesian, in the sense that the data would select the smoothness penalty. Another thought would be to impose structure on the level 1 covariance matrix. In general, for a p dimensional covariance matrix, there are $\frac{p*(p+1)}{2}$ entries to estimate. For our case ($p=300$), this results in there being 45,150 entries to estimate. Imposing a structure on this matrix, say for example a first order autoregressive structure would reduce the number of parameters to just two: the marginal variance and the correlation factor ρ . While the STAN model did not perform as well in this setting, perhaps it would do better in this structured covariance setting.

AUTHOR CONTRIBUTIONS

This is an author contribution text. This is an author contribution text.

ACKNOWLEDGMENTS

We gratefully acknowledge use of TMJ kinematics data obtained through NIH grant U01DE031512 (H. Yao and J. Lee, PIs).

FINANCIAL DISCLOSURE

None reported.

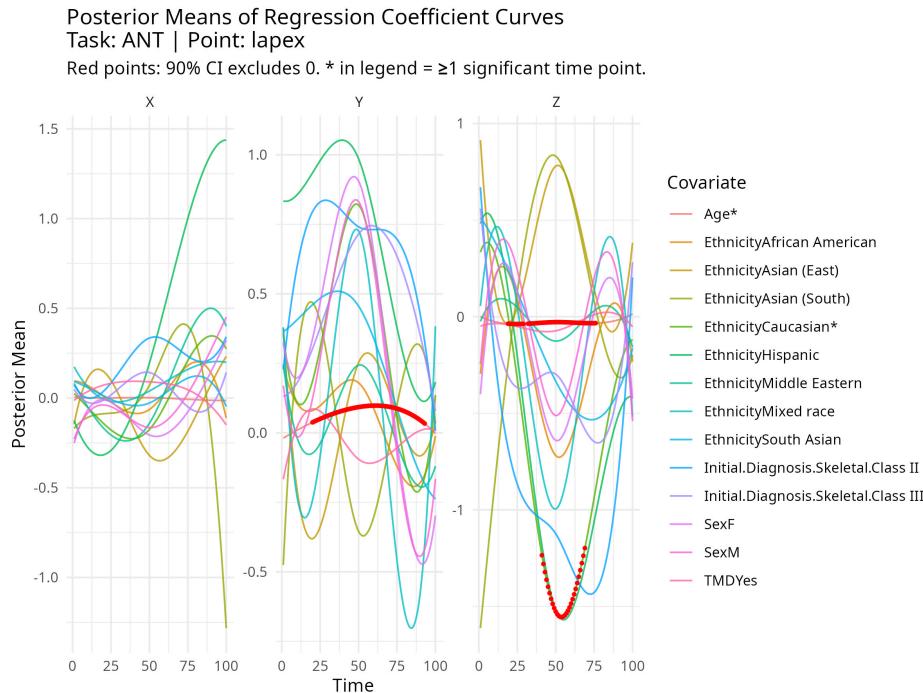


FIGURE 4 Results for Anterior Task, left apex tracking point.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

REFERENCES

1. Rao CR. Some statistical methods for comparison of growth curves. *Biometrics*. 1958;14(1):1–17. ISBN: 0006-341X Publisher: JSTOR.
2. Brumback BA, Rice JA. Smoothing Spline Models for the Analysis of Nested and Crossed Samples of Curves. *Journal of the American Statistical Association*. 1998;93(443):961–976. Publisher: [American Statistical Association, Taylor & Francis, Ltd.]doi: 10.2307/2669837
3. Nikolova B, Nikolov G, Todorov M. Curve fitting of sensors' characteristics. *Annual Journal of Electronics, ISSN*. 2009:1313–1842.
4. Fama EF. The behavior of stock-market prices. *The journal of Business*. 1965;38(1):34–105. ISBN: 0021-9398 Publisher: JSTOR.
5. Ramsay JO. When the data are functions. *Psychometrika*. 1982;47(4):379–396. doi: 10.1007/BF02293704
6. Ramsay J, Silverman B. *Functional Data Analysis*. Springer Series in StatisticsSpringer, 2005.
7. Reiss PT, Huang L, Mennes M. Fast function-on-scalar regression with penalized basis expansions.. *The international journal of biostatistics*. 2010;6(1):Article 28. Place: Germanydoi: 10.2202/1557-4679.1246
8. Morris JS. Functional regression. *Annual Review of Statistics and Its Application*. 2015;2(1):321–359. Publisher: Annual Reviews.
9. Goldsmith J, Kitago T. Assessing systematic effects of stroke on motorcontrol by using hierarchical function-on-scalar regression.. *Journal of the Royal Statistical Society. Series C, Applied statistics*. 2016;65(2):215–236. Place: Englanddoi: 10.1111/rssc.12115
10. Zhang Z. A Note on Wishart and Inverse Wishart Priors for Covariance Matrix. *Journal of Behavioral Data Science*. 2021;1. doi: 10.35566/jbds/v1n2/p2

SUPPORTING INFORMATION

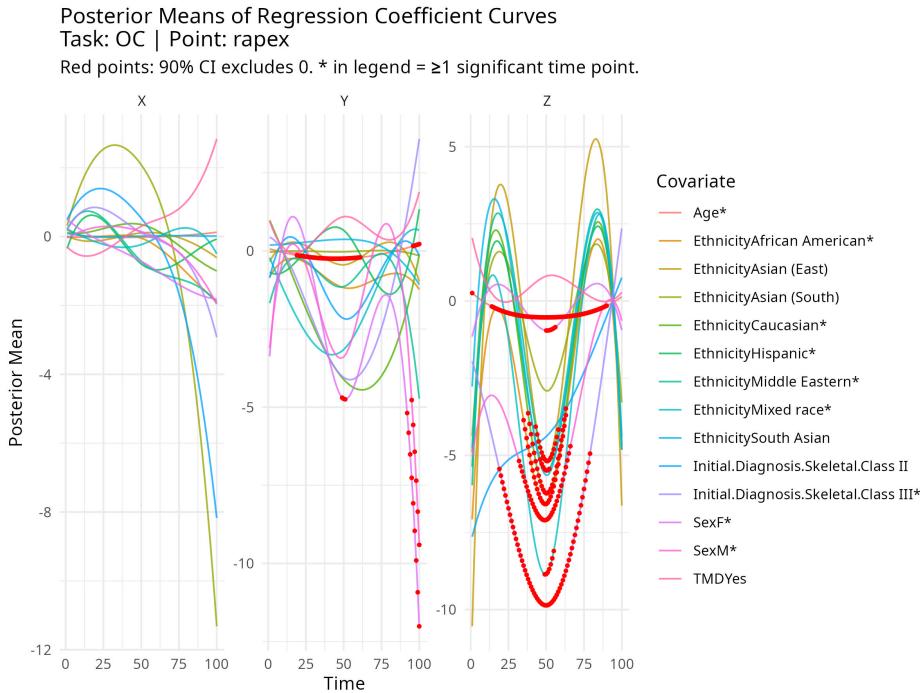
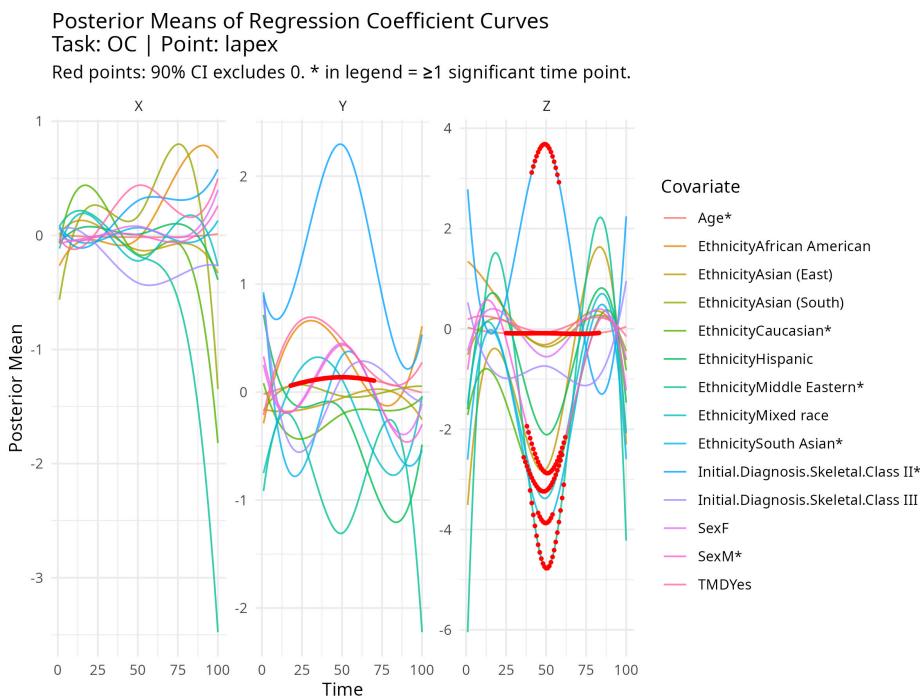
Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Swanson P, Lauritzen P, Erath C, and Mittal R. Functional Mixed-Effect Modeling of Multivariate Jaw Kinematics *J Comput Phys*. 2021;00(00):1–18.

	Overall
Number of Patients	40
Sex (%)	
M	15 (37.5)
F	25(62.5)
Age (mean (SD))	32.73 (14.81)
Ethnicity (%)	
African	1 (2.5)
African American	12 (30.0)
Asian (East)	1 (2.5)
Asian (South)	1 (2.5)
Caucasian	18 (45.0)
Hispanic	1 (2.5)
Middle Eastern	1 (2.5)
Mixed race	3 (7.5)
South Asian	2 (5.0)
Initial.Diagnosis.Skeletal. (%)	
Class I	12 (30.0)
Class II	6 (15.0)
Class III	22 (55.0)
TMD (%)	
Yes	16 (40.0)
No	24 (60.0)

TABLE A1 Summary of Covariates**APPENDIX****A TABLES****B EMPERICAL RESULTS**

Below are some more images from the emperical results:

**FIGURE B1** Results for Open/Close Task, Right Apex tracking point.**FIGURE B2** Results for Open/Close Task, Left Apex tracking point.

D	Kt_gen	Kt_sample	p	I	var_fixed	var_random	mean_stan.coverages	imse_stan	mean_gibbs.coverages	imse_gibbs	stan_times	gibbs_times
25	5	5	3	50	0.1000	0.1000	0.9040	0.0598	1.0000	0.0608	259.1143	3240.8970
25	5	10	3	50	0.1000	0.1000	0.9907	0.0638	1.0000	0.0607	1833.2448	7025.7435
25	5	5	10	50	0.1000	0.1000	0.8164	0.1386	0.9100	0.1262	299.9174	4234.6802
25	5	10	10	50	0.1000	0.1000	0.9808	0.0319	0.9997	0.0289	1113.6958	14491.6006
25	5	5	3	50	0.1000	0.1000	0.9280	0.0064	1.0000	0.0067	262.2866	3242.2934
25	5	10	3	50	0.1000	0.1000	0.9693	0.0069	1.0000	0.0070	915.1414	7037.9875
25	5	5	10	50	0.1000	0.1000	0.7984	0.0162	0.8704	0.0172	306.5938	4235.3996
25	5	10	10	50	0.1000	0.1000	0.9796	0.0039	0.9995	0.0039	1179.2835	14479.8795
25	5	5	3	50	0.1000	0.1000	0.9960	0.1054	1.0000	0.1036	277.2761	3224.6706
25	5	10	3	50	0.1000	0.1000	1.0000	0.1268	1.0000	0.1224	638.9297	7026.6621
25	5	5	10	50	0.1000	0.1000	0.8964	0.1487	0.9796	0.1501	316.9695	4221.2304
25	5	10	10	50	0.1000	0.1000	0.9916	0.0715	1.0000	0.0940	838.7632	14462.8522
25	5	5	3	50	0.1000	0.1000	0.9947	0.0116	1.0000	0.0117	280.4149	3234.2076
25	5	10	3	50	0.1000	0.1000	0.9840	0.0161	1.0000	0.0159	786.4246	7029.5817
25	5	5	10	50	0.1000	0.1000	0.8780	0.0178	0.9688	0.0211	318.1880	4229.4626
25	5	10	10	50	0.1000	0.1000	0.9884	0.0099	1.0000	0.0129	862.8026	14489.7824

TABLE A2 Summary of simulation partial results. Numeric results rounded to four decimal places.

coef	dim	time_points	task	tracking_point
Age	Y	18-93	ant	ci
Age	Z	15-84	ant	ci
EthnicityAfrican American	Z	42-62	ant	ci
EthnicityAsian (South)	Z	41-59	ant	ci
EthnicityCaucasian	Z	39-67	ant	ci
SexF	Z	46-52	ant	ci
SexM	Y	41-54	ant	ci
SexM	Z	50	ant	ci
Age	Y	20-93	ant	lapex
Age	Z	19-29, 32-76	ant	lapex
EthnicityCaucasian	Z	41-69	ant	lapex
Age	Y	18-87	ant	rapex
Age	Z	10-83	ant	rapex
EthnicityAfrican American	Z	40-62	ant	rapex
EthnicityAsian (South)	Z	1, 15-21, 39-63, 81-88	ant	rapex
EthnicityCaucasian	Y	42-53	ant	rapex
Initial.Diagnosis.Skeletal.Class III	Z	32-59	ant	rapex
Age	Z	41-68	latL	ci
EthnicityAsian (East)	Z	39-64	latL	ci
EthnicityAsian (South)	Z	41-62	latL	ci
EthnicityCaucasian	Z	42-65	latL	ci
Initial.Diagnosis.Skeletal.Class II	Z	43-59	latL	ci
Initial.Diagnosis.Skeletal.Class III	Z	43-60	latL	ci
TMDYes	Z	40-62	latL	ci
Age	Y	25-74	latL	lapex
Age	Z	20-83	latL	lapex
EthnicityAsian (East)	Z	38-66	latL	lapex
EthnicityAsian (South)	Z	47-55	latL	lapex
EthnicitySouth Asian	Z	39-63	latL	lapex
Initial.Diagnosis.Skeletal.Class III	Z	41-70	latL	lapex
SexF	Z	43-55	latL	lapex
Age	X	20-78	latL	rapex
Age	Z	23-86	latL	rapex
SexF	X	30-69	latL	rapex
SexM	X	32-68	latL	rapex
Age	Y	23-92	latR	ci
Age	Z	20-83	latR	ci
Initial.Diagnosis.Skeletal.Class II	Z	34-67	latR	ci
SexF	Y	40-52	latR	ci
SexF	Z	42-58	latR	ci
SexM	Z	46-53	latR	ci
EthnicityAsian (South)	X	90-100	latR	lapex
Age	X	14-86	latR	rapex
Age	Z	17-86	latR	rapex
SexF	X	37-59	latR	rapex
SexM	X	33-65	latR	rapex
Age	Y	16-76	oc	ci
Age	Z	14-85	oc	ci
EthnicityAsian (East)	Z	43-55	oc	ci
EthnicityCaucasian	Z	36-63	oc	ci
EthnicityMiddle Eastern	Z	44-56	oc	ci
Initial.Diagnosis.Skeletal.Class II	Z	37-58	oc	ci
SexM	Z	40-61	oc	ci
Age	Y	18-70	oc	lapex
Age	Z	25-83	oc	lapex
EthnicityCaucasian	Z	36-60	oc	lapex
EthnicityMiddle Eastern	Z	40-61	oc	lapex

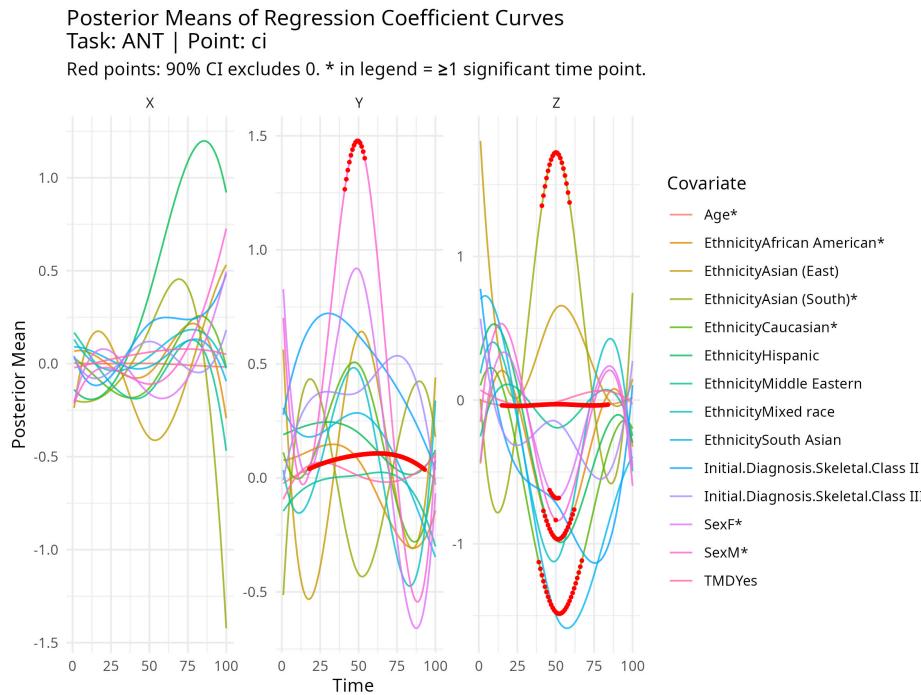


FIGURE B3 Results for Anterior Task, Central Incisor tracking point.

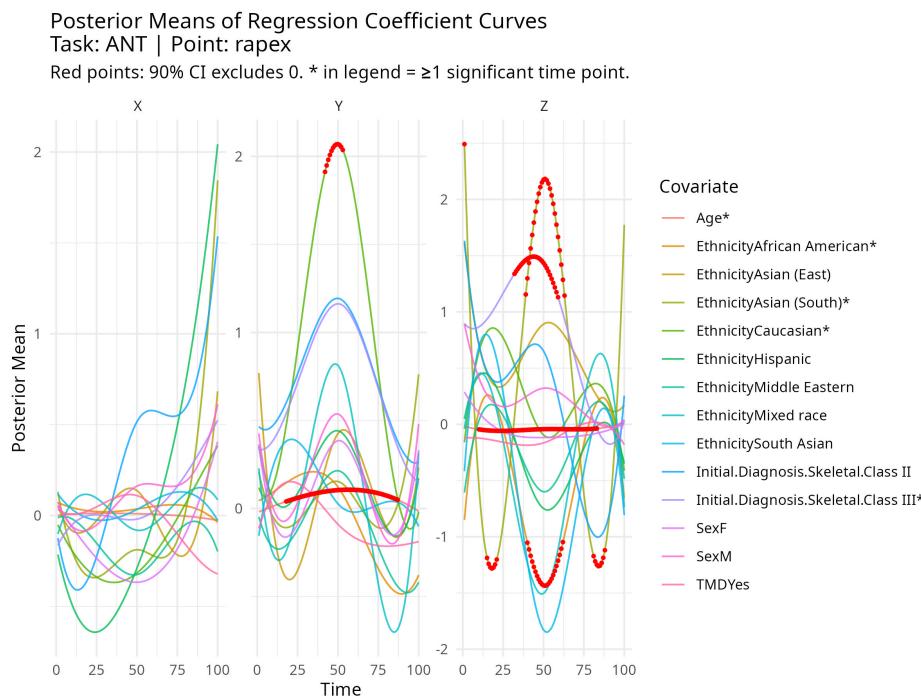


FIGURE B4 Results for Anterior Task, Right Apex tracking point.

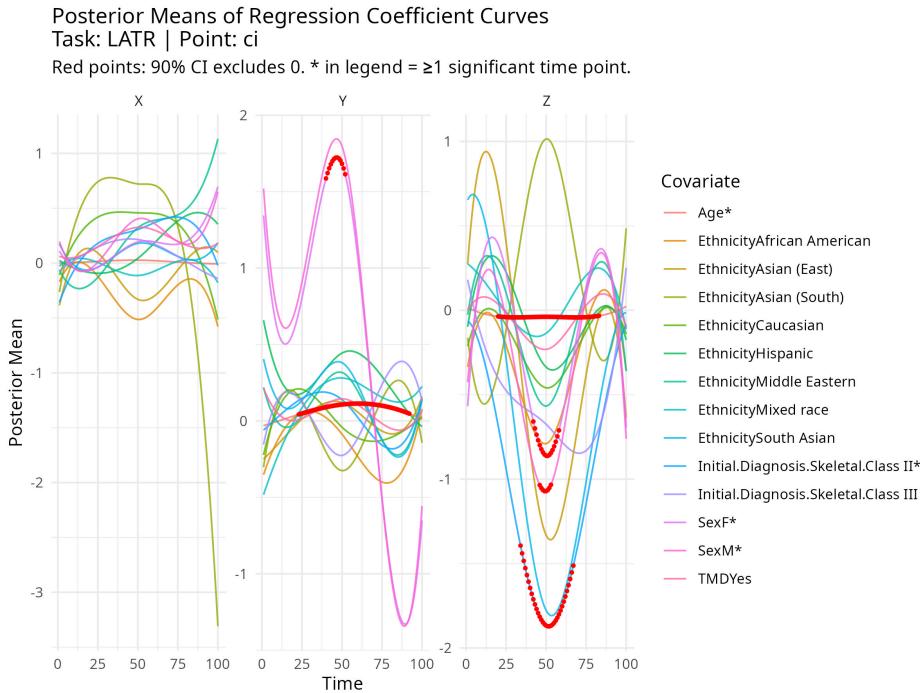


FIGURE B5 Results for Lateral Right Task, Central Incisor tracking point.

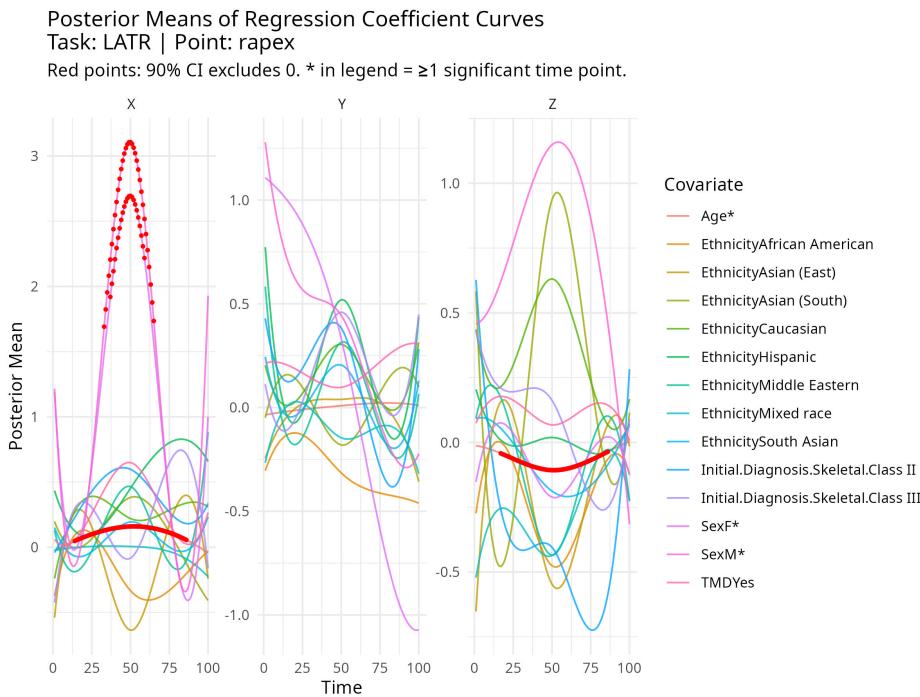


FIGURE B6 Results for Lateral Right Task, Right Apex tracking point.

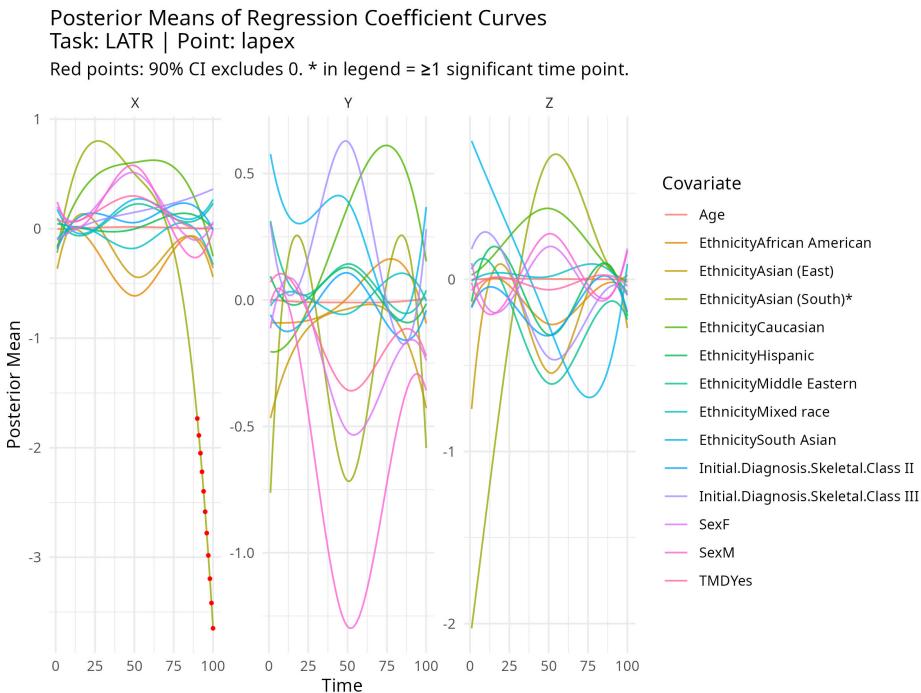


FIGURE B7 Results for Lateral Right Task, Left Apex tracking point.

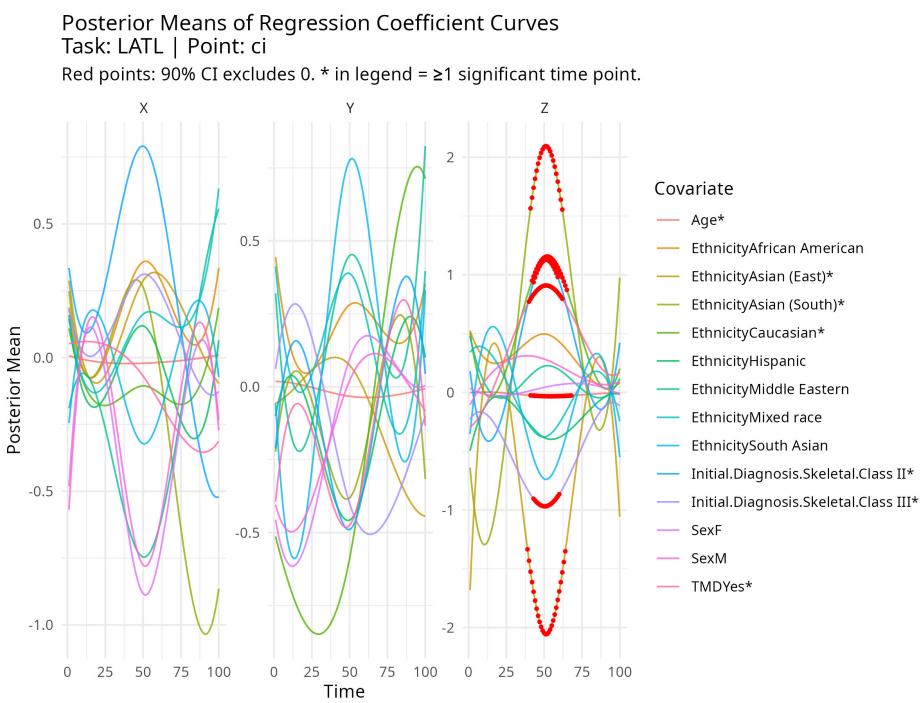
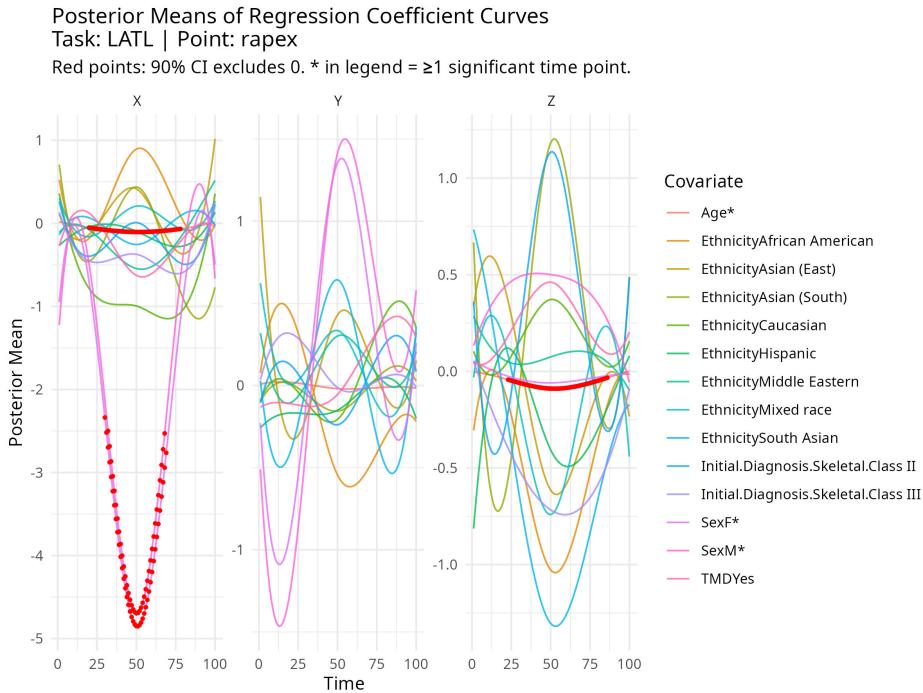
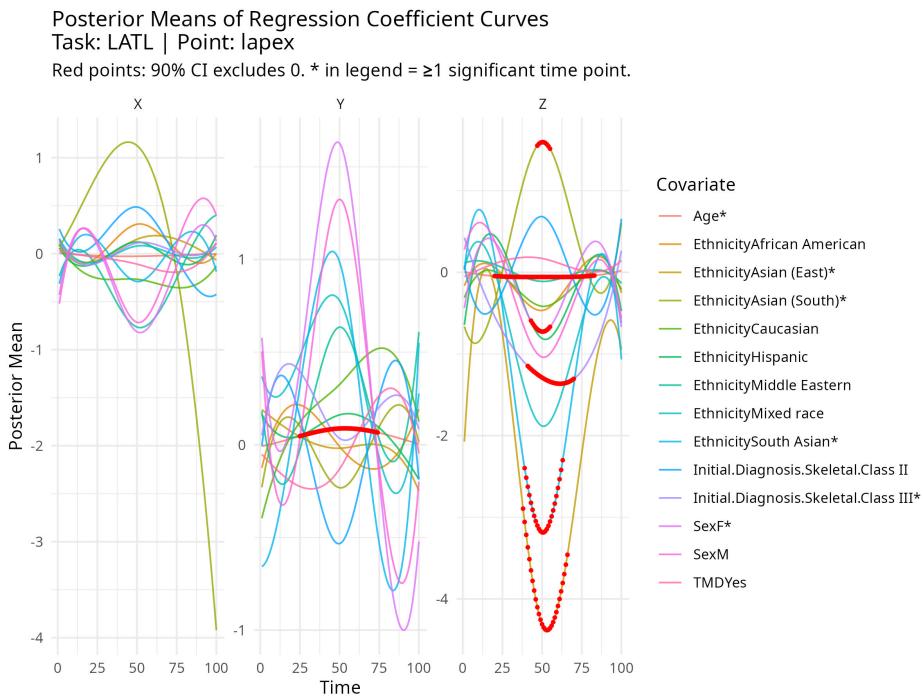


FIGURE B8 Results for Lateral Left Task, Central Incisor tracking point.

**FIGURE B9** Results for Lateral Left Task, Right Apex tracking point.**FIGURE B10** Results for Lateral Left Task, Left Apex tracking point.