




RESEARCH ARTICLE

Shape Mediation Analysis in Alzheimer's Disease Studies

Xingcai Zhou¹ | Miyeon Yeon²  | Jiangyan Wang¹ | Shengxian Ding²  | Kaizhou Lei² | Yanyong Zhao¹ | Rongjie Liu³ | Chao Huang⁴  | The Alzheimer's Disease Neuroimaging Initiative

¹Institute of Statistics and Data Science, Nanjing Audit University, Nanjing, China | ²Department of Statistics, Florida State University, Tallahassee, FL, USA |

³Department of Statistics, University of Georgia, Athens, GA, USA | ⁴Department of Epidemiology and Biostatistics, University of Georgia, Athens, GA, USA

Correspondence: Chao Huang (chaohuang@uga.edu)

Received: 5 September 2023 | **Revised:** 19 September 2024 | **Accepted:** 16 October 2024

Funding: C. Huang gratefully acknowledges financial support from the National Science Foundation Division of Mathematical Sciences (Grant No. DMS-1953087). X. Zhou gratefully acknowledges support from the National Natural Science Foundation of China (Grant No. 12171242). J. Wang gratefully acknowledges support from the National Natural Science Foundation of China (12271255) and Qinglan Project of Jiangsu Province.

Keywords: Alzheimer's disease | corpus callosum | mediation analysis | scalar-on-shape partial single index regression model | shape-on-scalar regression model

ABSTRACT

As a crucial tool in neuroscience, mediation analysis has been developed and widely adopted to elucidate the role of intermediary variables derived from neuroimaging data. Typically, structural equation models (SEMs) are employed to investigate the influences of exposures on outcomes, with model coefficients being interpreted as causal effects. While existing SEMs have proven to be effective tools for mediation analysis involving various neuroimaging-related mediators, limited research has explored scenarios where these mediators are derived from the shape space. In addition, the linear relationship assumption adopted in existing SEMs may lead to substantial efficiency losses and decreased predictive accuracy in real-world applications. To address these challenges, we introduce a novel framework for shape mediation analysis, designed to explore the causal relationships between genetic exposures and clinical outcomes, whether mediated or unmediated by shape-related factors while accounting for potential confounding variables. Within our framework, we apply the square-root velocity function to extract elastic shape representations, which reside within the linear Hilbert space of square-integrable functions. Subsequently, we introduce a two-layer shape regression model to characterize the relationships among neurocognitive outcomes, elastic shape mediators, genetic exposures, and clinical confounders. Both estimation and inference procedures are established for unknown parameters along with the corresponding causal estimands. The asymptotic properties of estimated quantities are investigated as well. Both simulated studies and real-data analyses demonstrate the superior performance of our proposed method in terms of estimation accuracy and robustness when compared to existing approaches for estimating causal estimands.

1 | Introduction

With the rapid growth of large-scale medical imaging studies, like the Alzheimer's Disease Neuroimaging Initiative (ADNI)

study and the UK Biobank (UKB) study [1, 2], there has been a significant surge in comprehensive research aimed at understanding the causal mechanisms underlying neurocognitive disorders such as Alzheimer's Disease (AD). One of the most

Xingcai Zhou and Miyeon Yeon are co-first authors.

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

important scientific questions is whether and how genetic exposure affects mental health through the changes in brain microstructures, typically treated as a causal mediation analysis problem [3]. Mediation analysis, a crucial tool in neuroscience, has been developed and widely adopted to elucidate the role of intermediary variables derived from neuroimaging data. These variables are situated along the path connecting exposures (or treatments) and clinical outcomes [4–7]. Both the direct effects of exposures on outcomes and the indirect effects mediated by these intermediary variables can be examined using statistical inference tools [8–10]. Typically, structural equation models (SEMs) are employed to investigate the influences of exposures on outcomes, with model coefficients being interpreted as causal effects [11–13]. Within the SEM framework, the primary focus lies in dissecting the total causal effects of exposures on outcomes into distinct direct and indirect components. Meanwhile, neuroimaging-related mediators in the SEM are usually represented as random variables or stochastic processes, depending on the characteristics of the extracted imaging features. These features can encompass univariate or multivariate variables (e.g., Region of Interest (ROI) based volumetric data [14, 15]), high dimensional variables (e.g., voxel-wised imaging signals [16]), and functions (e.g., fMRI or time series data [4, 17]).

In recent years, numerous significant scientific endeavors have aimed to examine the shapes of specific brain regions, revealing their pivotal roles in uncovering the causes of various mental disorders. One notable example is the shape of the corpus callosum (CC), the largest white matter structure in the human brain [18]. It has garnered increasing attention in Alzheimer’s Disease (AD) studies [19, 20] due to the significant atrophy observed in the CC subregion among AD patients [21] and mildly demented AD patients [22] in comparison to healthy controls. While existing SEMs have proven to be effective tools for mediation analysis involving various neuroimaging-related mediators, limited research has explored scenarios where these mediators are derived from the shape space. The shape is broadly defined to be a characteristic that is left after certain nuisance or shape-preserving transformations, such as rotations, translations, and scale, have been removed [23–25], with the result that shape representation spaces are nonlinear, high-dimensional, and have quotient space geometry [20]. Therefore, the conventional normality assumptions often applied to mediators in

existing SEMs cannot be directly extended to shape data. Furthermore, most existing SEMs assume linear relationships between mediators and clinical outcomes [4, 15, 16]. However, this linear relationship assumption may lead to substantial efficiency losses and decreased predictive accuracy in real-world applications, where constrained nonlinear relationships may indeed exist [26, 27].

This article aims to introduce a novel framework for Shape Mediation Analysis (SMA), designed to explore the causal relationships between genetic exposure and clinical outcomes, whether mediated or unmediated by shape-related factors while accounting for potential confounding variables. Figure 1a illustrates a Directed Acyclic Graph (DAG) depicting these causal connections.

To our knowledge, our SMA framework represents the first SEM tailored explicitly for shape mediators. Within our SMA framework, we initiate the process by applying the square-root velocity function (SRVF) [28] to extract elastic shape representations, which reside within the linear Hilbert space of square-integrable functions, denoted as $\mathbb{L}^2([0, 1])$. Subsequently, we introduce a two-layer shape regression model (2SRM). This model comprises two components: (i) a shape-on-scalar regression model, which explores the associations between elastic shape mediators, genetic exposure, and clinical confounding factors; and (ii) a scalar-on-shape partial linear single-index model, designed to characterize the nonlinear relationships among neurocognitive outcomes, elastic shape mediators, genetic exposure, and clinical confounders. The corresponding path diagram is depicted in Figure 1b. We employ a two-step estimation procedure to estimate unknown parameters within the 2SRM framework, along with the corresponding causal estimands. Additionally, we establish a bootstrap-based inference method to construct confidence intervals (CIs) and simultaneous confidence bands (SCBs) for these causal estimands. We also conduct a thorough investigation into the asymptotic properties of estimated quantities within 2SRM. Moreover, both simulated studies and real-data analyses demonstrate the superior performance of our proposed method in terms of estimation accuracy and robustness when compared to existing approaches for estimating causal estimands.

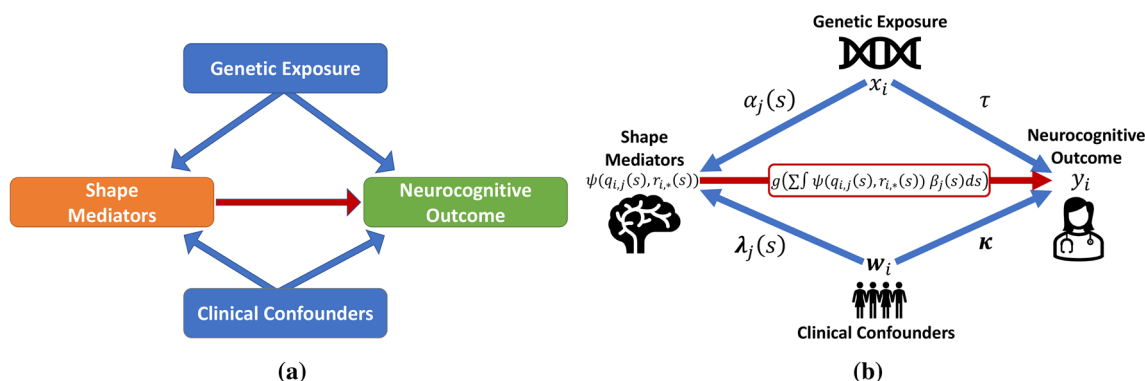


FIGURE 1 | (a) DAG illustrating the causal relationship among genetic exposure (blue), shape mediators (orange), clinical confounders (blue), and the neurocognitive outcome (green); (b) Path diagram including (i) a shape-on-scalar regression model and (ii) a scalar-on-shape partial linear single-index model to investigate the genetic causal effects on the neurocognitive outcome through the shape mediators.

The article is organized as follows. Section 2 introduces 2SRM and outlines the estimation and inference procedures for the parameters and causal estimands. Subsequently, we delve into exploring the theoretical properties of parameters of interest in 2SRM. In Section 3, simulation studies on synthetic curve data are presented to validate the finite performance of our proposed 2SRM across different scenarios. Finally, in Section 4, we apply our method to the CC shape dataset in the ADNI study.

2 | Methods

2.1 | Notation

Suppose we obtain the neuroimaging data, the genetic exposure, clinical confounders, and the neurocognitive outcome from n unrelated subjects. Instead of the whole brain image, we are interested in the contour of the ROI. For the i th subject, let L_i be a $n_v \times d$ matrix with n_v landmarks representing the contour of the ROI in \mathbb{R}^d , where $d = 2$ or 3 (in this article, we only consider the contour of the planar CC on the middle-sagittal slice, that is, $d = 2$: see Figure 2a,b; $x_i \in \mathbb{R}$ be the genetic exposure (e.g., selected causal SNP); $w_i \in \mathbb{R}^q$ be clinical confounders (e.g., gender, age, and APOE $\epsilon 4$); and y_i be the neurocognitive outcome (e.g., Alzheimer’s Disease Assessment Scale Cognitive Subscale 13 score (ADAS-13)). As shown in the causal DAG (Figure 1a), our goal is to identify the causal effect of genetic exposure x_i on the neurocognitive outcome y_i implicitly through the shape mediators L_i , with adjustment for the clinical confounders w_i .

2.2 | Elastic Shape Representation

Given the landmarks L_i from the contour of the planar CC, we first derive the coordinate functions, $f_i(s) \doteq (f_{i,1}(s), f_{i,2}(s))$, with $f_{i,j}(s) : [0, 1] \rightarrow \mathbb{R}, j = 1, 2$, in the x -axis and the y -axis, respectively (see Figure 2c). Since shapes are invariant to shape-preserving transformations, for example, rotation, translation, and scaling [23], $f_i(s)$ can be derived via removing these nuisance transformations from the landmarks L_i (detailed processing steps can be found in Srivastava and Klassen [28]).

After that, we derive the square-root velocity function (SRVF) representations for the i th subject:

$$q_i(s) : [0, 1] \rightarrow \mathbb{R}^2, q_i(s) = (q_{i,1}(s), q_{i,2}(s))$$

$$q_{i,j}(s) = \dot{f}_{i,j}(s) / \sqrt{|\dot{f}_{i,j}(s)|}, j = 1, 2 \quad (1)$$

According to the results in Srivastava et al. [29], if the functions in $f_i(s)$ are absolutely continuous, the corresponding SRVF representations are square integrable, that is, $q_i(s) \in \mathbb{L}^2([0, 1], \mathbb{R}^2), i = 1, \dots, n$. Under this representation, we can further determine the optimal registration (or re-parameterization) group action for each SRVF representation, which corresponds to the following optimization problem:

$$\gamma_{i,*}(s) = \operatorname{arginf}_{\gamma(s) \in \Gamma} \|\mu(s) - (q_i \circ \gamma(s)) \sqrt{\dot{\gamma}(s)}\|, i = 1, 2, \dots, n \quad (2)$$

where $\mu(s)$ is a template, such as the mean of $\{q_i(s), i = 1, \dots, n\}$, Γ includes all possible diffeomorphisms of $[0, 1]$ that preserve the boundaries, that is, $\Gamma = \{\gamma(s) : [0, 1] \rightarrow [0, 1] | \gamma(0) = 0, \gamma(1) = 1\}$, and the composition $q_i \circ \gamma(s)$ is a re-parameterization of $q_i(s)$. Then, we can obtain the aligned SRVF representations as follows:

$$\psi(q_{i,j}(s), \gamma_{i,*}(s)) = (q_{i,j} \circ \gamma_{i,*}(s)) \sqrt{\dot{\gamma}_{i,*}(s)}, j = 1, 2, i = 1, \dots, n \quad (3)$$

and the aligned SRVF representations of CC shapes are illustrated in Figure 2d.

2.3 | Shape Mediation Analysis

To conduct the mediation analysis with elastic shape mediators, we propose a 2-layer Shape Regression Model (2SRM). First, a shape-on-scalar regression model is introduced to investigate the relationship among the elastic shape mediators $\{\psi(q_{i,j}(s), \gamma_{i,*}(s)), s \in [0, 1]\}$, the genetic exposure x_i , and the clinical confounders w_i :

$$\psi(q_{i,j}(s), \gamma_{i,*}(s)) = x_i \alpha_j(s) + w_i^\top \lambda_j(s) + \eta_{i,j}(s) + \epsilon_{i,j}(s)$$

$$s \in [0, 1], i = 1, \dots, n, j = 1, 2 \quad (4)$$

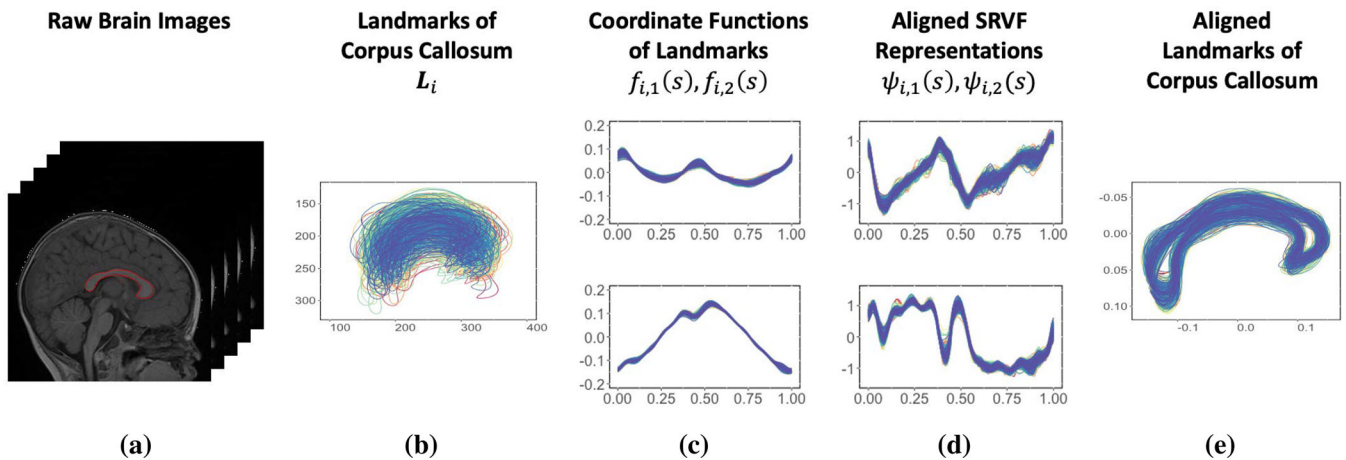


FIGURE 2 | Preprocessing procedures for shape mediator: (a) raw MRI brain images with 2D CC segmented at the middle-sagittal slice; (b) landmarks sampled along the 2D contour of CC; (c) coordinate functions of landmarks with rotation, translation, and scaling removed; (d) aligned SRVF representations of shape mediators; and (e) aligned landmarks by mapping the aligned SRVF representations back to the planar CC contour space.

where $\alpha(s) = (\alpha_1(s), \alpha_2(s))^T$ and $\lambda(s) = (\lambda_1^T(s), \lambda_2^T(s))^T$, including coefficient functions with continuous derivatives on $[0, 1]$, respectively represent the effects related to the genetic exposure and clinical confounders. The individual functions $\eta_i(s) = (\eta_{i,1}(s), \eta_{i,2}(s))^T$, independent of x_i , characterizes the subject-specific spatial variability and follows a Gaussian process (GP) with mean $\mathbf{0}$ and covariance function $\Sigma_{\eta_i}(s, t)$. The measurement error $\{\epsilon_{i,j}(s)\}_{j=1,2}$ are independent of $\{x_i, \eta_{i,j}(s)\}$ and Gaussian distributed with mean 0 and variance $\sigma_j^2(s)$, $j = 1, 2$. Second, to characterize the nonlinear relationship between the neurocognitive outcome y_i and the elastic shape mediators $\{\psi(q_{i,j}(s), \gamma_{i,*}(s))\}_{j=1}^2$, we consider a scalar-on-shape partial linear single-index model:

$$y_i = x_i \tau + \mathbf{w}_i^T \boldsymbol{\kappa} + g\left(\sum_{j=1}^2 \int_0^1 \psi(q_{i,j}(s), \gamma_{i,*}(s)) \beta_j(s) ds\right) + \delta_i, \quad i = 1, \dots, n \quad (5)$$

where $g(\cdot)$ is an unknown link function, and $\beta(s) = (\beta_1(s), \beta_2(s))^T$, including coefficient functions with continuous derivatives on $[0, 1]$, represent the effect related to the elastic shape mediators. In addition, the errors $\{\delta_i\}_{i=1}^n$ are independent and Gaussian distributed with mean 0 and constant variance σ_δ^2 . An illustration of 2SRM is presented in Figure 1b. Since the individual optimal re-parameterizations $\{\gamma_{i,*}(s)\}_{i=1}^n$ are already determined in the preprocessing step, for simplifying the notation, hereafter we denote $\psi_{i,j}(s) = \psi(q_{i,j}(s), \gamma_{i,*}(s))$ for $j = 1, 2$ and $\boldsymbol{\psi}_i(s) = (\psi_{i,1}(s), \psi_{i,2}(s))^T$ for $i = 1, \dots, n$.

2.4 | Causal Effect Estimands

Our aim is to quantify the causal effect of a genetic exposure (i.e., X) on the neurocognitive outcome (i.e., Y) mediated by the elastic shape mediator (i.e., $\boldsymbol{\Psi}$) given some clinical confounders (i.e., \mathbf{W}). Using the potential outcome framework [30, 31], we first formulate the causal estimands of interest, that is, the indirect effect and the direct effect. Let $\boldsymbol{\Psi}(s; x)$ denote the outcome of elastic shape mediators under genetic exposure x and $Y(x, \boldsymbol{\psi})$ the potential outcome of the neurocognitive outcome when the genetic exposure and elastic mediators are at the level x and $\{\boldsymbol{\psi}(s)\}_{s \in [0,1]}$ respectively. Under the framework of 2SRM, the average total causal effect (ATE) is decomposed as

$$ATE(x, x^*) = \mathbb{E}[Y(x, \boldsymbol{\psi}(s; x))] - \mathbb{E}[Y(x^*, \boldsymbol{\psi}(s; x^*))] \doteq AIE(x, x^*) + ADE(x, x^*) \quad (6)$$

where x and x^* are two distinct observations of genetic exposure (without loss of generality, we assume that $x > x^*$),

$$ADE(x, x^*) = \mathbb{E}[Y(x, \boldsymbol{\psi}(s; x^*))] - \mathbb{E}[Y(x^*, \boldsymbol{\psi}(s; x^*))] = (x - x^*) \tau \quad (7)$$

is the average direct effect (ADE) of genetic exposure on the neurocognitive outcome, and

$$AIE(x, x^*) = \mathbb{E}[Y(x, \boldsymbol{\psi}(s; x))] - \mathbb{E}[Y(x, \boldsymbol{\psi}(s; x^*))] \\ = \mathbb{E}\left[g\left(\int_0^1 \boldsymbol{\psi}(s; x)^\top \boldsymbol{\beta}(s) ds\right)\right]$$

$$-g\left(\int_0^1 \boldsymbol{\psi}(s; x^*)^\top \boldsymbol{\beta}(s) ds\right) \quad (8)$$

is the average indirect effect (AIE) of genetic exposure on the neurocognitive outcome. In particular, the AIE can be further derived as

$$AIE(x, x^*) = (x - x^*) C_{x, x^*, \beta} \sum_{j=1}^2 \int_0^1 \alpha_j(s) \beta_j(s) ds \quad (9)$$

where $C_{x, x^*, \beta} = \mathbb{E}[g(\int_0^1 \boldsymbol{\psi}(s; \tilde{x})^\top \boldsymbol{\beta}(s) ds)]$ and $\tilde{x} = tx + (1-t)x^*$ for some $t \in (0, 1)$. Here we are also interested in the spatial AIE (SAIE), the average indirect effect mediated through each grid s , which is defined as

$$SAIE(s, x, x^*) = (x - x^*) C_{x, x^*, \beta} \sum_{j=1}^2 \alpha_j(s) \beta_j(s) \quad (10)$$

which is consistent with the result in linear functional mediation analysis [4] when the nonlinear link function $g(\cdot)$ degenerates to an identity function. Furthermore, we impose the following causal assumptions to identify the ADE and AIE.

Assumption 1. There is no “exposure-outcome confounder”, that is,

$$Y(x, \boldsymbol{\psi}) \perp\!\!\!\perp X | \mathbf{W}$$

Assumption 2. There is no “exposure-mediator confounder”, that is,

$$\{\boldsymbol{\Psi}(s; x)\}_{s \in [0,1]} \perp\!\!\!\perp X | \mathbf{W}$$

Assumption 3. There is no “mediator-outcome confounder”, that is,

$$Y(x^*, \boldsymbol{\psi}) \perp\!\!\!\perp \boldsymbol{\Psi}(s; x) | X, \mathbf{W}$$

for all x, x^* and $\{\boldsymbol{\psi}(s)\}_{s \in [0,1]}$.

Assumptions 1 to 3 are extensions of the standard causal mediation assumptions [3] to the elastic shape mediators. Assumptions 1 and 2 state that the genetic exposure assignment is ignorable, as would be the case in a randomized experiment. Assumption 3 states that the neurocognitive outcomes $Y(x^*, \boldsymbol{\psi})$ are ignorable concerning the elastic shape intermediate outcomes $\boldsymbol{\Psi}(s; x)$, given $X = x$, as would be the case if subjects were randomly assigned to $\boldsymbol{\Psi}$ at both levels of X . In addition, we still need the stable unit treatment value assumption (SUTVA) [30, 32]. From Assumption 2, we have

$$\mathbb{E}[\boldsymbol{\Psi}(s; x) | \mathbf{W} = \mathbf{w}] = \mathbb{E}[\boldsymbol{\Psi}(s; x) | X = x, \mathbf{W} = \mathbf{w}] \\ = \mathbb{E}[\boldsymbol{\Psi}(s; X) | X = x, \mathbf{W} = \mathbf{w}]$$

and under Assumptions 1 to 3,

$$\mathbb{E}[Y(x, \boldsymbol{\psi}) | \mathbf{W} = \mathbf{w}] \\ = \mathbb{E}[Y(x, \boldsymbol{\psi}) | X = x, \mathbf{W} = \mathbf{w}] \\ = \mathbb{E}[Y(x, \boldsymbol{\psi}) | X = x, \{\boldsymbol{\Psi}(s; x)\}_{s \in [0,1]} = \{\boldsymbol{\psi}(s)\}_{s \in [0,1]}, \mathbf{W} = \mathbf{w}]$$

$$= \mathbb{E}[Y(X, \{\Psi(s; X)\}_{s \in [0,1]}) | X = x, \{\Psi(s; x)\}_{s \in [0,1]}]$$

$$= \{\psi(s)\}_{s \in [0,1]}, \mathbf{W} = \mathbf{w}$$

Therefore, if Assumptions 1 to 3 are valid, our 2SRM is correctly specified, and the causal estimands, that is, both ADE and AIE, can be estimated from the observed data.

2.5 | Model Estimation

For the purpose of estimating the parameters in 2SRM (4) and (5), the corresponding models can be rewritten as follows:

$$\psi_{i,j}(s_k) = \mathbf{u}_i^\top \boldsymbol{\theta}_j(s_k) + \eta_{i,j}(s_k) + \epsilon_{i,j}(s_k), \quad k = 1, \dots, n_o$$

$$i = 1, \dots, n, \quad j = 1, 2 \quad (11)$$

and

$$y_i = \mathbf{u}_i^\top \mathbf{v} + g\left(\int_0^1 \boldsymbol{\psi}_i^\top(s) \boldsymbol{\beta}(s) ds\right) + \delta_i, \quad i = 1, \dots, n \quad (12)$$

where $\mathbf{u}_i = (x_i, \mathbf{w}_i^\top)^\top$, $\boldsymbol{\theta}_j(\cdot) = (\alpha_j(\cdot), \boldsymbol{\lambda}_j^\top(\cdot))^\top$, $\boldsymbol{\beta}(\cdot) = (\beta_1^\top(\cdot), \beta_2^\top(\cdot))^\top$ and $\mathbf{v} = (\tau, \boldsymbol{\kappa}^\top)^\top$.

First, we consider the estimation procedure of the parameters of interest in (11) by using local linear kernel smoothing method [33], which can be divided into two steps [34], including (i) estimating varying coefficient functions $\{\boldsymbol{\theta}_j(\cdot)\}_{j=1}^2$ and (ii) smoothing individual functions $\{\eta_{i,j}(s)\}_{j=1}^2$.

Step 1.1: Estimating varying coefficient function $\boldsymbol{\theta}_j(\cdot)$. We employ multivariate local linear regression to derive the weighted least squares estimator of the coefficient function $\boldsymbol{\theta}_j(s)$. We apply the Taylor expansion for $\boldsymbol{\theta}_j(\cdot)$ at s as follows:

$$\boldsymbol{\theta}_j(s_k) \approx \boldsymbol{\theta}_j(s) + \dot{\boldsymbol{\theta}}_j(s)(s_k - s) = \mathbf{A}_j(s) \mathbf{z}_{H_\theta}(s_k - s)$$

where $\mathbf{A}_j(s) = [\boldsymbol{\theta}_j(s), H_\theta \dot{\boldsymbol{\theta}}_j(s)]$, $\mathbf{z}_{H_\theta}(s_k - s) = (1, H_\theta^{-1}(s_k - s))^\top$, and H_θ is a bandwidth. In practice, the generalized cross-validation (GCV) method [34, 35] is adopted to select the optimal bandwidth H_θ . In particular, we standardize all covariates to have mean zero and standard deviation one, thus, we may choose a common bandwidth H_θ for all covariates. Denote $K_{H_\theta}(s) = |H_\theta|^{-1} K(H_\theta^{-1}s)$, where $K(\cdot)$ is the kernel function and H_θ is the positive definite bandwidth diagonal matrix. Let $\mathbf{e}^{\otimes 2} = \mathbf{e}\mathbf{e}^\top$ for any vector, $\mathbf{C} \otimes \mathbf{D}$ be the Kronecker product of two matrices \mathbf{C} and \mathbf{D} , and $\text{vec}(\mathbf{A})$ be the vectorization of a matrix \mathbf{A} by vertically stacking the columns of the \mathbf{A} . We estimate $\boldsymbol{\theta}_j(s)$ via estimating $\mathbf{A}_j(s)$ by minimizing the following weighted least squares function:

$$\sum_{i=1}^n \sum_{k=1}^{n_o} [\psi_{i,j}(s_k) - \mathbf{u}_i^\top \mathbf{A}_j(s) \mathbf{z}_{H_\theta}(s_k - s)]^2 K_{H_\theta}(s_k - s)$$

which yields

$$\hat{\boldsymbol{\theta}}_j(s) = [I_{p+q} \otimes (1, \mathbf{0}_{1 \times d})] \text{vec}(\hat{\mathbf{A}}_j)$$

$$= (\mathbf{U}^\top \mathbf{U})^{-1} \mathbf{U}^\top \sum_{k=1}^{n_o} a_k(s, H_\theta) \boldsymbol{\psi}_j(s_k) \quad (13)$$

where $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_n)^\top$, $\boldsymbol{\psi}_j(s_k) = (\psi_{1,j}(s_k), \dots, \psi_{n,j}(s_k))^\top$, and

$$a_k(s, H_\theta) = [1, \mathbf{0}_{1 \times d}] \left[\sum_{k=1}^{n_o} K_{H_\theta}(s_k - s) \mathbf{z}_{H_\theta}(s_k - s) \otimes \mathbf{2} \right]^{-1}$$

$$\times K_{H_\theta}(s_k - s) \mathbf{z}_{H_\theta}(s_k - s)$$

Step 1.2: Smoothing individual functions $\eta_{i,j}(\cdot)$. We employ the local linear regression technique again to estimate all individual functions $\{\eta_{i,j}(s)\}$ by assuming certain smoothness conditions on $\eta_i(s)$ [34], that is,

$$\hat{\eta}_{i,j}(s) = \sum_{k=1}^{n_o} \tilde{K}_{H_\eta}(s_k - s) [\psi_{i,j}(s_k) - \mathbf{u}_i^\top \hat{\boldsymbol{\theta}}_j(s_k)] \quad (14)$$

where $\tilde{K}_{H_\eta}(s)$ is the empirical equivalent kernel with the optimal bandwidth H_η selected using the GCV method, that is,

$$\tilde{K}_{H_\eta}(s) = [1, 0] \left[\sum_{k=1}^{n_o} K_{H_\eta}(s) \mathbf{z}_{H_\eta}(s) \otimes \mathbf{2} \right]^{-1} K_{H_\eta}(s) \mathbf{z}_{H_\eta}(s)$$

Second, we consider the estimation procedure of the parameters of interest in (12) by the functional approximation and profile local linear regression techniques [36]. For $j = 1, 2$, we express $\boldsymbol{\beta}_j(s)$ as a linear combination of K_{β_j} -B-spline basis functions $\mathbf{Y}_j(s) \doteq (\mathbf{Y}_{1,j}(s), \dots, \mathbf{Y}_{K_{\beta_j},j}(s))^\top$, that is, $\boldsymbol{\beta}_j(s) = \sum_{k=1}^{K_{\beta_j}} b_{k,j} \mathbf{Y}_{k,j}(s)$, $j = 1, 2$, and therefore model (12) can be written as

$$y_i = \mathbf{u}_i^\top \mathbf{v} + g(\boldsymbol{\Phi}_i^\top \mathbf{b}) + \delta_i, \quad i = 1, \dots, n \quad (15)$$

where $\boldsymbol{\Phi}_i = (\zeta_{i,1,1}, \dots, \zeta_{i,K_{\beta_1},1}, \zeta_{i,1,2}, \dots, \zeta_{i,K_{\beta_2},2})^\top$ with $\zeta_{i,k,j} = \int_0^1 \psi_{i,j}(s) \mathbf{Y}_{k,j}(s) ds$ and $\mathbf{b} = (b_{1,1}, \dots, b_{K_{\beta_1},1}, b_{1,2}, \dots, b_{K_{\beta_2},2})^\top$. For identifiability we assume that the coefficient \mathbf{b} satisfies $\|\mathbf{b}\| = 1$ and the first element $b_{1,1} > 0$. For the transformed model (15), we are interested in the coefficient \mathbf{v} , the expansion coefficient \mathbf{b} , and the unknown link function $g(\cdot)$. We will estimate them in an iterative way. At the t th iteration, given the estimation of \mathbf{b} , we estimate \mathbf{v} and $g(\cdot)$ using kernel smoothing technique; given the estimation of $g(\cdot)$ and \mathbf{v} , the estimation of \mathbf{b} is updated using the profile local linear regression method. The corresponding estimation procedure is summarized in Algorithm 1.

2.6 | Estimation and Inference on Causal Estimands

First, we focus on the estimation procedures of ADE, AIE, and SAIE in (7–10). For the estimation of ADE, a plug-in estimator can be derived with the estimated parameter $\hat{\tau}$. However, this idea can not be applied to estimate AIE and SAIE due to the unknown constant $C_{x,x^*,\beta}$ involved in (9) and (10). To address this issue, we consider estimating the expectation term $\mathbb{E}[Y(x^{(a)}, \boldsymbol{\psi}(s; x^{(b)}))]$, in which two choices of the pair $(x^{(a)}, x^{(b)})$ are of interest, that is, (1, 1) and (0, 1), via resampling strategies [37]. We first fit the transformed outcome model (15) of 2SRM, conditional on exposure x , mediators $\boldsymbol{\psi}(s; x)$, and confounders \mathbf{w} , to the observed data, and derive the estimated outcome function

$$h(x, \boldsymbol{\psi}(s; x), \mathbf{w}) = \hat{\mathbb{E}}\{Y | X = x, \boldsymbol{\Psi}(s; x) = \boldsymbol{\psi}(s; x), \mathbf{W} = \mathbf{w}\}$$

$$= x \hat{\tau} + \mathbf{w}^\top \hat{\boldsymbol{\kappa}} + \hat{g}\left(\sum_{j=1}^2 \int_0^1 \boldsymbol{\psi}_j(s; x) \hat{\boldsymbol{\beta}}_j(s) ds\right) \quad (16)$$

ALGORITHM 1 | Estimation procedure for the transformed model (15).

input: data $\{u_i, y_i\}_{i=1}^n$, projection coefficients of elastic shape mediators $\{\Phi_i\}_{i=1}^n$, tolerance ϵ .

initialization: $t = 1$, $g^{(0)}(s) = s$, $\dot{g}^{(0)}(s) = 1$, and $\{v^{(0)}, b^{(0)}\} \rightarrow$ estimation of v and b through the least square method.

for $\|v^{(t)} - v^{(t-1)}\| \leq \epsilon$, $\|b^{(t)} - b^{(t-1)}\| \leq \epsilon$, and $\|g^{(t)}(\cdot) - g^{(t-1)}(\cdot)\| \leq \epsilon$ **do**

update $g^{(t-1)}(\cdot)$ and $\dot{g}^{(t-1)}(\cdot)$: estimate $g(\cdot)$ and $\dot{g}(\cdot)$ via local linear regression technique, which yields

$$\begin{aligned} & [g^{(t-1)}(\Phi_i^\top b^{(t-1)}), \dot{g}^{(t-1)}(\Phi_i^\top b^{(t-1)})] \\ &= \sum_{i=1}^n \tilde{K}_{h_g} \left((\Phi_i - \Phi_{i0})^\top b^{(t-1)} \right) [y_i - u_i^\top v^{(t-1)}] \end{aligned}$$

where $\Phi_{i0}^\top b^{(t-1)}$ is in a small neighborhood of $\Phi_i^\top b^{(t-1)}$ and $\tilde{K}_{h_g}(\cdot)$ is the empirical equivalent kernel with bandwidth h_g .

update $v^{(t)}$ and $b^{(t)}$: estimate v and b by minimizing the following profile least squares function

$$\begin{aligned} & \sum_{i=1}^n (y_i - u_i^\top v - g^{(t-1)}(\Phi_i^\top b^{(t-1)}) - \dot{g}^{(t-1)}(\Phi_i^\top b^{(t-1)})(\Phi_i^\top b \\ & - \Phi_i^\top b^{(t-1)}))^2 \end{aligned}$$

end for

output: the estimates of v , b , $g(\cdot)$, and reconstructed $\beta(s)$.

where the observed values of the confounders for each individual are assumed to be fixed and invariant to different values of the exposure x and counterfactual mediator $\psi(s; x)$. Next, we implement the resampling procedure as follows. For each individual whose observed exposure level is $X = x^{(a)}$, we randomly sample L counterfactual elastic shape mediators $\{\tilde{\Psi}_l(s; x^{(b)})\}_{l=1}^L$ from the observed group with exposure level $X = x^{(b)}$ with replacement. Then, the AIE can be estimated through the estimated expectation term:

$$\hat{\mathbb{E}}[Y(x^{(a)}, \psi(s; x^{(b)}))] = \frac{1}{n_a} \sum_{i: x_i = x^{(a)}} \frac{1}{L} \sum_{l=1}^L h(x_i, \tilde{\Psi}_l(s; x^{(b)}), \mathbf{w}_i) \quad (17)$$

where n_a is the number of subjects with observed treatment $X = x^{(a)}$. In (17), the inner average accounts for the variability in the sampled (counterfactual) mediator values while the outer average, across individuals with observed exposure level $X = x^{(a)}$, takes the empirical distribution of the confounders into account. Once we have the estimated AIE, the spatial AIE, $SAIE(s)$, can be estimated as follows. We fit the mediator model (11) of 2SRM, conditional on exposure x and confounders \mathbf{w} , to the observed data, yielding the corresponding estimated functional coefficients $\{\hat{\alpha}_j(s)\}_{j=1}^2$. Then, a plug-in estimator of $C_{x,x^*,\beta}$, $\hat{C}_{x,x^*,\beta}$, can be derived according to the definition of AIE in (9), that is, $\hat{C}_{x,x^*,\beta} = \widehat{AIE}(x, x^*) / \{(x - x^*) \sum_{j=1}^2 \int_0^1 \hat{\alpha}_j(s) \hat{\beta}_j(s) ds\}$. Thus, the estimated $SAIE(s)$ is obtained as

$$\widehat{SAIE}(s) = \hat{C}_{x,x^*,\beta} \sum_{j=1}^2 \hat{\alpha}_j(s) \hat{\beta}_j(s) \quad (18)$$

Second, we are interested in the inference procedures of ADE, AIE, and SAIE. In particular, for a given confidence level $(1 - \vartheta)$, we develop a bootstrap resampling method to construct the $(1 - \vartheta)$ CIs [38] of ADE and AIE, and the $(1 - \vartheta)$ SCB of $SAIE(s)$. Specifically, we first generate independent B bootstrap samples by randomly resampling observed subjects with replacement. Then, for each of the generated bootstrap samples, we utilize the proposed estimation procedures to get the estimated ADE sequence, $\{\widehat{ADE}^{(b)}\}_{b=1}^B$, the estimated AIE sequence, $\{\widehat{AIE}^{(b)}\}_{b=1}^B$, and the estimated SAIE sequence, $\{\widehat{SAIE}(s)^{(b)}\}_{b=1}^B$. For the ADE, its $(1 - \vartheta)$ CI is constructed as

$$\begin{aligned} & \left(\widehat{ADE} - t_{1-\vartheta/2}(n-1) \cdot \widehat{se}(\widehat{ADE}), \right. \\ & \left. \widehat{ADE} + t_{1-\vartheta/2}(n-1) \cdot \widehat{se}(\widehat{ADE}) \right) \quad (19) \end{aligned}$$

where $t_{1-\vartheta/2}(n-1)$ is the $(1 - \vartheta/2)$ quantile of Student's t distribution with $(n-1)$ degree of freedom and $\widehat{se}(\widehat{ADE})$ is the bootstrap-based estimated standard error of \widehat{ADE} . Utilizing the bias-corrected and accelerated CI, the $(1 - \vartheta)$ CI of AIE can be constructed as well. For the $(1 - \vartheta)$ SCB of $SAIE(s)$, we construct the sequence $\{\hat{R}^{(b)} = \sup_{s \in [0,1]} |\widehat{SAIE}(s) - \widehat{SAIE}(s)^{(b)}|\}_{b=1}^B$, and the $(1 - \vartheta)$ critical-value-based SCB [39] of $SAIE(s)$ is derived as

$$\left(\widehat{SAIE}(s) - \hat{q}_{1-\vartheta}, \widehat{SAIE}(s) + \hat{q}_{1-\vartheta} \right) \quad (20)$$

where $\hat{q}_{1-\vartheta}$ is the $(1 - \vartheta)$ empirical quantile of $\{\hat{R}^{(b)}\}_{b=1}^B$. The proposed bootstrap idea for constructing SCB has been used in several existing literature [34, 40–42], where the proposed method was demonstrated to outperform other SCB construction methods through both simulation studies. In particular, asymptotic theorems have been also established to lay the ground for the bootstrap-based method.

2.7 | Asymptotics of Estimation Procedure

We systematically investigate the asymptotic properties of all estimators proposed in Section 2. Specifically, Theorem 1 and Theorem 2 investigate the asymptotic properties of the estimated functional coefficients $\hat{\theta}_j$, $j = 1, 2$, the estimated link function $\hat{g}(\cdot)$, and the estimated B -spline coefficients \hat{b} , all of which provide the statistical guarantees in the estimation and inference of the causal estimands, AIE in (9) and SAIE in (10). In addition, Theorem 3 presents the asymptotic property of $\hat{\tau}$, which provides the statistical guarantee in the estimation and inference of ADE in (7). Assumptions used to facilitate the technical details and the proof and Theorems can be found in the Supporting Information.

Theorem 1. Under Assumptions (A1–A10) in the Supporting Information, the followings hold:

$$\begin{aligned} & n^{1/2} \left[(\mathbf{U}^{*\top} \mathbf{U}^*)^{-1/2} \text{vec} \left\{ \hat{\theta}_j(s) - \mathbb{E}(\hat{\theta}_j(s)) \right\} \right] \\ & j = 1, 2, s \in [0, 1] \end{aligned}$$

where $\mathbf{U}^* = (\mathbf{U}^\top \mathbf{U})^{-1} \mathbf{U}^\top$, weakly converges to a centered Gaussian process with covariance matrix $\Sigma_{\eta,jj} I_{1+q}$, in which $\Sigma_{\eta,jj}$ is the j th diagonal element in $\Sigma_\eta(s, s)$, $j = 1, 2$.

Theorem 2. Under Assumptions (A1–A10), one has

$$n^{1/2}h_g^{1/2} \left\{ \hat{g}(\Phi_{i0}^T \hat{\mathbf{b}}) - g(\Phi_{i0}^T \mathbf{b}) - 0.5u_2 \ddot{g}(\Phi_{i0}^T \mathbf{b}) h_g^2 \right\} \rightarrow_d N(0, \Psi_g),$$

where $\Psi_g = \int K^2(t)dt / f(\Phi_{i0}^T \mathbf{b})$.

Theorem 3. Under Assumptions (A1–A10), the followings hold:

$$(\hat{\tau} - \tau)^2 / (\hat{\sigma}_y^2 \hat{v}^2) \rightarrow_d \chi^2(1),$$

where $\hat{\sigma}_y^2 = \frac{1}{n} \sum_{i=1}^n \left(y_i - \hat{g}(\Phi_i^T \hat{\mathbf{b}}) + \Phi_i^T \hat{\mathbf{b}} \hat{g}(\Phi_i^T \hat{\mathbf{b}}) \right)^2$, and $\hat{v}^2 = [1, \mathbf{0}_{q+K_{\beta_1}+K_{\beta_2}}] (\mathbf{V}^T \mathbf{V})^{-1} [1, \mathbf{0}_{q+K_{\beta_1}+K_{\beta_2}}]^T$, $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n]^T$ with $\mathbf{v}_i = [\mathbf{u}_i^T, \hat{g}(\Phi_i^T \hat{\mathbf{b}}) \Phi_i^T]^T$.

3 | Simulation Studies

To examine the proposed methods, we first generated the elastic shape mediators from

$$\psi_{i,j}(s_k) = x_i \alpha_j(s_k) + \mathbf{w}_i^T \lambda_j(s_k) + \eta_{i,j}(s_k) + \epsilon_{i,j}(s_k) \quad (21)$$

$$k = 1, \dots, n_v, j = 1, 2, i = 1, \dots, n$$

$s_1 = 0 \leq s_2 \leq \dots \leq s_{n_v} = 1$, in which we independently simulated $\tilde{s}_k \sim U(0, 1)$ for $k = 2, \dots, n_v - 1$ and sorted them to obtain $\{s_k : k = 2, \dots, n_v - 1\}$. The exposure $x_i \sim Bin(1, 0.5)$, and the 3 confounders $\mathbf{w}_i \sim N(\mathbf{0}_3, \Sigma)$ where Σ is a 3×3 matrix with elements $\Sigma_{r,r'} = \rho^{|r-r'|}$ for $r, r' = 1, 2, 3$ and $\rho = 0.6$. To mimic the shape mediators extracted from the real data, we fit the model (21) to the ADNI CC contour data and adopted the estimated coefficient functions $\{\hat{\alpha}_j(s), \hat{\lambda}_j(s)\}_{j=1}^2$ as the true values. The $\eta_{i,j}(s)$ admits the Karhunen Loeve expansion as $\eta_{i,j}(s) = \sum_{k=1}^2 \xi_{i,j,k} \phi_{j,k}(s)$ with $\xi_{i,j,1} \sim N(0, 1)$, $\xi_{i,j,2} \sim N(0, 0.5)$, and the eigen functions $\{\phi_{j,k}(s)\}_{j,k}$ were determined via applying the functional PCA [43] to the estimated individual functions $\{\hat{\eta}_{i,j}\}_{i,j}$. In addition, the random noises $\{\epsilon_{i,j}(s)\}_{i,j}$ were grid-wisely independently generated from $N(0, 0.3^2)$. Without loss of generality, we assume that all the rotation, translation, and scaling transformations are already removed from the coordinate functions. Then, according to Section 2.2, the individual observed SRVF representations were constructed through stochastic warping functions, that is, $q_{i,j}(s) = (\psi_{i,j} \circ \gamma_i(s)) \sqrt{\dot{\gamma}_i(s)}$, where $\gamma_i(s) = (e^{a_i s} - 1) / (e^{a_i} - 1)$ with $a_i \sim N(0, 1)$, $i = 1, \dots, n$ [28]. Next, we took the generated elastic shape mediators $\boldsymbol{\psi}_i(s) = (\psi_{i,1}(s), \psi_{i,2}(s))^T$ from (21) as the predictors and generated the clinical outcome y_i from

$$y_i = x_i \tau + \mathbf{w}_i^T \boldsymbol{\kappa} + g\left(\int_0^1 \boldsymbol{\psi}_i(s)^T \boldsymbol{\beta}(s) ds\right) + \delta_i, \quad i = 1, \dots, n \quad (22)$$

where $g(u) = 4((u - 1)/\sqrt{2})^2$, $\tau = 1$, $\boldsymbol{\kappa} = (0.3, 0.3, 0.5)^T$, and δ_i are measurement errors independently following $N(0, 0.1^2)$. Each element in the coefficient function $\boldsymbol{\beta}(s) = (\beta_1(s), \beta_2(s))^T$ is specified as $\beta_j(s) = \sum_{k=1}^3 b_{j,k} \varphi_k(s)$, $j = 1, 2$, where $(b_{1,1}, \dots, b_{1,3}, b_{2,1}, \dots, b_{2,3})^T = \tilde{\mathbf{b}} / \|\tilde{\mathbf{b}}\|$ with elements in $\tilde{\mathbf{b}}$ independently following the uniform distribution $U(0, 1)$ and $\{\varphi_k(s)\}_{k=1}^3$ are B-spline basis function. Without special saying, we set $n = 500$, $n_v = 50$, and 100 simulation datasets were generated. We extracted the registered elastic shape representations and fitted the models (21) and (22) to the simulated data set. Table 1 presents the mean integrated absolute error (MIAE) and mean integrated square error (MISE) of estimated functions and mean absolute error (MAE) and root mean square error (RMSE) of estimated parameters. All four metrics are of small values, indicating our method's great performance in model estimation.

Next, we consider assessing the estimation performance of our method for the causal estimands, including ATE (6), ADE (7), and AIE (9). We included two other competing methods, that is, the linear functional mediation analysis (LFMA) and the multiple mediation analysis (MMA) [4, 44], for comparison. In the LFMA framework, since only observed functional mediators are taken into account, we considered the mediator model and the outcome model as follows:

$$\psi_{i,j}(s) = x_i \alpha_j(s) + \mathbf{w}_i^T \lambda_j(s) + \eta_{i,j}(s) + \epsilon_{i,j}(s), \quad j = 1, 2 \quad (23)$$

$$y_i = x_i \tau + \mathbf{w}_i^T \boldsymbol{\kappa} + \sum_{j=1}^2 \int_0^1 \psi_{i,j}(s) \beta_j(s) ds + \delta_i, \quad i = 1, \dots, n \quad (24)$$

which is a linear version of our model, that is, identity link function $g(\cdot)$, with registered elastic shape mediators. In the MMA framework, it treats the observed mediators as multivariate variables instead of functional representations. Thus, we consider the mediation analysis model with multiple mediators, that is,

$$\boldsymbol{\psi}_i = \boldsymbol{\alpha} x_i + \boldsymbol{\Lambda} \mathbf{w}_i + \boldsymbol{\epsilon}_i \quad (25)$$

$$y_i = x_i \tau + \mathbf{w}_i^T \boldsymbol{\kappa} + \boldsymbol{\psi}_i^T \boldsymbol{\beta} + \delta_i, \quad i = 1, \dots, n \quad (26)$$

where $\boldsymbol{\psi}_i = (\psi_{i,1}(s_1), \dots, \psi_{i,1}(s_{n_v}), \psi_{i,2}(s_1), \dots, \psi_{i,2}(s_{n_v}))^T$, $\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \Sigma_\epsilon)$ and $\delta_i \sim N(0, \sigma_\delta^2)$ are mutually independent, $i = 1, \dots, n$. For implementation, the Matlab code MVCM and the R packages, refund, PLSimCpp, and mma [34, 36, 44], were used to estimate coefficient functions and parameters. To estimate the causal estimands, we set the $L = 100$ for the number of random draws of mediators and $B = 500$ for the number of bootstrap replications. The corresponding estimation results for each competing method are reported in Table 2, where the computational time elapsed on a workstation with 28 cores and 64G memory is reported as well. It can be found that our

TABLE 1 | Estimation results based on our method from 100 replications.

| | $\boldsymbol{\alpha}(s)$ | $\boldsymbol{\lambda}(s)$ | $g(\cdot)$ | $\boldsymbol{\beta}(s)$ | | τ | $\boldsymbol{\kappa}_1$ | $\boldsymbol{\kappa}_2$ | $\boldsymbol{\kappa}_3$ |
|------|--------------------------|---------------------------|------------|-------------------------|------|--------|-------------------------|-------------------------|-------------------------|
| MIAE | 0.0731 | 0.0583 | 0.0624 | 0.0239 | MAE | 0.0737 | 0.0724 | 0.0721 | 0.0689 |
| MISE | 0.0084 | 0.0055 | 0.0101 | 0.0008 | RMSE | 0.3274 | 0.3413 | 0.3495 | 0.3230 |

Note: MIAE is the mean integrated absolute error, MISE is the mean integrated square error, MAE is the mean absolute error, and RMSE is the root mean square error.

TABLE 2 | Estimation results of the causal estimands based on three competing methods.

| | 2SRM | | | LFMA | | | MMA | | |
|----------------|---------|--------|---------|--------|---------|---------|---------|---------|---------|
| | ATE | ADE | AIE | ATE | ADE | AIE | ATE | ADE | AIE |
| Mean | -0.5854 | 1.0000 | -1.5855 | 4.3057 | -6.9573 | 11.2630 | -1.0049 | 0.2123 | -1.2172 |
| SD | 0.3025 | 0.1366 | 0.3033 | 2.2192 | 13.1947 | 14.5380 | 0.6861 | 0.3857 | 0.5840 |
| Bias | 0.0820 | 0.0119 | 0.0701 | 4.9731 | -7.9454 | 12.9186 | -0.3374 | -0.7758 | 0.4384 |
| Time (seconds) | 65 | | | 25 | | | 17 | | |

Note: SD is standard deviation, bias is estimation bias, and time is the computational time (in seconds).

TABLE 3 | Empirical coverage probabilities of (i) $(1 - \vartheta)$ CIs of ADE and AIE and (ii) $(1 - \vartheta)$ SCB of SAIE.

| ϑ | $n_o = 25$ | | | $n_o = 50$ | | |
|-------------|------------|-------|-------|------------|-------|-------|
| | ADE | AIE | SAIE | ADE | AIE | SAIE |
| 0.05 | 0.980 | 0.984 | 0.982 | 0.962 | 0.960 | 0.938 |
| 0.01 | 0.988 | 0.994 | 0.996 | 0.986 | 0.994 | 1.000 |

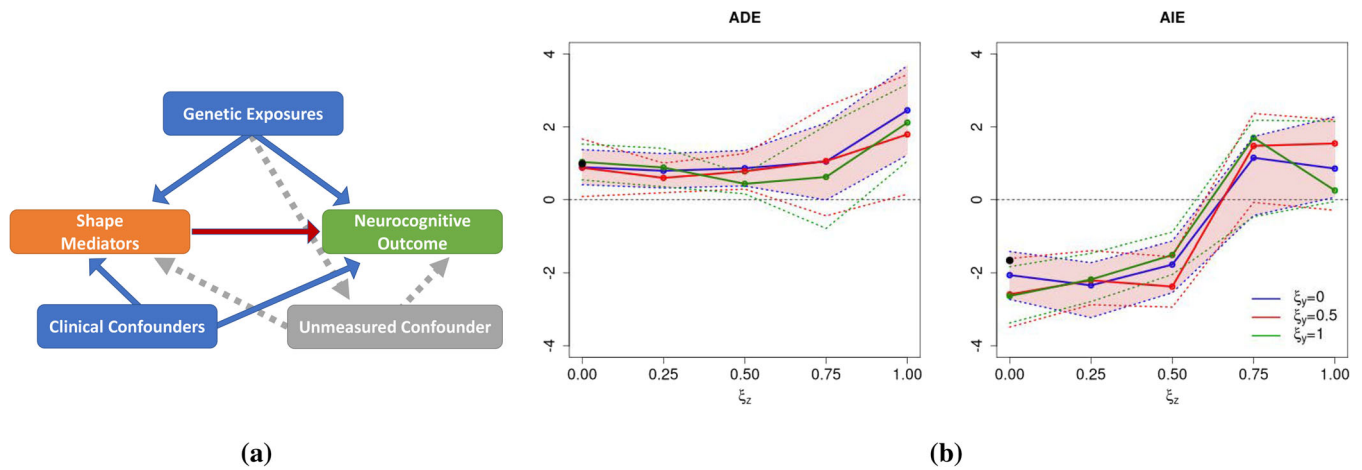


FIGURE 3 | (a) DAG of examples of violation to sequential ignorability. The arrows among genetic exposures, the elastic shape mediator, clinical confounders, and the neurocognitive outcome represent a causal relationship. The solid and broken lines represent measured and unmeasured relationships, respectively. (b) Sensitivity analysis for the true and estimated ADE and AIE by the sensitivity parameters ξ_z and ξ_y . The black points are the true ADE and AIE where $\xi_z = 0$. The color-coded solid lines are the estimated ADEs and AIEs and the color-coded broken lines are the 95% CIs. The blue, red, and green color-coded lines denote $\xi_y = 0, 0.50, 1$, respectively. The red shaded area shows the changes of CIs at each grid of ξ_z where $\xi_y = 0$.

proposed method outperforms both competing methods in terms of both the mean and variation of estimated causal estimands. In particular, compared to both two traditional methods, LFMA and MMA, our proposed 2SRM is successful in characterizing the true shape representations and the potential nonlinear relationship between neurocognitive outcomes and shape mediators.

We also evaluated the coverage probabilities of $(1 - \vartheta)$ CIs for ADE and AIE, and the $(1 - \vartheta)$ SCB for SAIE with 1000 bootstrap samples per replication. The 500 simulation datasets were generated from models (21) and (22) by randomly resampling observations with replacement for the number of grids $n_o = 25, 50$. Table 3 displays the empirical coverage probabilities for $\vartheta = 0.05, 0.01$, and the probabilities are getting closer to $1 - \vartheta$ when the number of grids n_o is getting larger.

Finally, we evaluated the effect of potential violations to sequential ignorability assumption using sensitivity analysis [45]. Sequential ignorability suggests there is no unmeasured exposure-mediator confounder, exposure-outcome confounder, and mediator-outcome confounder. Figure 3a explains cases of violation of the assumption by DAG compared with the valid assumption in Figure 1a. The arrows among genetic exposure, the elastic shape mediators, clinical confounders, and the neurocognitive outcome represent a causal relationship. The solid and broken lines represent measured and unmeasured relationships, respectively. The original models (21) and (22) expanded to (27) and (28) with an unmeasured confounder c_i and pre-specified sensitivity parameters ξ_z and ξ_y for the model-based approach:

$$\psi_{i,j}(s_k) = x_i \alpha_j(s_k) + \mathbf{w}_i^\top \lambda_j(s_k) + c_i \xi_z + \eta_{i,j}(s_k) + \epsilon_{i,j}(s_k) \quad (27)$$

$$j = 1, 2, k = 1, \dots, n_o$$

$$y_i = x_i\tau + \mathbf{w}_i^\top \boldsymbol{\kappa} + c_i\xi_y + g\left(\int_0^1 \boldsymbol{\psi}_i(s)^\top \boldsymbol{\beta}(s)ds\right) + \delta_i, \quad i = 1, \dots, n \quad (28)$$

where ξ_z and ξ_y measure the correlation between the unmeasured confounder and mediator and outcome, respectively.

Given the unobserved confounder $c_i \sim N(0, 1^2)$ and a grid of values of the sensitivity parameters $(\xi_z, \xi_y) \in \{0, 0.25, 0.50, 0.75, 1\} \times \{0, 0.50, 1\}$, we generated simulation datasets from the models (27) and (28) to estimate ADE and AIE for all combinations of (ξ_z, ξ_y) . Figure 3b presents the true and estimated ADE and AIE on the specified grid of (ξ_z, ξ_y) , where the black points indicate the true ADE and AIE where $\xi_z = 0$. The color-coded solid lines are the estimated ADEs and AIEs and the color-coded broken lines are the corresponding 95% CIs. The blue, red, and green color-coded lines denote $\xi_y = 0, 0.50, 1$, respectively, and the red shaded area displays the CIs where $\xi_y = 0$. When both ξ_z and ξ_y are zero, sequential ignorability is valid and there is no unmeasured confounder. When at least one in (ξ_z, ξ_y) is nonzero, sequential ignorability is invalid so that we can observe the change of estimates and CIs for ADE and AIE in case of potential violation to the assumption. The effect of ξ_y for ADE is relatively inappreciable compared with AIE at all grids of ξ_z . The 95% CIs become close to 0 so that both ADE and AIE have an insignificant trend when $\xi_z \geq 0.75$. The results of the sensitivity analysis imply our simulation has sensitive parameters ξ_z to the correlation between the unmeasured confounder and mediator for both ADE and AIE, and ξ_y to the correlation between the unmeasured confounder and outcome for AIE.

4 | Real Data Analysis

4.1 | Data Description

Data used in the preparation of this article were obtained from the ADNI database (<http://adni.loni.usc.edu>). The ADNI was launched in 2003 by the National Institute on Aging, National Institute of Biomedical Imaging and Bioengineering, Food and Drug Administration, private pharmaceutical companies and non-profit organizations as a \$60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians in developing new treatments and monitoring their effectiveness, as well as lessening the time and cost of clinical trials. The principal investigator of this initiative is Michael W. Weiner, MD, at the VA Medical Center and University of California, San Francisco. ADNI is the result of efforts of many coinvestigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The goal was to recruit 800 subjects, but the initial study (ADNI-1) has been followed by ADNI-GO and ADNI-2. To date, these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in

the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow-up duration of each group is specified in the protocols for ADNI-1, ADNI-2, and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see <http://www.adni-info.org>.

4.2 | Data Processing

We consider $n = 700$ MRI scans from both normal controls (NC) and individuals with MCI or AD in the ADNI-1 study. The scans, which were performed on a variety of 1.5 Tesla MRI scanners with protocols individualized for each scanner, include standard T1-weighted images obtained using volumetric 3-dimensional sagittal MPRAGE or equivalent protocols with varying resolutions. To obtain the contour of planar CC, we used *FreeSurfer* [46] to process each T1-weighted MRI, including motion correction, non-parametric non-uniform intensity normalization, affine transform to the MNI305 atlas, intensity normalization, skull-stripping, and automatic subcortical segmentation. Some quality control procedures were conducted on each output image data, consisting of two steps: (i) pial surfaces correction to remove any non-brain tissue and (ii) white matter surfaces correction to include any missing white matter and remove any errant grey matter. Then, through the package *CCSeg* [47], each T1-weighted MRI image and tissue segmentation results were used to extract the planar CC contour data on the midsagittal slice, which contains 100 landmarks (Figures 2a,b). Given the coordinate functions of landmarks (Figure 2c), we extracted the aligned SRVF shape representation. The resulting aligned SRVFs $\boldsymbol{\psi}_i(s)$ as shown in Figure 2d. In addition, the aligned CC landmarks were also derived by mapping the aligned SRVF representations back to the planar CC contour space with the optimal warping function (Figure 2e).

Besides the planar CC shape data, we also considered the subjects' genotype variables acquired by using the Human 610-Quad BeadChip (Illumina, Inc., San Diego, CA) in the ADNI database, which includes 620,901 SNPs. Quality control procedures included (i) call rate check per subject and SNP marker, (ii) gender check, (iii) sibling pair identification, (iv) the Hardy-Weinberg equilibrium test, (v) marker removal by MAF, and (vi) population stratification. The second line preprocessing steps included the removal of SNPs with (i) more than 5% missing values, (ii) MAF smaller than 5%, and (iii) Hardy-Weinberg equilibrium $p < 10^{-6}$. The remaining missing genotype variables were imputed as the modal value. After the quality control procedures, 700 subjects and 501,510 SNPs remained in the final data analysis.

4.3 | Data Analysis

Based on the aligned SRVFs, we applied the functional genome-wide association analysis (FGWAS) [40] to detect potential causal SNPs with significant p-values. After the correction for multiple testing by the wild bootstrap resampling techniques [40], one SNP, rs11719939, from chromosome 3, which is close to the AD-related high-risk gene, *ATP2B2* [48, 49], was detected with significant p-value (4.60E-07) and thus treated as the genetic exposure in our study. Typically a SNP consists of a major

allele (M) and a minor allele (m), and the genotype can be a major allele homozygote (MM), a heterozygote (Mm), or a minor allele homozygote (mm) [50]. Minor alleles are more likely to be risk alleles in the published GWASs on complex diseases, and AD (late onset) showed the highest risk allele frequencies from NHGRI-EBI Catalog data [51]. When we dichotomized SNP (0 = MM, 1 = Mm, 2 = mm) as minor alleles are included (1 = Mm and mm, 0 = MM), the odds ratio (OR) of rs11719939 for AD & MCI is 1.164. In addition, the confounding factors w_i include gender, age, handedness, APOE-4, and education level. All these variables are commonly considered as confounding factors in imaging genetics studies [40]. We also included the top 2 PC scores of all the SNPs as confounders for population stratification, which explained over 90% of the variance of the SNP data. The ADAS-13 score was adopted as the neurocognitive outcome y_i . Table 4 summarizes the demographic, clinical, and genetic information of all the 700 subjects.

First, we treated the relative volumetric data, that is, volumes of 5 subregions of CC (anterior, mid anterior, central, mid posterior, and posterior) divided by the total intracranial volume (ICV), as the mediators and considered the MMA approach for mediation analysis. All these volumetric data were calculated via the pre-processing software *FreeSurfer*. The genetic exposure, confounding factors, and the neurocognitive outcome are the same as we defined above. The estimated ADE is 0.166 and the estimated AIE is 0.386. Using 5000 bootstrap replications, the 95% CIs of the ADE and the overall AIE are, respectively, (−0.433, 0.815) and (−0.110, 0.926), while the 95% CIs of the marginal AIE for each of the five mediators are (−0.128, 0.135), (−0.057, 0.305), (−0.027,

0.388), (−0.041, 0.447), and (−0.098, 0.239), where neither significant direct nor (overall and marginal) indirect causal effects are detected when taking the non-shape information (volumetric data) as mediators.

Next, we considered the elastic shape representations as mediators and implemented our proposed 2SRM for mediation analysis. In particular, to investigate how the number of PCs and the registration step in elastic shape representations affect the mediation analysis, we considered three scenarios, that is, (i) registered shape mediators & top 2 PC scores were involved; (ii) unregistered shape mediators & top 2 PC scores were involved; (iii) registered shape mediators & top 5 PC scores were involved. For all the three scenarios, we chose 2 *B*-spline basis functions for each of the two coefficient functions $\{\beta_j(s)\}_{j=1}^2$. Table 5 shows the estimation and inference results of causal estimands in three scenarios. Some findings can be concluded as follows. First, both significant ADE and AIE are detected in Scenario (i), these significant findings by our 2SRM are consistent with existing literature. The SNP rs11719939 is related to the gene ATP2B2, which presses out calcium from the cytosol into the extracellular space [52]. Calcium in its ionized form (Ca^{2+}) plays a vital role in neuronal development, differentiation, and neuronal cell death, so that dysregulation of Ca^{2+} signals is associated with AD [49]. Compared to age-matched control brains, the Ca^{2+} sensitivity and activity of the plasma membrane Ca^{2+} -ATPases are reduced in the cortex of AD brains [48]. Second, compared to results in Scenario (i), no significant ADE is detected in Scenario (ii), which indicates that the elastic shape registration step in Section 2.2 is vital in shape mediation analysis. Third, compared to results in Scenario (i), no significant ADE is detected when the top 5 PC scores are involved

TABLE 4 | Demographic, clinical, and genetic information about ADNI data: SNP, gender, age (in years), handedness, APOE-4, education level (in years), and range of ADAS-13.

| Variable | AD | MCI | NC | Total |
|-------------------|--------------|--------------|--------------|--------------|
| SNP (Mm&mm/MM) | 74/82 | 150/196 | 81/117 | 305/395 |
| Gender (F/M) | 72/84 | 123/223 | 90/108 | 285/415 |
| Age (years) | [55.1, 90.9] | [54.4, 89.3] | [62.0, 89.6] | [54.4, 90.9] |
| Handiness (R/L) | 146/10 | 315/31 | 184/14 | 645/55 |
| APOe-4 (0/1/2) | 52/73/31 | 154/149/43 | 145/49/4 | 351/271/78 |
| Education (years) | [6.0, 20.0] | [6.0, 20.0] | [6.0, 20.0] | [6.0, 20.0] |
| ADAS-13 | [12.7, 50.0] | [3.0, 39.7] | [1.0, 19.7] | [1.0, 50.0] |

TABLE 5 | Estimation and inference results of ADE and AIE based on three methods, that is, 2SRM, LFMA, and MMA.

| Model | Scenario | ADE | | AIE | |
|-------|----------|-------|-----------------------------|-----------|-------------------------------------|
| | | Est | 95% CI | Est | 95% CI |
| 2SRM | (i) | 1.283 | (0.013, 2.553) ^a | 3.025e−02 | (9.301e−07, 4.636e−01) ^a |
| 2SRM | (ii) | 0.702 | (−0.489, 1.894) | 2.956e−02 | (7.606e−06, 3.436e−01) ^a |
| 2SRM | (iii) | 0.944 | (−0.370, 2.259) | 2.987e−07 | (8.511e−07, 4.952e−01) ^a |
| LFMA | | 0.538 | (−0.456, 1.531) | 5.384e−01 | (1.781e−01, 8.678e−01) ^a |
| MMA | | 0.136 | (−0.603, 0.701) | 1.687e−01 | (−4.919e−01, 1.574e+00) |

Note: The three scenarios refer to our 2SRM with (i) registered shape mediators and top 2 PC scores involved, (ii) unregistered shape mediators and top 2 PC scores involved, and (iii) registered shape mediators and top 5 PC scores involved.

^aDenotes the significant causal estimands regarding 95% CIs.

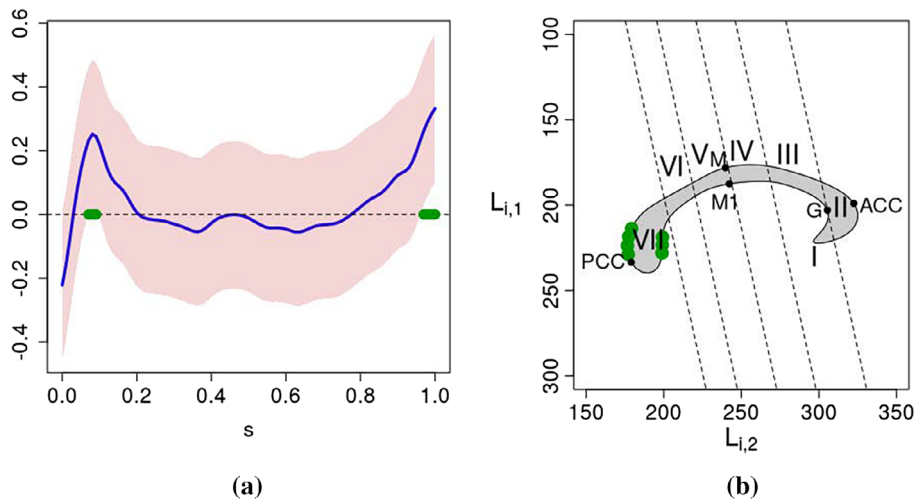


FIGURE 4 | Estimation and inference results of SAIE from 2SRM Scenario (i): (a) the estimation (blue line) and the corresponding 95% SCB (red shaded area) of $SAIE(s)$. The subregion with significant SAIE is highlighted in green; (b) the CC is parcellated by the seven regional subdivisions: (I) rostrum, (II) genu, (III) rostral body, (IV) anterior midbody, (V) posterior midbody, (VI) isthmus, and (VII) splenium. The significant subregion (green points) corresponds to the CC splenium after mapping back to the planar CC contour space.

in Scenario (iii). The possible reason is that, when more PCs are involved in the mediation analysis, potential over-fitting issues might exist for the population stratification adjustment and the genetic effect of top SNPs would be weakened. In comparison, we also conducted the mediation analysis using LFMA and MMA. The registered shape representations were adopted as functional mediators in the LFMA framework while treated as multivariate mediators in the MMA framework. The 95% CIs for ADE using LFMA and MMA from 5000 and 1000 bootstrap samples are $(-0.456, 1.531)$ and $(-0.603, 0.701)$, respectively, where no significant ADE was detected in either of these competing methods.

Finally, we investigated the SCB of the SAIE from 2SRM Scenario (i) using the bootstrap method to determine whether the SAIE is significantly different from 0 in some grid points within the interval $[0, 1]$. The 95% SCB for $SAIE(s)$ is presented in Figure 4a (red shaded area), which implies that the spatial average indirect causal effect is significantly nonzero for $s \in [8, 10], [97, 100]$ of the 100 grid points (highlighted in green). Furthermore, we mapped the inference result to the planar CC contour space. The significant subregion (green points in Figure 4a) corresponds to the CC splenium (subregion VII in Figure 4b), one of the seven regional subdivisions according to the segmentation results [18]. Our inference result on the SAIE is consistent with the existing literature, where significant atrophy of the CC splenium has been detected among the AD groups [21] and mildly demented AD subjects [22] compared with healthy controls.

5 | Discussion and Conclusion

In this study, we have proposed a novel shape mediation analysis framework to explore the causal relationships between genetic exposure and clinical outcomes, whether mediated or unmediated by shape-related factors while accounting for potential confounding variables. Within our framework, we have extracted the SRVF-based elastic shape representations and introduced the 2SRM to characterize the relationships among

neurocognitive outcomes, elastic shape mediators, genetic exposure, and clinical confounders. Both estimation and inference procedures have been established for unknown parameters along with the corresponding causal estimands. The asymptotic properties of estimated quantities have been also investigated. Both simulated studies and real-data analyses have shown the superior performance of our proposed method in terms of estimation accuracy and robustness when compared to existing approaches for estimating causal estimands. There are a couple of issues in our proposed method, both of which will be our future research topics. First, the linear assumption in the model (11) may not hold in practice. Thus, it is meaningful to extend it to a nonlinear version, such as the dynamic interaction effects $\alpha_j(s, \mathbf{w}_i^\top \lambda_j)$ and the partial linear single-index model $\psi_{i,j}(s) = x_i \alpha_j(s) + f(\mathbf{w}_i^\top \lambda_j(s)) + \eta_{i,j}(s) + \epsilon_{i,j}(s)$ with an unknown link function f , which will be more feasible for real neuroimaging datasets. Second, it is of great importance to extend our proposed model to 3D shape mediators. Although we can extract the elastic shape representations from the 3D objects for mediation analysis, the inverse of the elastic shape transformation back to the coordinate functional space is not unique, which will bring challenges in local inference of mediation analysis. We will put it as our future research topic. Furthermore, both simulation studies and real data analysis have demonstrated the important role of the elastic registration step. However, misalignment issues may exist in the current registration step, which would lead to opposite findings in the downstream mediation analysis. To handle this challenge, we can consider the strategy that conducts the registration and regression analysis simultaneously [20, 53], which will be our other future research direction.

Acknowledgments

C. Huang gratefully acknowledges financial support from the National Science Foundation Division of Mathematical Sciences (Grant No. DMS-1953087). X. Zhou gratefully acknowledges support from the

National Natural Science Foundation of China (Grant No. 12171242). J. Wang gratefully acknowledges support from the National Natural Science Foundation of China (12271255) and Qinglan Project of Jiangsu Province. The authors are grateful to the Editor, Associate Editor, and anonymous reviewers for their feedback which helped to improve the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. S. G. Mueller, M. W. Weiner, L. J. Thal, et al., "The Alzheimer's Disease Neuroimaging Initiative," *Neuroimaging Clinics of North America* 15, no. 4 (2005): 869–877.
2. C. Sudlow, J. Gallacher, N. Allen, et al., "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age," *PLoS Medicine* 12, no. 3 (2015): e1001779.
3. T. J. VanderWeele, *Explanation in Causal Inference: Methods for Mediation and Interaction* (New York, NY: Oxford University Press, 2015).
4. M. A. Lindquist, "Functional Causal Mediation Analysis With an Application to Brain Connectivity," *Journal of the American Statistical Association* 107, no. 500 (2012): 1297–1309.
5. K. J. Preacher, "Advances in Mediation Analysis: A Survey and Synthesis of New Developments," *Annual Review of Psychology* 66, no. 1 (2015): 825–852.
6. O. Y. Chén, H. Cao, H. Phan, et al., "Identifying Neural Signatures Mediating Behavioral Symptoms and Psychosis Onset: High-Dimensional Whole Brain Functional Mediation Analysis," *NeuroImage* 226 (2021): 117508.
7. J. J. Rijnhart, S. J. Lamp, M. J. Valente, D. P. MacKinnon, J. W. Twisk, and M. W. Heymans, "Mediation Analysis Methods Used in Observational Research: A Scoping Review and Recommendations," *BMC Medical Research Methodology* 21, no. 1 (2021): 226.
8. J. M. Robins and S. Greenland, "Identifiability and Exchangeability for Direct and Indirect Effects," *Epidemiology* 3, no. 2 (1992): 143–155.
9. J. M. Albert, "Mediation Analysis via Potential Outcomes Models," *Statistics in Medicine* 27, no. 8 (2008): 1282–1304.
10. J. Pearl, "Interpretation and Identification of Causal Mediation," *Psychological Methods* 19, no. 4 (2014): 459–481.
11. Y. T. Huang and W. C. Pan, "Hypothesis Test of Mediation Effect in Causal Mediation Model With High-Dimensional Continuous Mediators," *Biometrics* 72, no. 2 (2016): 402–413.
12. M. B. Sohn and H. Li, "Compositional Mediation Analysis for Microbiome Studies," *Annals of Applied Statistics* 13, no. 1 (2019): 661–681.
13. R. R. Zhou, L. Wang, and S. D. Zhao, "Estimation and Inference for the Indirect Effect in High-Dimensional Linear Mediation Models," *Biometrika* 107, no. 3 (2020): 573–589.
14. B. Caffo, S. Chen, W. Stewart, et al., "Are Brain Volumes Based on Magnetic Resonance Imaging Mediators of the Associations of Cumulative Lead Dose With Cognitive Function?," *American Journal of Epidemiology* 167, no. 4 (2008): 429–437.
15. Y. Zhao and L. Li, "Multimodal Data Integration via Mediation Analysis With High-Dimensional Exposures and Mediators," *Human Brain Mapping* 43, no. 8 (2022): 2519–2533.
16. O. Y. Chén, C. Crainiceanu, E. L. Ogburn, B. S. Caffo, T. D. Wager, and M. A. Lindquist, "High-Dimensional Multivariate Mediation With Application to Neuroimaging Data," *Biostatistics* 19, no. 2 (2018): 121–136.
17. Y. Zhao and X. Luo, "Granger Mediation Analysis of Multiple Time Series With an Application to Functional Magnetic Resonance Imaging," *Biometrics* 75, no. 3 (2019): 788–798.
18. S. F. Witelson, "Hand and Sex Differences in the Isthmus and Genu of the Human corpus Callosum: A Postmortem Morphological Study," *Brain* 112, no. 3 (1989): 799–835.
19. E. Cornea, H. Zhu, P. Kim, and J. G. Ibrahim, "Regression Models on Riemannian Symmetric Spaces," *Journal of the Royal Statistical Society Series B (Statistical Methodology)* 79, no. 2 (2017): 463–482.
20. C. Huang, A. Srivastava, and R. Liu, "Geo-FARM: Geodesic Factor Regression Model for Misaligned Pre-Shape Responses in Statistical Shape Analysis," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition IEEE Computer Society* (Nashville, TN: IEEE, 2021), 11496–11505.
21. P. J. Wang, A. J. Saykin, L. A. Flashman, et al., "Regionally Specific Atrophy of the corpus Callosum in AD, MCI and Cognitive Complaints," *Neurobiology of Aging* 27, no. 11 (2006): 1613–1617.
22. I. K. Lyoo, A. Satlin, C. K. Lee, and P. F. Renshaw, "Regional Atrophy of the corpus Callosum in Subjects With Alzheimer's Disease and Multi-Infarct Dementia," *Psychiatry Research: Neuroimaging* 74, no. 2 (1997): 63–72.
23. F. L. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology* (Cambridge: Cambridge University Press, 1991).
24. C. G. Small, *The Statistical Theory of Shape* (New York, NY: Springer, 1996).
25. D. G. Kendall, D. Barden, T. K. Carne, and H. Le, *Shape and Shape Theory* (Chichester: Wiley, 1999).
26. X. Luo, L. Zhu, and H. Zhu, "Single-Index Varying Coefficient Model for Functional Responses," *Biometrics* 72, no. 4 (2016): 1275–1284.
27. J. Li, C. Huang, and H. Zhu, "A Functional Varying-Coefficient Single-Index Model for Functional Response Data," *Journal of the American Statistical Association* 112, no. 519 (2017): 1169–1181.
28. A. Srivastava and E. Klassen, *Functional and Shape Data Analysis* (New York, NY: Springer, 2016).
29. A. Srivastava, W. Wu, S. Kurtek, E. Klassen, and J. S. Marron, "Registration of Functional Data Using Fisher-Rao Metric," arXiv:1103.3817v2 [math.ST] (2011).
30. D. B. Rubin, "Bayesian Inference for Causal Effects: The Role of Randomization," *Annals of Statistics* 6, no. 1 (1978): 34–58.
31. D. B. Rubin, "Causal Inference Using Potential Outcomes: Design, Modeling, Decisions," *Journal of the American Statistical Association* 100, no. 469 (2005): 322–331.
32. D. B. Rubin, "Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment," *Journal of the American Statistical Association* 75, no. 371 (1980): 591–593.
33. J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications* (London, UK: Chapman and Hall, 1996).
34. H. Zhu, R. Li, and L. Kong, "Multivariate Varying Coefficient Model for Functional Responses," *Annals of Statistics* 40, no. 5 (2012): 2634–2666.
35. J. T. Zhang and J. Chen, "Statistical Inferences for Functional Data," *Annals of Statistics* 35, no. 3 (2007): 1052–1079.
36. H. Liang, X. Liu, R. Li, and C. L. Tsai, "Estimation and Testing for Partially Linear Single-Index Models," *Annals of Statistics* 38, no. 6 (2010): 3811–3836.

37. W. W. Loh, B. Moerkerke, T. Loeys, and S. Vansteelandt, "Nonlinear Mediation Analysis With High-Dimensional Mediators Whose Causal Structure Is Unknown," *Biometrics* 78, no. 1 (2022): 46–59.
38. K. Jung, J. Lee, V. Gupta, and G. Cho, "Comparison of Bootstrap Confidence Interval Methods for GSCA Using a Monte Carlo Simulation," *Frontiers in Psychology* 10 (2019): 2215.
39. J. L. Montiel Olea and M. Plagborg-Møller, "Simultaneous Confidence Bands: Theory, Implementation, and an Application to SVARs," *Journal of Applied Econometrics* 34, no. 1 (2019): 1–17.
40. C. Huang, P. Thompson, Y. Wang, et al., "FGWAS: Functional Genome Wide Association Analysis," *NeuroImage* 159 (2017): 107–121.
41. H. Zhu, J. G. Ibrahim, N. Tang, et al., "A Statistical Analysis of Brain Morphology Using Wild Bootstrapping," *IEEE Transactions on Medical Imaging* 26, no. 7 (2007): 954–966.
42. H. Zhu, K. Chen, X. Luo, Y. Yuan, and J. Wang, "FMEM: Functional Mixed Effects Models for Longitudinal Functional Responses," *Statistica Sinica* 29, no. 4 (2019): 2007–2033.
43. F. Yao, "Functional Principal Component Analysis for Longitudinal and Survival Data," *Statistica Sinica* 17 (2007): 965–983.
44. Q. Yu and B. Li, "MMA: An R Package for Mediation Analysis With Multiple Mediators," *Journal of the Operational Research Society* 5, no. 1 (2017): 11.
45. S. Zeng, E. C. Lange, E. A. Archie, F. A. Campos, S. C. Alberts, and F. Li, "A Causal Mediation Model for Longitudinal Mediators and Survival Outcomes With an Application to Animal Behavior," *Journal of Agricultural, Biological and Environmental Statistics* 28, no. 2 (2023): 197–218.
46. B. Fischl, "FreeSurfer," *NeuroImage* 62, no. 2 (2012): 774–781.
47. C. Vachet, B. Yvernault, K. Bhatt, et al., "Automatic corpus Callosum Segmentation Using a Deformable Active Fourier Contour Model," in *Medical Imaging 2012: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 8317 (San Diego, California: SPIE, 2012), 79–85.
48. M. Berrocal, D. Marcos, M. R. Sepúlveda, M. Pérez, J. Ávila, and A. M. Mata, "Altered Ca^{2+} Dependence of Synaptosomal Plasma Membrane Ca^{2+} -ATPase in Human Brain Affected by Alzheimer's Disease," *FASEB Journal* 23, no. 6 (2009): 1826–1834.
49. M. J. Berridge, "Calcium Signalling Remodelling and Disease," *Biochemical Society Transactions* 40, no. 2 (2012): 297–309.
50. N. Horita and T. Kaneko, "Genetic Model Selection for a Case–Control Study and a Meta-Analysis," *Meta Gene* 5 (2015): 1–8.
51. T. Kido, W. Sikora-Wohlfeld, M. Kawashima, et al., "Are Minor Alleles More Likely to Be Risk Alleles?," *BMC Medical Genomics* 11, no. 1 (2018): 3.
52. S. Guo, Y. Zhou, C. Xing, et al., "The Vasculome of the Mouse Brain," *PLoS One* 7, no. 12 (2012): e52665.
53. K. Ahn, J. D. Tucker, W. Wu, and A. Srivastava, "Regression Models Using Shapes of Functions as Predictors," *Computational Statistics & Data Analysis* 151 (2020): 107017.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.