

RESEARCH ARTICLE

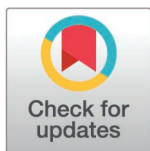
Double shrinkage transfer causal learning: An application to alzheimer's disease

Yunxin Shi¹*, Lulu Pan¹, Yu Hu, Yongfu Yu*, Guoyou Qin¹*

Department of Biostatistics, Key Laboratory of Public Health Safety of Ministry of Education, NHC Key Laboratory for Health Technology Assessment, School of Public Health, Fudan University, Shanghai, China

* These authors are Joint first authors.

* yu@fudan.edu.cn (YY); gyqin@fudan.edu.cn (GQ)



OPEN ACCESS

Citation: Shi Y, Pan L, Hu Y, Yu Y, Qin G (2026) Double shrinkage transfer causal learning: An application to alzheimer's disease. PLoS Comput Biol 22(8): e1014706. <https://doi.org/10.1371/journal.pcbi.1014706>

Editor: Rahul Singh, CSIR-IHBT: Institute of Himalayan Bioresource Technology CSIR, INDIA

Received: February 20, 2026

Accepted: August 11, 2026

Published: August 24, 2026

Copyright: © 2026 Shi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The R code for implementing DSTCL, generating the simulated data, and reproducing the analyses is publicly available on GitHub at <https://github.com/LESWinnie/DSTCL>. The simulated datasets can be generated using the code provided in the repository. The data used in the real-data

Abstract

Amyloid-Beta 42 (ABeta42) is a key biomarker of cerebral amyloidosis in Alzheimer's disease, and estimating its causal effect on subsequent cognitive outcomes is important for understanding disease progression in racial minority populations, such as African Americans. However, causal effect analyses in such populations are often challenged by limited sample sizes, leading to biased or unstable estimates when only target-population data are used. Existing transfer causal learning methods borrow information from a larger related source domain, but may be unstable when source-domain estimation is high-dimensional or when irrelevant covariates introduce noise into transferred nuisance models. We propose a double shrinkage transfer causal learning (DSTCL) estimator that regularizes both the initial source domain estimation and the source-target parameter differences, and then estimates causal effects using a doubly robust framework. We evaluated DSTCL in simulations under varying source sample sizes, inter-domain similarities, covariate dimensions, and other settings, and then applied it to the Alzheimer's Disease Neuroimaging Initiative dataset using African American participants as the target domain and non-Hispanic White participants as the source domain. In simulations, DSTCL achieved small bias, low mean squared error, and stable confidence interval coverage, generally outperforming competing methods. In the ADNI application, lower baseline ABeta42, reflecting greater amyloid burden, was associated with worse 12-month cognitive performance among African American participants, with DSTCL producing the narrowest confidence interval among the methods compared. These findings suggest that DSTCL provides a practical framework for transfer causal inference in biomedical studies with limited target population.

application were obtained from the ADNI database (<https://adni.loni.usc.edu/>). Under the ADNI Data Use Agreement, participant-level data cannot be redistributed by the authors, but qualified researchers may apply for access at <https://adni.loni.usc.edu/data-samples/adni-data/>. The authors had no special access privileges.

Funding: This work was supported by National Natural Science Foundation of China (GQ, No. 82473724; YY, No. 82273730) (<https://www.nsfc.gov.cn/>), Shanghai Municipal Science and Technology Major Project (GQ, ZD2021CY001), Shanghai Talent Programs (YY, BJKJ2024050), the Shuguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission (YY). The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. No authors received a salary from any of the funders.

Competing interests: The authors have declared that no competing interests exist.

Author summary

African Americans continue to be included in relatively small numbers in many studies of Alzheimer's disease. We sought to improve the investigation of disease-related biomarkers and later cognitive health in populations for which only limited data are available. Lower levels of amyloid-beta 42 in the fluid surrounding the brain and spinal cord generally reflect greater amyloid accumulation in the brain, a hallmark of Alzheimer's disease. However, drawing reliable conclusions is difficult when a study includes few participants but many potentially relevant characteristics. We developed a statistical approach that learns useful information from a larger, related population while retaining important differences in the smaller population and reducing the influence of irrelevant information. In the simulations, our approach generally provided more accurate and stable results than existing methods. We then applied it to data from the Alzheimer's Disease Neuroimaging Initiative, using African Americans as the target domain and non-Hispanic Whites as the larger source domain. Under assumptions needed for causal interpretation, our findings suggested that lower baseline amyloid-beta 42 was associated with worse cognitive performance after 12 months among African Americans. These findings suggest that our approach provides a practical framework for biomedical studies when target population data are limited.

1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by a gradual decline in memory and cognitive function, with its prevalence increasing worldwide in parallel with population aging [1,2]. Cerebrospinal fluid (CSF) amyloid-Beta 42 (ABeta42) is a core biomarker of cerebral amyloidosis and plays a central role in characterizing AD severity and disease progression trajectories [3,4]. Previous studies have shown that amyloid burden and clinical progression patterns differ across racial groups, with African Americans experiencing a substantially higher disease burden than non-Hispanic Whites [5,6]. These disparities motivate a more precise characterization of the relationship between amyloid biomarkers and disease severity in African Americans, to inform population-specific understanding of disease evolution and potential intervention strategies. However, progress in addressing these questions is constrained by fundamental data limitations. Major AD cohorts exhibit substantial imbalance in population representation, largely driven by limited sample sizes of African American participants. For example, in the Alzheimer's Disease Neuroimaging Initiative (ADNI), which includes data from 59 North American sites and several thousand participants, African Americans account for only about 10% of the total sample [7,8].

Meanwhile, growing evidence underscores the importance of integrating molecular-level data to better capture AD mechanisms and heterogeneity. Genome-wide DNA methylation (DNAm) profiles reflect systemic and brain-related pathological processes

and are widely used as covariates to improve confounding control in etiologic analyses [9,10]. Epigenome-wide meta-analyses further demonstrate that AD-related DNAm changes are broadly shared across cohorts, with a limited number of locus-specific differences [11,12]. While incorporating DNAm into causal analyses can substantially improve confounding control, it also increases model complexity and places greater demands on stable estimation, particularly when the target population sample size is limited. In such a setting, methods that estimate all parameters using only target domain data often perform poorly: classical doubly robust estimators that rely on matrix inversion may fail numerically, while regularization or machine learning alternatives can produce substantially biased and unstable causal effect estimates [13–16].

Transfer learning offers a promising avenue to mitigate these limitations. Its central idea is to improve performance in a target dataset by borrowing information from external source data, and it has demonstrated substantial potential in applications such as drug discovery and metabolic mechanism studies [17–19]. Künzel et al. were among the first to systematically investigate how transfer learning can be used for causal effect estimation under a shared covariate space, and referred to this as the transfer causal learning (TCL) problem [19]. Building on this, Wei et al. proposed the I_1 -TCL method [20], which obtains initial parameter estimates via maximum likelihood in the source domain and then uses target data to estimate parameter differences, thereby achieving transfer learning of nuisance parameters and plugging the resulting estimates directly into causal effect estimators. The I_1 -TCL method typically assumes a high degree of similarity in covariate structures between the source and target domains. However, when a traditional maximum likelihood method is used in the source domain at the initial stage, the resulting estimates often retain nonzero effect estimates for covariates with little true relevance to the outcome, thereby transmitting instability into the estimation of target domain parameters, potentially amplifying estimation error and undermining the accuracy and stability of causal effect estimates.

To address this issue, we propose a double shrinkage transfer causal learning (DSTCL) method that improves upon existing TCL approaches by applying shrinkage both to the initial source domain estimation and to the source-target parameter differences, yielding more stable and reliable causal effect estimates. We conduct systematic simulation studies to evaluate the performance of DSTCL across a range of scenarios varying in source sample size, inter-domain similarity, and covariate dimension, and apply the method to the ADNI dataset to estimate and quantify the causal effect of baseline ABeta42 status on subsequent cognitive function among African American participants. Our findings not only enhance understanding of AD-related biomarkers and disease progression in the African American population, but also provide a transferable statistical framework for causal inference in studies of racial minority populations and other biomedical studies.

2. Materials and methods

2.1. Ethics Statement

This study used simulated data and de-identified secondary data obtained from the ADNI database (<https://adni.loni.usc.edu/>). Ethics approval was not applicable to the simulations because they did not involve human participants or animals. The ADNI study was approved by the Institutional Review Boards of participating institutions, and written informed consent was obtained from all participants or their legally authorized representatives. The present secondary analysis was conducted under the ADNI Data Use Agreement, and no additional local ethics approval was required.

2.2. Notation and assumptions

For simplicity, each individual has two potential outcomes: $Y_i(1)$ if treated, and $Y_i(0)$ if untreated. Under standard causal assumptions for observational data, including consistency, the stable unit treatment value assumption, unconfoundedness, and positivity [13,21], the causal effect of interest is identifiable as $\Delta = \mathbb{E}(Y(1)) - \mathbb{E}(Y(0))$.

Let Z_i indicate exposure status and $\mathbf{X}_i$ denote the vector of observed covariates. The propensity score (PS) is defined as the conditional probability of receiving exposure given covariates, $e(\mathbf{X}_i) = P(Z_i = 1 | \mathbf{X}_i)$. The outcome regression (OR) function is considered as: $\mu_z(\mathbf{x}) = \mathbb{E}(Y|Z=z, \mathbf{X} = \mathbf{x})$, $z \in \{0, 1\}$.

To enhance robustness against potential model misspecification, the DR estimator has been proposed in the standard augmented inverse probability weighted form:

$$\hat{\Delta}_{DR} = \frac{1}{n} \sum_{i=1}^n \left[\hat{\mu}_1(\mathbf{X}_i) - \hat{\mu}_0(\mathbf{X}_i) + \frac{Z_i \{Y_i - \hat{\mu}_1(\mathbf{X}_i)\}}{\hat{e}(\mathbf{X}_i)} - \frac{(1 - Z_i) \{Y_i - \hat{\mu}_0(\mathbf{X}_i)\}}{1 - \hat{e}(\mathbf{X}_i)} \right].$$

The key property of the DR estimator is that as long as either the PS model or the OR model is correctly specified, the DR estimator remains consistent.

2.3. Proposed method

We consider two related domains, a source domain S and a target domain T . Following the transfer causal learning framework, the source and target domains are assumed to share the same covariate space and to have sufficient covariate overlap in applications, while they are allowed to have different feature distributions, treatment assignment mechanisms, and causal effects.

Let the target domain dataset be $\{(\mathbf{x}_{i,t}, z_{i,t}, y_{i,t})\}_{i=1}^{n_t}$, where $\mathbf{x}_{i,t} \in \mathbf{R}^p$ denotes the covariates, $z_{i,t} \in \{0, 1\}$ is a binary exposure variable, and $y_{i,t} \in \mathbf{R}$ is a continuous outcome variable. Similarly, suppose the source domain dataset is $\{(\mathbf{x}_{i,s}, z_{i,s}, y_{i,s})\}_{i=1}^{n_s}$, with $n_t < n_s$.

Our goal is to estimate the causal effect in the target domain, where the limited sample size makes traditional estimation biased and unstable if relying solely on target data. Because exposure mechanisms differ across domains, direct pooling is invalid, and we therefore adopt a transfer learning approach to borrow information from the source domain while preserving causal validity.

2.3.1 Transfer learning for propensity score model parameters. Let the PS models in the two domains be defined as:

$$P(Z_t = 1 | \mathbf{X}_t) = e_t(\mathbf{X}_t^T \beta_t), \quad P(Z_s = 1 | \mathbf{X}_s) = e_s(\mathbf{X}_s^T \beta_s),$$

where $e_t(\cdot)$, $e_s(\cdot)$ are known link functions, and $\beta_t, \beta_s \in \mathbf{R}^p$ are domain-specific parameters. In this study, we used the logit link for propensity score modeling in both the source and target domains because the exposure is binary in both domains. This choice was made for simplicity and interpretability, and identical link functions are not required by the h-transferability assumption. Following Wei et al., knowledge transfer does not require the same link function across domains as long as the link functions are known [20]. Meanwhile, the source-target parameter difference is defined as $\delta_\beta = \beta_t - \beta_s$. The success of transfer learning relies on the sparsity assumption: if $\|\delta_\beta\|_0 \leq h$ for some small constant h , then the exposure assignment mechanisms across domains differ only in a limited subset of covariates. This property is termed **h-transferability**, under which source domain information can be effectively adapted to the target domain.

The estimation proceeds in two steps:

1. Initial estimation in the source domain

The initial parameter estimation of β_s is obtained by penalized maximum likelihood with abundant source data:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n_s} \left[-z_{i,s} \mathbf{x}_{i,s}^T \beta + \ln(1 + e^{\mathbf{x}_{i,s}^T \beta}) \right] + \lambda_b \|\beta\|_1 \right\}$$

where $\lambda_b > 0$ is selected via cross-validation.

2. Bias correction using target domain data

The sparse inter-domain difference δ_β is estimated via ℓ_1 -regularization:

$$\hat{\delta}_\beta = \underset{\Delta_\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n_t} \left[-z_{i,t} \mathbf{x}_{i,t}^T (\hat{\beta}_s + \Delta_\beta) + \ln \left(1 + e^{\mathbf{x}_{i,t}^T (\hat{\beta}_s + \Delta_\beta)} \right) \right] + \lambda_{ps} \|\Delta_\beta\|_1 \right\}$$

where $\lambda_{ps} > 0$ is a tuning parameter selected via cross-validation. Wei et al. demonstrated the theoretical validity of the Lasso-based transfer estimator under appropriate regularity conditions.

Then, the final PS parameter estimate is:

$$\hat{\beta}_t = \hat{\beta}_s + \hat{\delta}_\beta$$

Algorithm 1 Double Shrinkage Transfer Causal Learning Estimation

1. **Input** target data $\mathcal{D}_t$, source data $\mathcal{D}_s$
2. **Initial estimation in the source domain**
 $\hat{\theta}_s = \underset{\theta_s}{\operatorname{argmin}} \mathcal{L}(\theta_s; \mathcal{D}_s) + \lambda_s \|\theta_s\|_1$
3. **Bias correction using target domain data**
 $\hat{\delta}_\theta = \underset{\delta_\theta}{\operatorname{argmin}} \mathcal{L}(\hat{\theta}_s + \delta_\theta; \mathcal{D}_t) + \lambda_\theta \|\delta_\theta\|_1$
4. **Final estimation in the target domain**
 $\hat{\theta}_t = \hat{\theta}_s + \hat{\delta}_\theta$
5. **Output** $\hat{\theta}_t$

2.3.2. Transfer learning for outcome model parameters. Analogously, the OR models are specified separately for each group

$$\mu_{z,t}(\mathbf{x}_t) = \mathbb{E}[Y_t | Z_t = z, \mathbf{X}_t = \mathbf{x}_t] = m_{z,t}(\mathbf{x}_t; \alpha_{z,t}),$$

$$\mu_{z,s}(\mathbf{x}_s) = \mathbb{E}[Y_s | Z_s = z, \mathbf{X}_s = \mathbf{x}_s] = m_{z,s}(\mathbf{x}_s; \alpha_{z,s}),$$

where $m_{z,t}(\cdot)$ and $m_{z,s}(\cdot)$ are functions with known forms, whereas $\alpha_{z,t}$ and $\alpha_{z,s}$ represent the unknown parameters. For simplicity, we parameterize the conditional expectations in both domains via linear regression models as follows:

$$\mu_{z,t}(\mathbf{x}_t) = \mathbb{E}[Y_t | Z_t = z, \mathbf{X}_t = \mathbf{x}_t] = \mathbf{x}_t^T \alpha_{z,t}$$

$$\mu_{z,s}(\mathbf{x}_s) = \mathbb{E}[Y_s | Z_s = z, \mathbf{X}_s = \mathbf{x}_s] = \mathbf{x}_s^T \alpha_{z,s}$$

As long as the forms of $m_{z,t}(\cdot)$ and $m_{z,s}(\cdot)$ are known, the learned components remain independent of the assumption that the “means” follow linear regression conditions.

For OR models, the source-target parameter difference is given by:

$$\delta_{z,\alpha} = \alpha_{z,t} - \alpha_{z,s}$$

The joint transportability of the model parameters similarly requires sparsity of $\delta_{z,\alpha}$. For $z \in \{0, 1\}$, denote $n_{z,s} = \#\{i : z_{i,s} = z\}$ and $n_{z,t} = \#\{i : z_{i,t} = z\}$, where $\#\{\cdot\}$ represents the cardinality of a set. This allows us to apply the aforementioned regularized transfer learning techniques, leveraging source domain data to estimate the OR parameters in the target domain:

$$\hat{\alpha}_{z,s} = \underset{\alpha}{\operatorname{argmin}} \left[\frac{1}{n_{z,s}} \sum_{z_{i,s}=z} (y_{i,s} - \mathbf{x}_{i,s}^T \alpha)^2 + \lambda_s \|\alpha\|_1 \right]$$

$$\hat{\delta}_{z,\alpha} = \underset{\Delta_{z,\alpha}}{\operatorname{argmin}} \left[\frac{1}{n_{z,t}} \sum_{z_{i,t}=z} (y_{i,t} - \mathbf{x}_{i,t}^T (\hat{\alpha}_{z,s} + \Delta_{z,\alpha}))^2 + \lambda_{or} \|\Delta_{z,\alpha}\|_1 \right]$$

$$\hat{\alpha}_{z,t} = \hat{\alpha}_{z,s} + \hat{\delta}_{z,\alpha}$$

where $\lambda_s, \lambda_{or} > 0$ are selected via cross-validation.

2.3.3 Estimation of causal effect. With transfer-learned parameters in PS and OR models, we construct the DSTCL estimator as follows, with detailed steps in Algorithm 2:

$$\hat{\Delta}_{\text{DSTCL}} = \frac{1}{n_t} \sum_{i=1}^{n_t} \left[\hat{\mu}_{1,\text{DSTCL}}(\mathbf{x}_i) - \hat{\mu}_{0,\text{DSTCL}}(\mathbf{x}_i) + \frac{Z_i \{Y_i - \hat{\mu}_{1,\text{DSTCL}}(\mathbf{x}_i)\}}{\hat{e}_{\text{DSTCL}}(\mathbf{x}_i)} - \frac{(1 - Z_i) \{Y_i - \hat{\mu}_{0,\text{DSTCL}}(\mathbf{x}_i)\}}{1 - \hat{e}_{\text{DSTCL}}(\mathbf{x}_i)} \right]$$

Algorithm 2 Proposed DSTCL Estimator of causal effect

1. **Input** target data $\{(\mathbf{x}_{i,t}, z_{i,t}, y_{i,t})\}_{i=1}^{n_t}$, source data $\{(\mathbf{x}_{i,s}, z_{i,s}, y_{i,s})\}_{i=1}^{n_s}$

2. **Transfer learning for propensity score**

$$\hat{\beta}_s = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n_s} \left[-z_{i,s} \mathbf{x}_{i,s}^T \beta + \ln(1 + e^{\mathbf{x}_{i,s}^T \beta}) \right] + \lambda_b \|\beta\|_1 \right\}$$

$$\hat{\delta}_\beta = \underset{\Delta_\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n_t} \left[-z_{i,t} \mathbf{x}_{i,t}^T (\hat{\beta}_s + \Delta_\beta) + \ln(1 + e^{\mathbf{x}_{i,t}^T (\hat{\beta}_s + \Delta_\beta)}) \right] + \lambda_{ps} \|\Delta_\beta\|_1 \right\}$$

$$\hat{\beta}_t = \hat{\beta}_s + \hat{\delta}_\beta$$

3. **Transfer learning for outcome regression**

$$\hat{\alpha}_{z,s} = \underset{\alpha}{\operatorname{argmin}} \left[\frac{1}{n_{z,s}} \sum_{z_{i,s}=z} (y_{i,s} - \mathbf{x}_{i,s}^T \alpha)^2 + \lambda_s \|\alpha\|_1 \right]$$

$$\hat{\delta}_{z,\alpha} = \underset{\Delta_{z,\alpha}}{\operatorname{argmin}} \left[\frac{1}{n_{z,t}} \sum_{z_{i,t}=z} (y_{i,t} - \mathbf{x}_{i,t}^T (\hat{\alpha}_{z,s} + \Delta_{z,\alpha}))^2 + \lambda_{or} \|\Delta_{z,\alpha}\|_1 \right]$$

$$\hat{\alpha}_{z,t} = \hat{\alpha}_{z,s} + \hat{\delta}_{z,\alpha}$$

4. **Construct DR estimator in target domain**

$$\hat{\Delta}_{\text{DSTCL}} = \frac{1}{n_t} \sum_{i=1}^{n_t} \left[\hat{\mu}_{1,\text{DSTCL}}(\mathbf{x}_i) - \hat{\mu}_{0,\text{DSTCL}}(\mathbf{x}_i) + \frac{Z_i \{Y_i - \hat{\mu}_{1,\text{DSTCL}}(\mathbf{x}_i)\}}{\hat{e}_{\text{DSTCL}}(\mathbf{x}_i)} - \frac{(1 - Z_i) \{Y_i - \hat{\mu}_{0,\text{DSTCL}}(\mathbf{x}_i)\}}{1 - \hat{e}_{\text{DSTCL}}(\mathbf{x}_i)} \right]$$

5. **Output** $\hat{\Delta}_{\text{DSTCL}}$

The proposed DSTCL framework assumes that the source and target domains share the same covariate space and relies on standard causal assumptions, including consistency, unconfoundedness, and positivity, as well as transfer-learning assumptions including sufficient source-target covariate overlap and sparsity of source-target parameter differences. These assumptions are consistent with existing transfer learning theory, where theoretical guarantees generally require source-target similarity, sparsity, and regularity conditions [26,27]. In our setting, these assumptions allow information from the source domain to improve parameter estimation in the target domain while preserving the target domain causal estimand.

All l_1 -regularized nuisance models were fitted using glmnet with $\alpha = 1$, and tuning parameters were selected by cross-validation. We used five-fold cross-validation whenever feasible, with binomial deviance for propensity score models and mean squared error for outcome regression models. In DSTCL, tuning was performed separately for each nuisance component, and the source domain initial estimation step and target domain correction step used separate cross-validation procedures. The primary analysis used λ_{1se} , with λ_{min} considered in sensitivity analyses.

3. Simulation

3.1. Simulation design

We evaluated the proposed method using simulated data with a target sample size $n_t = 50$ and a source sample size $n_s = 1000$, and $p = 200$ covariates in both domains. For each individual i in both the source and target domains, the covariates $\mathbf{x}_i$ are generated from a multivariate normal distribution:

$$\mathbf{x}_i \sim N(\mathbf{0}_p, \Sigma), \Sigma = (\Sigma_{jk}), \Sigma_{jk} = 0.2^{|j-k|+1}, i = 1, 2, \dots, n.$$

The design matrix is constructed as $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T$.

A binary variable is generated based on these covariates, with the true PS models specified as follows:

$$\text{logit}\{P(Z_t = 1 | \mathbf{X}_t)\} = \beta_{0,t} + \mathbf{X}_t^T \beta_{1,t}, \text{logit}\{P(Z_s = 1 | \mathbf{X}_s)\} = \beta_{0,s} + \mathbf{X}_s^T \beta_{1,s}.$$

The parameters are set as $\beta_{1,t} = (\underbrace{-0.3, \dots, -0.3}_5, \underbrace{0.6, \dots, 0.6}_{45}, \underbrace{0, \dots, 0}_{150})^T$ with $\beta_{0,t} = -0.2$ and

$\beta_{1,s} = (\underbrace{0.3, \dots, 0.3}_5, \underbrace{0.6, \dots, 0.6}_{45}, \underbrace{0, \dots, 0}_{150})^T$ with $\beta_{0,s} = 0.2$. The true coefficient differences in the PS models between source and target domains are $\delta_\beta = (\underbrace{-0.4, -0.6, \dots, -0.6}_5, \underbrace{0, \dots, 0}_{195})$, implying the transportability parameter $h = 6$.

The outcome variable Y is defined as:

$$Y_{z,t} = \alpha_{0,z,t} + \mathbf{X}_t^T \alpha_{1,t} + \varepsilon_{z,t}, Y_{z,s} = \alpha_{0,z,s} + \mathbf{X}_s^T \alpha_{1,s} + \varepsilon_{z,s}, z \in \{0, 1\}.$$

where the residuals in both groups across domains follow a normal distribution $N(0, 0.5^2)$. Additionally, for the target domain, the parameters are set as $\alpha_{1,t} = (\underbrace{-0.3, -0.3, -0.3, 0.3, 0.3}_5, \underbrace{0.8, \dots, 0.8}_{25}, \underbrace{-0.8, \dots, -0.8}_{20}, \underbrace{0, \dots, 0}_{150})^T$,

with intercept $\alpha_{0,1,t} = -1$ and $\alpha_{0,0,t} = 1$, yielding a true causal effect of $\Delta_t = -2$. For the source domain, the parameters are set as $\alpha_{1,s} = (\underbrace{0.3, 0.3, 0.3, -0.3, -0.3}_5, \underbrace{0.8, \dots, 0.8}_{25}, \underbrace{-0.8, \dots, -0.8}_{20}, \underbrace{0, \dots, 0}_{150})^T$, with intercept $\alpha_{0,1,s} = 1$ and

$\alpha_{0,0,s} = -1$, then the true causal effect is $\Delta_s = 2$. Then the true coefficient differences between two domains are $\delta_{z,\alpha} = (2 - 4z, \underbrace{-0.6, -0.6, -0.6, 0.6, 0.6}_5, \underbrace{0, \dots, 0}_{195})$, implying the transportability parameter for the outcome model is also 6.

To assess robustness under model misspecification, we considered two regimes: a PS-correct setting, where the exposure model is correctly specified but the outcome model includes a nonlinear sine term, and an OR-correct setting, where the outcome model is correct but the exposure model is misspecified by omitting covariates and adding nonlinearity. Simulations were repeated 200 times, with performance metrics evaluated based on these simulated datasets. We also recorded computational runtime and convergence status for each simulation replication.

We compared the following 8 methods:

1. **Crude difference (Crude_diff)**: The unadjusted mean difference in outcomes between the treated and control groups in the target domain.
2. **Known parameters (Known_param)**: A doubly robust estimator in which the true PS and OR model parameters are plugged in, serving as a benchmark.
3. **Generalized linear model (GLM)**: The DR estimator with nuisance parameters estimated in the target domain using traditional logistic regression for PS and linear regression for OR model.
4. **Least absolute shrinkage and selection operator (Lasso)**: The DR estimator with parameters estimated in the target domain via Lasso regularization.
5. **Double Machine Learning (DML)**: A cross-fitted causal learning framework that estimates the PS and OR model using machine learning methods.
6. **Targeted Maximum Likelihood Estimation (TMLE)**: A doubly robust procedure that updates an initial OR model using information from PS to target the parameter of interest; both the initial PS and OR model are fitted with Lasso regularization.
7. **l_1 -Transfer causal learning estimator (l_1 -TCL)**: A transfer learning approach in which models are first initialized using MLE in source domain, then bias-corrected in target domain via Lasso, and finally substituted into the DR estimator.
8. **Double Shrinkage Causal Transfer Learning (DSTCL)**: Our proposed transfer causal learning estimator.

Furthermore, the evaluation of estimator performance was based on the following criteria:

1. **Bias**: The Bias of overall causal effect was calculated across 200 replications as: $Bias = \frac{1}{200} \sum_{i=1}^{200} (\hat{\theta}_i - \theta)$, where θ denotes the true value of causal effect, and $\hat{\theta}_i$ represents the estimated value obtained in the i -th replication.
2. **Standard deviation (SD)**: The standard deviation of the estimator across 200 replications was computed as: $SD = \sqrt{\frac{1}{200-1} \sum_{i=1}^{200} (\hat{\theta}_i - \bar{\theta})^2}$, where $\bar{\theta} = \frac{1}{200} \sum_{i=1}^{200} \hat{\theta}_i$ denotes the average estimate across the 200 replications.
3. **Mean squared error (MSE)**: The mean squared error of the estimator was assessed over 200 replications as:

$$MSE = \frac{1}{200} \sum_{i=1}^{200} (\hat{\theta}_i - \bar{\theta})^2$$

4. **Coverage probability of the 95% percentile confidence interval (Coverage)**: For each of the 200 replications, a 95% percentile bootstrap confidence interval was constructed based on 500 bootstrap resamples, defined as $CI_i = [\hat{\theta}_{i,0.025}, \hat{\theta}_{i,0.975}]$. Coverage was then evaluated by checking whether the true parameter θ fell within the confidence interval, with $Cover_i = \begin{cases} 1, & \text{if } \theta \in CI_i \\ 0, & \text{otherwise} \end{cases}$. Finally, the overall coverage probability was obtained as:

$$95\% \text{ Cov} = \frac{1}{200} \sum_{i=1}^{200} Cover_i$$

3.2. Simulation scenarios

Scenario 1. Effect of source domain sample size: In this scenario, we investigated how gradual increases in source domain sample size affected the performance of the DSTCL estimator. Target domain sample size was fixed at $n_t = 50$, while source domain sample size varied as $n_s \in \{500, 1000, 2000\}$.

Scenario 2. Effect of inter-domain similarity: To evaluate the effect of inter-domain similarity, we further varied $h \in \{6, 11, 16\}$, which quantifies the number of nonzero elements in parameter difference vectors. Smaller h corresponds to stronger similarity between source and target domains, while larger h reflects weaker transferability. All other settings remained unchanged.

Scenario 3. Effect of the number of covariates: In this setting, we examined the sensitivity of different estimators to increasing dimensionality by systematically varying the covariate size. Specifically, we considered $p \in \{50, 100, 200\}$ while holding all other simulation parameters fixed.

We also conducted several sensitivity analyses under representative settings to evaluate robustness to the penalty-selection rule, extreme propensity scores, propensity-score trimming, and mild covariate shift between the source and target domains. Specifically, we compared λ_{1se} with λ_{min} , increased the propensity-score strength parameter to generate probabilities closer to 0 or 1, compared no trimming with propensity score truncation to $[0.01, 0.99]$ and $[0.05, 0.95]$, and introduced mild target domain covariate mean shift. Details are provided in the Supporting Information.

3.3. Simulation results

In the primary simulation setting with $n_t = 50$, $n_s = 1000$, $p = 200$, $h = 6$, the proposed DSTCL estimator showed performance close to the oracle benchmark and clearly superior to other methods (Table 1). When both the PS and OR models

Table 1. The performance across different methods in primary simulation setting.

	Estimator	Bias (SD)	MSE	Coverage
Both the PS and OR models are correct	Crude_diff	0.42 (0.87)	0.93	0.92
	Known_param	0.00 (0.33)	0.11	0.94
	Glm	—	—	—
	Lasso	0.45 (0.82)	0.87	0.78
	DML1	0.28 (0.79)	0.69	0.92
	DML2	0.25 (0.83)	0.76	0.92
	TMLE	0.34 (0.84)	0.82	0.99
	I_1 -TCL	—	—	—
	DSTCL	0.06 (0.26)	0.07	0.94
Only the PS model is correct	Crude_diff	1.22 (0.45)	1.69	0.26
	Known_param	0.19 (1.60)	2.58	0.90
	Glm	—	—	—
	Lasso	1.21 (0.44)	1.65	0.06
	DML1	1.20 (0.54)	1.72	0.28
	DML2	1.19 (0.52)	1.68	0.27
	TMLE	0.89 (0.51)	1.05	0.73
	I_1 -TCL	—	—	—
	DSTCL	0.12 (0.24)	0.07	0.92
Only the OR model is correct	Crude_diff	0.36 (0.86)	0.87	0.94
	Known_param	0.02 (0.32)	0.10	0.90
	Glm	—	—	—
	Lasso	0.41 (0.79)	0.79	0.80
	DML1	0.53 (0.84)	0.97	0.87
	DML2	0.52 (0.83)	0.96	0.88
	TMLE	0.52 (0.92)	1.12	0.96
	I_1 -TCL	—	—	—
	DSTCL	0.07 (0.25)	0.07	0.96

<https://doi.org/10.1371/journal.pcbi.1014706.t001>

were correctly specified, DSTCL exhibited negligible Bias and the smallest SD among all data-driven estimators, leading to the lowest MSE and coverage close to 95%. In contrast, Lasso, DML1 (Lasso-randomforest), DML2 (SVM-randomforest), and TMLE all displayed either larger absolute Bias, inflated SD, or both, which translated into higher MSEs, while I_1 -TCL failed to yield an estimate. Similar patterns were observed when only the PS or OR model was correctly specified. DSTCL consistently achieved small Bias and relatively tight variability, with MSE values comparable to or smaller than those of the oracle estimator and coverage remaining near 0.95. In this primary simulation setting, DSTCL required approximately 1.71 seconds per main fit and 13.96 seconds per individual bootstrap refit on average; thus, 500 bootstrap resamples required approximately 6,980 seconds. No DSTCL estimation failures were observed across the 200 simulation replications.

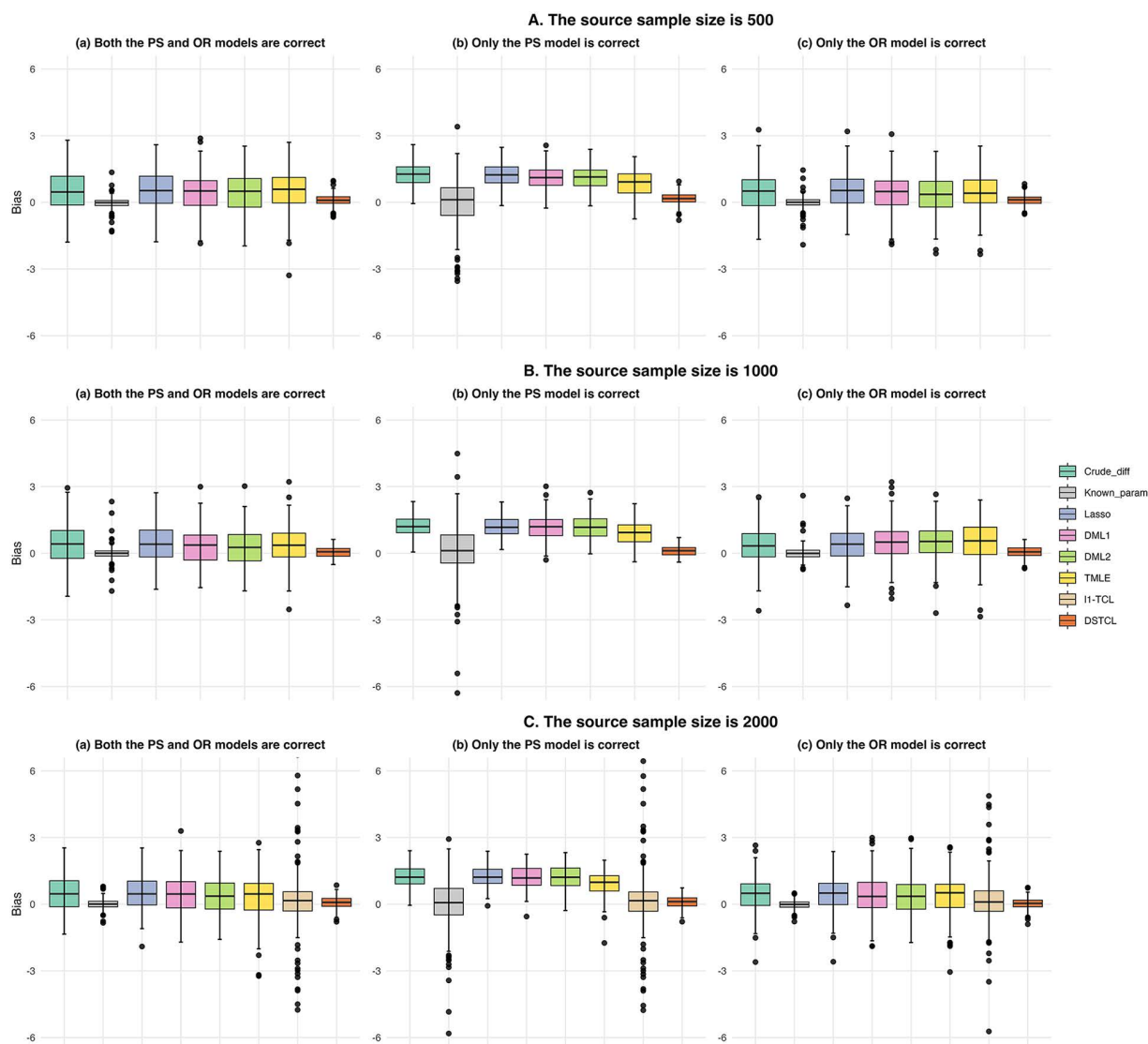


Fig 1. Boxplots of Bias and SD for varying the sample size in source domain (Scenario 1).

<https://doi.org/10.1371/journal.pcbi.1014706.g001>

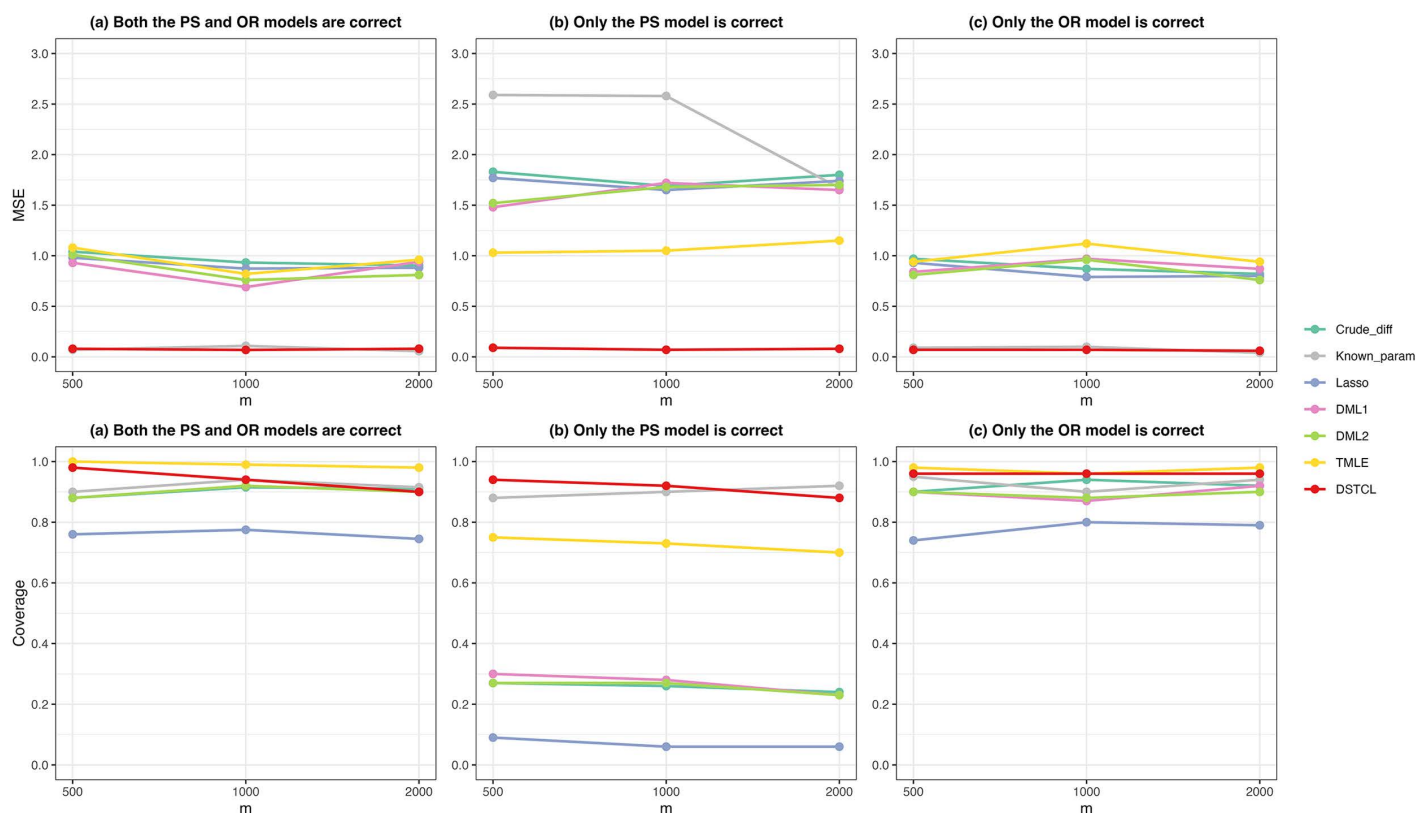


Fig 2. Line plots of MSE and Coverage for varying the sample size in source domain (Scenario 1). I_1 -TCL was omitted from the line plots because complete simulation and bootstrap results were not available across all settings. Available I_1 -TCL results are reported in the corresponding Supporting Information tables.

<https://doi.org/10.1371/journal.pcbi.1014706.g002>

As n_s increased from 500 to 1000 and 2000, the proposed DSTCL Bias remained very close to zero and SD changed only modestly, leading to uniformly low MSE (Figs 1 and 2 and Table A in S1 Appendix). Across these settings, its bias ranged from 0.04 to 0.17, its MSE ranged from 0.06 to 0.09, and its coverage ranged from 0.88 to 0.98. In contrast, the crude difference, Lasso, DML1, DML2, and TMLE estimators exhibited persistently larger absolute Bias and/or inflated variability, yielding substantially higher MSEs even when either the PS or OR model was correctly specified. Moreover, I_1 -TCL failed to produce valid estimates in a non-negligible fraction of replicates when the source-domain sample size was 500 or 1000, but converged reliably when $n_s = 2000$, indicating that its numerical stability requires a sufficiently large source dataset. These results indicate that DSTCL can reliably exploit auxiliary information from source domain even when the source sample size is moderate, and that it does not rely on a very large source cohort to achieve clear improvements.

In Scenario 2, we examined the impact of inter-domain similarity by varying h (Figs 3 and 4 and Table B in S1 Appendix). When $h = 6$, corresponding to relatively strong similarity between domains, DSTCL again achieved near-oracle Bias and the lowest MSE among all estimators. As h increased to 11 and 16, the absolute Bias and variability of DSTCL increased only modestly and remained much smaller than those of target-only estimators. Across the three model-specification scenarios, DSTCL had bias ranging from 0.06 to 0.19, MSE ranging from 0.07 to 0.21, and coverage ranging from 0.92 to 0.98. In contrast, complete I_1 -TCL results were not available across these settings. These findings suggest that DSTCL provides robust inference when cross-domain parameter differences are relatively large.

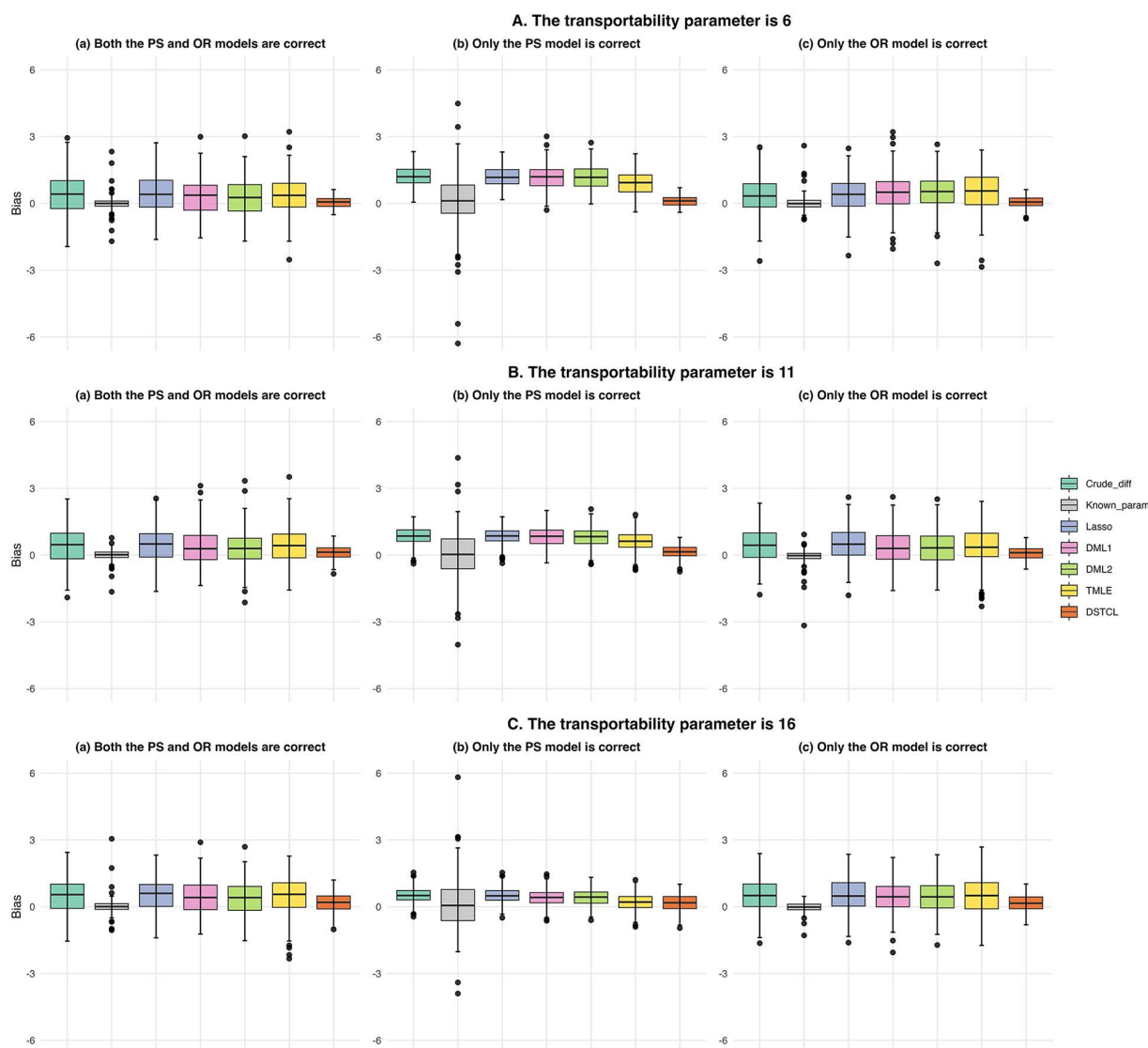


Fig 3. Boxplots of Bias and SD for varying the h-transferability (Scenario 2).

<https://doi.org/10.1371/journal.pcbi.1014706.g003>

When the covariate dimension was increased, the performance of the DSTCL estimator was essentially unchanged (Figs 5 and 6 and Table C in S1 Appendix). Across the three model-specification scenarios, its bias ranged from 0.04 to 0.12, its MSE ranged from 0.06 to 0.08, and its coverage ranged from 0.91 to 0.97. Although some competing estimators achieved comparable or lower MSE in the PS-correct setting when $p=50$, their performance varied more substantially across covariate dimensions and model-specification scenarios. In most settings, the competing data-driven estimators exhibited greater variability or larger MSE than DSTCL. In contrast, DSTCL consistently maintained low MSE and coverage close to 95%.

Taken together, these simulations suggest that the proposed DSTCL estimator performs well when the target population has a limited sample size. Under different data generating mechanisms, DSTCL achieved small Bias and low MSE while maintaining approximately 95% coverage.

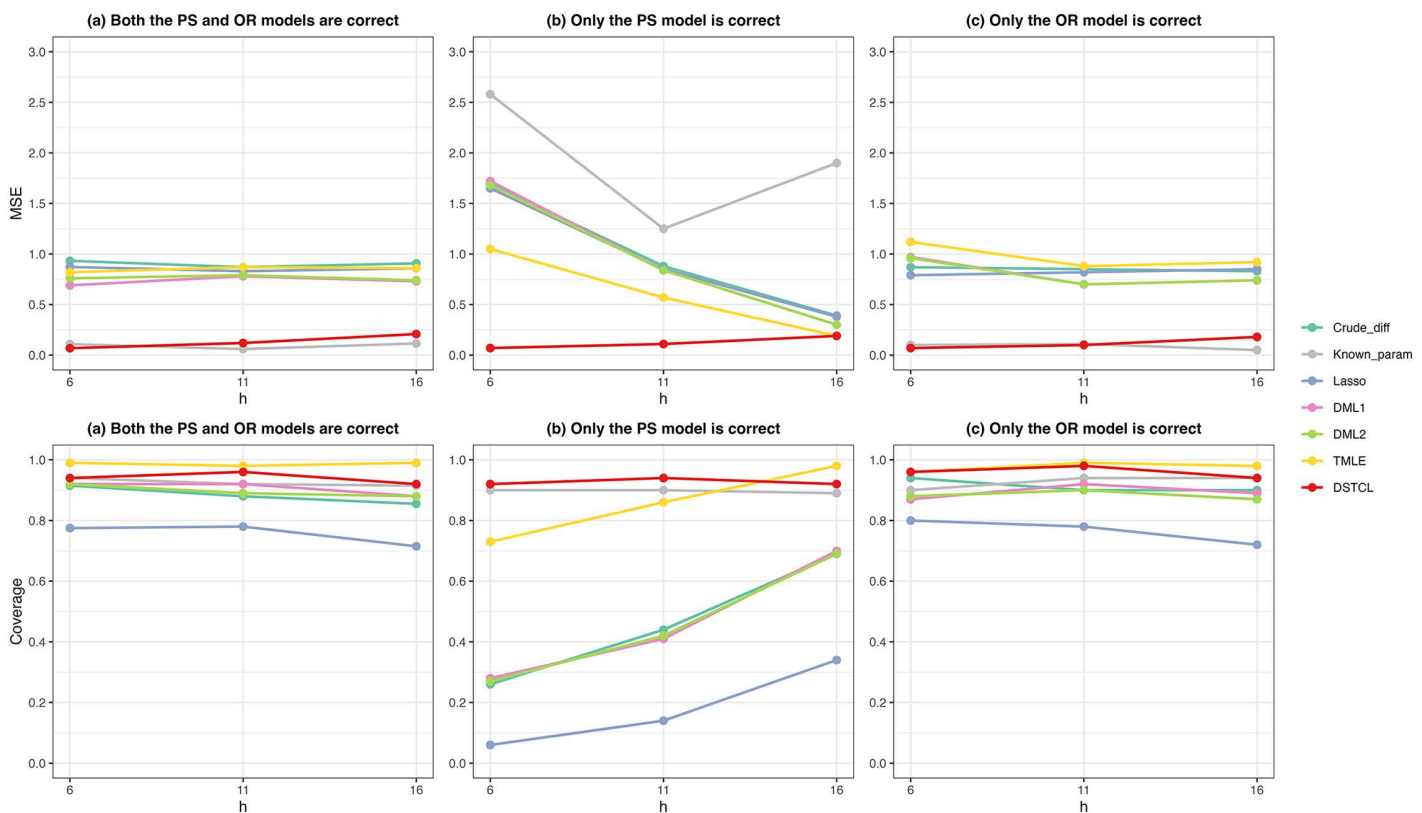


Fig 4. Line plots of MSE and Coverage for varying the h -transferability (Scenario 2). I_1 -TCL was omitted from the line plots because complete simulation and bootstrap results were not available across all settings. Available I_1 -TCL results are reported in the corresponding Supporting Information tables.

<https://doi.org/10.1371/journal.pcbi.1014706.g004>

Additional sensitivity analyses showed that the main conclusions were not highly sensitive to the penalty-selection rule, propensity-score trimming, extreme propensity-score settings, or mild covariate shift. Results obtained using $\lambda_{\min}$ were broadly comparable to those from the primary λ_{1se} analysis, although some deterioration was observed in the PS-correct setting. Under the most extreme propensity-score setting, DSTCL showed some expected performance degradation but remained relatively stable, with bias ranging from 0.11 to 0.18, MSE ranging from 0.11 to 0.17, and coverage ranging from 0.92 to 0.97. Propensity-score trimming did not materially alter the bias or MSE, although coverage changed modestly in some settings. Under mild covariate shift, bias ranged from 0.05 to 0.15, MSE ranged from 0.08 to 0.10, and coverage ranged from 0.92 to 0.96. Computational diagnostics showed no DSTCL estimation failures across the 200 replications at any evaluated covariate dimension. The mean main-fit time ranged from 0.86 to 2.01 seconds, while the mean time per individual bootstrap refit ranged from 1.03 to 16.04 seconds and increased with the covariate dimension. Detailed sensitivity and computational results are reported in the Supporting Information (**Tables D-H in S1 Appendix**).

4. Real-data application

4.1. Data description

Participants in this study were drawn from ADNI, a long-term, multicenter longitudinal study designed to identify clinical, genetic, and biochemical markers associated with AD progression. ADNI-1 was launched in 2003 and enrolled individuals

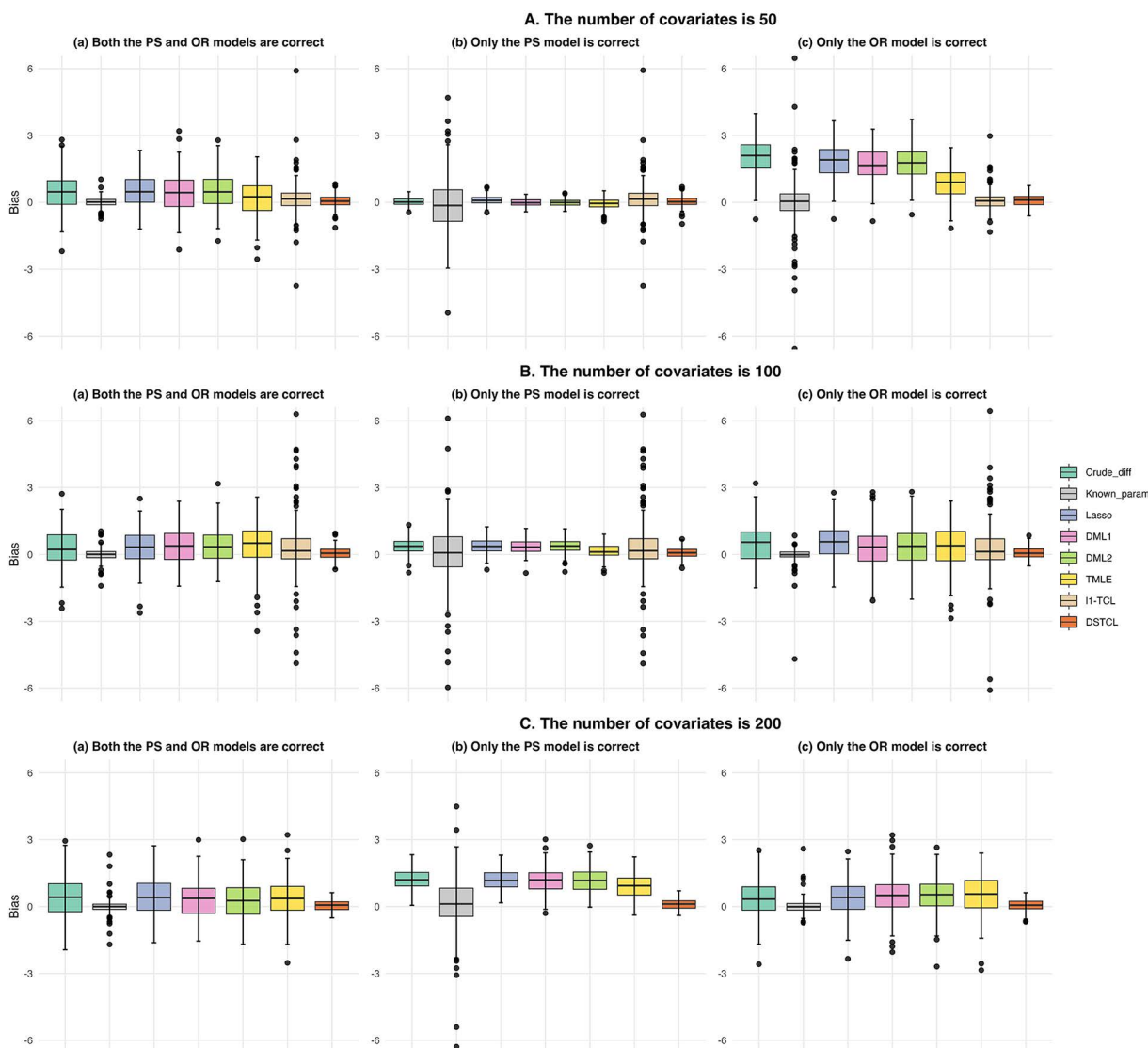


Fig 5. Boxplots of Bias and SD for varying the number of covariates (Scenario 3).

<https://doi.org/10.1371/journal.pcbi.1014706.g005>

aged 55–90 years from 63 sites across the United States and Canada; subsequent phases (ADNI-GO, ADNI-2, and ADNI-3) followed existing participants and recruited additional cohorts. Details of the ADNI study design have been described elsewhere [22–24]. The ADNI database provides access to demographic information, questionnaire data, cerebrospinal fluid (CSF) biomarkers, and multi-omics measurements including DNA methylation (DNAm).

We examined baseline demographic and genetic covariates and focused on baseline CSF ABeta42 status as the primary exposure. Following Hansson et al. [25], who derived Elecsys assay-specific ABeta42 cut points with high concordance to amyloid PET and validated them in ADNI, we used 977 pg/mL as the threshold. Participants were classified into low (≤ 977 pg/mL) and high (> 977 pg/mL) ABeta42 groups, with status treated as a binary exposure. We then estimated the causal effect of baseline ABeta42 on AD severity at 12 months, measured by ADAS-11, which ranges from 0 to 70, with higher scores indicating more severe cognitive impairment.

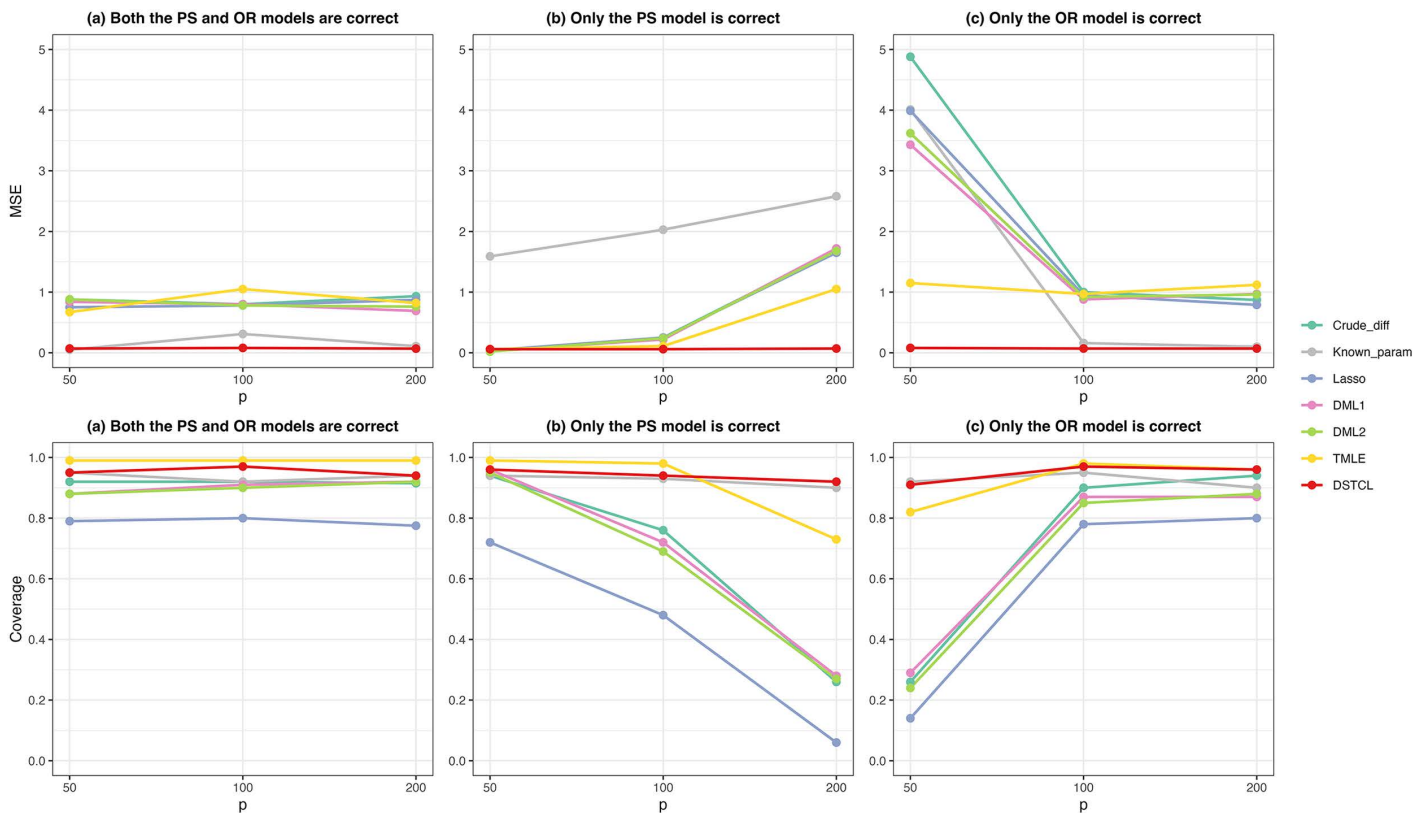


Fig 6. Line plots of MSE and Coverage for varying the number of covariates (Scenario 3). I_1 -TCL was omitted from the line plots because complete simulation and bootstrap results were not available across all settings. Available I_1 -TCL results are reported in the corresponding Supporting Information tables.

<https://doi.org/10.1371/journal.pcbi.1014706.g006>

Prior studies have shown that DNAm alterations differ systematically between cognitively normal individuals and AD patients and are associated with CSF biomarkers including ABeta42, suggesting that amyloid-related processes in the central nervous system are at least partly reflected in the peripheral epigenome [9]. Accordingly, to mitigate potential confounding, we included baseline genome-wide peripheral blood DNAm markers as candidate covariates and applied a standardized preprocessing pipeline: (i) excluding probes with detection P-values > 0.05; (ii) removing sex-chromosome and sex-related probes; (iii) discarding probes harboring single nucleotide polymorphisms (SNPs) at the CpG site; (iv) removing cross-reactive probes; and (v) averaging DNAm levels across repeated measurements for the same individual. After quality control, 865,859 CpG sites, age, gender, and education were retained. Given the extreme covariate dimensionality, we further conducted an epigenome-wide association study (EWAS) and selected the top 100 CpG sites ranked by Bonferroni-adjusted P-values to serve as the covariate set in the subsequent causal analysis.

The final analytical sample comprised 1,269 eligible participants, including 49 African American and 1,220 non-Hispanic White individuals. In the analysis, African Americans were defined as the target domain with a limited sample size, whereas non-Hispanic Whites were defined as the source domain with a substantially larger sample size. All participants had complete exposure and outcome data. For partially missing covariates, we applied chained-equation imputation with predictive mean matching, implemented with the mice package, to generate one completed analytic dataset. Building on our simulation results demonstrating the limited applicability and instability of conventional estimators, we compared Lasso, DML, TMLE, I_1 -TCL, and DSTCL methods. After adjusting for age, sex, years of education, and the selected CpG

sites, we estimated the causal effect of interest and its 95% confidence interval. The confidence intervals were obtained via bootstrap with 500 resamples.

4.2. Results

[Table 2](#) shows that the source and target domains were characterized using the same measured covariates, although their covariate distributions were not identical. African American participants were slightly younger and more often female, and they tended to have higher ABeta42 levels. We further assessed source-target comparability using standardized mean differences and distribution comparison tests for demographic, clinical, and selected DNAm covariates. These diagnostics showed that age was well balanced and education was acceptably balanced, whereas sex, ABeta42 status, and a subset of selected CpG covariates showed moderate source-target differences ([Table 1 in S1 Appendix](#)).

Consistent with the observed source-target comparability in measured covariates, principal component analysis of the CpG sites demonstrated substantial overlap between African American and non-Hispanic White participants in the PC1-PC2 space, with closely aligned group centers and largely overlapping confidence ellipses ([Fig 7](#)). This pattern indicates the presence of a shared global structure in the DNAm feature space used for modeling, despite potential population-specific variation at individual loci. Taken together, these diagnostics suggest that the two domains were not

Table 2. Characteristics of African American and Non-Hispanic White participants.

	African American (n=49)	Non-Hispanic White (n=1,220)	P value
Age, median (IQR)	62.0 (60.0-66.0)	63.0 (58.0-67.0)	0.744
Gender, %			0.009
Male	18 (36.7)	693 (56.8)	
Female	31 (63.3)	527 (43.2)	
Education, median (IQR)	16.0 (14.0-18.0)	16.0 (14.0-18.0)	0.189
ABeta42, median (IQR)	1001.0 (565.6-1605.0)	843.2 (594.2-1396.8)	0.568
ADAS-11, median (IQR)	7.0 (4.3-10.0)	9.0 (5.3-14.3)	0.025

<https://doi.org/10.1371/journal.pcbi.1014706.t002>

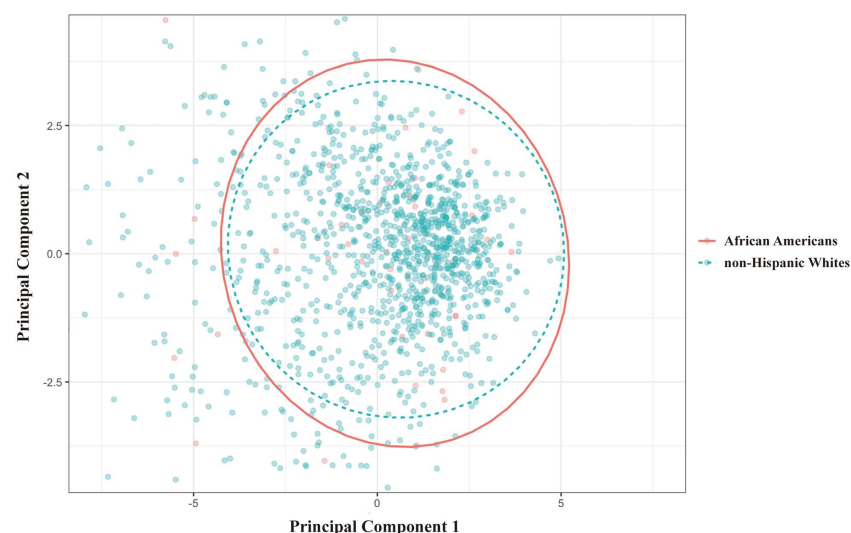


Fig 7. Results of principal component analysis of top 100 CpG sites in two populations.

<https://doi.org/10.1371/journal.pcbi.1014706.g007>

identical but shared a broadly comparable measured covariate space. Therefore, transfer learning was considered plausible in this application, while the results should be interpreted together with these source-target distributional diagnostics.

To evaluate the stability of the real-data DSTCL estimate in the small target sample, we examined propensity-score overlap and effective-sample-size diagnostics. In the primary analysis, the propensity scores in the target domain ranged from 0.30 to 0.52, with no values below 0.05 or above 0.95. The effective sample size was 46.78 out of 49, suggesting good overlap between exposure groups and limited instability from inverse probability weighting (**Table J in S1 Appendix**).

As shown in **Table 3**, all methods yielded broadly similar negative estimates of the effect of high versus low baseline ABeta42 status on ADAS-11 scores in African Americans under the stated causal assumptions. As an imputation sensitivity analysis, we repeated the DSTCL analysis using a simple single-imputation strategy, with median imputation for continuous variables and mode imputation for categorical variables. The estimate remained negative, consistent with the primary analysis, although the bootstrap 95% CI was wider under simple imputation, suggesting reduced stability with the simpler imputation strategy (**Table K in S1 Appendix**). Lower ABeta42 levels indicate greater amyloid burden. Therefore, these negative estimates should be interpreted according to the ABeta42 coding, such that lower ABeta42, corresponding to higher amyloid burden, is associated with higher ADAS-11 scores and worse cognitive performance. Furthermore, the target-only Lasso produced wide confidence intervals, reflecting substantial variability in the small target sample. By contrast, the other methods yielded more stable estimates with 95% confidence intervals that exclude the null, and the proposed DSTCL estimator achieved the narrowest 95% CI, indicating the highest precision among all approaches.

5. Discussion

In this study, we proposed a double shrinkage transfer causal learning method for causal effect estimation. Compared with existing approaches, DSTCL improves performance by simultaneously regularizing the initial source-domain estimation and the source-target difference parameters, thereby improving the accuracy and stability of parameter estimates in the target domain and yielding more reliable and stable causal effect estimates.

Through extensive simulation studies covering a range of data generating mechanisms, DSTCL achieved finite-sample performance close to that of an ideal oracle estimator with prior knowledge. Compared with other methods, DSTCL generally yielded the smallest Bias and MSE, together with relatively high coverage. In the real-data application, we applied DSTCL to the ADNI dataset to estimate the causal effect of baseline CSF ABeta42 status on 12-month ADAS-11 cognitive scores among African Americans, adjusting for age, sex, years of education, and a set of DNAm covariates. DSTCL and the compared methods consistently suggested that lower baseline CSF ABeta42, reflecting greater cerebral amyloid burden, was associated with worse 12-month cognitive performance among African Americans, supporting a positive relationship between amyloid pathology and cognitive impairment. Meanwhile, DSTCL produced the narrowest 95% confidence interval among all methods, indicating high statistical precision. From a biological perspective, reduced CSF ABeta42 levels typically indicate increased cerebral amyloid plaque deposition, which in turn triggers a cascade of downstream pathological changes and ultimately leads to progressive impairment of multiple cognitive domains, including memory, attention, and executive function.

Table 3. Causal effect estimates and 95% CIs among African American participants.

Methods	Estimate	95% CI
Lasso	-5.02	(-11.23, -0.50)
DML1	-7.56	(-12.29, -3.46)
DML2	-8.12	(-11.94, -3.49)
TMLE	-6.05	(-15.32, -2.15)
I ₁ -TCL	-5.24	(-9.13, -1.31)
DSTCL	-6.33	(-9.59, -2.48)

<https://doi.org/10.1371/journal.pcbi.1014706.t003>

Several limitations should be noted. Firstly, DSTCL relies on a sparse parameter-shift assumption, under which only a limited subset of model parameters differ between source and target domains. While this assumption appears reasonable in our setting, it may not hold in applications involving markedly different diseases or highly heterogeneous populations, and future work should explore more flexible notions of cross-domain similarity and corresponding penalty structures. In addition, because DSTCL applies l_1 regularization in both the source-domain estimation and source-target correction steps, double shrinkage may introduce a bias-variance trade-off; we mitigated this risk through cross-validation and regularization sensitivity analyses. Secondly, although DSTCL does not require identical marginal covariate distributions across domains, its performance still depends on sufficient source-target covariate overlap and the plausibility of sparse parameter differences. In the ADNI application, distributional diagnostics indicated both shared structure and measurable differences between African American and non-Hispanic White participants. Moreover, because the target domain included only 49 African American participants, finite-sample instability and overfitting cannot be completely ruled out, despite favorable propensity-score overlap and effective-sample-size diagnostics. Thirdly, establishing a complete formal theory for DSTCL remains challenging because the method combines transfer learning, two l_1 -regularized steps, and a doubly robust estimating equation. Existing studies suggest that such theory requires source-target similarity, sparsity, and careful control of regularized nuisance estimation errors [15,26,27]. Formal consistency and asymptotic normality results remain an important direction for future work. Finally, the current study focused on binary exposure and continuous outcomes and used an EWAS-based screening step to select CpG sites, which may not fully preserve the methylome's correlation structure. Future work should extend DSTCL to more general exposure and outcome types and explore alternative dimension-reduction strategies. Because the ADNI application was based on observational data, the estimated effects should also be interpreted cautiously and conditional on standard causal assumptions, as residual confounding may remain despite adjustment for measured covariates. Notwithstanding these limitations, our results indicate that DSTCL offers a practical and robust approach for causal effect estimation in biomedical studies.

In conclusion, DSTCL provides a stable transfer causal learning framework for causal effect estimation when the target population has a limited sample size. The simulation studies showed favorable finite-sample performance, and the ADNI application suggested that lower baseline CSF ABeta42, reflecting greater amyloid burden, was associated with worse 12-month cognitive performance among African American participants.

Supporting information

S1 Appendix. Supplementary simulation and real-data analyses. This appendix contains supplementary simulation methods, main simulation results, sensitivity analyses, computational diagnostics, and supplementary analyses of the ADNI data, including the following tables: Table A. Results across methods as the source-domain sample size varies (Scenario 1). Table B. Results across methods as h-transferability varies (Scenario 2). Table C. Results across methods as the covariate dimension varies (Scenario 3). Table D. Regularization sensitivity analysis for DSTCL. Table E. Sensitivity analysis under increasingly extreme true propensity scores. Table F. Propensity-score trimming sensitivity analysis for DSTCL. Table G. Covariate-shift sensitivity analysis for DSTCL. Table H. Computational time and convergence diagnostics for DSTCL as the covariate dimension increases. Table I. Variable-level source-target distribution diagnostics in the ADNI application. Table J. Propensity-score and effective-sample-size diagnostics in the ADNI analyses. Table K. DSTCL estimates from the ADNI imputation analyses.
(DOCX)

Author contributions

Conceptualization: Yunxin Shi, Lulu Pan, Yu Hu, Yongfu Yu, Guoyou Qin.

Data curation: Yunxin Shi, Lulu Pan.

Formal analysis: Yunxin Shi, Lulu Pan, Yu Hu.

Funding acquisition: Yongfu Yu, Guoyou Qin.

Investigation: Yunxin Shi, Lulu Pan, Yu Hu.

Methodology: Yunxin Shi, Lulu Pan, Yu Hu, Yongfu Yu, Guoyou Qin.

Project administration: Yongfu Yu, Guoyou Qin.

Resources: Yongfu Yu, Guoyou Qin.

Software: Yunxin Shi, Lulu Pan.

Supervision: Yongfu Yu, Guoyou Qin.

Validation: Yunxin Shi.

Visualization: Yunxin Shi.

Writing – original draft: Yunxin Shi, Lulu Pan.

Writing – review & editing: Yunxin Shi, Lulu Pan.

References

- Scheltens P, De Strooper B, Kivipelto M, Holstege H, Ch  telat G, Teunissen CE. Alzheimer's disease. *Lancet*. 2021;397(10284):1577–90.
- Matthews KA, Xu W, Gaglioti AH, Holt JB, Croft JB, Mack D, et al. Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged ≥ 65 years. *Alzheimers Dement*. 2019;15(1):17–24.
- Tosun D, Schuff N, Truran-Sacrey D, Shaw LM, Trojanowski JQ, Aisen P, et al. Relations between brain tissue loss, CSF biomarkers, and the ApoE genetic profile: a longitudinal MRI study. *Neurobiol Aging*. 2010;31(8):1340–54. <https://doi.org/10.1016/j.neurobiolaging.2010.04.030> PMID: 20570401
- Wang J, Fan D-Y, Li H-Y, He C-Y, Shen Y-Y, Zeng G-H, et al. Dynamic changes of CSF sPDGFR β during ageing and AD progression and associations with CSF ATN biomarkers. *Mol Neurodegener*. 2022;17(1):9. <https://doi.org/10.1186/s13024-021-00512-w> PMID: 35033164
- Morris JC, Schindler SE, McCue LM, Moulder KL, Benzinger TLS, Cruchaga C, et al. Assessment of Racial Disparities in Biomarkers for Alzheimer Disease. *JAMA Neurol*. 2019;76(3):264–73. <https://doi.org/10.1001/jamaneurol.2018.4249> PMID: 30615028
- 2021 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2021;17(3):327–406. <https://doi.org/10.1002/alz.12328> PMID: 33756057
- Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. Recent publications from the Alzheimer's Disease Neuroimaging Initiative: Reviewing progress toward improved AD clinical trials. *Alzheimers Dement*. 2017;13(4):e1–85. <https://doi.org/10.1016/j.jalz.2016.11.007> PMID: 28342697
- Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, et al. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement*. 2005;1(1):55–66. <https://doi.org/10.1016/j.jalz.2005.06.003> PMID: 17476317
- Zhang W, Young JI, Gomez L, Schmidt MA, Lukacsovich D, Varma A, et al. Distinct CSF biomarker-associated DNA methylation in Alzheimer's disease and cognitively normal subjects. *Alzheimers Res Ther*. 2023;15(1):78. <https://doi.org/10.1186/s13195-023-01216-7> PMID: 37038196
- Shireby GL, Davies JP, Francis PT, Burrage J, Walker EM, Neilson GWA, et al. Recalibrating the epigenetic clock: implications for assessing biological age in the human cortex. *Brain*. 2020;143(12):3763–75. <https://doi.org/10.1093/brain/awaa334> PMID: 33300551
- Galanter JM, Gignoux CR, Oh SS, Torgerson D, Pino-Yanes M, Thakur N, et al. Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *Elife*. 2017;6:e20532. <https://doi.org/10.7554/eLife.20532> PMID: 28044981
- Zhang L, Silva TC, Young JI, Gomez L, Schmidt MA, Hamilton-Nelson KL, et al. Epigenome-wide meta-analysis of DNA methylation differences in prefrontal cortex implicates the immune processes in Alzheimer's disease. *Nat Commun*. 2020;11(1):6114. <https://doi.org/10.1038/s41467-020-19791-w> PMID: 33257653
- Rubin DB. Causal inference using potential outcomes: design, modeling, decisions. *J Am Stat Assoc*. 2005;100(469):322–31.
- Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics*. 2005;61(4):962–73. <https://doi.org/10.1111/j.1541-0420.2005.00377.x> PMID: 16401269
- Chernozhukov V, Chetverikov D, Demirer M, Duflo E, Hansen C, Newey W. Double/debiased machine learning for treatment and structural parameters. *Econom J*. 2018;21(1):C1–68.
- Schuler MS, Rose S. Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies. *Am J Epidemiol*. 2017;185(1):65–73. <https://doi.org/10.1093/aje/kww165> PMID: 27941068

17. Hosna A, Merry E, Gyalmo J, Alom Z, Aung Z, Azim MA. Transfer learning: a friendly introduction. *J Big Data*. 2022;9(1):102. <https://doi.org/10.1186/s40537-022-00652-w> PMID: [36313477](#)
18. Kim HE, Cosa-Linan A, Santhanam N, Jannesari M, Maros ME, Ganslandt T. Transfer learning for medical image classification: a literature review. *BMC Med Imaging*. 2022;22(1):69. <https://doi.org/10.1186/s12880-022-00793-7> PMID: [35418051](#)
19. Künzel SR, Stadie BC, Vemuri N, Ramakrishnan V, Sekhon JS, Abbeel P. Transfer learning for estimating causal effects using neural networks. 2018. <https://arxiv.org/abs/1808.07804>
20. Wei S, Zhang H, Moore R, Kamaleswaran R, Xie Y. Transfer learning for causal effect estimation. *arXiv*. 2023. <https://arxiv.org/abs/2305.09126>
21. Pearl J. An introduction to causal inference. *Int J Biostat*. 2010;6(2):Article 7. <https://doi.org/10.2202/1557-4679.1203> PMID: [20305706](#)
22. Saykin AJ, Shen L, Yao X, Kim S, Nho K, Risacher SL, et al. Genetic studies of quantitative MCI and AD phenotypes in ADNI: Progress, opportunities, and plans. *Alzheimers Dement*. 2015;11(7):792–814. <https://doi.org/10.1016/j.jalz.2015.05.009> PMID: [26194313](#)
23. Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology*. 2010;74(3):201–9. <https://doi.org/10.1212/WNL.0b013e3181cb3e25> PMID: [20042704](#)
24. Jack CR Jr, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging*. 2008;27(4):685–91. <https://doi.org/10.1002/jmri.21049> PMID: [18302232](#)
25. Hansson O, Seibyl J, Stomrud E, Zetterberg H, Trojanowski JQ, Bittner T, et al. CSF biomarkers of Alzheimer's disease concord with amyloid- β PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts. *Alzheimers Dement*. 2018;14(11):1470–81. <https://doi.org/10.1016/j.jalz.2018.01.010> PMID: [29499171](#)
26. Li S, Cai TT, Li H. Transfer Learning for High-Dimensional Linear Regression: Prediction, Estimation and Minimax Optimality. *J R Stat Soc Series B Stat Methodol*. 2022;84(1):149–73. <https://doi.org/10.1111/rssb.12479> PMID: [35210933](#)
27. Tian Y, Feng Y. Transfer Learning under High-dimensional Generalized Linear Models. *J Am Stat Assoc*. 2023;118(544):2684–97. <https://doi.org/10.1080/01621459.2022.2071278> PMID: [38562655](#)