

A One-shot Federated Learning algorithm to quantify racial disparities in graft failure rates with kidney transplant using US registry data from 35,497 patients across 201 hospitals

Authors: Jiayi Tong^a, Yishan Shen^b, Alice Xu^c, Chongliang Luo^d, Mackenzie Edmondson^e, Ruowang Li^f, Di Wang^g, Kevin He^g, David A. Asch^{h,i}, Yong Chen^a

a. Perelman School of Medicine, The University of Pennsylvania, Philadelphia, PA, USA

b. Applied Mathematics and Computational Science, The University of Pennsylvania, Philadelphia, PA, USA

c. St. John's School, Houston, TX, USA

d. Division of Public Health Sciences, Department of Surgery, Washington University in St. Louis, St. Louis, MO, USA

e. Merck, Philadelphia, PA, USA

f. Department of Computational Biomedicine, Cedars-Sinai Medical Center, Los Angeles, CA, USA

g. Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA

h. Division of General Internal Medicine, University of Pennsylvania, Philadelphia, PA, USA

i. Leonard Davis Institute of Health Economics, Philadelphia, PA, USA

Background

Kidney transplant is a renal replacement therapy for eligible patients with end-stage renal disease (ESRD)¹⁻³. Unfortunately, the racial disparities in receipt of a transplanted kidney are observed for the Black across states⁴. Black patients are recognized to have lower graft survival rates compared with White patients⁵⁻⁸. Therefore, there is a great need in understanding, quantifying, and reducing the racial disparities in access to transplant and post-transplant outcomes for kidney transplant patients.

Site of care has been considered as a major contributor to disparities in kidney transplants due to differences in time on the transplant waiting list, access to live donor kidney transplants, care coordination with the donor organ procurement system, and acute rejection rates⁹. Since racial groups tend to live in certain areas, they may be subject to varying transplant care, leading to the difference in the outcomes (e.g., kidney graft failure)¹⁰.

To study the potential association between the site of care and racial disparity in kidney transplant graft failure with multi-site data, a counterfactual modeling method can be implemented¹¹. By considering the underlying distribution in which patients of different races receive care, we can use a counterfactual model to simulate patients from a certain transplant center getting admitted to another. For example, with the proportions of White patients at the studied centers, we can simulate the reassignments for Black patients with the same proportions within the hospitals hypothetically. Under this scenario, the Black patients are pretended to attend the hospitals in the same distribution as the White patients. The observed and counterfactual kidney graft failure rates can be calculated, and the hypothesis is that the counterfactual graft failure rates would be lower than the observed graft failure rates for the Black patients if there exist site-of-care-associated racial disparities for the Black patients in

kidney transplants.

This counterfactual modeling has been recently implemented in the study by Asch et al (2020)¹² to explore racial disparities for patients hospitalized with COVID-19 infection. In this study, a Health Quality Forum certified generalized linear mixed model (GLMM) was fitted on a centralized multi-site data to characterize the odds of the event with the adjustments for both fixed (e.g., the effects of patient- and hospital-level factors) and random effects (e.g., hospital-specific effects). In studying the effect of site of care on an event rate (e.g., COVID-19 mortality rate, kidney graft failure rate), the hospital-specific effects are essential in the calculation of counterfactual event rate. However, when the multi-site data cannot be centralized, the utilization of centralized modeling is not achievable, leading to challenges in estimating the hospital-specific effects. When the patient-level data are stored in a decentralized format and only aggregated data are allowed to be shared across hospitals, a decentralized algorithm for GLMM is in critical need.

We proposed a framework of federated learning algorithm to investigate the association between the site of care and racial disparities in kidney graft failure for Black patients. We termed this framework as dGEM-disparity, which stands for decentralized algorithm for Generalized linear mixed Effect Model for disparity quantification. dGEM-disparity is a federated learning framework to effectively integrate heterogeneous multi-site data by fitting generalized mixed effect models. It allows us to investigate the site-of-care-associated racial disparity by considering the patients who are hypothetically admitted to other hospitals.

Methods

There are three main steps in dGEM-disparity as shown in **Figure 1**. Let K denote the number of total hospitals. In Step I, each of the K sites estimates the common effect $\hat{\beta}_k$ and variance $I_k^{-1}(\hat{\beta}_k)$ with its own patient-level data. The common effects are combined to obtain a meta-estimate $\hat{\beta}$ of the common effect. The meta-estimate $\hat{\beta}$, then, is broadcasted to each site for the estimation of hospital effect $\hat{\gamma}_k^*$ and its variance \hat{s}_k^* . In Step II, a hospital-level calibration through meta-regression is conducted with hospital-level characteristics Z_k within each hospital and the estimated hospital effects $\{\hat{\gamma}_1^*, \dots, \hat{\gamma}_K^*\}$ are shared among all hospitals. In Step III, each site calculates the counterfactual event rates and the calibrated hospital effects $\{\hat{\gamma}_1^{**}, \dots, \hat{\gamma}_K^{**}\}$ for disparity quantification.

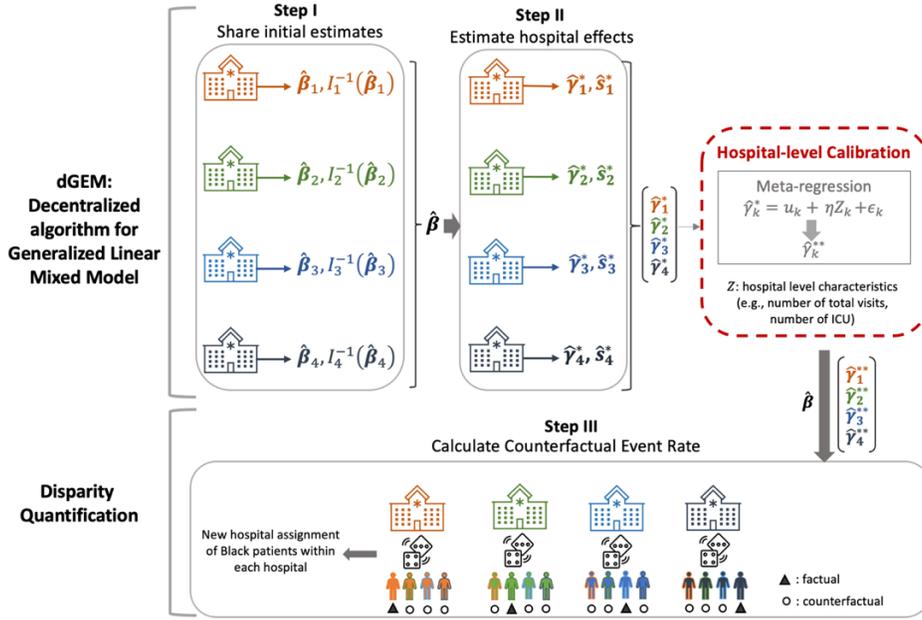


Figure 1. Workflow of the proposed decentralized algorithm for generalized mixed effect models framework for disparity quantification -- dGEM-disparity. This framework including three main steps: initial estimation of the common effect (i.e., association between outcome and covariates), estimation of hospital-effects with the meta-estimator combined from Step I, following with a hospital-level calibration, and the calculation of the counterfactual rates for disparity quantification.

The event rate in Step III can be calculated as

$$\hat{p}_k(\mathbf{X}_{kj}) = \text{expit}(\hat{\gamma}_k^{**} + \mathbf{X}_{kj}^T \hat{\beta}),$$

which is a factual event rate if $k = k^*$, and a counterfactual event rate if $k \neq k^*$ (i.e., pretending the patient from hospital k^* are admitted to hospital k). In the distributed format where patient-level data (i.e., \mathbf{X}_{kj}) cannot be shared across site, the calculation of the counterfactual event rate is achieved by transferring hospital-specific effects $\hat{\gamma}_k^{**}$ to protect the patient-level information.

The major aim of developing dGEM-disparity is to investigate what if the Black patients had been admitted to different hospitals for kidney transplant as the White patients. Therefore, we conducted a simulation with the centralized data to facilitate the above idea. In the simulation, we supposed that the Black patients had been admitted to hospitals in the same distribution as White patients while retaining their sociodemographic and clinical characteristics. The simulation procedure is depicted in a high-level conception in **Figure 2**. The proportions of White patients for all hospitals are calculated as $\{\hat{w}_1, \dots, \hat{w}_K\}$. Within each hospital, the Black patients are randomly assigned to a hospital by a multinomial distribution with probabilities $\{\hat{w}_1, \dots, \hat{w}_K\}$. Hospital assignments are then used to calculate mortality risk estimates for each Black patient in their (counter)factual hospital, as shown in **Figure 2**. Within each hospital, the patient-level rate estimates of the Black patients with newly assigned hospital are summed up, and the overall counterfactual event rate estimate is obtained using these sums. The simulation procedure is replicated multiple times to quantify the uncertainty.

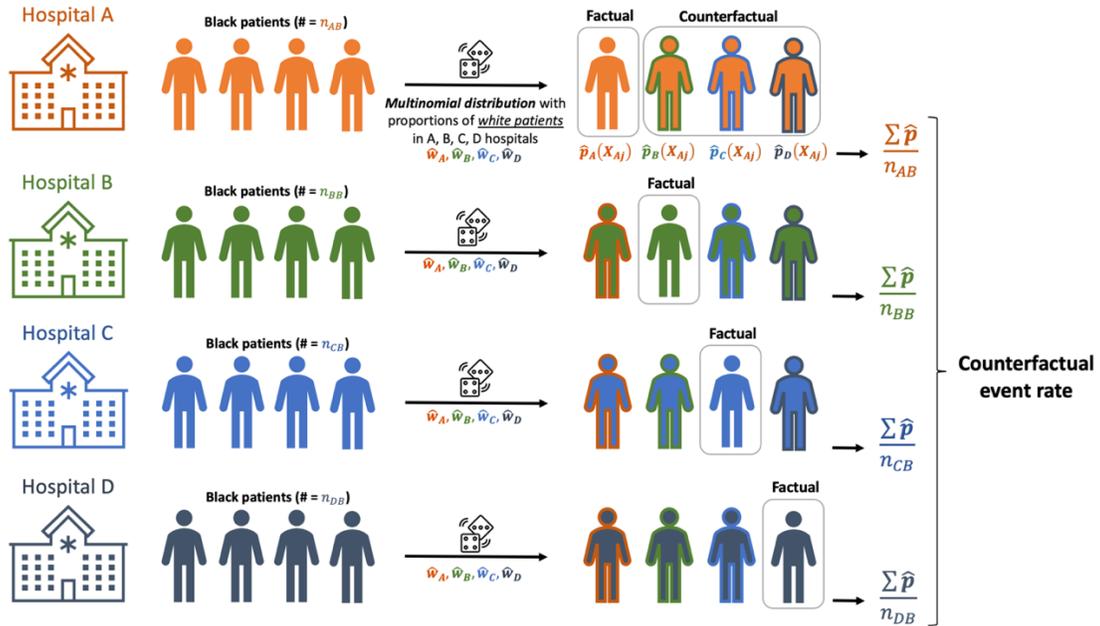


Figure 2. Overview of the calculation of counterfactual kidney transplant failure risks. Within each hospital, the black patients are reassigned with a hospital based on a multinomial distribution with the proportions of White patients in the hospitals. (Counter)factual rates are calculated within each site for Black patients, and then summed together to get the overall counterfactual event rate.

Results

The following Figure 3 summarizes the results using the registry data collected from the U.S. Organ Procurement and Transplantation Network (OPTN).

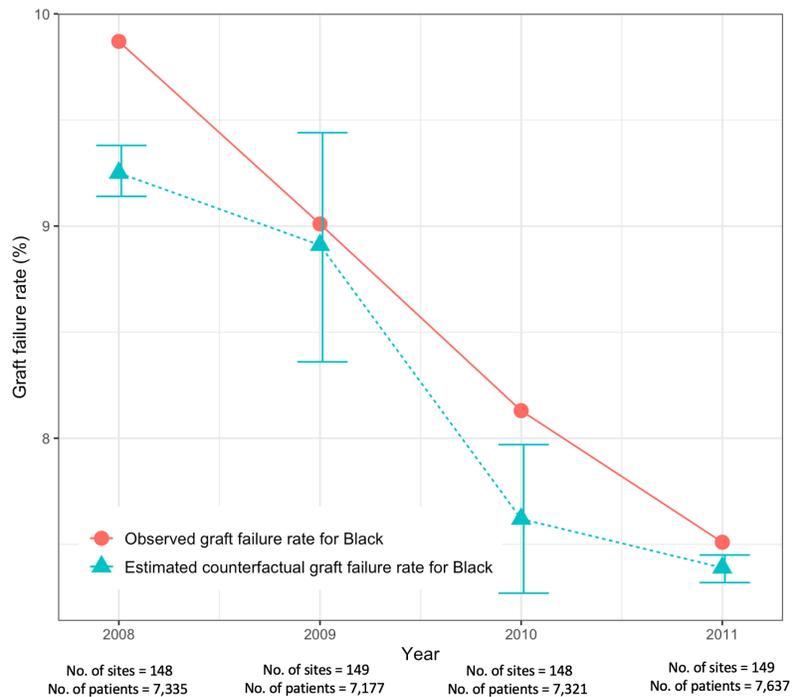


Figure 3. Results of the racial disparity in kidney transplant graft failure rate study with US registry data. Two rates are compared within each year (from 2008 to 2011), including the observed rate for Black patients, and the estimated *counterfactual* event rate for Black patients with the proposed dGEM-disparity framework. At the bottom of each year, the total numbers of sites and patients for analysis are listed.

With the proposed framework, where the Black patients were reassigned to the centers as the distributions of White patients, the estimated counterfactual graft failure rates (blue triangle) of the Black patients *consistently* declined compared with the observed rates (red circle) across four consecutive years.

Conclusion

We presented the real-world application of the proposed dGEM-disparity framework. Our framework provides the estimates of fixed and random effects to further obtain the counterfactual event rates to investigate the association between the attending hospital and racial disparities. The results of the data analysis show if the Black patients in the study cohort are admitted to hospitals with the same distributions as the White patients, the event risk would be reduced. dGEM-disparity can be generalized to investigate other mediation effects associated with access to healthcare.

References/Citations

- 1 *Kidney Transplant* | NIDDK. n.d. URL: <https://www.niddk.nih.gov/health-information/kidney-disease/kidney-failure/kidney-transplant> (Accessed May 27, 2022).
- 2 *Kidney transplantation* - Wikipedia. n.d. URL: https://en.wikipedia.org/wiki/Kidney_transplantation (Accessed May 27, 2022).
- 3 *Kidney transplant* - Mayo Clinic. n.d. URL: <https://www.mayoclinic.org/tests-procedures/kidney-transplant/about/pac-20384777> (Accessed May 27, 2022).
- 4 Kulkarni S, Ladin K, Haakinson D, Greene E, Li L, Deng Y. Association of racial disparities with access to kidney transplant after the implementation of the new kidney allocation system. *JAMA Surg* 2019;**154**:618–25.
- 5 Malek SK, Keys BJ, Kumar S, Milford E, Tullius SG. Racial and ethnic disparities in kidney transplantation. *Transplant International* 2011;**24**:419–24.
- 6 Hariharan S, Israni AK, Danovitch G. Long-term survival after kidney transplantation. *New England Journal of Medicine* 2021;**385**:729–43.
- 7 Hariharan S, Johnson CP, Bresnahan BA, Taranto SE, McIntosh MJ, Stablein D. Improved graft survival after renal transplantation in the United States, 1988 to 1996. *New England Journal of Medicine* 2000;**342**:605–12.
- 8 Patzer RE, Mohan S, Kutner N, McClellan WM, Amaral S. Racial and ethnic disparities in pediatric renal allograft survival in the United States. *Kidney Int* 2015;**87**:584–92.
- 9 Hasnain-Wynia R, Baker DW, Nerenz D, Feinglass J, Beal AC, Landrum MB, *et al.* Disparities in health care are driven by where minority patients seek care: examination of the hospital quality alliance measures. *Arch Intern Med* 2007;**167**:1233–9.
- 10 Liu A, Woodside KJ, Augustine JJ, Sarabu N. Racial disparity in kidney transplant survival relates to late rejection and is independent of steroid withdrawal. *Clinical Transplantation* 2018;**32**:e13381.
- 11 George EI, Ročková V, Rosenbaum PR, Satopää VA, Silber JH. Mortality rate estimation and standardization for public reporting: Medicare’s hospital compare. *J Am Stat Assoc* 2017;**112**:933–47.
- 12 Asch DA, Islam MN, Sheils NE, Chen Y, Doshi JA, Buresh J, *et al.* Patient and hospital factors associated

with differences in mortality rates among Black and White US Medicare beneficiaries hospitalized with COVID-19 infection. *JAMA Netw Open* 2021;4:e2112842–e2112842.