

ANTI-RACISM METHODS FOR BIG DATA RESEARCH: LESSONS LEARNED FROM THE HIV TESTING, LINKAGE, & RETENTION IN CARE (HIV TLR) STUDY

Chandra L. Ford, PhD, MPH^{1,2}; Lois M. Takahashi, PhD, MS^{1,5};
P. Paul Chandanabhumma, MPH^{1,2}; Maria Elena Ruiz, PhD^{1,4};
William E. Cunningham, MD, MPH^{1,3}

Public Health Critical Race Praxis (PHCRP) contributes three functional elements to health equity studies: a race conscious orientation; an antiracism lexicon based on Critical Race Theory (CRT); and an integrated, reflexive approach. Few big data studies employ all three functional elements. Therefore, this article describes the application of PHCRP to the *Human Immunodeficiency Virus Testing, Linkage and Retention in care (HIV TLR)* study (N=3,476,741), which connects multiple large datasets to electronic medical records to examine contextual determinants of racial/ethnic disparities in HIV care continuum outcomes in southern California. As HIV TLR demonstrates, PHCRP's innovative tools and strategies help big data research maintain fidelity to CRT. *Ethn Dis.* 2018; 28 (Suppl 1):261-266; doi:10.18865/ed.28.S1.261

Keywords: Critical Race Theory; Health Services Research; Social Epidemiology; Methods; Electronic Medical Records; Big Data

¹Center for the Study of Racism, Social Justice & Health, Fielding School of Public Health at UCLA, Los Angeles, CA

²Department of Community Health Sciences, Fielding School of Public Health at UCLA, Los Angeles, CA

³Department of Health Services Research & General Internal Medicine, Geffen School of Medicine at UCLA, Los Angeles, CA

⁴School of Nursing, University of California at Los Angeles (UCLA), Los Angeles, CA

⁵Price School of Public Policy, University of Southern California, Los Angeles, CA

INTRODUCTION

As detailed elsewhere in this special issue, Public Health Critical Race Praxis (PHCRP) tailors Critical Race Theory (CRT), which is the anti-racism scholarship approach that originated in the field of Law, for health equity research. PHCRP contributes three functional elements to the research process: a race conscious orientation; an antiracism lexicon borrowed primarily from CRT; and a systematic, reflexive approach (self-critique).¹ PHCRP's four foci (contemporary racial relations; knowledge production; conceptualization and measurement; and action) represent key domains in which racialization may influence research. In most critical race projects, researchers draw on the racial consciousness or lexical elements to analyze qualitative data. Their findings illuminate the salience of racism

Address correspondence to Chandra L. Ford, PhD, MPH; Associate Professor, Department of Community Health Sciences; Founding Director, *Center for the Study of Racism, Social Justice & Health*; Jonathan & Karin Fielding School of Public Health; University of California at Los Angeles; Box 951772; 650 Charles E. Young Dr., South; Los Angeles, CA 90095-1772. 310.825.6063; clford@ucla.edu

and its potential health implications among socially marginalized populations. Social epidemiology and health services research increasingly use electronic medical records (EMRs) and other big data to estimate associations between social inequalities such as racism and health disparities; however, they rarely employ all three functional elements of PHCRP. Doing so presents unique challenges, principal among which are integrating anti-racism into the research process.

We address these knowledge gaps by describing the application of PHCRP to the Human Immunodeficiency Virus Testing, Linkage and Retention in care (HIV TLR) study, which we believe can inform social epidemiology, health services research and other studies involving large datasets. With each new project, public health critical race theorists (ie, healthcrits) seek to understand, "How is racialization operating here?" Racialization refers to the mechanisms that produce inequalities in populations and ascribe racial meanings to them. In addition to influencing the exposure disease relationship being studied, racialization may influence research processes and the conclusions drawn from the research.^{2,3} This article outlines how we addressed

three PHCRP foci (contemporary racial relations; knowledge production; conceptualization and measurement), and how we integrated the PHCRP principles of race consciousness and critical self-reflection into the research process. We conclude with lessons learned and implications for future PHCRP research.

THE HIV TLR STUDY

Background

Detection of and treatment for HIV infection have improved over the past decade, yet fewer than 50% of all people living with HIV (PLWH) have received a diagnosis, entered HIV care and achieved viral suppression (VS), ie, <200 mL copies of HIV RNA.⁴ The US HIV Care Continuum (HCC) initiative was launched in response to this trend. It guides and monitors progress toward the nation's 2020 HIV prevention goals to increase the proportions of PLWH who know their HIV status, are engaged in comprehensive HIV care, and have achieved viral suppression.⁵ The HCC outcomes represent the five sequential steps PLWH must take to achieve viral suppression. The first step is HIV testing to detect infection early. For those who test positive, the next steps are prompt linkage to HIV care, retention in care, and adherence to prescribed medication regimens.⁶

US native- and foreign-born Blacks (henceforth, African Americans) experience disparities along the HCC, including delayed receipt of diagnoses and poorer survival rates than members of other groups.⁷

These disparities persist even when African Americans engage in risk behaviors at similar rates as Whites,^{8,9} or engage in preventive behaviors (eg, HIV testing) at higher rates than Whites.¹⁰ While most studies examining the disparities target individual-level compositional factors (eg, attitudes, behaviors), upstream contextual factors (eg, segregation) may offer better explanations

*This article outlines
how we addressed
three PHCRP foci
(contemporary
racial relations;
knowledge production;
conceptualization and
measurement), and
how we integrated the
PHCRP principles of race
consciousness and critical
self-reflection into the
research process.*

for the persistent disparities.¹¹⁻¹³

To understand how racialization might contribute to these disparities, the HIV TLR study: 1) draws on conceptualization and measurement (PHCRP Focus 3) to reframe "race" as a social construct rather than a biological attribute; 2) esti-

mates contemporary racial relations (PHCRP Focus 1) by highlighting dimensions of the social context of care relative to disparities in HCC outcomes and VS; and 3) integrates race consciousness and critical self-awareness (PHCRP Functional Elements 1 and 3) by assessing the *a priori* assumptions of the research team.

Conceptual Model

The Behavioral Model for Vulnerable Populations,¹⁴ a health services research framework, is the backbone of the study's quantitative analytic model. It posits that environmental factors within and outside of the health care setting, and individual-level predisposing, enabling and need factors influence patient behaviors in health care settings, which in turn influence their clinical outcomes. The CRT principles, racial realism and the primacy of racism, guided the adaptation of this model. According to our model, even in settings such as primary care where care is ostensibly colorblind and available to all patients equally, racialization operates across every socioecologic level in level-specific ways. Drawing on Gilmore's definition of racism as "the state-sanctioned and/or extralegal production and exploitation of group-differentiated vulnerability to premature death,"¹⁵ we conceptualize structural racism (ie, embedded in culture and context, and difficult to discern at the individual level) and institutional racism (ie, embedded in policy and program practices) as reinforcing and sustaining HCC disparities.¹⁶ Racism manifests as systematic racial/ethnic differences in access, care or outcomes.

Table 1. Integrating Public Health Critical Race Praxis (PHCRP) into the research process

PHCRP Focus	Approach	Timeline
Race conscious orientation	<ol style="list-style-type: none"> 1. Group discussions eliciting racial salience of colorblind concepts 2. Study team completed confidential questionnaires assessing <i>a priori</i> assumptions before each study aim 3. Group discussions regarding racial considerations in the data, variables, analyses, and interpretations 	<ol style="list-style-type: none"> 1. Year 1 onward 2. Year 1 at initial study meeting; thereafter, prior to each analysis 3. Years 2-4 continuously
Focus 1. Contemporary racial relations	<ol style="list-style-type: none"> 1. Framed the study aims and research questions in proposal 2. Interdisciplinary literature reviews and reports in the lay media 	<ol style="list-style-type: none"> 1. Prior to initiating study 2. Year 1 onward
Focus 2. Knowledge production	<ol style="list-style-type: none"> 1. Written into proposal as critique of conventional focus 2. Considered CRT-informed hypotheses post-hoc alternatives 	<ol style="list-style-type: none"> 1. Prior to initiating study 2. After each analysis (ongoing)
Focus 3. Concepts and measures	<ol style="list-style-type: none"> 1. Written into proposal based on prior research and CRT 2. Extensive, critical exploration of available measures 3. Qualitatively confirmed the measures capture CRT constructs 	<ol style="list-style-type: none"> 1. Prior to initiating study 2. Years 2-3 - After the exploratory portion of each analysis
Focus 4. Evidence-based action	<ol style="list-style-type: none"> 1. Formed a community advisory board to ensure the relevance of findings for diverse communities, inform dissemination of findings, and plan future research 	<ol style="list-style-type: none"> 1. Established in Year 3 (ongoing)

DESIGN AND METHODS

This was a retrospective cohort study conducted among adults enrolled in a large managed care organization (MCO) in southern California. We examined HIV testing-related outcomes among primary care patients and post-HIV diagnosis outcomes (ie, linkage to HIV/AIDS care, retention in care, medication adherence, and VS) among patients newly diagnosed as HIV-positive during the study period.

Population and Samples

The primary care sample (N=3,476,741) comprised all adult MCO enrollees presenting for routine preventive care (ie, primary care) at least once during the study

period, 2008-2012. Overall inclusion criteria were aged ≥ 18 years at the index visit, which was either the last visit at which HIV testing occurred or, if no test, simply the last visit in a calendar year; presenting for care at least once in a calendar year; and, no prior positive HIV test result. To focus on racial rather than socioeconomic determinants, the sample was limited to persons continuously enrolled in the MCO for at least 12 months prior to the index visit.

The HIV care sample comprised all adult MCO enrollees newly diagnosed with HIV infection during the study and presenting for HIV care at least once during this period. For this sample, only six months of continuous enrollment in the MCO prior to the index visit, which was the

first HIV care visit following HIV diagnosis, was required. Different HIV screening guidelines apply to pregnant women; therefore, women who were pregnant in the past 12 months were excluded from both samples.

Data

Data on each patient's medical history and assigned providers were extracted from the EMRs. Detailed professional and demographic information on the providers were obtained from the MCO's employee files. Data on health care context were derived from the EMRs, the census, and other area-based data sources. To maintain patient confidentiality, researchers at the MCO created study-specific codes for each patient, linked the multiple data-

sets, removed all personal identifiers, and conducted quality control checks before transferring the data beyond the firewall for the analysis.

Statistical Analyses

We used random effects multi-level modeling to estimate the effects of the hypothesized racialized context of care variables relative to each HCC outcome, while controlling for patients' self-reported race/ethnicity, other demographic and clinical factors (eg, recent STD diagnosis), and non-racial/ethnic contextual factors (eg, practice characteristics). Each analysis explored potential effect measure modification (ie, interaction) and confounding by racial/ethnic and ostensibly non-racial, socioeconomic variables, and accounted for any observed non-independence.

IMPLEMENTING PHCRP

To examine the hypothesized racialized context of care and HCC disparities, the study's principal investigator (PI) incorporated PHCRP constructs into the research proposal and conceptual model. A multidisciplinary, racially/ethnically diverse research team, which included physicians, nurses, social scientists and others, was established based on content area and methodological expertise. Researchers do not have to be racial/ethnic minorities in order to use PHCRP; our team included investigators with varying levels of expertise in CRT and experience living in social margins. We began integrating the four-stage process (Table 1) while writing the proposal (eg, in develop-

ing the research questions). In year 1, we initiated racial consciousness raising. During regular team meetings, we drew on extant literature, theory, and experiential knowledge explicitly drawn from our diverse backgrounds to refine the criteria and protocols for abstracting data from patients' medical records, census data and other geocoded datasets, and compiling the complex datasets. These were critically engaged, self-reflexive discussions focused on clarifying potential racial considerations.

Critical Self-Reflection and the Social Construction of Knowledge: Clarifying Prior Assumptions

Although the systematic nature of science enhances its reliability, this does not automatically render it bias-free.¹⁻³ According to PHCRP, researchers' *a priori* assumptions may unwittingly influence a study's approaches (eg, how variables are operationalized) and conclusions (eg, how racial effects are interpreted) if the social construction of knowledge is not recognized.³ Although the HIV TLR's hypotheses were evidence-based, individuals might nevertheless endorse findings more readily if the findings align with their *a priori* assumptions. To apply the functional element of critical self-reflection, each member of the study team completed brief, confidential questionnaires assessing their own potential *a priori* assumptions relative to each set of research questions at the initial meeting and immediately before each study aim. The items assessed the kinds of racial/ethnic patterns, if any, they ex-

pected to see in the findings, as well as their certainty about and bases for these expectations (eg, content area expertise, the study's findings to date, just a "hunch"). Each person self-administered the paper and pencil (Aim 1) and online (Aims 2 and 3) questionnaires. As explained at the time, there were no right or wrong answers and individual results were not disclosed to other team members. A non-research staff member collected the completed questionnaires and filed them confidentially for future use. Upon completion of the analyses (which are ongoing), we plan to use these priors as qualitative checks to enhance the rigor and transparency with which we interpret the data and derive conclusions.

Operationalizing "Race" as a Social Construct

According to our model, patient outcomes reflect, in part, racial dynamics in health care contexts that are not necessarily apparent to providers and patients.¹⁷ We operationalized "race" as a social construct to capture social ways race is meaningful (PHCRP Focus 3: Conceptualization and Measurement). This challenges most prior research, which treats it as an individual trait potentially causing the disparities (PHCRP Focus 2: Knowledge Production). We hypothesized that the care any individual patient receives is influenced not only by patient factors, including clinical indicators, but also by the salience of race/ethnicity in the context of care (PHCRP Focus 1: Contemporary Racial Relations). Thus, the racial/ethnic composition of a provider's patient population may constitute a ra-

cializing frame¹⁸ through which providers (consciously or unconsciously) perceive, assess, diagnose, and treat a patient. To measure this dimension of the racial/ethnic context of care, we first identified the provider and clinic of record for each patient and obtained the reported race/ethnicity of every patient seen by that provider or at that clinic from baseline until the date of the patient's index visit, then calculated the cumulative proportions of patients seen. The social meanings (and HCC implications) of race/ethnicity differ for Whites, African Americans, and Latinos/as; therefore, we calculated the cumulative proportions of African American, Latino/a, and White patients, respectively, in each provider's patient population over the specified period. The multi-level statistical models controlled for each patient's race/ethnicity and other covariates, while estimating the contribution of this "race" as social construct variable relative to each HCC outcome and VS. Our results indicate whether, regardless of a patient's self-reported race/ethnicity, an outcome varies based on the racial/ethnic composition of the patient population in which the patient obtains care, controlling for known determinants (eg, recent STI/STD diagnosis).

LESSONS LEARNED

The HIV TLR study helps establish the feasibility of PHCRP's three functional elements and four foci for social epidemiology, health services research and other studies involving big data. PHCRP introduces methods for integrating racial con-

sciousness into the research process, identifying implicit assumptions, which studies rarely do explicitly,³ and operationalizing abstract racial concepts such as "race" as a social construct or the social construction of knowledge. These steps enable healthcrits to respond empirically to the question, "How might racialization contribute to HCC disparities?"

Although the study's hypotheses were evidence-based, some team members expressed greater confidence in hypotheses aligned more closely with their *a priori* assumptions. Fortunately, assessing our priors enables us to bring specific assumptions to light and account for them in drawing conclusions.

There were several challenges. Using medical records to capture racism effects is complicated by the main purpose of EMR data, which is to facilitate diagnosis and care based on biological, not social, constructions of personhood, including race/ethnicity. Further, EMRs inherently reflect an organization's infrastructure, established guidelines, and budgetary constraints. We circumvented some of these challenges by supplementing the EMR data with data from other sources (eg, census) and constructing additional variables (eg, patient-provider racial/ethnic concordance), though these strategies did not address the challenges (eg, difficulty matching observations across datasets) that arise from using linked datasets.

While nearly all team members completed the self-administered questionnaires as instructed (eg, avoided doing literature searches to answer the items), a few feared reporting a "wrong" answer. Such as-

essment is critical and necessary in PHCRP research, but some researchers may need multiple experiences to become accustomed to this.

There were important strengths. The approaches used to operationalize PHCRP concepts quantify the racial context of care in a large dataset, and identify *a priori* assumptions representing novel, theory-informed contributions to the field. Logistically, using web-based software was an efficient way to survey members of the research team confidentially. While subjectivities are implicit in all research,² assessing them empirically enables us to identify and account for them. In fact, examining research team members' assumptions as part of the study is one of the most unique and innovative aspects of this application of PHCRP.

The study's next steps include completing the analyses, evaluating the assumptions relative to the results, and engaging the Community Advisory Board (CAB) to disseminate the findings¹⁹ and develop an evidence-based action plan (PHCRP Focus 4: Action).

CONCLUSIONS

A paradigm shift is underway.²⁰ PHCRP contributes innovative tools and strategies for conducting big data research with fidelity to CRT. The HIV TLR study demonstrates the feasibility of PHCRP for conceptualizing racial constructs, measuring racism effects, and identifying implicit assumptions.

ACKNOWLEDGMENTS

This work was funded by the National Institute of Nursing Research

5R01NR14789-03. We thank Kia Skrine Jeffers, Chi-hong Tseng as well as Rulin Hechter and colleagues at Kaiser Permanente Southern California Research Division for comments on an early draft. We thank Vanessa Medina for administrative assistance. The views expressed in this article are those of the authors and may not reflect those of the National Institutes of Health.

CONFLICT OF INTEREST

No conflicts of interest to report.

AUTHOR CONTRIBUTIONS

Research concept and design: Ford, Takahashi, Cunningham; Acquisition of data: Ford, Takahashi, Cunningham; Data analysis and interpretation: Ford, Takahashi, Chandanabhumma, Ruiz, Cunningham; Manuscript draft: Ford, Takahashi, Chandanabhumma, Ruiz, Cunningham; Acquisition of funding: Ford, Takahashi, Cunningham; Supervision: Ford

REFERENCES

1. Ford CL, Airhihenbuwa CO. The public health critical race methodology: praxis for antiracism research. *Soc Sci Med*. 2010;71(8):1390-1398. <https://doi.org/10.1016/j.socscimed.2010.07.030> PMID:20822840
2. Harper S, King NB, Meersman SC, Reichman ME, Breen N, Lynch J. Implicit value judgments in the measurement of health inequalities. *Milbank Q*. 2010;88(1):4-29. <https://doi.org/10.1111/j.1468-0009.2010.00587.x> PMID:20377756
3. Zuberi T, Bonilla-Silva E, eds. *White Logic, White Methods: Racism and Methodology*. Lanham, MD: Rowman & Littlefield Publishers; 2008.
4. HIV Continuum of Care. *U.S., 2014, Overall and by Age, Race/Ethnicity, Transmission Route and Sex*. [press release] Atlanta, GA: Centers for Disease Control and Prevention; 2017.
5. White House Office of National AIDS Policy. *National HIV/AIDS Strategy for the United States: Updated to 2020*. Washington, DC: The White House; July 23, 2015.
6. Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clin Infect Dis*. 2011;52(6):793-800. <https://doi.org/10.1093/cid/ciq243> PMID:21367734
7. Whiteside YO, Cohen SM, Bradley H, Skarbinski J, Hall HI, Lansky A; Centers for Disease Control and Prevention (CDC). Progress along the continuum of HIV care among blacks with diagnosed HIV- United States, 2010. *MMWR Morb Mortal Wkly Rep*. 2014;63(5):85-89. PMID:24500286
8. Biello KB, Niccolai L, Kershaw TS, Haiqun L, Ickovics J. Residential racial segregation and racial differences in sexual behaviours: an 11-year longitudinal study of sexual risk of adolescents transitioning to adulthood. *J Epidemiol & Comm Health*. 2013;67(1):28-34.
9. Hallfors DD, Iritani BJ, Miller WC, Bauer DJ. Sexual and drug behavior patterns and HIV and STD racial disparities: the need for new directions. *Am J Public Health*. 2007;97(1):125-132. <https://doi.org/10.2105/AJPH.2005.075747> PMID:17138921
10. Kaiser Family Foundation. *HIV Testing in the United States*. Menlo Park, CA: The Henry J. Kaiser Family Foundation; December 2012.
11. Fennie KP, Lutfi K, Maddox LM, Lieb S, Trepka MJ. Influence of residential segregation on survival after AIDS diagnosis among non-Hispanic blacks. *Annals of Epidemiology*. 2015;25(2):113-119, 119 e111. <https://doi.org/10.1016/j.annepidem.2014.11.023>
12. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. *Public Health Rep*. 2001;116(5):404-416. [https://doi.org/10.1016/S0033-3549\(04\)50068-7](https://doi.org/10.1016/S0033-3549(04)50068-7) PMID:12042604
13. King WD, Wong MD, Shapiro MF, Landon BE, Cunningham WE. Does racial concordance between HIV-positive patients and their physicians affect the time to receipt of protease inhibitors? *J Gen Intern Med*. 2004;19(11):1146-1153. <https://doi.org/10.1111/j.1525-1497.2004.30443.x> PMID:15566445
14. Gelberg L, Andersen RM, Leake BD. The Behavioral Model for Vulnerable Populations: application to medical care use and outcomes for homeless people. *Health Serv Res*. 2000;34(6):1273-1302. PMID:10654830
15. Gilmore RW. *Golden Gulag: Prisons, Surplus, Crisis, and Opposition in Globalizing California*. Berkeley, CA: University of California Press; 2007.
16. Omi M, Winant H. *Racial Formation in the United States from the 1960's to the 1990's*. New York: Routledge; 1994.
17. van Ryn M, Fu SS. Paved with good intentions: do public health and human service providers contribute to racial/ethnic disparities in health? *Am J Public Health*. 2003;93(2):248-255. <https://doi.org/10.2105/AJPH.93.2.248> PMID:12554578
18. Watkins-Hayes C, Pittman-Gay L, Beaman J. 'Dying from' to 'living with': framing institutions and the coping processes of African American women living with HIV/AIDS. *Soc Sci Med*. 2012;74(12):2028-2036. <https://doi.org/10.1016/j.socscimed.2012.02.001> PMID:22480872
19. Airhihenbuwa CO, Ford CL, Iwelunmor JI. Why culture matters in health interventions: lessons from HIV/AIDS stigma and NCDs. *Health Educ Behav*. 2014;41(1):78-84. <https://doi.org/10.1177/1090198113487199> PMID:23685666
20. Thomas SB, Quinn SC, Butler J, Fryer CS, Garza MA. Toward a fourth generation of disparities research to achieve health equity. *Annu Rev Public Health*. 2011;32(1):399-416. <https://doi.org/10.1146/annurev-publ-health-031210-101136> PMID:21219164