

Algorithm-Assisted Decision Making and Racial Disparities in Housing: A Study of the Allegheny Housing Assessment Tool

Lingwei Cheng¹, Cameron Drayton¹, Alexandra Chouldechova¹, Rhema Vaithianathan²*

¹Heinz College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, PA,

²Auckland University of Technology, New Zealand

{lingweic, cdrayton, achoulde}@andrew.cmu.edu, rhema.vaithianathan@aut.ac.nz

Abstract

The demand for housing assistance across the United States far exceeds the supply, leaving housing providers the task of prioritizing clients for receipt of this limited resource. To be eligible for federal funding, local homelessness systems are required to implement assessment tools as part of their prioritization processes. The Vulnerability Index Service Prioritization Decision Assistance Tool (VI-SPDAT) is the most commonly used assessment tool nationwide. Recent studies have criticized the VI-SPDAT as exhibiting racial bias (Cronley 2022; Petry et al. 2021), which may lead to unwarranted racial disparities in housing provision. In response to these criticisms, some jurisdictions have developed alternative tools, such as the Allegheny Housing Assessment (AHA), which uses algorithms to assess clients' risk levels. Drawing on data from its deployment, we conduct descriptive and quantitative analyses to evaluate whether replacing the VI-SPDAT with the AHA affects racial disparities in housing allocation. We find that the VI-SPDAT tended to assign higher risk scores to white clients and lower risk scores to Black clients, and that white clients were served at a higher rates pre-AHA deployment. While post-deployment service decisions became better aligned with the AHA score, and the distribution of AHA scores is similar across racial groups, we do not find evidence of a corresponding decrease in disparities in service rates. We attribute the persistent disparity to the use of Alt-AHA, a survey-based tool that is used in cases of low data quality, as well as group differences in eligibility-related factors, such as chronic homelessness and veteran status. We discuss the implications for housing service systems seeking to reduce racial disparities in their service delivery.

1 Introduction

On any given night, roughly half a million people in the United States are experiencing homelessness (U.S. Department of Housing and Urban Development 2022). Black and African American persons are disproportionately represented among the homeless population, with lifetime rates of homelessness estimated to be more than three times higher for non-Hispanic Blacks compared with non-Hispanic whites (Fusaro, Levy, and Shaefer 2018). Some

but not all of this disparity is explained by racial differences in income, incarceration rates, and exposure to traumatic events (Day 2015). Providing long-term supportive housing is the single most effective intervention for reducing homelessness (Evans, Philips, and Ruffini 2019).

As a requirement for participating in programs funded by the United States Department of Housing and Urban Development (HUD), local homelessness systems must implement assessment tools that prioritize the most vulnerable homeless clients for supportive services. The most common housing assessment tool is an interview-based instrument called the Vulnerability Index Service Prioritization Decision Assistance Tool (VI-SPDAT) (Community Solutions 2016). Problematically, recent studies suggest that the VI-SPDAT is poorly validated (Brown et al. 2018) and racially biased (Cronley 2022; Petry et al. 2021). Partly in response to such criticism, homelessness systems across the country have been exploring alternative assessment tools, including ones based on predictive risk models.

We study one such model developed and deployed by the Allegheny County Department of Human Services (ACDHS). Allegheny County is a medium-sized US county in Pennsylvania that contains the city of Pittsburgh. Allegheny County's Continuum of Care (CoC) receives HUD funding to coordinate and provide homelessness services such as shelters, supportive housing, outreach, engagement and assessment, and prevention strategies. ACDHS is responsible for managing HUD funding and operating the Coordinated Entry system that assesses and refers clients to CoC providers (Allegheny County Executive Office 2023).

In 2020, the county deployed a tool called the Allegheny Housing Assessment (AHA) to "standardize the assessment process and move away from the interview-based VI-SPDAT approach to an automated assessment based on administrative data that can better predict the likelihood of future adverse outcomes" (Vaithianathan and Kithulgoda 2020). In this paper, we study the transition from VI-SPDAT to the predictive risk model-based risk assessment, AHA, and evaluate the AHA tool's impact on racial disparities in access to supportive housing services. Specifically, we want to understand:

1. How did the prioritization for housing change when the VI-SPDAT was replaced with the AHA?
2. How does AHA impact the racial disparities in housing

*We thank the Allegheny County Department of Human Services for their invaluable support and feedback.

Copyright © 2024, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

service enrollment rates?

3. How does the use of the alternative survey assessment tool (Alt-AHA) impact racial disparities?

While there have been extensive evaluations of the performance, ethics, and algorithmic fairness of the AHA models (Vaithianathan and Kithulgoda 2020; Eticas Research and Consulting 2020; Allegheny County Analytics 2020), we are among the first to study the impact of its deployment¹.

Using extensive descriptive and quantitative analysis, we find that prior to the deployment of AHA, VI-SPDAT tended to assign higher risk scores to white clients and lower risk scores to Black clients. While there was no disparity in service rates conditioned on VI-SPDAT scores, the net effect was that white clients were served at a higher rate. Following the adoption of AHA, there was an increase in service to Black clients, especially those at medium- and high-risks. However, these increases paralleled those for white clients, and thus did not result in a reduction in racial disparities. While there is no evidence of racial disparity among clients assessed with AHA only, there is a significant racial disparity in service rates among clients assessed with Alt-AHA. White clients were about twice as likely as Black clients to be assessed with Alt-AHA, and conditional on being assessed were much more likely to be served. The differences in Alt-AHA assessment rates is due in large part to differences in data quality, which is significantly lower for white clients. Disparities in service rates across groups following an Alt-AHA assessment are at least partly attributable to other explanatory factors that affect eligibility, such as chronic homelessness, disability, and veteran status.

Our study contributes to the literature on algorithmic fairness and human-AI interaction in several important ways. First, to our knowledge, it is the first study to investigate how algorithm-assisted decision-making affects housing disparities using real data from a deployed tool. Second, we directly study algorithm-assisted decision outcomes of an automated risk assessment tool in the presence of a manual assessment option (the Alt-AHA). Lastly, by identifying important differences between the widely used VI-SPDAT and a machine learning-based alternative (AHA), our study has important policy implications for housing providers.

2 Background

In response to longstanding housing shortfalls, in 2008 federal rules began incentivizing housing providers to use risk assessment tools to ensure that the most vulnerable clients are prioritized for supportive housing. To receive federal funding, HUD requires housing providers to use tools that assess a client’s housing and service needs and eligibility, and score their risk vulnerability (U.S. Department of Housing and Urban Development 2017).

In this section we describe the most widely used tool, the VI-SPDAT, and then detail the development and deployment of the AHA tool that has since replaced the VI-SPDAT in ACDHS operations.

¹We are aware of other ongoing work, but none that has been published.

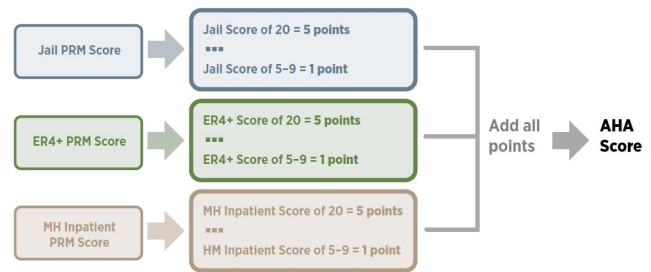


Figure 1: Mapping of Predictive Risk Model Scores to AHA Scores (Vaithianathan and Kithulgoda 2020)

VI-SPDAT

The VI-SPDAT is a tool that combines OrgCode’s SPDAT (Service Prioritization Decision Assistance Tool) Prescreen tool and Community Solutions’ VI (Vulnerability Index). It features a survey for Coordinated Entry (CE) workers to assess client risk and need at intake. Gaining widespread adoption, it was used in 39 states and countries by 2015 (National Alliance to End Homelessness 2022) and remained in use in Allegheny County until 2020.

Recent research has concluded that the VI-SPDAT is poorly validated (Brown et al. 2018) and racially biased (Cronley 2022; Petry et al. 2021; Institute for Research on Poverty 2021). Moreover, the assessment questions can be invasive, asking about drug use, risk-taking behavior, and mental health diagnoses. Focus groups conducted in Allegheny County showed that former clients had concerns about the long and detailed 45 minute interview the tool entailed, which specifically opened the potential to re-traumatize. Despite the invasive questions, a 2018 analysis by ACDHS found that the VI-SPDAT was not effective at identifying clients who are at the greatest risk of adverse outcomes associated with housing instability, and the scores did not significantly predict risk of return to homeless services. Partly in response to the criticism, by December 2020 the maintainers of VI-SPDAT had announced that they would phase out the tool (National Alliance to End Homelessness 2022).

The Allegheny Housing Assessment

In 2014, ACDHS issued a request for proposals (RFP) to explore the use of predictive decision support tools in their operations. The RFP resulted in a research partnership between ACDHS and the Center for Social Data Analytics at the Auckland University of Technology (CSDA). By 2018, CSDA began working on a machine learning-based risk assessment model that would eventually become the Allegheny Housing Assessment (AHA).

The AHA tool is based on three separate machine learning models, each trained to predict the likelihood of a given adverse event occurring in the 12 months following a housing services assessment: (i) a jail stay; (ii) a Medicaid-funded inpatient hospital stay for a behavioral health event; and (iii) 4 or more Medicaid-funded Emergency Room events. Unlike the interview-based VI-SPDAT, the AHA relies not

on self-reported answers to interview questions, but rather on demographic and county service utilization data stored in the Allegheny County Data Warehouse (The Allegheny County Department of Human Services 2018). Features considered by the models are drawn from child welfare services, county jail, courts, juvenile probation, behavioral and physical health, previous interactions with homeless services, previous interactions with assisted housing, demographics (other than race), neighborhood data from the American Community Survey, and household characteristics.

Each of the three models produces a risk score between 1-20, and the three scores are then combined into a single AHA score. Figure 1 below summarizes the combination step. Scores from the three separate models are binned, summed and then converted to deciles to produce a final AHA score between 1-10, with 10 being the highest risk.

The Alt-AHA Tool

In certain cases, new clients may lack sufficient data to run the AHA tool. For this reason, ACDHS complements the AHA with Alt-AHA, an alternative risk assessment based on a subset of VI-SPDAT questions identified by tool developers as being most predictive of risk. Alt-AHA was validated on the same outcomes as AHA, and calibrated to ensure that clients scoring 10 on Alt-AHA have a similar risk of adverse outcomes as those scoring 10 on AHA.

To help guide the use of Alt-AHA, each client on a referral receives a “data quality indicator.” The indicator is 1 if the Data Warehouse has a unique ID for the individual that has existed for at least 90 days and 0 otherwise.

CE workers are required to use the Alt-AHA when: (i) the quality indicator is less than 1; or (ii) someone moved to Allegheny County within the last 90 days; or (iii) someone has not consistently been living in Allegheny County; or (iv) someone is fleeing domestic violence and thus may have been unable to utilize services. Alt-AHA is sometimes used outside of these four cases. When CE workers use the Alt-AHA, they must record their reason for doing so.

An AHA score is always produced, irrespective of whether an Alt-AHA is conducted. The final score used in prioritizing clients, hereafter referred to as “*posted risk score*”, is the greater of the AHA and Alt-AHA scores. The next section describes how these scores are used as part of the broader intake process.

Housing Intake Process

The full client intake process is illustrated in Figure 2. Clients can access the Coordinated Entry system for homeless services by contacting Allegheny Link or through referrals by the Office of Community Services (OCS) Field Units. For clients identified as homeless via the Current Living Situation Screening, CE workers then conduct an assessment using AHA or Alt-AHA (prior to August 2020, VI-SPDAT).

After screening and scoring, CE workers assess clients’ eligibility for specific housing programs. This entails screening for chronic homeless status, which can prioritize them for housing (HUD 2015). The combination of (Alt-)AHA scores and special circumstances (e.g. veteran status, fleeing from domestic abuse, or disability status) dictates program

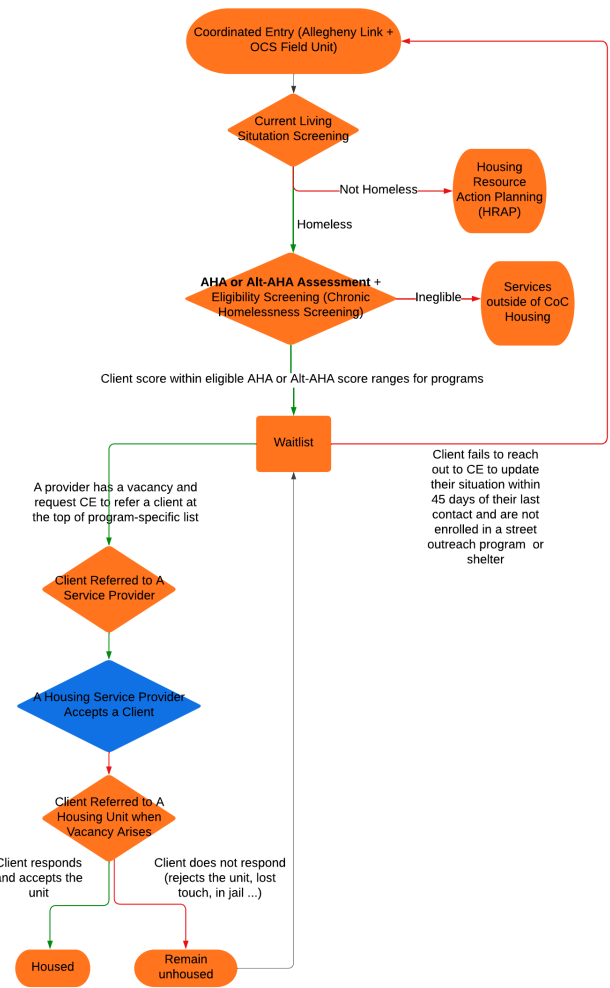


Figure 2: DHS Housing Services Intake Process

eligibility. The CE coordinator then places eligible clients on waitlists for the appropriate housing types. If the CE cannot reach a client for updates every 45 days from their last contact, and the client is not enrolled in a street outreach program or shelter, they are removed from the waitlist.

Once a provider has an available slot and the capacity to serve and find a client housing, they request a referral from ACDHS off the waitlist. The housing provider then enrolls clients into their programs, offering case management services and housing assistance. Importantly, this does not guarantee housing; clients might not accept the offer due to personal preferences or circumstances such as hospitalization or incarceration. Despite the AHA tool’s key role in housing allocation, human factors like maintaining the waitlist and locating clients still play a crucial part in the process, thus influencing when clients are referred. However, CE workers’ ability to alter a client’s waitlist position is limited by strict eligibility rules. This is in contrast to other algorithm-assisted decision-making settings such as bail or child maltreatment screening, where human decision-

makers have considerably more discretion.

3 Related Literature

Algorithm-assisted decision-making is increasingly used in the public sector, including in homelessness prevention. Predictive models utilize proxy outcomes, administrative data, and target specific sub-populations to efficiently allocate limited resources in crucial settings. Various studies have used these models to predict outcomes such as high-cost users (Toros and Flaming 2018), first-time homelessness (Von Wachter et al. 2019), returns to homelessness (Von Wachter et al. 2019; Kube, Das, and Fowler 2019), shelter entries (Shinn et al. 2013), and chronic homelessness (Messier, John, and Malik 2022; VanBerlo et al. 2021). Additional research focuses on specific groups, predicting risks such as daily sheltering arrangements for homeless youth (Suchting et al. 2020) and their susceptibility to substance abuse (Tabar et al. 2020; Dou et al. 2021), as well as assessing homelessness risks among veterans (Koh et al. 2022; Byrne, Montgomery, and Fargo 2019). The AHA tool differs from these other approaches by focusing on long-term risks associated with homelessness, providing algorithmic scores rather than binary risk assessments to prioritize clients on housing waitlists managed by ACDHS.

Recent studies on predictive tools in housing service allocation reveals several significant challenges. The first concern is the accuracy of administrative data in reflecting clients’ past circumstances and urgent needs (Kuo et al. 2023; Karusala et al. 2019). Research by Karusala et al. (2019) involving interviews with policymakers and case-workers shows mixed feelings about the objectivity and accuracy of the VI-SPDAT. Factors like clients’ trust in case-workers and their privacy preferences impact what they disclose, which in turn affects scoring. The scoring also depends on the coordinators’ ability to prompt questions and gauge the effects of a client’s mood and mental status. Thus, the fairness of these tools may suffer from discrepancies in what is intended to be measured versus what is actually being measured and assessed in the AI model (Jacobs 2021). Karusala et al. (2019) note that tools like AHA, which rely on administrative rather than self-reported data, address mis-measurement attributable to faulty self-reporting, but remain susceptible to other biases in data collection.

The second concern focuses on the use of proxy outcomes in prediction, which may not fully capture the subjective experiences of unhoused individuals Kuo et al. (2023). Research shows that ML systems might de-prioritize crucial values like efficiency and privacy that are essential for protecting homeless individuals (Showkat et al. 2023). Building trust in these systems at provider and bureaucratic levels is crucial for their support, as noted by Saxena et al. (2021).

Debate around fairness in housing services also persists. Kube, Das, and Fowler (2019) highlight a trade-off between fairness and efficiency, where imposing fairness constraints can increase re-entry rates. Karusala et al. (2019) note a division between service providers, who prioritize vulnerability over race, and policymakers, who seek to address racial disparities in housing access. Some researchers advocate for using reparative algorithms to rectify historical inequities (So

et al. 2022). Our study specifically evaluates racial disparities in the allocation of housing.

ACDHS commissioned an independent equity evaluation (Eticas Research and Consulting 2020) of the models prior to deployment, which found that the models operated consistently across racial sub-groups.

4 Methodology

Research Questions

In the study, we seek to understand how CE workers interact with the AHA tool and how the AHA tool impacts racial disparity in housing allocation. More specifically, we ask three Research Questions (RQ’s):

1. How did the prioritization for housing change when the VI-SPDAT was replaced with the AHA?
2. How does AHA impact the racial disparities in housing service enrollment rates?
3. How does the use of the alternative survey assessment tool (Alt-AHA) impact racial disparities?

We study changes in *service rates*. An assessment is “served” if it results in an enrollment with a Permanent Supportive Housing (PSH), Rapid Rehousing (RR), or Transitional Housing (TH) provider. This means the client can start receiving comprehensive case management services, such as mental health and substance use treatment, and housing assistance. As discussed in § 2, enrollment does not mean they are assigned to or have moved into housing yet. We do not distinguish between housing types, as all three provide services that address housing instability barriers. This study is not a comprehensive evaluation of the impact of the tool on overall welfare—only on racial disparities.

Descriptive Analysis and Empirical Model

We conduct descriptive analyses to examine high-level trends, and then more rigorously study the research questions using logistic regressions to control for client characteristics. We refer readers to the Appendix section in (Cheng et al. 2024) for all supplementary materials.

Table 5 in the Appendix (Cheng et al. 2024) summarizes key variables and their notation. The distinction between demographics and housing eligibility characteristics is further discussed in Appendix E (Cheng et al. 2024).

To answer RQ1, we use retrospectively generated AHA scores for the pre- and post-deployment periods to investigate changes in service rates. We conduct the following regression on clients who were assessed only via AHA, focusing on β_3 . *Post* is equal to 1 if the assessment was conducted after AHA deployment, and 0 if before.

$$\begin{aligned} Served_{i,t} = & \beta_0 + \beta_1 Post_{i,t} + \beta_2 Score + \beta_3 Score * Post_{i,t} \\ & + \beta_4 Black_{i,t} + \beta_5' X + \gamma_t + \delta_t + \epsilon_{i,t} \end{aligned} \quad (1)$$

β_3 captures the change in association between service decisions and AHA scores from pre to post. We control for client race (*Black* indicator variable which equals to 1 if the client is Black and 0 if white), as well as other controls, X ,

which include age, gender, veteran status, chronic homelessness status, disability status, if a client is a domestic violence survivor, if the client is the head of the household, the size of the household, the type of household (Single or Households without Children). Time-fixed effects γ_t and δ_t represent assessment year and month, respectively, and capture yearly and seasonal variations in service provision. We use robust standard errors, clustered by client ID, to account for multiple assessments per client and for among-client variation.

We study RQ2 with the following specification, using data only on clients where AHA (not Alt-AHA) was used.

$$\begin{aligned} Served_{i,t} = & \beta_0 + \beta_1 Post_{i,t} + \beta_2 Black_{i,t} + \beta_3 Score_{i,t} \\ & + \beta_4 RiskGroup_{i,t} + \beta_5 Black * Post_{i,t} \\ & + \beta_6 RiskGroup * Post_{i,t} + \beta_7 RiskGroup * Black_{i,t} \\ & + \beta_8 RiskGroup * Black * Post_{i,t} \\ & + \beta_9'X + \gamma_t + \delta_t + \epsilon_{i,t} \end{aligned} \quad (2)$$

This model allows for heterogeneity in the association between service decisions, race, and AHA deployment across risk group strata. The risk strata are defined as low-risk (AHA scores 1-4), medium-risk (scores 5-8), and high-risk (scores 9-10). We present a simplified short-hand equation here and include an expanded equation in Appendix D (Cheng et al. 2024) for clarification.

Two coefficients of interest, β_5 and β_8 , measure changes in service rates. β_5 reflects how service rates change for the low-risk group post-AHA. β_8 shows the additional service rate change for Black clients in the high-risk group, compared to white clients, pre- and post-AHA deployment, relative to the low-risk group. Essentially, it shows the impact of high risk status on racial disparities post-AHA. As a robustness check, we test alternative specifications in Appendix E (Cheng et al. 2024), treating AHA scores as continuous or binary for high-risk. The results are consistent with the main findings presented in this paper.

For RQ3, we use logistic regression on data from clients assessed with Alt-AHA, using the following specification where the Alt-AHA score, if higher, is used as the posted score for housing prioritization.

$$\begin{aligned} Served_{i,t} = & \beta_0 + \beta_1 Black_{i,t} + \beta_2 PostedScore_{i,t} \\ & + \beta_4 UseAltScore + \beta_5 QualityIndicator \quad (3) \\ & + \beta_6'X + \gamma_t + \delta_t + \epsilon_{i,t} \end{aligned}$$

QualityIndicator assesses data quality on a 0 to 1 scale. All other covariates X and time fixed effects remain consistent. The coefficient of interest, β_1 , evaluates potential racial disparities in service rates for clients prioritized based on Alt-AHA after controlling for other factors.

5 Data

We extract housing assessment data from the County's Homelessness Management Information System (HMIS) for assessments conducted between January 1, 2018 and March 29, 2022. This data is linked with enrollment data for clients in PSH, RR, or TH services during the study period.

We retrospectively (re-)run the AHA tool to produce scores for pre- and post-AHA deployment periods. We also obtain VI-SPDAT scores from the pre-deployment period.

We exclude assessments with a race other than white or Black/African American, as they represent only 3.2% of the data. We also exclude 36.1% of the assessments associated with families or youth household types, focusing instead on single adult household types. Lastly, we exclude clients who are enrolled with programs that do not require AHA scores. The final dataset contains 6,542 assessments, in which 66.1% were from the pre-AHA period. 399 Alt-AHA assessments were conducted, with 370 using the Alt-AHA scores as their posted scores.

Table 1 summarizes pre- and post-deployment client characteristics. We see that the majority of single homeless clients are male, and over 90% report a disability. Post deployment, the prevalence of chronically homeless clients rises. When comparing Black and white clients, their assessment scores, ages, and likelihood of being domestic violence survivors remain similar pre- and post-deployment. However, they diverge in key eligibility criteria: white clients exhibit higher rates of chronic homelessness and disability, whereas Black clients are more likely to be veterans.

Before deployment, Black clients have a lower average VI-SPDAT score (7.5) than white clients (8.4) on a 1-16 scale. Post-deployment, Black clients average a slightly higher AHA score (7.1) compared to white clients (6.6) on a 1-10 scale. White clients are twice (24.5%) as likely to be assessed with Alt-AHA compared to Black clients (11.8%), and to have those scores used for prioritization (20.7% vs. 8.0% for Black clients). This significant difference may be linked to consistently higher data quality indicators for Black clients, with scores of 0.96 vs. 0.87 pre-deployment and 0.91 vs. 0.82 post-deployment.

Finally, white clients were served at a higher rate than Black clients before AHA deployment, 17.6% compared to 14.5%. The disparity persists post-deployment.

6 Descriptive Analysis

We compare risk scores and service rates from VI-SPDAT, AHA, and Alt-AHA for high-level insight into our research questions. In Appendix A (Cheng et al. 2024) we provide a brief analysis of concurrent events, particularly the impact of COVID-19 and lockdown orders, surrounding the AHA deployment. While we note some variation in service rates by race, these changes are not statistically significant.

Score Comparisons

AHA vs. VI-SPDAT Scores Because the VI-SPDAT was no longer conducted after the AHA was deployed, this comparison is based on data during the pre-deployment period. VI-SPDAT scores range from 1-16 where 1 is the lowest and 16 is the highest risk, and AHA scores range from 1-10. The left panel on Figure 3 shows that there is a weak positive correlation between the retrospectively generated AHA scores and the unconverted VI-SPDAT scores (Spearman's $\rho = 0.218$, $p < 0.0001$).

To facilitate a more direct comparison of the scores for the remainder of the analysis, we bin the VI-SPDAT scores

	Pre-deployment		Post-deployment	
	Black/AA	White	Black/AA	White
% Female	33.1 (47.1)	32.1 (46.7)	35.1 (47.7)	37.3 (48.4)
% Male	59.9 (49.0)	62.1 (48.5)	64.9 (47.8)	62.6 (48.4)
% Unknown G	7.0 (25.5)	5.8 (23.3)	0.1 (3.0)	0.1 (3.1)
Age	46.3 (12.3)	45.7 (12.0)	46.1 (12.2)	44.0 (11.7)
% CH	25.2 (43.4)	29.8 (45.8)	35.4 (47.8)	38.6 (48.7)
% Disability	91.2 (28.4)	95.2 (21.5)	91.0 (28.7)	93.9 (24.0)
% Veteran	9.1 (28.8)	7.8 (26.8)	7.8 (26.8)	6.4 (24.5)
% DV Survivor	11.5 (32.0)	11.4 (31.8)	13.5 (34.2)	13.6 (34.3)
AHA Score	6.5 (2.7)	6.3 (2.7)	7.1 (2.5)	6.6 (2.8)
VI-SPDAT Score	7.5 (2.5)	8.4 (3.1)	-	-
Alt-AHA Score	-	-	5.1 (2.7)	6.1 (2.5)
% Use Alt-AHA	-	-	11.8 (32.3)	24.5 (43.0)
% Use Alt Score	-	-	8.0 (27.1)	20.7 (40.5)
Quality Indicator	0.96 (0.20)	0.87 (0.32)	0.91 (0.28)	0.82 (0.38)
Served	14.5 (35.2)	17.6 (38.1)	19.5 (39.7)	23.3 (42.3)
N	2,176	2,147	1,141	1,078

Standard deviation in parentheses

Unknown G = Unknown Gender, CH = Chronic Homeless

DV = Domestic Violence

Table 1: Summary Statistics by Race and Deployment Status

into a 1-10 range. Scores 1-4 are mapped to 1, 13-16 are mapped to 10, and values in-between are mapped in a one-to-one correspondence to 2:9. This mapping is monotonic, and is chosen to ensure that there are comparable numbers of clients in the lowest and highest score bins for both the AHA and (post-mapping) VI-SPDAT. Figure 3 (right) shows a heatmap of scores after mapping. After conversion, the AHA and VI-SPDAT Scores remain significantly weakly positively correlated (Spearman’s $\rho = 0.217$, $p < 0.0001$)

AHA vs. Alt-AHA Scores Figure 4 compares AHA and Alt-AHA scores for clients assessed with Alt-AHA post-deployment. Clients with initial low AHA scores (1-3) typically received higher Alt-AHA scores. As a business rule, the *higher* of the two scores is used as the score (“posted score”) for prioritization. Alt-AHA and AHA scores do not appear correlated (Spearman’s $\rho = -0.02$, $p = 0.64$).

Figure 5 shows a breakdown of Alt-AHA use by race. We observe that white clients are more likely than Black clients to be assessed via Alt-AHA (left panel) and are also more likely to then have their Alt-AHA scores used (right panel). These differences are, at least, partly attributed to lower data quality indicators for white clients, as shown in Table 1. CE workers provide reasons for disagreeing with AHA scores,

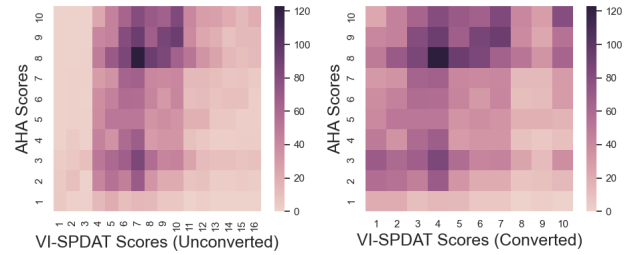


Figure 3: AHA vs. VI-SPDAT Scores Prior to AHA Implementation. Original VI-SPDAT scores (left), converted VI-SPDAT scores (right).

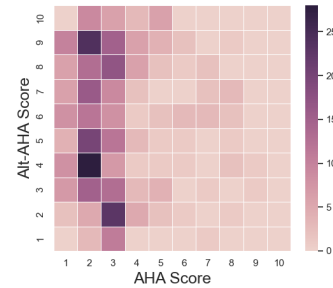


Figure 4: AHA vs. Alt-AHA Scores for clients who were assessed with Alt AHA. Clients with low AHA scores (1-3) typically received higher Alt-AHA scores.

which we discuss in Appendix B (Cheng et al. 2024). There we find that white clients are more often not residing consistently in Allegheny County, have lower data quality, and have lower service utilization than Black clients.

Score Distribution by Race

Figure 6 shows the (frequency) distribution of pre-deployment VI-SPDAT and AHA scores and post-deployment AHA and Alt-AHA scores². We present within-racial-group frequencies instead of counts so that pre-vs-post comparisons are not confounded by variation in the number of clients being assessed over time. We see that pre-deployment, VI-SPDAT scores were lower for Black clients than for white clients, with white clients being more likely to receive VI-SPDAT scores in the 8-10 range.

Conditional on VI-SPDAT score, Figure 8 shows that Black and white clients were equally likely to receive housing services. However, at least in part because VI-SPDAT scores tended to be lower for Black clients, overall service rates were significantly lower for Black clients than white clients pre-AHA (see Figure 7).

Unlike the VI-SPDAT, in both the pre- and post-deployment period, Black clients were approximately equally likely to receive AHA scores in the 8-10 range. The third panel of Figure 6 shows the subsample where AHA

²A cumulative density version of these plots is shown in Figure 11 of Appendix C (Cheng et al. 2024).

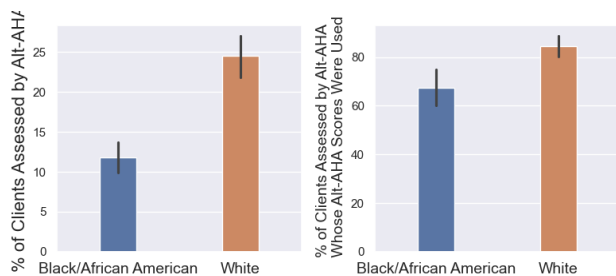


Figure 5: Percent of Clients Assessed by Alt-AHA by Race (left); and Percent of Clients Assessed by Alt-AHA whose Alt-AHA Scores were used for Prioritization by Race (right). White clients are twice as likely to be assessed via Alt-AHA than Black clients. Of those assessed by Alt-AHA, white clients' Alt-AHA scores are also more likely to be used for prioritization.

scores determined prioritization—this combines clients assessed only via AHA and those assessed via Alt-AHA but whose AHA score was higher. Here, Black and white clients exhibit similar score distributions.

Second, transitioning from VI-SPDAT to AHA, we observe a significant shift in the score distribution, with a larger percentage of clients across both races now classified as higher risk, whereas previously, most fell into the low to medium risk category. Post-AHA deployment, there is an increased representation of Black clients in the medium and high-risk categories, compared to the VI-SPDAT.

Finally, in the case of clients assessed via Alt-AHA whose Alt-AHA scores were used for prioritization (where the posted scores are identical to their Alt-AHA scores), the distribution shows a slightly higher percentage of Black clients in the lower risk range and a slightly lower percentage in the high risk range compared to white clients. Given the small sample size, the difference between the distributions is unlikely to be statistically significant.

In summary, we observe a notable disparity under VI-SPDAT, with a higher percentage of Black clients categorized as lower risk and white clients as higher risk. Post-deployment, both Black and white clients receive similar scores under AHA, with Black clients showing an increased presence in the medium to high-risk categories compared to their earlier VI-SPDAT scores. White clients tend to receive higher risk scores when assessed by Alt-AHA.

Service Rates by Race

According to Figure 7, pre-AHA, white clients were served at a significantly higher rate than Black clients (t -statistic = -2.72 , P -value = 0.006). Post-deployment, the service rates increased for both groups, but white clients continued to be served at a significantly higher rate than Black clients (t -statistic = -2.15 , P -value = 0.03). When we break down post-deployment assessments by whether the Alt-AHA scores were used, we find that the two groups are served at similar rate when AHA scores were used. The racial disparity in service rates appears to be driven mainly

by the sub-group of assessments where an Alt-AHA was used. For clients assessed via Alt-AHA, on aggregate, there is a statistically significant difference (t -statistic = -2.72 , P -value = 0.007) between Black and white clients, with white clients (30.0%) about twice more likely to be enrolled with a provider than Black clients (15.4%). However, the Alt-AHA sample size is small, and the aggregate comparison neglects risk scores and other eligibility factors like veteran status and chronic homelessness. Additionally, data quality for white clients is generally lower with Alt-AHA. We will address these issues more rigorously in § 7.

Figure 8 shows service rates conditioned on AHA and VI-SPDAT scores. We note that in the pre-AHA period, VI-SPDAT scores are strongly correlated with the likelihood of enrolling with a provider. In contrast, conditioned on the retroactively calculated AHA risk scores, the service rates appear uncorrelated with AHA scores. Post-deployment, we observe that service rates aligned more closely with AHA scores, with higher scores more likely to be served.

Furthermore, conditioned on VI-SPDAT scores, we observe no racial disparities in service rates during the pre-deployment period. However, as mentioned previously, this lack of association does not imply an absence of racial disparity as VI-SPDAT tended to underestimate the risk of Black clients. If we consider the proxy harms that AHA measures, when CE workers assessed risks using VI-SPDAT without access to AHA scores in the pre-deployment period, white clients were served at higher rates than Black clients among those whose AHA risk score was 9 or 10, as shown in the first panel to the left in Figure 8. Post-deployment, the introduction of AHA scores appears to have closed this disparity for clients in the high-risk category. We will test this hypothesis more rigorously in § 7.

When examining the third and fourth panels in Figure 8, which depict service rates for AHA and Alt-AHA assessments, respectively, we observe no significant gap in service for assessments conducted via AHA. While there is a significant overall difference between Black and white clients assessed via Alt-AHA, it disappears when conditioned on Alt-AHA scores—likely due to small sample sizes.

Conclusions of Descriptive Analysis

Addressing RQ1, prior to AHA deployment, service rates are highly associated with VI-SPDAT scores, and largely uncorrelated with AHA scores. However, post-AHA deployment, service rates begin to correlate strongly with AHA scores, especially for clients assessed exclusively with AHA.

Addressing RQ2, both pre- and post-deployment, white clients were served at a significantly higher rate than Black clients. While there is no overall disparity in service rates in clients assessed by AHA only, service rates for clients assessed by Alt-AHA are higher for white clients than Black clients. There is no racial disparity conditioned on VI-SPDAT scores pre-AHA. Yet, retrospective AHA scoring identify two key service rate disparities. (1) VI-SPDAT tended to assign higher risk scores to white clients compared to Black clients. By contrast, AHA scores are more comparable across groups. (2) Under VI-SPDAT, white clients were served at significantly higher rates than Black clients



Figure 6: AHA and VI-SPDAT Score Distributions by Deployment Period and Race, Normalized by Race. Normalization within each racial group shows the proportion of individuals receiving each risk score. Pre-AHA deployment, VI-SPDAT classified more Black clients as low-risk and more white clients as high-risk, while AHA would have classified more white clients as low-risk. Post-AHA deployment, the score distribution shifted towards higher risk levels. Both AHA and Alt-AHA scored Black and white clients at comparable rates.

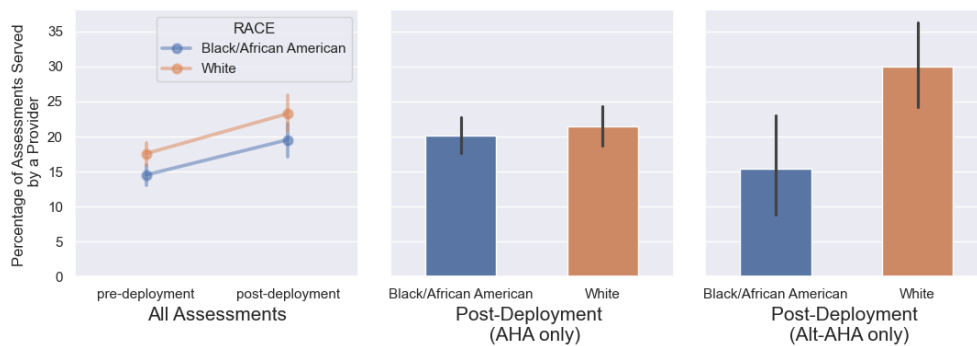


Figure 7: Percent of Assessments Served by a Provider by Race Pre- and Post-AHA deployment. Pre-deployment, providers served white clients at a higher rate than Black clients. Post-deployment, service rates increased for both, but white clients continued to be served at higher rates. AHA-assessed clients are served at similar rates by race. White Alt-AHA assessed clients are served at significantly higher rates.

at the same retrospective AHA risk scores, particularly in the high-risk category. The deployment of AHA reduced the service rate disparity among high-risk clients.

Finally, for RQ3, we find that post-AHA deployment, white clients were more likely to be assessed via Alt-AHA, their Alt-AHA scores were more likely to be higher than their AHA scores compared to Black clients, and they were more likely to be served. However, this may be due to differences in data quality, demographics, and eligibility criteria.

7 Quantitative Results

In this section, we further interrogate the findings from the descriptive analysis using logistic regressions to control for additional explanatory variables.

RQ1

Because we retrospectively (re-)calculated both pre- and post-deployment AHA scores, we can examine whether the switch from VI-SPDAT to AHA resulted in service decisions becoming more closely aligned with AHA scores. Table 2 shows the results of testing this hypothesis using Equa-

tion (1). In the table columns we progressively control for demographics and eligibility criteria, and time fixed effects.

Comparing the *AHA Score* coefficient across columns (1) and (2), we see that once eligibility criteria are controlled for, there is no statistically significant association between decisions and the AHA score *pre*-deployment. On the other hand, the coefficient of *AHA Score * Post* is consistently statistically significant across columns, indicating that service rates became significantly positively associated AHA scores *post*-deployment. Across all columns, the coefficients on *Black* consistently show that Black clients are statistically significantly less likely to be served compared to white clients.

RQ2

As discussed in § 6, when CE workers relied solely on VI-SPDAT without access to AHA scores, they tended to serve white clients at a higher rate within the AHA score range 9 and above. We further examine the change in service rate by risk groups using the Equation 2 allowing for heterogeneous treatment effect. We present the results in Table 3.

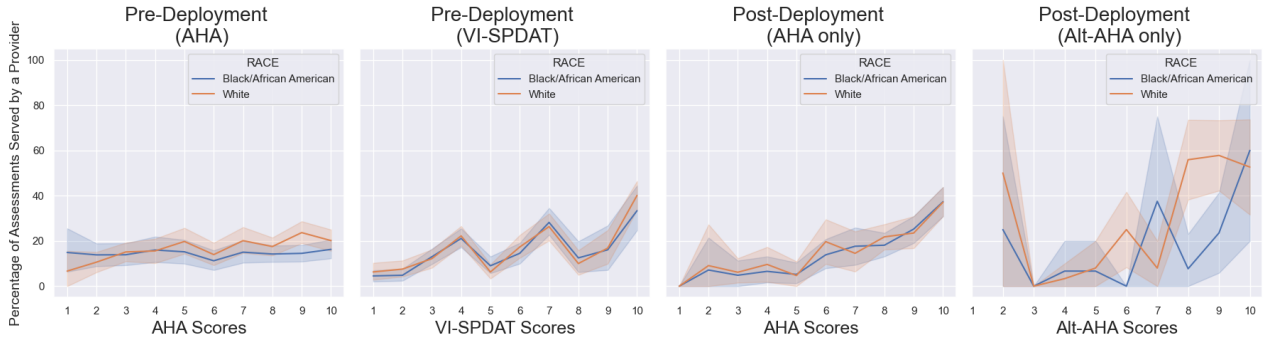


Figure 8: Percent of Assessments Served by Race for Pre- and Post-deployment Periods, Conditioned on Scores.

	(1)	(2)	(3)
Post	-2.01*** (0.33)	-2.01*** (0.42)	-2.24*** (0.46)
AHA Score	0.05** (0.02)	-0.02 (0.02)	-0.01 (0.02)
AHA Score * Post	0.28*** (0.04)	0.31*** (0.04)	0.31*** (0.04)
Black	-0.19** (0.07)	-0.15 (0.08)	-0.16* (0.08)
D & E Crit.		✓	✓
Quarter FE			✓
Yearly FE			✓
N	6,143	6,143	6,143
Pseudo. R ²	0.03	0.28	0.28

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Service Rate by AHA Implementation

We find that that the coefficients of High Risk * Post are consistently statistically significant across the specifications, indicating that high risk white clients are more likely to be served post-AHA deployment compared to pre. The coefficients of High Risk * Black * Post and Medium Risk * Black * Post are not statistically significant for either specification that controls for demographics and eligibility criteria. We therefore do not have sufficient evidence to conclude that the change in service rates post-deployment is different for Black clients than for white clients for any of the risk strata.

RQ3

As for the use of Alt-AHA, across all models shown in Table 4, higher posted scores and quality indicator scores are consistently associated with a higher likelihood of receiving services through Alt-AHA usage. Conversely, the use of Alt-AHA alone does not significantly impact service likelihood.

When considering the effect of race and posted score (the higher of Alt-AHA or AHA scores that is used for prioritization for this sub-population) in column (1) and (2) and (3) on service rates, Black clients are significantly less likely

to receive services through Alt-AHA compared to white clients, even after adjusting for differences in data quality (in (3)). Once demographic, eligibility and time fixed effects are added to the regression, the race coefficient is no longer statistically significant. However, this is more due to an increase in the standard error of the coefficient estimate than a clear decrease in the magnitude of the coefficient. The evidence of higher service rates for white clients under Alt-AHA is therefore borderline.

Conclusion of Quantitative Analysis

Our quantitative analysis confirms that post-AHA deployment, service decisions become more closely aligned with AHA scores. Overall, we do not find evidence that the AHA deployment significantly bridges the disparity in service rates between Black and white clients. The post-deployment change in service rates for medium and high risk Black clients is similar to that observed for white clients. Lastly, while there is borderline evidence that white clients are more likely to be served when Alt-AHA is used, ultimately the sample size underpinning that analysis is too small to draw strong conclusions.

8 Limitations

Despite modeling AHA risk scores within actual decision-making contexts, the complex journey from initial assessment to final service enrollment requires us to make certain assumptions in our interpretation.

First, it is essential to acknowledge the possibility that the effects identified in our study are in part attributable to changes during the COVID period related to housing policy and client composition that influenced service allocations. We discuss the effects of the stay-at-home order and other COVID-era impacts in Appendix §A (Cheng et al. 2024).

In order to conclude that the effects we identified are attributable to the deployment of AHA rather than other unmeasured factors, we need to assume that there are no “unobservables” that both (i) changed at the same time as AHA was deployed, and (ii) were relied upon by CE workers in their service decisions. It could be problematic for our study design if, for instance, some information gathered during the assessment process but not reflected in the administrative data influenced CE worker decisions. We were not able to

	(1)	(2)	(3)
Black	0.09 (0.16)	0.15 (0.19)	0.16 (0.19)
Post	-0.63 (0.36)	-0.66 (0.46)	-0.83 (0.49)
AHA Score	0.12** (0.04)	0.06 (0.05)	0.06 (0.05)
High Risk	-0.22 (0.30)	-0.26 (0.35)	-0.25 (0.35)
Medium Risk	-0.16 (0.22)	-0.06 (0.25)	-0.05 (0.25)
High Risk * Black	-0.52* (0.21)	-0.47 (0.26)	-0.51 (0.26)
Medium Risk * Black	-0.37 (0.20)	-0.42 (0.24)	-0.42 (0.24)
High Risk * Post	1.12** (0.39)	1.41** (0.46)	1.34** (0.46)
Medium Risk * Post	0.54 (0.39)	0.72 (0.46)	0.65 (0.46)
Black * Post	-0.46 (0.54)	-0.54 (0.64)	-0.59 (0.64)
High Risk * Black * Post	0.87 (0.58)	0.85 (0.69)	0.93 (0.69)
Medium Risk * Black * Post	0.60 (0.59)	0.62 (0.69)	0.65 (0.70)
D & E Crit.		✓	✓
Quarter FE			✓
Yearly FE			✓
N	6,143	6,143	6,143
Pseudo. R ²	0.03	0.28	0.28

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Service Rate by Race Risk Group, and AHA Implementation

identify alternative observational study designs that would not be susceptible to potential unobserved confounding.

9 Discussion

We present the first evaluation of the impact of replacing the VI-SPDAT with the Allegheny Housing Assessment on housing service outcomes. Specifically, we examine how CE decisions change post-deployment and the effects of those changes on racial disparities in service rates. While earlier studies of AHA considered the predictive performance and group-fairness properties of the tool’s constituent models, they relied on pre-deployment data and did not study the tool’s influence on decisions.

We find that white clients were significantly more likely than Black clients to receive high VI-SPDAT scores but were approximately equally likely to receive high AHA scores. However, later analyses do not enable us to conclude that the shift to AHA-based prioritization significantly reduced the racial disparities in service decisions present under the VI-SPDAT policies. Pre-AHA, white clients were served at higher rates than Black clients, both overall and conditional on AHA score. We find that racial disparities largely persisted post-AHA, particularly in the subset of clients as-

	(1)	(2)	(3)	(4)	(5)
Black	-0.79** (0.28)	-0.61* (0.31)	-0.69* (0.32)	-0.57 (0.37)	-0.66 (0.38)
Posted Score		0.48*** (0.07)	0.50*** (0.08)	0.50*** (0.10)	0.54*** (0.12)
Use Alt AHA			-0.08 (0.40)	-0.30 (0.52)	-0.45 (0.53)
QI			1.04*** (0.28)	0.84* (0.36)	0.77* (0.36)
D & E Crit.				✓	✓
Quarter FE					✓
Yearly FE					✓
N	399	399	399	371	371
Pseudo. R ²	0.02	0.18	0.21	0.41	0.42

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

QI = Quality Indicator.

disability != 1 predicts failure perfectly;

disability omitted and 28 obs not used.

Table 4: Alt-AHA Service Rate

sessed with Alt-AHA. Our analysis indicates that the persistent disparities are at least in part attributable to differences in data quality and eligibility criteria across groups.

This study has several policy implications for CoC providers considering a switch from the VI-SPDAT to predictive models like the AHA tool or other survey-based assessment tools such as the Alt-AHA. Our results show that predictive models like AHA *can* reduce or altogether eliminate differences in the prioritization score distributions. This, however, *does not* guarantee that overall disparities in service decisions will be reduced as a result. Whereas we found no significant post-deployment disparity in service rates among clients assessed with AHA, disparities among clients assessed with Alt-AHA were at least as large as those existing pre-deployment. This finding underscores difficulties with anticipating the impact of a predictive risk model deployment in settings where the model cannot be run due to poor data quality in a substantial number of cases.

As discussed extensively in the literature review, there is an important concern about whether administrative data can accurately capture clients’ past and current needs (Kuo et al. 2023). Efforts to improve data accuracy and enhance cultural competence among service providers are critical for addressing disparities effectively. Policymakers must consider the strengths and weaknesses of risk assessments derived from administrative data and predictive modeling. Effectively integrating AI with CE workers’ expertise could be pivotal in reducing racial disparities in housing service provision.

Future research could examine how CE workers utilize Alt-AHA or diverge from standard AHA procedures to support their clients. Such insights could clarify additional risks identified during interviews and guide improvements in algorithm iterations, enhancing decision-making in social services.

Researcher Positionality Statement

All authors have previously conducted evaluations of algorithmic tools in either child welfare or housing services deployed by the ACDHS for academic research. The second author has interned part-time at the ACDHS, gaining first-hand experience with their operational processes and tools. The last author led the development of the AHA tool, and the third author provided feedback on the methodology used in model development and validation. As a result, the research team collectively has a deep understanding of the design, development, validation and deployment process surrounding the AHA tool.

We recognize that this experience and familiarity also present the potential for conflicts of interest. In conducting the research described in this paper we maintained a firm commitment to academic integrity and transparency. As per the terms of our data use agreement, ACDHS was provided with the submission version of this manuscript and given the opportunity to review and provide feedback. The data use agreement does not enable ACDHS to block the publication. Through this review process we received helpful feedback that helped improve the clarity of the manuscript, led to a better-motivated model specification for RQ1, and also resulted in several robustness checks now presented in the Appendix (Cheng et al. 2024).

Our collective expertise includes backgrounds in public policy, statistics, and health economics, enabling a multidisciplinary approach to this evaluation. We are motivated by a shared goal of improving the efficacy and fairness of algorithmic tools in social services. Our work is informed not only through collaborations with government agencies, but also through engagement with stakeholders most likely to be impacted by the technologies.

References

- Allegheny County Analytics. 2020. AHA DHS Response to Eticas. Allegheny County Analytics Website.
- Allegheny County Executive Office. 2023. Allegheny County Continuum of Care. *DHS Advisory Boards*.
- Brown, M.; Cummings, C.; Lyons, J.; Carrión, A.; and Watson, D. P. 2018. Reliability and validity of the Vulnerability Index-Service Prioritization Decision Assistance Tool (VI-SPDAT) in real-world implementation. *Journal of Social Distress and the Homeless*, 27(2): 110–117.
- Byrne, T.; Montgomery, A. E.; and Fargo, J. D. 2019. Predictive modeling of housing instability and homelessness in the Veterans Health Administration. *Health Services Research*, 54(1): 75–85.
- Cheng, L.; Drayton, C.; Chouldechova, A.; and Vaithianathan, R. 2024. Algorithm-Assisted Decision Making and Racial Disparities in Housing: A Study of the Allegheny Housing Assessment Tool. arXiv:2407.21209.
- Community Solutions. 2016. Vulnerability Index-Service Prioritization Decision Assistance Tool (VI-SPDAT). <https://pehgc.org/wp-content/uploads/2016/09/VI-SPDAT-v2.01-Single-US-Fillable.pdf>.
- Cronley, C. 2022. Invisible intersectionality in measuring vulnerability among individuals experiencing homelessness – critically appraising the VI-SPDAT. *Journal of Social Distress and Homelessness*, 31(1): 23–33.
- Day, J. E. 2015. *Racial disparities and place effects for housing instability and homelessness in the US*. The University of Utah.
- Dou, Z.-Y.; Barman-Adhikari, A.; Fang, F.; and Yadav, A. 2021. Harnessing Social Media to Identify Homeless Youth At-Risk of Substance Use. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(17): 14748–14756.
- Eticas Research and Consulting. 2020. Eticas Assessment. Website.
- Evans, W. N.; Philips, D. C.; and Ruffini, K. J. 2019. Reducing and preventing homelessness: A review of the evidence and charting a research agenda.
- Fusaro, V. A.; Levy, H. G.; and Shaefer, H. L. 2018. Racial and ethnic disparities in the lifetime prevalence of homelessness in the United States. *Demography*, 55(6): 2119–2128.
- HUD. 2015. Homeless Emergency Assistance and Rapid Transition to Housing (HEARTH): Defining Chronically Homeless Final Rule. *HUD Exchange*.
- Institute for Research on Poverty. 2021. Understanding the Role of Caseworker Bias in Racial and Ethnic Inequities in Human Services. <https://www.irp.wisc.edu/wp/wp-content/uploads/2021/09/Memo-2-Caseworker-Bias-09272021pdf.pdf>. Accessed: December 7, 2023.
- Jacobs, A. Z. 2021. Measurement as governance in and for responsible AI. *arXiv preprint arXiv:2109.05658*.
- Karusala, N.; Wilson, J.; Vayanos, P.; and Rice, E. 2019. Street-Level Realities of Data Practices in Homeless Services Provision. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW).
- Koh, K. A.; Montgomery, A. E.; O’Brien, R. W.; Kennedy, C. J.; Luedtke, A.; Sampson, N. A.; Gildea, S. M.; Hwang, I.; King, A. J.; Petriceks, A. H.; Petukhova, M. V.; Stein, M. B.; Ursano, R. J.; and Kessler, R. C. 2022. Predicting Homelessness Among U.S. Army Soldiers No Longer on Active Duty. *American Journal of Preventive Medicine*, 63(1): 13–23.
- Kube, A.; Das, S.; and Fowler, P. J. 2019. Allocating Interventions Based on Predicted Outcomes: A Case Study on Homelessness Services. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01): 622–629.
- Kuo, T.-S.; Shen, H.; Geum, J.; Jones, N.; Hong, J. I.; Zhu, H.; and Holstein, K. 2023. Understanding Frontline Workers’ and Unhoused Individuals’ Perspectives on AI Used in Homeless Services. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI ’23. New York, NY, USA: Association for Computing Machinery. ISBN 9781450394215.
- Messier, G.; John, C.; and Malik, A. 2022. Predicting chronic homelessness: The importance of comparing algorithms using client histories. *Journal of Technology in Human Services*, 40(2): 122–133.

- National Alliance to End Homelessness, H. R. I. 2022. NextGenTools VISPDAT Brief. https://endhomelessness.org/wp-content/uploads/2022/08/NextGenTools_VISPDATBrief_08-30-22.pdf. Accessed: December 7, 2023.
- Petry, L.; Hill, C.; Vayanos, P.; Rice, E.; Hsu, H.-T.; and Morton, M. 2021. Associations Between the Vulnerability Index-Service Prioritization Decision Assistance Tool and Returns to Homelessness Among Single Adults in the United States. *Cityscape*, 23(2): pp. 293–324.
- Saxena, D.; Badillo-Urquiola, K.; Wisniewski, P. J.; and Guha, S. 2021. A Framework of High-Stakes Algorithmic Decision-Making for the Public Sector Developed through a Case Study of Child-Welfare. *Proc. ACM Hum.-Comput. Interact.*, 5(CSCW2).
- Shinn, M.; Greer, A. L.; Bainbridge, J.; Kwon, J.; and Zuiderveen, S. 2013. Efficient targeting of homelessness prevention services for families. *American journal of public health*, 103(S2): S324–S330.
- Showkat, D.; Smith, A. D. R.; Lingqing, W.; and To, A. 2023. “Who is the Right Homeless Client?”: Values in Algorithmic Homelessness Service Provision and Machine Learning Research. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI ’23. New York, NY, USA: Association for Computing Machinery. ISBN 9781450394215.
- So, W.; Lohia, P.; Pimplikar, R.; Hosoi, A.; and D’Ignazio, C. 2022. Beyond Fairness: Reparative Algorithms to Address Historical Injustices of Housing Discrimination in the US. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, FAccT ’22, 988–1004. New York, NY, USA: Association for Computing Machinery. ISBN 9781450393522.
- Suchting, R.; Businelle, M. S.; Hwang, S. W.; Padhye, N. S.; Yang, Y.; and Santa Maria, D. M. 2020. Predicting Daily Sheltering Arrangements among Youth Experiencing Homelessness Using Diary Measurements Collected by Ecological Momentary Assessment. *International Journal of Environmental Research and Public Health*, 17(18).
- Tabar, M.; Park, H.; Winkler, S.; Lee, D.; Barman-Adhikari, A.; and Yadav, A. 2020. Identifying Homeless Youth At-Risk of Substance Use Disorder: Data-Driven Insights for Policymakers. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD ’20, 3092–3100. New York, NY, USA: Association for Computing Machinery. ISBN 9781450379984.
- The Allegheny County Department of Human Services. 2018. Allegheny County Data Warehouse. Published on July 2018.
- Toros, H.; and Flaming, D. 2018. Prioritizing Homeless Assistance Using Predictive Algorithms: An Evidence-Based Approach. *Cityscape*, 20(1): 117–146.
- U.S. Department of Housing and Urban Development. 2017. Notice CPD-17-01: Notice Establishing Additional Requirements for a Continuum of Care Centralized or Coordinated Assessment System. <https://www.hud.gov/sites/documents/17-01CPDN.PDF>.
- U.S. Department of Housing and Urban Development. 2022. The 2022 Annual Homeless Assessment Report (AHAR) to Congress: Part 1. <https://www.huduser.gov/portal/sites/default/files/pdf/2022-AHAR-Part-1.pdf>.
- Vaithianathan, R.; and Kithulgoda, C. I. 2020. Using Predictive Risk Modeling to Prioritize Services for People Experiencing Homelessness in Allegheny County.
- VanBerlo, B.; Ross, M. A.; Rivard, J.; and Booker, R. 2021. Interpretable machine learning approaches to prediction of chronic homelessness. *Engineering Applications of Artificial Intelligence*, 102: 104243.
- Von Wachter, T.; Bertrand, M.; Pollack, H.; Rountree, J.; and Blackwell, B. 2019. Predicting and preventing homelessness in Los Angeles. *California Policy Lab and University of Chicago Poverty Lab*.