Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)

Jeremy D. Goldhaber-Fiebert, PhD and Lea Prince, PhD, Stanford University

MARCH 26, 2023

EXECUTIVE SUMMARY

Goal

The impact evaluation assesses how implementation of the screening score (LASSO) within Allegheny County's Office of Children, Youth, and Families helped to:

Improve short-term outcomes for referrals by call-screeners (increasing the fraction of "higher risk" children who screen-in and their likelihood of further action taken upon investigation if screened-in)

Improve longer-term outcomes for referrals by call-screeners (decreasing subsequent referral rates and removal rates within 6-months; but being more likely to have a subsequent screen-in for those who initially screened-in and to have further action taken upon investigation and being more likely to have a subsequent screen-out and not to have further action taken upon investigation for those who initially screened-out)

Reduce disparities in terms of the above outcomes for similar children across racial groups

Approach

Following on our previous impact evaluation of Allegheny County's AFST tool, the present Phase 2 evaluation uses individual-level multivariable regressions employing generalized linear models and time-to-event models to achieve its assessment goals. The study primarily compares outcomes for children involved in GPS referral episodes in the 17 months after the full implementation of the LASSO tool (October 27, 2019 through March 31, 2021) (~20,000 children) to outcomes for children involved in GPS referral episodes in the period before implementation of any tool, primarily January 1, 2015 through July 31, 2016 (~31,000 children). As a second point of comparison, the study also evaluates changes in outcomes across these periods for children involved in CPS referral episodes for whom the LASSO tool was not part of the referral decisions. Further details appear in the Methods sections below.

Findings and Interpretation

Summaries of the overall findings for short-term and longer-term outcomes from the multivariable regression analyses are shown in **Table 1** with summaries of overall findings for analyzes of differences in outcomes by race group shown in **Table 2**. It is critical to note that the findings and their interpretation have two important caveats: 1) the effects of the implementation of the LASSO tool itself cannot be fully disentangled from the effects of accompanying policy and practice changes; 2) how the LASSO tool and the policy and practices changes may have impacted outcomes if COVID-19 had not occurred is also very difficult to determine.

Short-term outcomes

The LASSO tool and surrounding policy and practice changes increased GPS episode screen-in probabilities by about 2 percentage points, with the vast majority of this overall increased due to substantial increases in the probability of screening in for children involved in GPS episodes who were classified by the LASSO tool as High Risk Protocol. Despite the upwards shift in the average score of children involved in GPS screen-in episodes, the probability that they accepted for services declined by 8 percentage points with the implementation of the LASSO tool and surrounding policy and practice changes. However, at least part of the decrease may have been due to the policy and practice changes given that the probability of accepting for service after CPS episodes also declined.

Longer-term outcomes

The LASSO tool and surrounding policy and practice changes showed evidence of predictive accuracy and safety for 180 days after an index GPS episode. They decreased the 180-day risk of subsequent episodes after both index GPS screen-in episodes by 1.2 percentage points and index GPS screen-out episodes by 1.4 percentage points. Subsequent episodes in the 180-day period after an index GPS screen-in episode were more likely to screen in (3.8 percentage points) whereas subsequent episodes after an index GPS screen-out episode were less likely to screen in (2.8 percentage points). Furthermore, the subsequent 180-day risk of an episode that screened-in and accepted for service after an index GPS screen-in increased 2.8 percentage points while the risk declined after an index GPS screen-out by 2.7 percentage points. The 180-day risk of home removal after an index GPS screen-in episodes decreased by 2.7 percentage points with the implementation of LASSO and surrounding policy and practice changes. Home removals rates after an index GPS screen-out were much lower than for GPS screen-in episodes in the Pre-AFST period (approximately one third), and while they were even lower in the LASSO period (<1%), they could not be analyzed in full multivariable regression model given the low rates.

At least part of the changes in the longer-term outcomes for GPS episodes may have been attributable tool given differences in these 180-day outcomes after an index CPS episode. In comparison to index GPS screen-in episodes, after CPS index episodes, the 180-day risks for subsequent episodes did not decline significantly with the implementation of the LASSO tool and surrounding policy and practice changes. While it increased the 180-day risk of subsequent episodes that screened in for index CPS episodes by nearly the same amount as for index GPS screen-in episodes, it did not have a statistically significant effect on subsequent 180-day risks of accepting for services. Finally, the magnitude of the decline in home removals within 180 days of an index CPS episode (1.5 percentage points) was smaller than those observed for an index GPS episode.

Differences in outcomes between White children and Black/African American children

LASSO and surrounding policy and practice changes induced changes in short-term outcomes for both White children and Black/African American children that were consistent in terms of direction but and differential in terms of magnitude such that racial differences in outcome levels generally narrowed.

LASSO and surrounding policy and practice changes induced changes in longer-term 180-day outcomes after index GPS screen-in episodes and after index GPS screen-out episodes for both White children and Black/African American children that were consistent in terms of direction but and differential in terms of magnitude such that some racial differences in some outcome narrowed while others widened. In comparison, the race-specific changes in 180-day outcomes after index CPS episodes were inconsistent in terms of direction and differential in terms of magnitude such that most racial differences in some outcomes did not change significantly or widened. This suggests that the implementation of the LASSO tool and surrounding policy and practice changes improved 180-day predictive accuracy and safety for both White children and Black/African American children and generally narrowed but did not eliminate racial differences in these outcomes relative to their Pre-AFST magnitudes.

CONTENTS

METHODS

Overview

This Phase 2 impact evaluation a predictive risk modeling tool for Allegheny County focuses on the introduction of an updated screening score (LASSO) produced by the tool for use by Allegheny County Child Welfare Office's intake office as part of its decision-making process for children involved in GPS referrals combined with
a set of policy and practice changes. The evaluation analyzes how the LASSO tool and surrounding policy and practice changes affected several important short- and long-term outcomes as well as racial disparities in those outcomes in comparison to such outcomes for children prior to the use of any such tool (prior both to implementation of the original screening score [AFST] and to adoptions of the LASSO screening score). Short-term outcomes address how the LASSO tool and surrounding policy and practice changes improved accuracy of referrals by intake staff and ensured that children at high risk received further investigations and appropriate services. Analyses of longer-term outcomes address how the use of the LASSO tool and surrounding policy and practice changes for an index referral impacted rates of subsequent referrals, their likelihood of screening-in for investigations and attachment to services, and changes in home removal rates over 180 days.

Previously, we conducted an impact evaluation of the AFST tool and the policy and practice changes that accompanied its implementation.[1] The Phase 2 evaluation differs in several ways. Like the first evaluation/report, we focus on impacts on outcomes related to accuracy and safety as well as impacts on racial disparities in those outcomes. Unlike the first report, in addition to shorter term outcomes, we analyze longer-term, 180-day outcomes including home removal. Methodologically, we focus on individual-level multivariable analyses using generalized linear models and time-to-event models and do not use interrupted time series analyses as we did in the Phase 1 analysis.

The present phase 2 evaluation is substantially more challenging than the phase 1 evaluation because of two key differences: the timing of the LASSO implementation and COVID-19. Unlike the phase 1 evaluation, the time gap between the comparison pre-AFST period when no screening score was used and the use of the LASSO screening score is substantial. Since much could have changed in addition to the implementation during the intervening period, it is importance to employ statistical models that adjust for multiple covariates to account for changes in composition of the at-risk population, the policy environment, experience with the AFST tool, and other factors that may impact outcomes between the two analyzed periods. Of note, the COVID-19 pandemic occurred during the period in which the LASSO screening tool was implemented and used. The pandemic disrupted in-person schooling, employment of parents and caregivers, the delivery of social services, and the amount of time spent at home and in other locations. All of these COVID-related impacts could have also affected both the risks and outcomes measured in addition to potential effects on those outcomes due to

---

1   https://www.alleghenycountyanalytics.us/
    wp-content/uploads/2019/05/Impact-
    Evaluation-from-16-ACDHS-26_
    PredictiveRisk_Package_050119_FINAL-6.pdf

the LASSO screening tool as well as policies and practices surrounding it. Hence, we introduced controls into our analyses to account for shifts in outcomes due to COVID-19. Furthermore, we included comparisons to outcomes for children involved in CPS referrals for whom neither the AFST tool nor the LASSO tool were directly part of the investigatory and subsequent decisions.

The sections below describe the LASSO and its implementation, the outcome measures used, the policy changes the evaluation accounts for, the data used in the evaluation, and the analytic approach chosen to perform the evaluation along with its underlying rationale.

Implementation of the LASSO Scoring Tool and Other Changes Occurring with the Implementation of the LASSO Scoring Tool

The analysis focuses on two policy periods. The LASSO policy period (October 27, 2019 to October 2, 2021) and the Pre-AFST period (January 1, 2015–July 31, 2016).

The Pre-AFST period begins following a set of amendments to the State of Pennsylvania's existing Child Protective Services Law which became effective on December 31, 2014 and which had the effect of altering a number of features of referrals to Allegheny County call center.[2] The period ends on the day before the first version of the risk scoring tool (AFST) was implemented. The pre-AFST period is used as a counterfactual— to make a best attempt at measuring outcomes related to screening decisions in the absence of any screening tool and policy and practice changes that accompanied it.

The LASSO period starts following a set of substantial systematic changes to earlier versions of the tool on October 26, 2019. These changes among other things altered which children are flagged under the Low Risk Protocol (one of the categories of risk used in this analyses) (see **Appendix section A.1**). Hence, we analyze the LASSO data starting on October 27, 2019. We analyze referrals through October 2, 2020. Note that our data include referrals through March 31, 2021 so we are able to observe subsequent referrals or removals up to 180 days following the initial referral date.

Because the start of the global COVID-19 pandemic impacted the normal operation of the Allegheny County call center (both in terms of referrals coming into the call center [**Appendix Figure 1**] and in call-center logistics) along with its impacts on many other parts of daily life, we include a control for referrals coming in on or after March 15, 2020 in our multivariate analyses.

---

2   Child Protective Services Act, P.L.1240,
    No.206, 23 PA §§6301-6386 (2015)

Unit of Analysis

Unlike the phase 1 evaluation which analyzed each referral separately, the unit of analysis for the current evaluation is the "episode". Allegheny County's Office of Children, Youth, and Families receives referrals involving one or more child in a household. Each referral can be a "general protective service" (GPS) referral or a "child protective service" (CPS) referral.[3] Because, multiple referrals can occur within a short period of time and because the types of referrals and decisions about whether or not to screen them in for further investigation can be related to one another and because rules about how referrals that are close together are coded as screened-in or screened-out have changed over time, we defined an episode as a referral or cluster of referrals within 14 days of one another. We coded each episode:

> CPS episode: 1+ referral(s) within the episode is a CPS referral (note: CPS is always a screen-in)
>
> GPS screen-in: the episode contains no CPS referrals and 1+ screen-ins among the GPS referrals
>
> GPS screen-out: the episode contains only GPS screen-out referral(s)

In both the pre-AFST and LASSO time periods, >95% of episodes in the dataset contain only 1 referral. The vast majority of episodes with more than 1 referral consisted of 2 referrals. The mix of referral-types within GPS screen-in episodes and CPS episodes does not change substantially between policy periods (**Appendix Figures A2 and A3**).

Outcomes and Questions

We selected outcomes to measure underlying effects of the LASSO screening tool implementation in terms of the accuracy and safety of the call screening process for children involved in referrals and disparities across children's race in terms of accuracy and safety. We examined both short and longer-term outcomes. Specifically, the analysis examines how the implementation of the LASSO screening tool may have impacted these multiple outcomes, including:

---

3   Allegations fall under the state of Pennsylvania's Child Protective Service (CPS) statutes (23 Pa.C.S. § 6303) or General Protective Service (GPS) statutes (23 Pa.C.S. § 6334). CPS referrals include those made for child abuse, including physical and sexual abuse. CPS referrals must be investigated and require more urgent response times, often overlap with law enforcement and medical investigations, and lead to a determination of whether abuse occurred (that may result in perpetrators being registered in the state's ChildLine registry). GPS referrals include referrals made when there is a risk of harm. For example, neglect, truancy, and substance use by parents would all fall under GPS referrals. GPS referrals may be investigated or screened out without further assessment, at the discretion of call screening staff. GPS investigations assess for risk and safety to ensure well-being of children and provide families with any supports they may need. GPS investigations cannot result in registry with the state's ChildLine registry. Both CPS and GPS referrals can result in a family having a case opened at the end of an investigation for ongoing services and supports. In the pre-AFST period, approximately 21 percent of DHS referrals were CPS referrals and 79 percent were GPS referrals. In the LASSO period, approximately 18 percent of DHS referrals were CPS referrals and 82 percent were GPS referrals.

Short-Term Accuracy Outcomes (Overall and by LASSO risk category)

1. How did the probability of a GPS episode screening-in change with the implementation of the LASSO tool?

   A child is considered to be in this category if a referral episode that includes the child is screened-in

2. How did the probability of accepting for service for GPS screen-in episodes change with the implementation of the LASSO tool?

3. For comparative purposes, how did the probability of accepting for service for CPS episodes change with the implementation of the LASSO tool?

   A child is considered to be in this category if a referral episode that includes the child is screened-in and upon investigation the disposition is further action taken. "Further action" is defined by the referral service decision and occurs when a referral accepts for service or connects to either an open case or connects to a closed case and is re-opened for service. This outcome is computed for all children.

LASSO-related changes in these outcomes are intended to measure how LASSO tool implementation impacted two feature of accuracy: a) are otherwise similar children more or less likely to be screened-in for investigation and does this occur because the tool identifies those as higher (lower) risk who are then more (less) likely to screen-in? b) are otherwise similar children that screen-in for investigation more likely to have further action taken and does this occur because the tool identifies and helps to screen in higher risk children (and screen out lower risk children)?

Longer-Term Accuracy Outcomes (Overall and by LASSO risk category)

4. How did the probability of having any episode within 180 days after a GPS screen-in episode change with the implementation of the LASSO tool?

5. How did the probability of having screen-in episode (GPS or CPS) within 180 days after a GPS screen-in episode change with the implementation of the LASSO tool?

6. How did the probability of having screen-in episode (GPS or CPS) that accepts for service within 180 days after a GPS screen-in episode change with the implementation of the LASSO tool?

7. How did the probability of a home removal within 180 days after a GPS screen-in episode change with the implementation of the LASSO tool?

8. For comparative purposes, how did the probability of having any episode within 180 days after a CPS episode change with the implementation of the LASSO tool?

9. For comparative purposes, how did the probability of having screen-in episode (GPS or CPS) within 180 days after a CPS episode change with the implementation of the LASSO tool?

10. For comparative purposes, how did the probability of having screen-in episode (GPS or CPS) that accepts for service within 180 days after a CPS episode change with the implementation of the LASSO tool?

11. For comparative purposes, how did the probability of a home removal within 180 days after a CPS episode change with the implementation of the LASSO tool?

LASSO-related changes in these outcomes are intended to measure how LASSO tool implementation impacted longer-term accuracy: are otherwise similar children that screen-in for investigation who are higher (lower) risk then subsequently more (less) likely to have episodes, episodes that screen-in, accept for service, or have home removal? If the LASSO tool helps to identify children for whom investigation and services for the index episode can be effective, then one may expect lower rates of subsequent episodes, especially those that require another screen-in, services, or result in removal. In contrast, since CPS episodes do not involve the use of the LASSO tool, we would expect either no changes in similar outcomes or else only changes in outcomes related to other factors including some of the policy and practice changes that accompanied the LASSO tool implementation.

Longer-Term Safety Outcomes (Overall and by LASSO risk category)

12. How did the probability of having any episode within 180 days after a GPS screen-out episode change with the implementation of the LASSO tool?

13. How did the probability of having screen-in episode (GPS or CPS) within 180 days after a GPS screen-out episode change with the implementation of the LASSO tool?

14. How did the probability of having screen-in episode (GPS or CPS) that accepts for service within 180 days after a GPS screen-out episode change with the implementation of the LASSO tool?

15. How did the probability of a home removal within 180 days after a GPS screen-out episode change with the implementation of the LASSO tool?

LASSO-related changes in these outcomes are intended to measure how LASSO tool implementation impacted longer-term safety: are otherwise similar children that screen-out who are higher (lower) risk then subsequently more (less) likely to have episodes, episodes that screen-in, accept for service, or have home removal? If the LASSO tool helps to identify children that do not require investigation and services for the index episode, then one may expect lower rates of subsequent episodes, especially those that require screen-in, services, or result in removal.

An analogy can be made between these accuracy/safety outcomes and the more general concepts and language of screening test assessment. In general screening test terminology, test accuracy is measured based on sensitivity (i.e., true positive fractions) which is the percentage of those subjects with the underlying condition who test positive and specificity (i.e., the true negative fraction) which is the percentage of those subjects without the underlying condition who test negative. Ultimately, the ideal is to have a test with a high positive predictive value, the fraction of test positives that have the underlying condition and the fraction of test negatives who do not have the underlying condition. In our context, accuracy for screen-ins (i.e., test positives) is the fraction of children who screen-in that have further action taken upon investigation (an indicator of the underlying condition). Likewise, accuracy for screen-outs (i.e., test negatives) is the fraction of children who screen-out that have no further episodes over 180 days (an indicator of the absence of the underlying condition). The hypothesized effect of the LASSO tool would be to increase both of these.

Disparities Outcomes

For questions 1–15, we examine how changes in probabilities differed between black and white children.

LASSO-related changes in these outcomes are intended to measure how LASSO tool implementation may have helped to standardize system responses which in turn may have reduced differences between outcomes for otherwise similar black and white children.

While other/undetermined race categories are controlled for in the main analysis for the short- and long-term accuracy and safety outcomes, the disparities analysis is limited to stratification by white and Black/African American race; these two categories comprises over 90% of children included in referrals.[4]

Data

All analyses use de-identified data relating to those involved in referrals to Allegheny County's call center. The data consist of demographics, allegation, call-screening decisions, investigations, accept for service decisions, and removals for all children included in all referrals to the call center. Additionally, the data track previous referrals and investigations with child welfare and other child-serving systems from August 1, 2013 through March 31, 2021.

The analytic dataset focuses on outcomes (described above) for children under the age of 18 years at the time of the first referral in an episode. Children >17 years of age at the time of the first referral were excluded (although we account for 18-year-old children in subsequent episodes).

Because there was a practice change related to investigations on the 60th day (see **Appendix section A.1**), we exclude episodes in which the time from referral to service decision is ≥60 days (0% of GPS screen-out episodes by definition; 0.5% of GPS screen-in episodes; 2.1 % of CPS episodes).

Because we use outcomes from CPS episodes as a further comparator for analyses of outcomes from GPS screen-in episodes, we drop episodes that are designated CPS, but are mixed CPS-GPS (a small fraction of the CPS episodes in both the Pre-AFST and LASSO periods, see Appendix Figures 2 and 3).

Overview of Analytic Approach

For the outcomes, we first report a comparison of unadjusted population means for the Pre-AFST Period (January 1, 2015 through July 31, 2016) and the LASSO Period (October 27, 2019 through March 31, 2021).

For short-term outcomes, we perform individual-level multivariable regression analyses to estimate the impact of the LASSO implementation and surrounding policy and practice changes on the predicted level of each

---

4   A child was coded as "Black/African American" if his/her race was Black, African American or mixed Black or African American, at the time of the referral. For outcomes which incorporate re-referrals, race was coded based on the race recorded in the index referral.

outcome (e.g., probabilities of an episode screening-in). Results from the multivariable analyses are reported as the predicted probability that an outcome will occur. Predicted probabilities are averaged predicted values for each child in the dataset with all variables held at their observed value.

For longer-term outcomes, we performed multivariable time-to event (survival) analyses to estimate the rates of an outcome occurring (e.g., subsequent episodes of various types and home removal). The model was a generalized gamma survival model, chosen for its flexibility in underlying assumptions regarding changes in the baseline hazard over follow-up time. Children are considered "at-risk" while they are 18 years old or younger and they are not removed from their home (as determined by dates of removal/return in the data set). We estimate cumulative 180-day risks for each outcome by measuring the area under the average adjusted survival curve at 180 days for each policy period.

We also examine all outcomes vary across policy periods within each LASSO risk score category. We generate predicted probabilities for all children as described above except that we hold score category at a fixed value score and all other independent variables at their observed values.

Finally, for all outcomes, to examine LASSO's impact for children in different risk categories or of different races (Black or white),[5] we estimate the predicted probability of an outcome for all children by holding race category at a fixed value and all other independent variables at their observed values.

We perform all analyses using Stata (v14) software.

Covariates and Standard Errors
For the models involving multivariate adjustments, definitions of the included covariates are as follows. Child characteristics at the time of the first referral in an episode include the child's age group at referral (<4 years, 4 to 6, 7 to 12, or 13 to 17), race, legal sex, and risk score category (low risk protocol [LR], low, medium, high, high risk protocol [HR])[6] based on the maximum score across all referrals within a given episode (using the LASSO algorithm to create risk score for Pre-AFST children). Indicators are included to identify episodes in which the referrals are attached to an open case (ACT) or if the referrals are truancy allegations. Finally, to absorb level differences by month of the year and to account for change in call volume related to the COVID-19 pandemic (**Appendix Figure 1**), we control for the month during which the first referral in an episode was recorded, a binary indicator of referrals taking place after March 15, 2020, and the interaction of these two. We also control for episode history — the number of previous episodes since the start of the year 2013 divided by the number of days at risk over the same period. Because we expect some correlation in outcomes among individual children appearing more than once in the dataset, we cluster the standard errors at the child-level.

---

5   The race of child was coded as noted above. For outcomes which incorporate re-referrals, race was coded based on the race recorded in the index referral.

6   See Appendix section A.1 for protocol and risk score category definitions, and their policy implications.

Description of Specific Methods

First, the simplest comparison we perform is the comparison of unadjusted means of outcomes and covariates in the Pre-AFST and LASSO periods, testing whether they are statistically different from one another. Comparisons of levels of outcomes that show statistically significant differences between the two policy periods suggest that the LASSO tool and the accompanying policy and practice changes may have had an impact. However, comparisons of levels of covariates that show statistically significant differences between the two periods suggest that case-mix is changing in potentially important ways over time and hence the multivariable adjustment is particularly important to assess whether observed changes in outcomes are in fact (at least partially) attributable to changes in case mix.

Second, we used multivariable individual-episode-level regression analyses to assess short-term outcomes while adjusting for child and referral characteristics. Specifically, we estimate Generalized Linear Models (*glm* in Stata) with a logit link enabling greater flexibility in the distribution of the error term than a standard logit model. We run our analysis at the level of the episode for any child involved in any call (i.e., not only the child for which the call was made).

The dependent variables in our specifications are the short-term outcomes: Screen-in or Accept-for-service (conditional on screening in). The independent variables are a binary indicator of the policy period (0 if pre-AFST or 1 if LASSO) and all of the covariates described above. We interact policy with risk score category because we hypothesize that the impact of LASSO may be different at different risk levels and also because we are interested specifically in how LASSO impacted the short-term outcomes for children of different risk.

The following GLM model specification represents the formalizing of this approach (but where we have not presented the clustering at the child level across episodes to aid in readability):

$$
\begin{aligned}
Z_i = &\; \beta\, LASSO + \sum \beta\, ScoreCat_i + \sum (\beta\, LASSO * ScoreCat_i) + \sum \beta\, AgeCat_i \\
&+ \sum \beta\, GenderCat_i + \sum \beta\, RaceCat_i + \beta\, PriorEpisodeRate_i \\
&+ \sum \beta\, EpisodeMonth_i + \beta\, COVIDPeriod_i \\
&+ \sum (\beta\, COVIDPeriod_i * EpisodeMonth_i) + \beta\, Active_i + \beta\, LASSO_i * Active_i \\
&+ \beta\, Truancy_i
\end{aligned}
$$

$$
Q_i = \frac{1}{1 + e^{(-Z_i)}}
$$

$$
Y_i \sim Bern(Q_i)
$$

where *i* represents a specific episode for a specific child and LASSO is the indicator variable of the policy period, *ScoreCat* is the LASSO risk category of the child at the time of the episode, *AgeCat* is the age category of the child at the time of the episode, *GenderCat* is the gender category of the child at the time of the episode, *RaceCat* is the race category of the child at the time of the episode, *PriorEpisodeRate* represents the number of prior episodes in which the child was involved since 2013 divided by the total time the child was at risk for

episodes, *EpisodeMonth* indicates the month of the year in which the episode occurred, *COVIDPeriod* indicates whether the episode occurred on a date on or after the start of the COVID-19 pandemic, *Active* represents whether the episode attached to an active family, and *Truancy* represents whether the episode involved truancy.

To examine differences across race (Black/African American versus White), we also interact policy and risk score category with race category. Thus the only differences with the GLM specification above is that we limit the *RaceCat* to those who are either White or Black/African American (now a binary indicator where 0 = White and 1 = Black/African American), and we add an interaction term between *RaceCat* and LASSO policy period ($LASSO * RaceCat_i$).

We use the margins command in Stata to compute the predicted level of each outcome for all children, by risk score category, by race, and by race and risk score category, both Pre-AFST and post LASSO implementation. These analyses do not evaluate Pre-AFST or LASSO time trends analyses but rather, they focus on estimates of the average effect of the LASSO adjusting for evolving case mix over time. The predictive margins presented in tables and figures of the results can be interpreted as the average outcome if all children in the sample were in either the Pre-AFST or the LASSO period, holding all other control variables as they happen to be.

Third, we use a time-to event model (survival analysis) to assess longer-term outcomes while adjusting for the same set of characteristics and indicators and with the same policy-risk score- race interactions described above. In time-to-event analysis, we can analyze the length of time until an outcome occurs. In this case, we estimate the time until a subsequent episode occurs, the time until a subsequent episode that screens in occurs, the time until a subsequent episode that screens in and accepts for services occurs, and the time until home removal occurs. A child is "at-risk" in the model for up to 180 days. At-risk time can be shorter if a child turns 18 or is removed from the household (for all outcomes other than time to home removal). Using time-series analyses, we can construct a "survival curve", which indicates the probability of remaining under observation without a subsequent event occurring (i.e., episode or removal) over time. We can then estimate the probability of an outcome occurring by measuring the area under the survival curve at 180 days. We estimate a Generalized Gamma model, which allows non-constant, non-monotonic underlying hazard (i.e., as time passes since an index episode, the rate of subsequent events can both increase and decrease). These analyses use the same specifications with the same control variables as described in the equations above.

Note on Interpretation of Findings Given Analytic Approach
Our analyses report on results describing differences in Pre-AFST and LASSO period outcomes. These results inform us about how the implementation of the LASSO tool along with concomitant policy and practice changes as well as changes in call-screener or supervisor behavior or knowledge performs compared to a system prior to the implementation of any risk tool or related training/experience or policy or practice changes (see **Appendix Section A.1**). Because this counterfactual becomes increasingly imperfect in the phase 2 evaluation due to the now substantial time-gap between the Pre-AFST and LASSO policy periods relative to this approach for the phase 1 evaluation, we use changes in outcomes in CPS episodes as a second point of comparison to changes in outcomes for GPS screen-in episodes. Recall that CPS episodes always screen-in for investigation. Thus, the LASSO tool is not directly relevant for CPS episodes and their subsequent outcomes. Hence, if we observe changes in outcomes between pre-AFST and LASSO periods for the CPS episodes, this implies that there are

factors at play other than the LASSO tool itself which may be driving (at least partially) changes that we observe between pre-AFST and LASSO period outcomes in the GPS episodes. Unfortunately, we are not able to draw firm conclusions about what these other factors might be – for example, the policy and practice changes that accompanied the LASSO tool may impact CPS-related outcomes as well (e.g., investigations and attach for service processes and decisions), the LASSO tool itself may induce change in downstream decisions in some systematic way (so we cannot simply subtract out the CPS changes as a counterfactual), or there could be external (unobservable) factors which are different between the two periods but that impact both CPS and GPS episodes. While a design which randomized contemporaneous GPS episodes to the use of the LASSO tool would allow us to estimate the effects directly, randomization in this situation was not possible. Therefore, we report our results but interpret them cautiously and conservatively.

RESULTS

Descriptives

Case Mix Is Different across the Two Policy Periods
The main analysis included 31,379 children involved in GPS episodes in the Pre-AFST period and 20,114 children in the LASSO period. Tables 3a and 3b show the demographic characteristics of these children along with other referral characteristics. These tables also highlight the importance of evaluating outcomes adjusting for case mix given the statistically significant differences in the characteristics of children across the two periods. While the mix of legal sex stayed roughly similar, the fraction who were female or other sex increased. Likewise, while there were no major shifts in the racial distribution of the sample, the fraction who were Black/African American race increased and the fractions that were Other race or Unable to Determine decreased by a factor of nearly 2. Similarly, the age distribution appeared overall similar though there was a slight shift upwards in the fraction of children aged 13-17 and downwards in the fractions in the < 4 years and 4–6 years age groups.

In terms of referral and history related characteristics, differences tended to be larger/more appreciable, the LASSO score category of children shifted towards a higher predicted risk population with a smaller fraction in the low risk protocol (LR) and low and medium categories and a larger fraction in high and high risk protocol (HR) categories. The fraction of children with a referral attaching to an active family increased and the fraction flagged as a referral due to truancy declined.

Unadjusted Differences in Outcomes across the Two Policy Periods
Numerous short-term and longer-term outcomes changed between the Pre-AFST and LASSO periods in comparisons that were not adjusted for changes in covariates (Tables 4–6). Outcome changes could be observed overall as well as for subgroups of children defined by their LASSO-classified risk level. Differences in outcomes for children across race groups also changed. Many of these differences were statistically significant.

Multivariable Adjusted Analyses

Adjusted Differences in Short-Term Outcomes across the Two Policy Periods
LASSO and surrounding policy and practice changes increased GPS episode screen-in probabilities, especially for children who were High Risk Protocol (HR): Use of the LASSO tool and accompanying changes to policy and practice for GPS episodes resulted in a 2 percentage point increase in the probability of screen-in

(3.8% increase). The pattern of change across children of different risk scores as classified by the LASSO tool was consistent – children who were at lower risk were less likely to screen-in while those in High Risk Protocol were substantially more likely to screen-in (**Table 1, Figure 1, Appendix Tables B1a and B1b**).

LASSO and surrounding policy and practice changes decreased Accepting for Service probabilities for children involved in GPS episodes that screened in across LASSO-classified risk levels: The probability of accepting for service after a GPS screen-in declined by 8 percentage points (20% decrease) despite a shift to a higher risk distribution of children involved in GPS episodes that screened in. Declines were consistent across the LASSO-classified risk gradient of children (**Table 1, Figure 2, Appendix Tables B2a and B2b**).

At least part of the decrease in accepting for service probabilities may have been due to other policy and practice changes given that accepting for service after CPS episodes also declined across LASSO-classified risk levels: The magnitude of the decline for CPS episodes was 3.8 percentage points (18.4% decrease) was half the GPS absolute percentage point decline but nearly equal in terms of relative magnitude of decline given that there were lower CPS Pre-AFST probabilities of accepting for service (**Table 1, Figure 3, Appendix Tables B3a and B3b**).

Adjusted Differences in Longer-Term Outcomes across the Two Policy Periods
The 180-day risks of subsequent episodes decreased after index GPS episodes with the implementation of LASSO and surrounding policy and practice changes, with the smaller number of subsequent episodes following a GPS screen-in being more likely to screen-in and accept for services and with those following a GPS screen-out being less likely to screen-in and to accept for services, suggesting reasonable predictive accuracy and safety: Overall subsequent 180-day risks of referral episodes declined after both GPS screen-in and screen-out episodes (1.2 and 1.4 percentage points respectively [4.4% and 5.4% decreases respectively]). However, the subsequent 180-day risk of an episode that screened-in after an index GPS screen-in increased 3.8 percentage points (20.5% increase) while the risk declined after an index GPS screen-out by 2.8 percentage points (15.9% decrease). Furthermore, the subsequent 180-day risk of an episode that screened-in and accepted for service after an index GPS screen-in increased 2.8 percentage points (34.1% increase) while the risk declined after an index GPS screen-out by 2.7 percentage points (41.1% decrease) (**Table 1, Figures 4-9, Appendix Tables C1a-C6b**). Thus, with the LASSO tool, GPS screen-in episodes were followed by a lower rate of episodes but ones that were more likely to screen-in and be actionable, whereas GPS screen-out episodes were followed by a lower rate of episodes and ones that were less likely to screen-in or be actionable. In comparison, after CPS index episodes the 180-day risks for subsequent episodes, episodes screening-in, and episodes accepting for service either did not change or increased, suggesting the role of the LASSO tool in these changes (**Table 1, Figures 10-12, Appendix Tables C7a-C9b**).

The 180-day risk of home removal after an index GPS screen-in episode decreased with the implementation of LASSO and surrounding policy and practice changes (home removals rates after a GPS screen-out were very low and had insufficient sample size for full multivariable analysis of their changes): Overall subsequent 180-day risks of home removal after an index GPS screen-in declined 2.7 percentage points (33.4% decrease), with most of the decline for children that the LASSO tool classified as medium risk, high risk, or High Risk Protocol at their index episode. This result provides additional evidence for longer-term safety (**Table 1, Figure 13, Appendix Tables C10a and C10b**).

At least part of the decrease in the 180-day risk of home removal may have been due to other policy and practice changes given that home removal after CPS episodes also declined: The magnitude of the decline for CPS episodes was 1.5 percentage points (33.3% decrease) which was nearly half the GPS absolute percentage point decline but nearly equal in terms of relative magnitude given lower CPS pre-AFST 180-day probabilities of home removal (**Table 1, Figure 13, Appendix Tables C11a and C11b**).

Adjusted Differences in Outcomes between White Children and Black/African American Children across the Two Policy Periods

LASSO and surrounding policy and practice changes induced changes in short term outcomes for both White children and Black/African American children that were consistent in direction and differential in magnitude such that racial differences in outcome levels generally narrowed: Probabilities of screening in after a GPS episode were increased with the implementation of LASSO and accompanying policy and practice changes for both White children and Black/African American children. While a significant difference remained between the groups in the LASSO period it was narrower than in the Pre-AFST period (**Table 2, Appendix Figure D1, Appendix Tables D1a and D1b**). Probabilities of accepting for service after a GPS screen-in declined with the implementation of LASSO and accompanying policy and practice changes for both White children and Black/African American children. There was no longer a significant difference between the groups in the LASSO period (**Table 2, Appendix Figure D2, Appendix Tables D2a and D2b**). In comparison, probabilities of accepting for service after a CPS episode declined with the implementation of LASSO and accompanying policy and practice changes for both groups. The significant difference between groups for CPS episodes in the Pre-AFST period remained in the LASSO period (unlike for GPS screen-ins) though there was non-significant narrowing of this difference between the groups for CPS episodes (**Table 2, Appendix Figure D3, Appendix Tables D3a and D3b**).

LASSO and surrounding policy and practice changes induced changes in subsequent outcomes within 180 days of an index GPS screen-in episode for both White children and Black/African American children that were consistent in direction and differential in magnitude such that some racial differences in outcome levels widened while others narrowed: For both groups, probabilities of subsequent episodes within 180 days of an index GPS screen-in episode declined even as both the probability of a subsequent screen-in episode and the probability of accepting for service after a screen-in episode within 180 days of an index GPS screen-in episode increased. This implied the predictive accuracy increased with LASSO for both race groups. However, these changes were differential in magnitude for White children and Black/African American such that across-group differences in some outcomes widened (**Table 2, Appendix Figures D4-D6, Appendix Tables D4a-D6b**). In comparison, for these same outcomes within 180 days of an index CPS episode, changes in outcomes were not always consistent in direction nor magnitude (**Table 2, Appendix Figures D7-D9, Appendix Tables D7a-D9b**).

LASSO and surrounding policy and practice changes induced changes in subsequent outcomes within 180 days of an index GPS screen-out episode for both White children and Black/African American children that were consistent in direction and differential in magnitude: For both groups, probabilities of subsequent episodes within 180 days of an index GPS screen-out episode declined along with declines in both the probability of a subsequent screen-in episode and the probability of accepting for service after a screen-in episode. This implied that safety increased with LASSO for both race groups. These changes were differential in magnitude for White children and Black/African American (**Table 2, Appendix Figures D10-D12, Appendix Tables D10a-D12b**).

LASSO and surrounding policy and practice changes reduced home removals for both White children and Black/ African American children within 180 days of an index GPS screen-in episode such that the pre-existing racial difference this outcome was narrowed and was no longer significant: While there was a significant difference in the 180-day probability of home removal after a GPS screen-in episode in the Pre-AFST period between White children (7.3 percentage points) and Black/African American children (8.7 percentage points), the levels declined for both groups (5.9 percentage points for White children; 5.0 percentage points for Black/African American children) in the LASSO period, and the difference between the two groups was no longer significant (**Table 2, Appendix Figure D13, Appendix Tables D3a and D13b**).

DISCUSSION, CONCLUSIONS, AND IMPLICATIONS

We evaluated the impact of the LASSO screening score implementation in conjunction with policy and practice changes within Allegheny County's Office of Childre, Youth, and Families in terms of its short- and longer-term impacts on outcomes related to accuracy and safety and on racial differences in these outcomes for children involved in GPS referrals. Overall, our analyses showed that the LASSO screening score and surrounding policy and practice changes modestly increased the probability of a GPS episode screening in especially for children classified by the LASSO tool as High Risk Protocol. However, both children involved in GPS screen-in episodes and CPS episodes across the risk spectrum were less to accept for services, suggesting the role of policy and practice changes in impacting this outcome. The LASSO tool and surrounding policy and practice changes appeared to modestly increase the predictive accuracy of GPS screen-in decisions over the subsequent 180 days without negatively affecting safety. Overall rates of subsequent episodes declined; however subsequent episodes following an index GPS screen-in episode were more likely to screen-in and to accept for service while subsequent episodes following an index GPS screen-out episode were less likely to screen-in and to accept for services. These longer-term impacts on accuracy and safety appeared to be related to the LASSO tool itself as index CPS episodes did not have similar changes to their subsequent episodes over 180 days. Likewise, rates of home removals in the 180 days after an index GPS screen-in episode also declined, though, as this also occurred for index CPS episodes, at least part of this change may have been due to surrounding policy and practice changes. In general, the impacts of the LASSO tool implementation and surrounding policy and practice changes were consistent for White children and for Black/African American children. They tended to be different in magnitude such that for many of the GPS episode short and longer-term outcomes, differences were reduced from those observed in the Pre-AFST period.

As the previous summary of findings implies, cleanly separating out the effects of the LASSO tool itself from the surrounding policy and practice changes is extremely difficult. Ultimately, however, the combined effect may be the most relevant from a systems implementation evaluation perspective since such tools are unlikely to be implemented without such accompanying changes.

It should also be emphasized that several other features of this phase 2 evaluation make its interpretation challenging and call for caution when attributing changes as causal effects. First, is that there is a roughly 3 year gap between the end of the comparator (Pre-AFST) period and the start of the LASSO period — given this gap, other changes due to the AFST or to secular trends over that time may make this compactor more challenging to interpret as a counterfactual. While we control for case mix and other known policy change differences, such

controls are imperfect. Even more importantly, COVID-19 occurred in the middle of the LASSO period. Given the level of disruption to all parts of society due to the pandemic, it is impossible to known what would have happened had the pandemic not occurred. While we control the pandemic period in our analyses, such controls are unlikely to capture the pandemic's full effects.

One of the outcomes we examined was the effect of the implementation of the LASSO and surrounding policy changes on differences in outcomes between White children and Black/African American children. It is important to note that true underlying rates of neglect and maltreatment for each of these groups is unknown and hence increases/decreases in a given measured system outcome (e.g., screen-ins) of one subgroup relative to another in principle could represent either a widening or a narrowing of a disparity (e.g., in terms of children experiencing actual neglect or maltreatment having the referral investigated). One reassuring finding in this regard is that the general consistency of direction of change in outcomes for White children and for Black/African American children accompanied narrowing of differences in many of the outcomes for GPS episodes (for which the LASSO tool is most directly relevant) that were not generally observed across these groups for CPS episodes. Even so, proper interpretation of such results critically depends on the ability of the LASSO tool to detect risk of actual neglect or maltreatment in each group and for workers to act accordingly.

The goal of the evaluation of the effects of the LASSO tool and surrounding policies was to provide a set of measures that are meaningful and important. However, the evaluation makes no claim or judgement about the relative importance of one outcome related to another. We encourage interpretation of findings across outcomes in a holistic way and with reference to the stated goals and constraints of child-serving systems.

In conclusion, our evaluation of the effects of implementing the LASSO tool and surrounding policy and practice changes shows modest improvements in the predictive accuracy of GPS screen-ins without any negative changes to safety. The effects appear to often occur as a combination of the tool itself and the policy and practice changes. Their interpterion is challenged by the advent of the COVID-19 pandemic — it is impossible to know for sure what impacts they may have had in the absence of the pandemic. Evaluations of large-scale, real-world system changes like the LASSO tool implementation must contend with the reality of implementation, the myriad other changes and adjustments that occur, and the gap between process measures and the underlying concepts (e.g., accuracy, safety) that they are intended to represent. Ultimately, Allegheny County and other systems considering the use tools like LASSO will need to consider how such metrics relate to their core goals – doing so from a holistic assessment of a range of short and longer-term outcomes is advisable. In this case and with these caveats in mind, a holistic assessment the LASSO tool and surrounding policy and practice changes supports the conclusion that it modestly improved a range of outcomes and modestly narrowed differences in such outcomes across race groups.

TABLES

TABLE 1. Summary of Change in Outcomes during the LASSO Period Compared to the Pre-AFST Period from the Multivariable Regression Analyses

OUTCOME	EPISODE TYPE	ALL	BY RISK GROUP				
			LR	LOW	MED	HIGH	HR
SHORT-TERM							
Screen-in	GPS	**Up**	**Down**	**Down**	**Down**	Down	**Up**
Accept for Service	GPS screen-in	**Down**	Down	**Down**	**Down**	**Down**	**Down**
Accept for Service	CPS	**Down**	**Down**	**Down**	Down	**Down**	**Down**
LONGER-TERM							
Episode within 180 days	GPS screen-in	**Down**	Down	**Down**	**Down**	**Down**	Up
SI episode within 180 days	GPS screen-in	**Up**	**Up**	**Up**	**Up**	**Up**	**Up**
Accept within 180 days	GPS screen-in	**Up**	**Up**	Down	**Up**	**Up**	**Up**
Removal within 180 days	GPS screen-in	**Down**	Down	Down	**Down**	**Down**	**Down**
Episode within 180 days	GPS screen-out	**Down**	Down	Down	**Down**	Down	**Down**
SI episode within 180 days	GPS screen-out	**Down**	**Down**	**Down**	**Down**	Down	**Down**
Accept within 180 days	GPS screen-out	**Down**	Down	**Down**	**Down**	**Down**	**Down**
Removal within 180 days	GPS screen-out	N/A	N/A	N/A	N/A	N/A	N/A
Episode within 180 days	CPS	Up	**Down**	**Up**	Up	**Up**	Down
SI episode within 180 days	CPS	**Up**	**Down**	**Up**	**Up**	**Up**	**Up**
Accept within 180 days	CPS	Up	**Down**	Down	Down	**Up**	**Down**
Removal within 180 days	CPS	**Down**	**Down**	**Down**	Up	**Down**	**Down**

Up: Probability of Outcome Increased in the LASSO Period compared to otherwise similar children in the Pre-AFST period

Down: Probability of Outcome Decreased in the LASSO Period compared to otherwise similar children in the Pre-AFST period

Darker shading and bolding means statistically significant at p<0.05; lighter shading and non-bolding means not statistically significant

n/a: Insufficient sample size and outcomes to estimate the multivariable regression analyses for this outcomeSI: Screen-in

TABLE 2. **Summary of Change in Differential Outcomes between White children and Black/African American children during the LASSO Period Compared to the Pre-AFST Period from the Multivariable Regression Analyses**

OUTCOME	EPISODE TYPE	PRE-AFST		LASSO		CHANGE		DIFFERENTIAL CHANGE	ORDER SWITCH
		WHITE	BLACK	WHITE	BLACK	WHITE	BLACK		
SHORT TERM									
Screen-in	GPS	**44.1**	**49.1**	**47.1**	**50.5**	3.0	1.4	1.6	No
Accept for Service	GPS screen-in	**37.1**	**42.7**	31.7	32.1	-5.4	-10.6	**5.2**	No
Accept for Service	CPS	**18.4**	**22.4**	**14.9**	**18.3**	-3.5	-4.1	0.4	No
LONGER-TERM									
Episode within 180 days	GPS screen-in	28.7	28.9	28.2	27.6	-.05	-1.3	0.7	Yes
SI episode within 180 days	GPS screen-in	18.7	19.4	23.7	22.8	5.0	3.4	**1.6**	Yes
Accept within 180 days	GPS screen-in	7.5	8.8	11.4	10.7	3.9	1.9	**2.1**	Yes
Removal within 180 days	GPS screen-in	**7.3**	**8.7**	5.9	5.0	-1.4	-3.7	**2.3**	Yes
Episode within 180 days	GPS screen-out	**28.1**	**27.0**	**25.4**	**26.4**	-2.7	-0.6	**2.2**	Yes
SI episode within 180 days	GPS screen-out	18.4	18.4	**14.7**	**16.2**	-3.7	-2.2	1.5	No
Accept within 180 days	GPS screen-out	5.9	7.7	4.0	4.4	-1.9	-3.3	1.4	No
Removal within 180 days	GPS screen-out	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Episode within 180 days	GPS	**26.6**	**24.6**	26.0	26.1	-0.6	1.5	2.1	Yes
SI episode within 180 days	GPS	17.0	16.1	19.4	20.7	2.4	4.65	**2.3**	Yes
Accept within 180 days	GPS	**5.8**	**7.4**	6.2	6.6	0.4	-0.8	1.2	No
Removal within 180 days	GPS	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

The Pre-AFST and LASSO columns show the levels of each outcome for White children and Black/African American children in each policy period. Bolded pairs of numbers within the Pre-AFST or LASSO columns show when the White/Black difference in an outcome was significant in that respective policy period.

The Change columns changes for White children and Black/African American children from the Pre-AFST to the LASSO period. For example, a -5.4 means that the LASSO period's outcome level was 5.4 percentage points lower than in the Pre-AFST level of the outcome.

The Differential Change column shows the absolute magnitude of the difference between the race-specific Pre-AFST-to-LASSO changes in the outcome level. Those colored green indicate that the White/Black gap narrowed in the LASSO period compared to the Pre-AFST period, while those colored yellow indicate that the gap widened. Bold indicates that the difference in the Pre-AFST-to-LASSO changes was significant.

The Order Switch column indicates whether the level of one race group's outcome was higher in the Pre-AFST period but due to differential change was lower in the LASSO period.

SI: Screen-in

TABLE 3A. **Summary statistics, child characteristic of children in GPS episodes**

	PRE-AFST			LASSO			P-VALUE*
	MEAN	95% CI		MEAN	95% CI		
LEGAL SEX							
Female	48.8%	48.2%	49.3%	49.1%	48.4%	49.8%	0.037
Male	50.9%	50.3%	51.4%	50.4%	49.7%	51.1%	
Other	0.3%	0.3%	0.4%	0.5%	0.4%	0.6%	
RACE							
White	42.1%	41.6%	42.6%	41.4%	40.7%	42.1%	0.000
Black	47.2%	46.7%	47.8%	52.6%	51.9%	53.3%	
Other	4.3%	4.1%	4.5%	2.8%	2.5%	3.0%	
Unable to Determine	6.4%	6.1%	6.6%	3.2%	3.0%	3.5%	
AGE-GROUP							
< 4 years	23.1%	22.6%	23.6%	22.9%	22.3%	23.5%	0.007
4–6 years	17.9%	17.5%	18.3%	17.0%	16.5%	17.5%	
7–12 years	34.8%	34.3%	35.4%	34.8%	34.2%	35.5%	
13–17 years	24.1%	23.7%	24.6%	25.3%	24.7%	25.9%	
LASSO SCORE CATEGORY							
LR	12.4%	12.0%	12.7%	8.7%	8.3%	9.0%	0.000
LOW	14.8%	14.4%	15.2%	10.0%	9.6%	10.4%	
MED	24.9%	24.5%	25.4%	22.7%	22.1%	23.2%	
HIGH	25.6%	25.1%	26.1%	27.1%	26.5%	27.7%	
HR	22.0%	21.5%	22.5%	25.4%	24.8%	26.0%	
Attach to active family	12.8%	12.4%	13.2%	15.1%	14.6%	15.6%	0.000
Flag for court truancy	1.6%	1.5%	1.7%	0.0%	0.0%	0.1%	0.000

*Sample sizes are 31,379 (Pre-AFST) and 20,114 (LASSO). The child is coded as Black if their race is recorded "Black or African American" or "Black or African American" combined with another race. *P-value is following a Pearson's chi-square test for the equality of distribution across policy periods. Flag for court truancy controls for policy differences in the handling of cases where children were processed for truancy issues. Attach to active family flag controls for policy differences related to the decisions made on active family cases.*

TABLE 3B. **Summary statistics, child characteristic of children in CPS episodes**

	PRE-AFST			LASSO			P-VALUE*
	MEAN	95% CI		MEAN	95% CI		
LEGAL SEX							
Female	49.6%	48.5%	50.6%	50.9%	49.4%	52.4%	0.167
Male	50.3%	49.3%	51.4%	48.9%	47.4%	50.4%	
Other	0.1%	0.0%	0.2%	0.2%	0.0%	0.3%	
RACE							
White	42.2%	41.2%	43.2%	39.8%	38.3%	41.3%	0.000
Black	47.8%	46.7%	48.8%	55.2%	53.7%	56.7%	
Other	4.3%	3.9%	4.8%	3.4%	2.8%	3.9%	
Unable to Determine	5.7%	5.2%	6.1%	1.7%	1.3%	2.0%	
AGE-GROUP							
< 4 years	18.5%	17.7%	19.3%	21.3%	20.1%	22.5%	0.001
4–6 years	17.6%	16.8%	18.4%	16.3%	15.2%	17.4%	
7–12 years	36.4%	35.4%	37.4%	34.2%	32.8%	35.7%	
13 –17 years	27.5%	26.6%	28.4%	28.1%	26.8%	29.5%	
LASSO SCORE CATEGORY							
LR	16.2%	15.4%	17.0%	10.1%	9.2%	11.0%	0.000
LOW	19.2%	18.4%	20.0%	12.4%	11.4%	13.4%	
MED	24.5%	23.6%	25.4%	22.8%	21.6%	24.1%	
HIGH	21.0%	20.1%	21.8%	23.7%	22.5%	25.0%	
HR	19.1%	18.3%	19.9%	25.0%	23.7%	26.3%	
Active family	16.8%	16.0%	17.6%	18.8%	17.6%	20.0%	0.005
Flag for court truancy	n/a						

*Sample sizes are 8,845 (Pre-AFST) and 4,290 (LASSO). The child is considered "Black or African American" if their race is recorded as "Black or African American" or "Black or African American" combined with another race. *P-value is following a Pearson's chi-square test for the equality of distribution across policy periods. Attach to active family flag controls for policy differences related to the decisions made on active family cases.*

TABLE 4. Means of short-term outcomes

	PRE-AFST				LASSO				P-VALUE*
	TOTAL COUNT	MEAN OUTCOME	95% CI		TOTAL COUNT	MEAN OUTCOME	95% CI		
SCREEN-IN, GIVEN GPS EPISODE									
All children	31,379	45.36%	44.81%	45.91%	20,114	49.73%	49.04%	50.42%	0.000
White	13,210	42.23%	41.39%	43.08%	8,332	43.16%	42.10%	44.22%	0.181
Black/African American	14,824	48.95%	48.14%	49.75%	10,582	56.62%	55.68%	57.57%	0.000
LR	3,883	31.29%	29.83%	32.75%	1,741	19.47%	17.61%	21.33%	0.000
Low	4,657	36.31%	34.93%	37.69%	2,015	26.85%	24.91%	28.79%	0.000
Med	7,825	46.53%	45.42%	47.64%	4,560	39.96%	38.53%	41.38%	0.000
High	8,033	51.92%	50.83%	53.02%	5,444	52.94%	51.61%	54.27%	0.247
HR	6,904	50.77%	49.59%	51.95%	5,113	81.85%	80.79%	82.91%	0.000
ACCEPT FOR SERVICES, GIVEN GPS SCREEN-IN									
All children	14,234	36.80%	36.01%	37.59%	10,002	38.74%	37.79%	39.70%	0.002
White	5,579	32.28%	31.05%	33.51%	3,596	36.10%	34.53%	37.67%	0.000
Black/African American	7,256	40.88%	39.75%	42.01%	5,992	40.24%	39.00%	41.48%	0.456
LR	1,215	21.48%	19.17%	23.79%	339	15.93%	12.01%	19.84%	0.025
Low	1,691	27.50%	25.37%	29.63%	541	17.19%	14.00%	20.38%	0.000
Med	3,641	32.16%	30.64%	33.68%	1,822	23.71%	21.76%	25.66%	0.000
High	4,171	39.46%	37.98%	40.95%	2,882	38.24%	36.46%	40.01%	0.300
HR	3,505	48.16%	46.50%	49.81%	4,185	50.73%	49.21%	52.24%	0.025
ACCEPT FOR SERVICES, GIVEN CPS SCREEN-IN									
All children	8,845	16.87%	16.09%	17.65%	4,290	17.27%	16.14%	18.40%	0.563
White	3,733	12.38%	11.32%	13.43%	1,708	11.30%	9.80%	12.80%	0.258
Black/African American	4,227	21.65%	20.40%	22.89%	2,367	22.01%	20.34%	23.68%	0.731
LR	1,432	4.05%	3.03%	5.07%	433	1.15%	0.14%	2.17%	0.003
Low	1,696	5.84%	4.72%	6.95%	532	3.20%	1.70%	4.69%	0.017
Med	2,167	13.80%	12.34%	15.25%	979	13.69%	11.53%	15.84%	0.934
High	1,854	21.36%	19.49%	23.23%	1,018	15.03%	12.83%	17.23%	0.000
HR	1,689	37.77%	35.46%	40.09%	1,074	39.29%	36.37%	42.22%	0.424

Counts are based on episodes, not individual children, so some children may be represented more than once. The counts of unique children for GPS episodes; GPS screen-in; CPS screen-in are 23,555; 12,256; 7,956 for Pre-AFST and 16,010; 8101; 3,978 for LASSO. Note that counts of children by race and score category may not add up to count of all children due to omitted categories or missing data.

*P-value is the two-sided p-value based on a two-sample t-test of the equality of means.

TABLE 5. Means of longer-term outcomes; subsequent episode within 180 days

	PRE-AFST			LASSO			P-VALUE*
	TOTAL COUNT	MEAN OUTCOME	95% CI	TOTAL COUNT	MEAN OUTCOME	95% CI	
SUBSEQUENT EPISODE, GIVEN GPS SCREEN-IN							
All children	14,234	23.38%	22.69% 24.08%	10,002	30.02%	29.13% 30.92%	0.000
White	5,579	24.16%	23.04% 25.29%	3,596	29.00%	27.52% 30.49%	0.000
Black/African American	7,256	24.21%	23.23% 25.20%	5,992	31.48%	30.30% 32.65%	0.000
LR	1,215	16.38%	14.29% 18.46%	339	18.58%	14.42% 22.75%	0.338
Low	1,691	17.62%	15.80% 19.44%	541	19.59%	16.24% 22.95%	0.300
Med	3,641	21.53%	20.20% 22.87%	1,822	20.80%	18.94% 22.67%	0.534
High	4,171	25.08%	23.76% 26.39%	2,882	27.24%	25.61% 28.86%	0.042
HR	3,505	28.47%	26.98% 29.97%	4,185	39.57%	38.09% 41.05%	0.000
SUBSEQUENT EPISODE, GIVEN GPS SCREEN-OUT							
All children	17,145	25.57%	24.92% 26.22%	10,112	23.82%	22.99% 24.65%	0.001
White	7,631	26.09%	25.11% 27.08%	4,736	22.87%	21.67% 24.06%	0.000
Black/African American	7,568	27.56%	26.56% 28.57%	4,590	27.17%	25.88% 28.46%	0.635
LR	2,668	15.52%	14.14% 16.89%	1,402	17.26%	15.28% 19.24%	0.151
Low	2,966	15.00%	13.72% 16.29%	1,474	17.77%	15.82% 19.73%	0.017
Med	4,184	27.15%	25.80% 28.50%	2,738	24.95%	23.32% 26.57%	0.041
High	3,862	29.52%	28.08% 30.96%	2,562	30.95%	29.16% 32.74%	0.220
HR	3,399	36.75%	35.12% 38.37%	928	36.42%	33.32% 39.52%	0.856
SUBSEQUENT EPISODE, GIVEN CPS							
All children	8,845	21.68%	20.83% 22.54%	4,290	27.48%	26.15% 28.82%	0.000
White	3,733	21.46%	20.14% 22.77%	1,708	24.41%	22.38% 26.45%	0.015
Black/African American	4,227	23.49%	22.21% 24.77%	2,367	30.08%	28.23% 31.93%	0.000
LR	1,432	13.20%	11.44% 14.95%	433	12.24%	9.14% 15.34%	0.603
Low	1,696	10.79%	9.31% 12.27%	532	15.60%	12.51% 18.69%	0.003
Med	2,167	21.55%	19.82% 23.28%	979	25.74%	23.00% 28.48%	0.010
High	1,854	24.76%	22.79% 26.72%	1,018	33.01%	30.11% 35.90%	0.000
HR	1,689	36.53%	34.23% 38.83%	1,074	41.71%	38.76% 44.67%	0.006

Counts are based on episodes, not individual children, so some children may be represented more than once. The counts of unique children for GPS screen-in; GPS screen-out; CPS are 12,256; 14,360; 7,956 for Pre-AFST and 8,101; 9,106; 3,978 for LASSO. Note that counts of children by race and score category may not add up to count of all children due to omitted categories or missing data.

*P-value is the two-sided p-value based on a two-sample t-test of the equality of means.

TABLE 6. **Means of longer-term outcomes; removal within 180 days**

	PRE-AFST			LASSO				P-VALUE*
	TOTAL COUNT	MEAN OUTCOME	95% CI		TOTAL COUNT	MEAN OUTCOME	95% CI		
REMOVAL, GIVEN GPS SCREEN-IN								
All children	14,234	7.67%	7.23%	8.11%	10,002	6.78%	6.29%	7.27%	0.009
White	5,579	6.49%	5.84%	7.14%	3,596	6.84%	6.02%	7.67%	0.508
Black/African American	7,256	8.63%	7.98%	9.27%	5,992	6.69%	6.06%	7.33%	0.000
LR	1,215	4.12%	3.00%	5.23%	339	2.65%	0.93%	4.37%	0.214
Low	1,691	3.61%	2.72%	4.50%	541	2.77%	1.38%	4.16%	0.352
Med	3,641	4.81%	4.11%	5.50%	1,822	3.46%	2.62%	4.30%	0.021
High	4,171	8.94%	8.08%	9.81%	2,882	6.42%	5.52%	7.31%	0.000
HR	3,505	12.35%	11.26%	13.44%	4,185	9.34%	8.46%	10.22%	0.000
REMOVAL, GIVEN GPS SCREEN-OUT								
All children	17,145	2.46%	2.22%	2.69%	10,112	0.95%	0.76%	1.14%	0.000
White	7,631	1.77%	1.47%	2.06%	4,736	0.49%	0.29%	0.68%	0.000
Black/African American	7,568	3.30%	2.90%	3.71%	4,590	1.55%	1.19%	1.90%	0.000
LR	2,668	0.37%	0.14%	0.61%	1,402	0.29%	0.01%	0.56%	0.643
Low	2,966	0.37%	0.15%	0.59%	1,474	0.27%	0.01%	0.54%	0.591
Med	4,184	1.31%	0.97%	1.66%	2,738	0.80%	0.47%	1.14%	0.047
High	3,862	2.72%	2.21%	3.23%	2,562	1.41%	0.95%	1.86%	0.000
HR	3,399	7.06%	6.20%	7.92%	928	2.80%	1.74%	3.87%	0.000
REMOVAL, GIVEN CPS								
All children	8,845	4.19%	3.78%	4.61%	4,290	4.29%	3.68%	4.90%	0.801
White	3,733	2.79%	2.26%	3.31%	1,708	2.99%	2.18%	3.79%	0.681
Black/African American	4,227	5.65%	4.96%	6.35%	2,367	5.20%	4.30%	6.09%	0.434
LR	1,432	0.70%	0.27%	1.13%	433	0.46%	0.01%	1.10%	0.590
Low	1,696	1.83%	1.19%	2.47%	532	1.69%	0.59%	2.79%	0.837
Med	2,167	3.18%	2.44%	3.92%	979	4.19%	2.93%	5.44%	0.156
High	1,854	5.72%	4.66%	6.78%	1,018	3.83%	2.65%	5.01%	0.027
HR	1,689	9.18%	7.80%	10.56%	1,074	8.19%	6.55%	9.84%	0.374

Counts are based on episodes, not individual children, so some children may be represented more than once. The counts of unique children for GPS screen-in; GPS screen-out; CPS are 12,256; 14,360; 7,956 for Pre-AFST and 8,101; 9,106; 3,978 for LASSO. Note that count of children by race and score category may not add up to count of all children due to omitted categories or missing data.

*P-value is the two-sided p-value based on a two-sample t-test of the equality of means.

FIGURES

Short-term outcomes

FIGURE 1: **Adjusted Analysis: probability of a screen-in in a GPS episode**
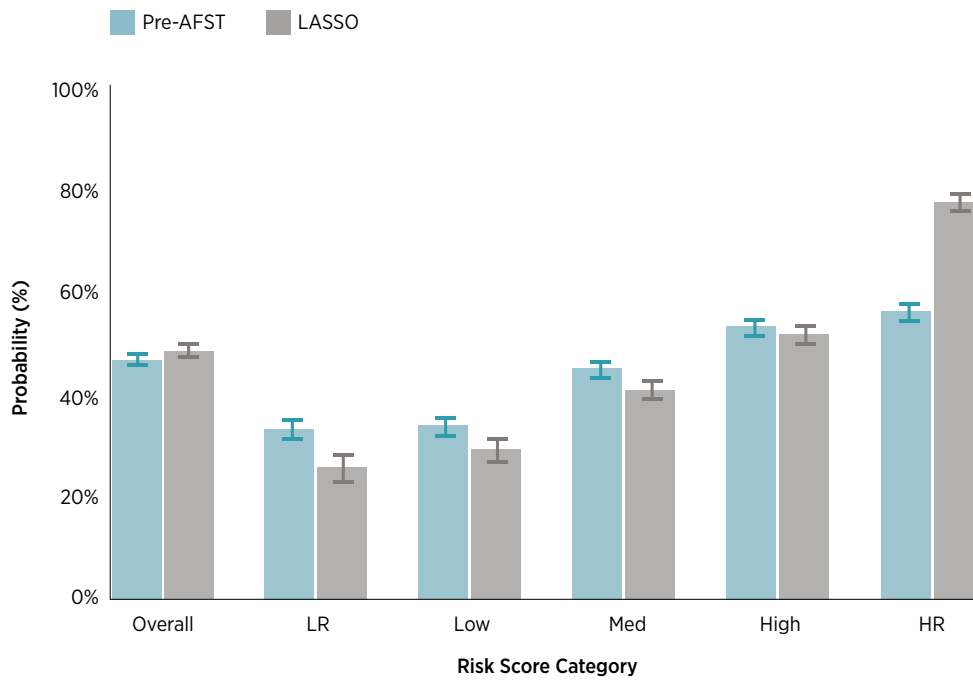
FIGURE 2: **Adjusted Analysis: probability of accept for service in a GPS episode screen-in**
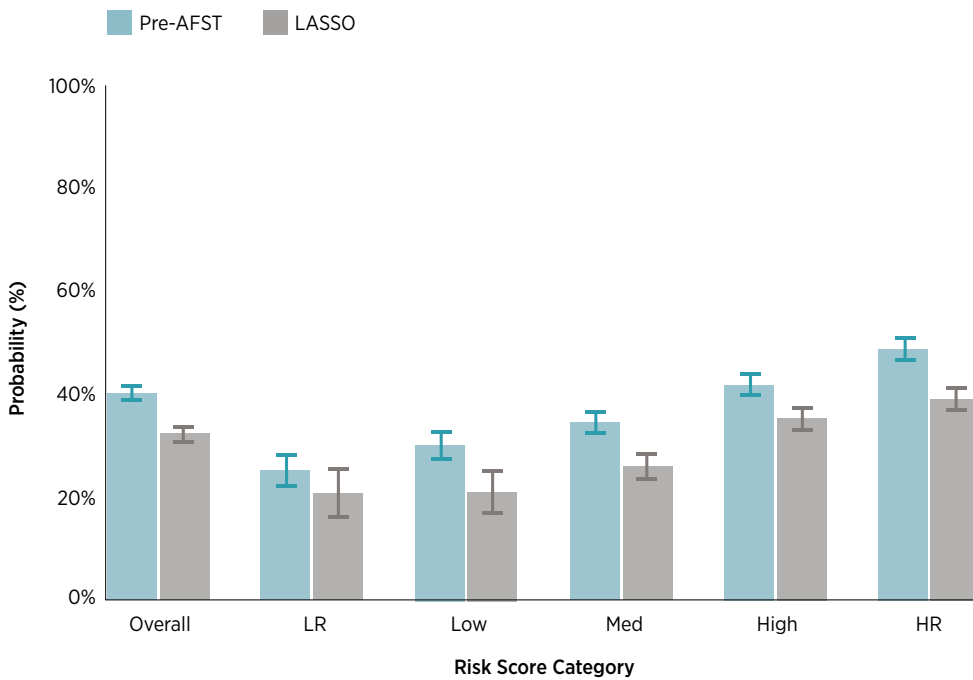
FIGURE 3: **Adjusted Analysis: probability of accept for service in a CPS episode**
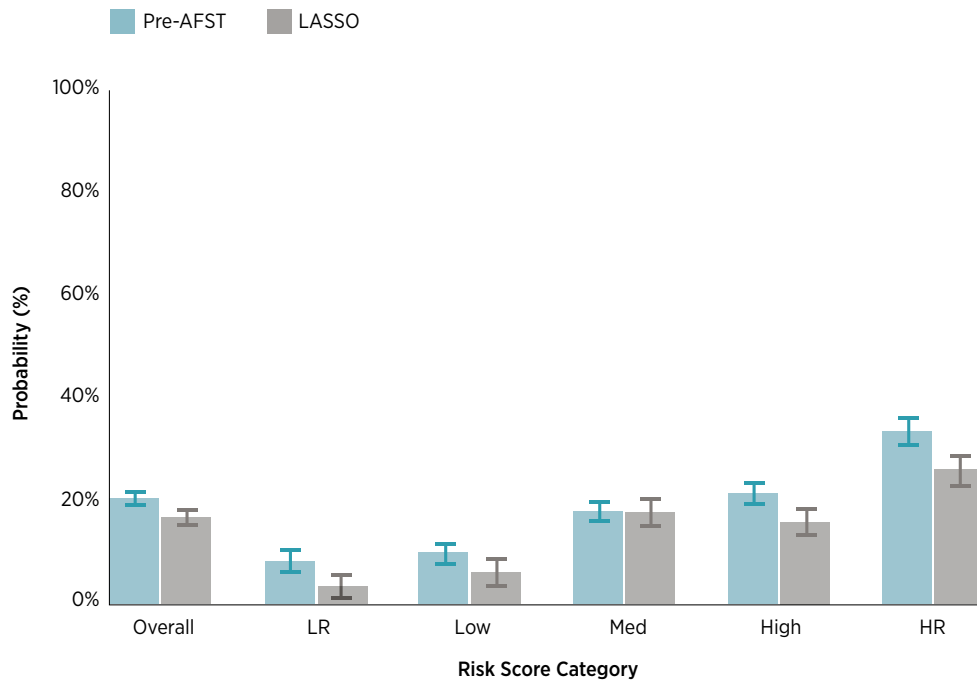
Longer-term outcomes

FIGURE 4. **Adjusted Analysis: probability of another episode within 180 days of a GPS screen-in episode**
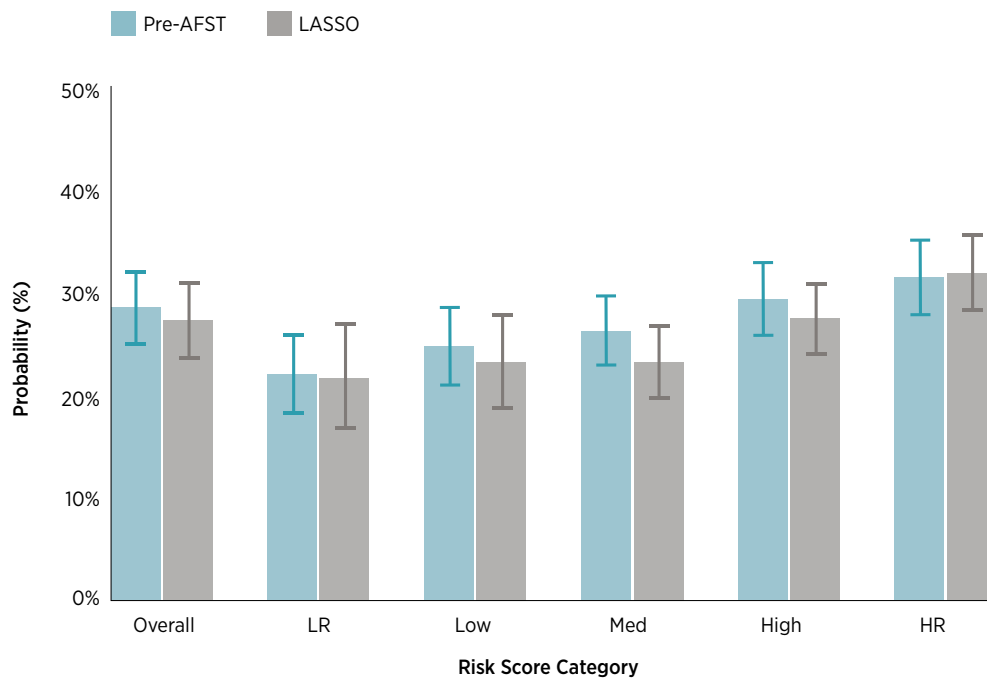
FIGURE 5. **Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-in episode**
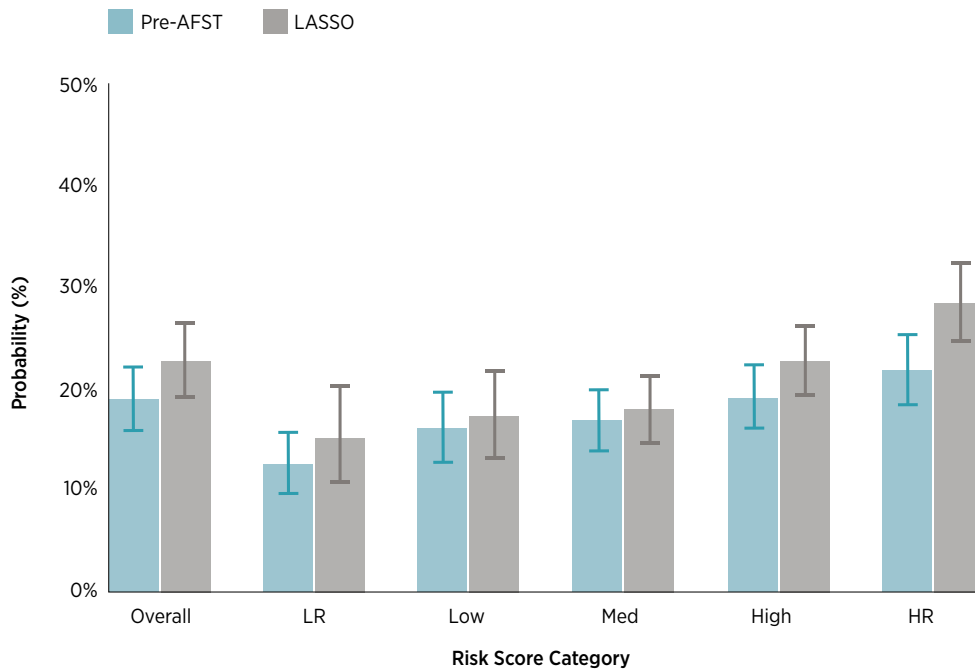
FIGURE 6. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-in episode**
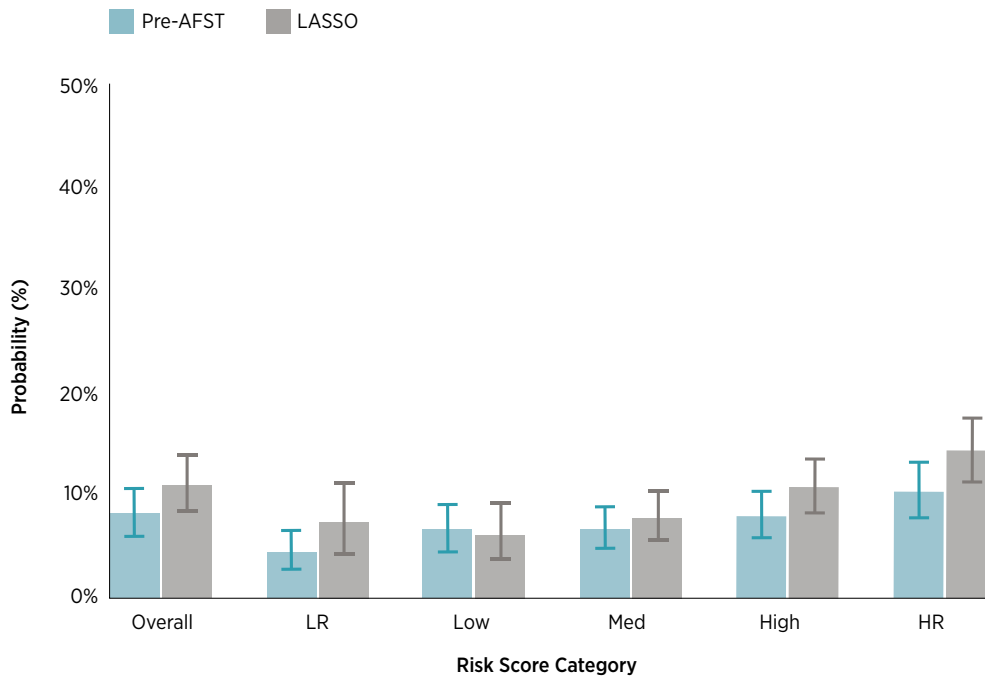
FIGURE 7. **Adjusted Analysis: probability of another episode within 180 days of a GPS screen-out episode**
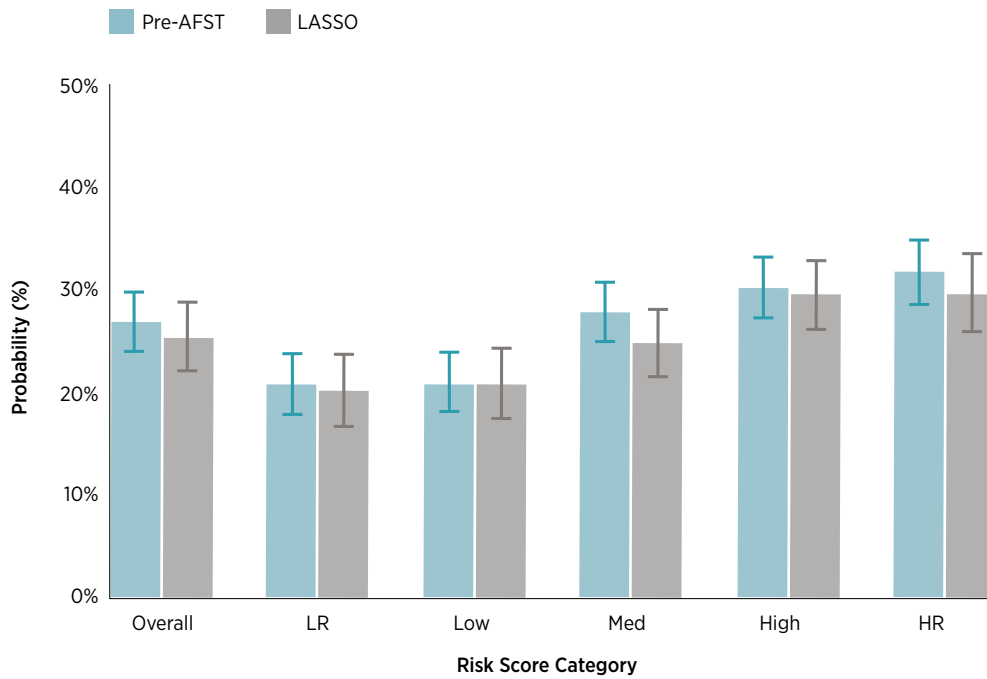
FIGURE 8. **Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-out episode**
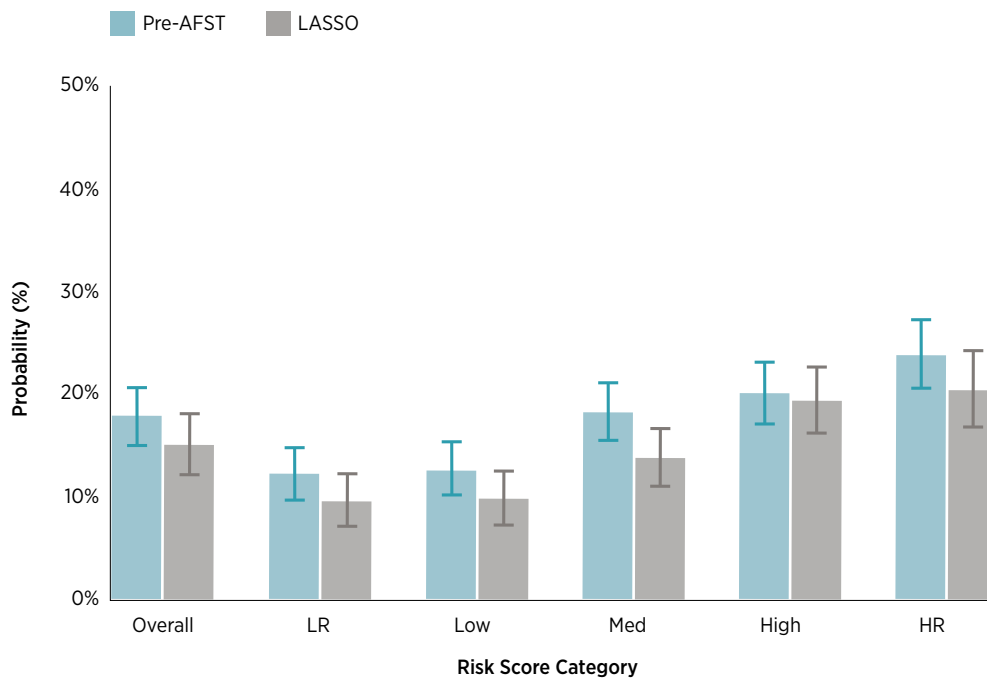
FIGURE 9. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-out episode**
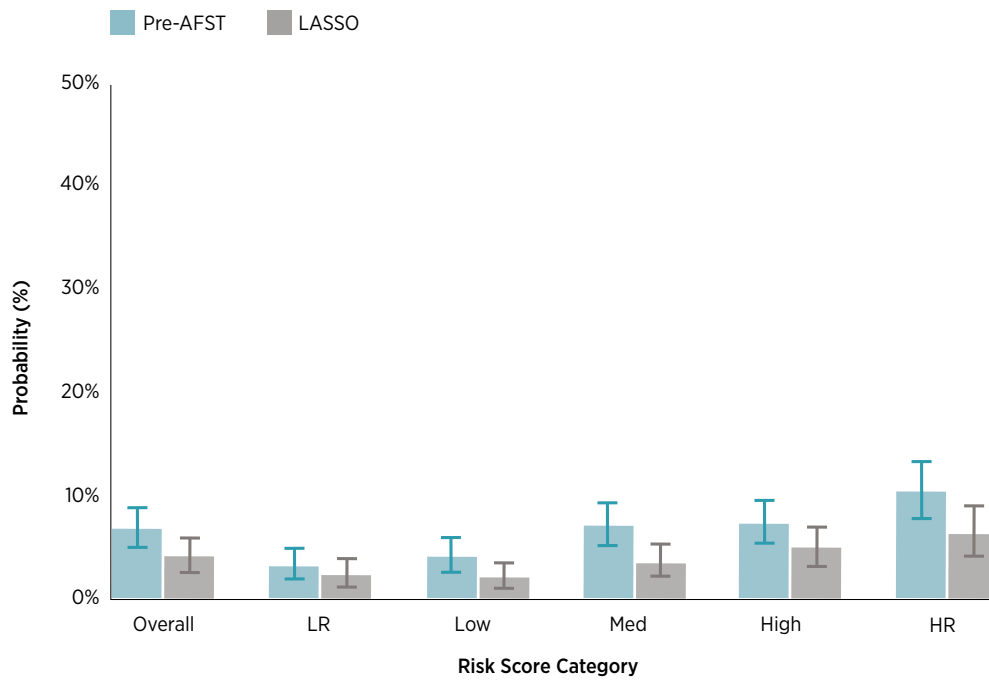
FIGURE 10. **Adjusted Analysis: probability of another episode within 180 days of a CPS episode**
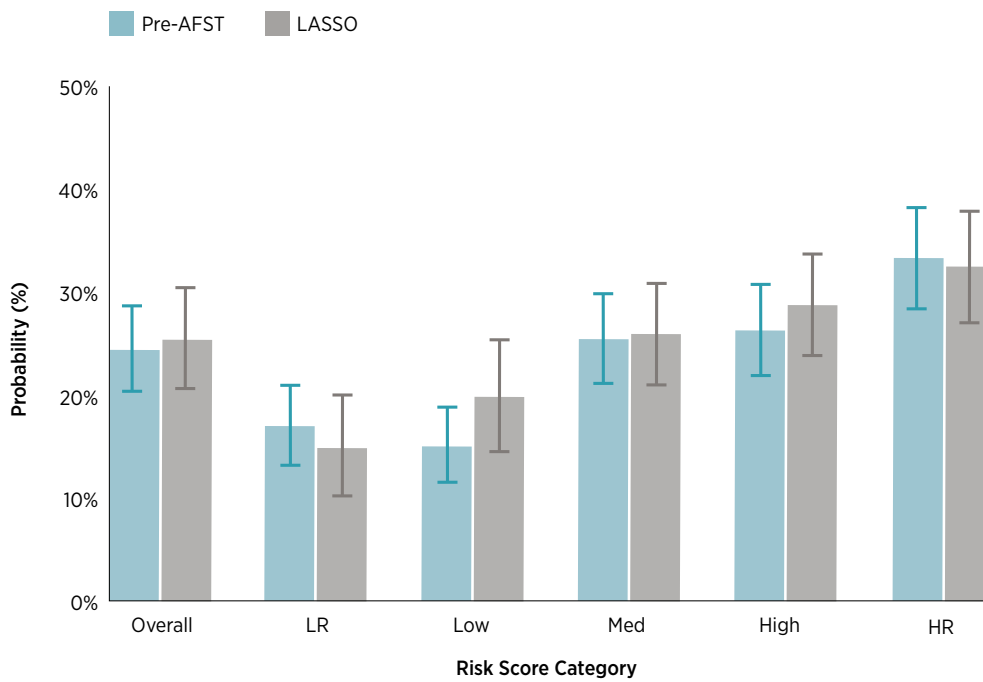
FIGURE 11. Adjusted Analysis: probability of another episode that screens-in within 180 days of a CPS episode
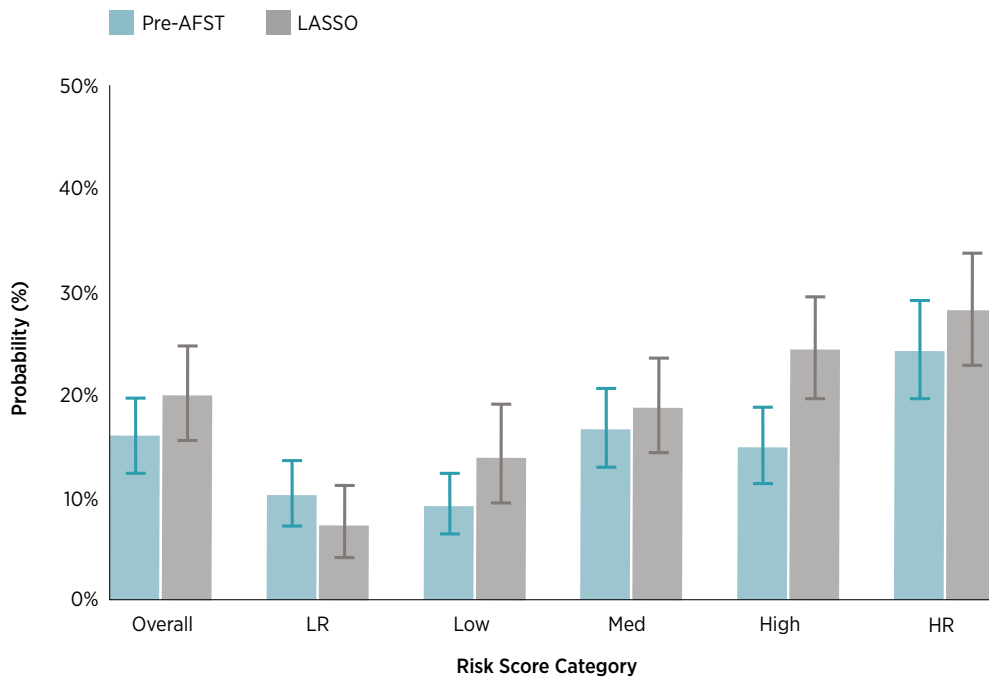
FIGURE 12. Adjusted Analysis: probability of another episode that accepts for service within 180 days of a CPS episode
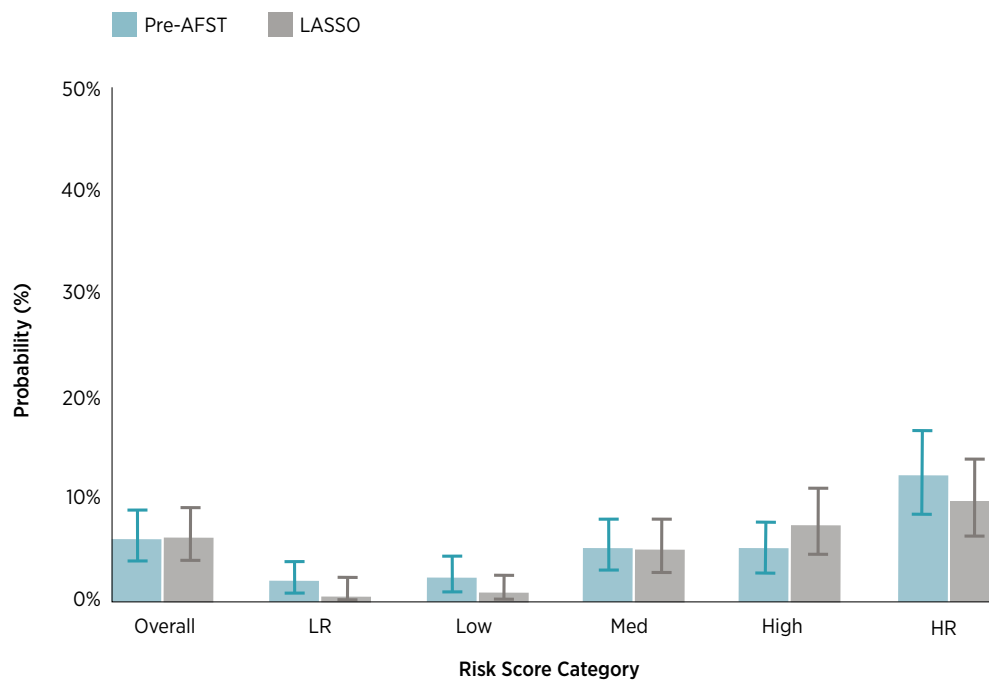
FIGURE 13. **Adjusted Analysis: probability of home removal within 180 days of a GPS screen-in episode**
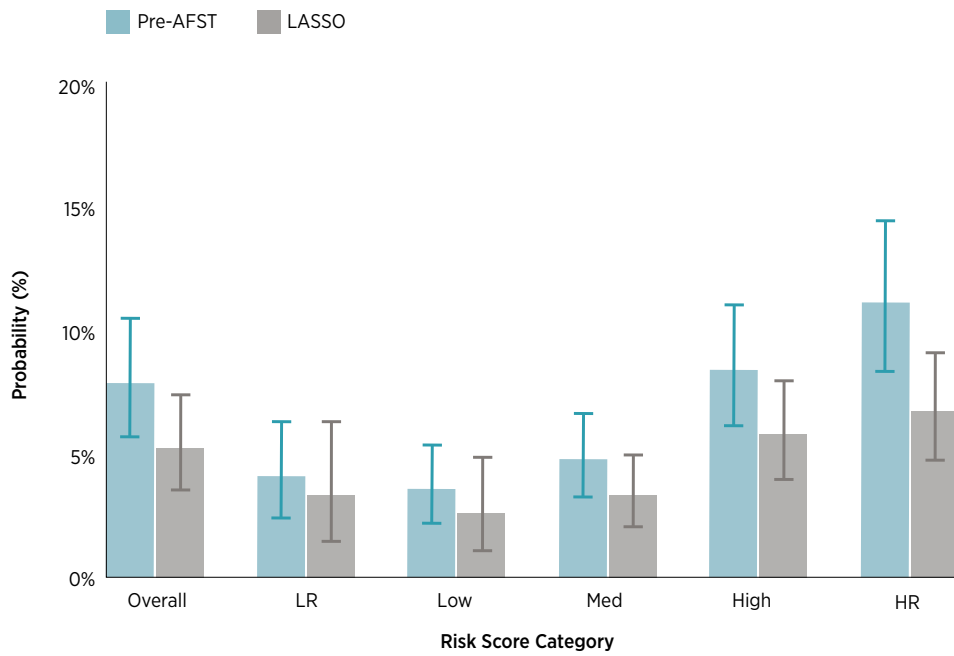
FIGURE 14. **Adjusted Analysis: probability of home removal within 180 days of a CPS episode**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *37*
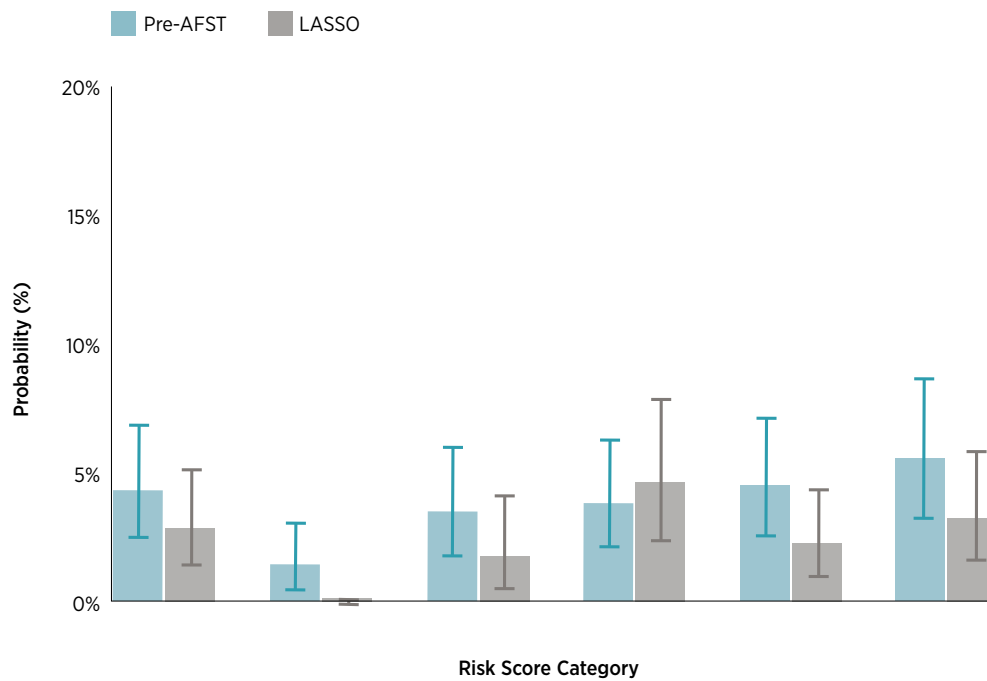
APPENDIX A

APPENDIX SECTION A

A.1. Description of policy and practice changes over time change

The following details key events surrounding implementation of risk scoring tools and other policy and practice changes relevant for the present analyses.

Rationale for the start of the Pre-AFST period in the analysis

On January 1, 2015, new mandated reported legislation went into effect in Pennsylvania. This resulted in a considerable increase in referral rates that persisted over time, as well as potential shifts in the underlying risk levels of the new referrals coming into the call center.

Also, in January 2015, Allegheny County began maintaining information regarding referrals that were otherwise expunged in the statewide database, to assist in future risk and safety assessments and research. Prior to 2015, referrals (particularly CPS referrals) that were expunged in the statewide database were fully and systematically expunged in the county databases as well. This policy change was codified by Pennsylvania law in 2018. As a result of the prior policy and the 2015 change, CPS data prior to 2015 are almost entirely unavailable, while data are nearly complete from 1/1/2015 onward.

For these reasons, although we have some degree of data back through 2013, we focus on the referrals occurring from January 2015 until the implementation of the first screening tool (AFST).

The AFST tool and policy and practice changes that occurred

Our phase 1 analysis evaluated the effects of the AFST tool and surrounding policy and practices changes relative to the Pre-AFST period. The AFST tool was implemented August 1, 2016 and was updated at various points (e.g., November 30, 2016).

As the AFST tool was implemented, so too was the practice of field screening changed from sending a staff member out for referrals involving children under 7 to only sending out for children under 4. Likewise, a "one caseworker" model was adopted across all offices. This change meant that the worker screening in referrals would be doing so for cases that remained in their own caseload for further investigation.

In 2018, call screening "data entry specialists" began working to increase intake processing capacity and to help prevent data entry lag time from undermining the relevance of the AFST's input – particularly at times when high referral volume was causing intake to informally triage incoming GPS referrals in terms of initial perceived severity. Around the same time, the intake office's field screening unit was trained and given access to the AFST score.

The implementation of LASSO versions and policy and practice changes that occurred

The first iteration of the AFST was replaced by a new version featuring a LASSO model on November 29, 2018. Along with this implementation there were changes to the score categories with the addition of new categories for the lowest (Low-Risk Protocol) and the highest (High-Risk Protocol) scoring referrals. These were intended to regularize and standardize the majority of screening out and screening in decisions made for these groups. The first few weeks after implementation saw fixes and updates to the LASSO tool.

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *38*

APPENDIX A

In early 2019, hundreds of "notices" of court truancy hearings abruptly began to arrive in the CYF intake office. There was initial uncertainty as to whether or how these notices should be processed and responded to by CYF. Initially they were handled like typical incoming CYF referrals to be screened and possibly launch an investigation. At first, almost all of them were screened in, but after further organizational discussion the intake office reversed course and began screening them out. By the summer of 2019, such notices were essentially no longer being sent to call screening to be treated as CYF referrals. The end result of this episode has been that an indicator was developed to identify and exclude these notices from analyses when deemed appropriate.

On March 15, 2019, a practice change occurred regarding the "60-day auto-accept" policy. Before this date, investigations that reached day 60 were automatically accepted for services and opened as a new CYF case. Based on changes in state policy, this policy reversed and investigations that reached day 60 automatically were now not accepted for services. The initial aftermath of this change saw a noticeable downtick in the acceptance rate.

In the Spring and Summer of 2019, initiatives within the call screening unit aimed to increase same-day processing of referrals and make the LASSO score timelier and more relevant for referrals that would have otherwise been likely to screen-out prior to receiving the score.

On July 1, 2019, a major change occurred in the handling of referrals for families who already had an active CYF case or investigation ongoing. Prior to September 2017, referrals on active families were typically screened-out. September 2017 saw the implementation of a new construct — an "ACT" — which represented a scaled-down investigation of sorts. From September 2017 onward, active family referrals would not be screened-out, but rather screened-in for investigation under this new entity type. However, on July 1, 2019, the "ACT" construct was retired, and active family referrals instead would be screened-in for typical, full CYF investigations unless they met a narrow set of criteria indicating that the new referral was essentially identical to an already processed allegation. As a result of all of these changes, the analytics team implemented a new indicator in October 2019 aiming to identify and exclude active family referrals when appropriate to help narrow down referrals where intake had screening "discretion," and to allow for analytical consistent across policy periods.

On July 23, 2019, a new version of the LASSO tool launched which incorporated new data elements and had modified logic to account for data lags more appropriately. The model was re-estimated and new weights were applied.

In late August 2019, policy changes were approved including an expansion of the Low-Risk Protocol eligibility and the elimination of the rules related to field screening for 4 prior referrals that were not investigated in the previous 2 years. The technical implementation of this policy change was implemented on October 26, 2019 within the LASSO tool and supporting system logic.

Based on the timeline and various changes above, our analysis of the most current LASSO period begins on October 27, 2019 and includes controls in the regression for score categories including low risk protocol, age groups related to cutoffs described above, flags for referrals related to truancy, and flags for referrals related to referrals on already-active families. However, it should also be clear that in addition to the integration of the LASSO screening score into the call screening process, substantial policy and practice differences existed relative to the Pre-AFST period.

AFST Protocols and Risk Score Categories, and their significance in practice

Some AFST scores initiate a "High-Risk Protocol" (HRP) or "Low-Risk Protocol" (LRP) designation that is displayed to the screener instead of a numeric score, based on a combination of the referral's maximum AFST score and the ages of the children on the referral. The HRP includes scores of 18 or more where any child on the referral is age 16 or under. When generated, this protocol defaults a referral to be screened in for investigation, unless overridden by an intake supervisor. The LRP is initiated for scores of 12 or less where all children on the referral are age 7+. (Note: Prior to the build on 10/26/2019, the thresholds were scores of 10 or less, and where all children on the referral were age 12+.) When generated, the LRP defaults a referral to be screened out, unless overridden by an intake supervisor.

It is important to note that, while these protocols do activate default actions, intake supervisors always have ultimate override discretion – and both High- and Low-Risk Protocol referrals do frequently have their default actions overridden.

In the AFST visualization, displayed numeric scores are also categorized into a higher-level "risk score category." Scores of 1-9 display in the "Low" category, 10-14 are grouped as "Medium," and 15-20 are grouped as "High." The cutoffs for these groupings were established based on the underlying distribution of subsequent placement for each numeric score in the historic modeling data. For AFST scores falling outside of the two established protocols above, neither the numeric score (1-20), nor the "risk score category" grouping (Low, Medium, High) are associated with any direct policies or defaults, and neither dictate screening decisions other than through communicating the AFST's assessment of relative placement risk for the intake workers to consider in their screening decision.

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *40*

APPENDIX A

A.2. Changes in episode rates during COVID-19

The following figure (**Appendix Figure A1**) illustrates the month number of episodes in the Pre-AFST period and the LASSO period. The vertical solid red line divides the 2 periods. The month in which the COVID-19 pandemic led to lockdowns and other public health measures (March 2020) is marked with a vertical red dashed line. Episode rates during LASSO but after the start of the pandemic are noticeably lower than those during LASSO but prior to the beginning of the pandemic.

FIGURE A1: **Counts of number of episodes by month for Pre-AFST and LASSO**

Partial Months Excluded from Count

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* *41*
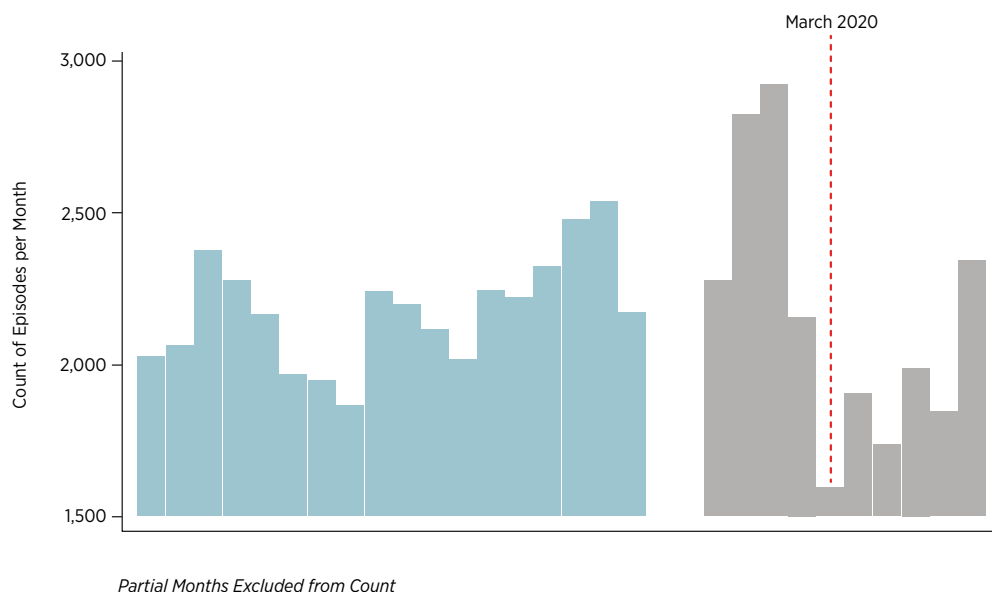
APPENDIX A

A.3. Relationship of referrals to episodes

The following figure (**Appendix Figure A2**) shows that: 1) most episodes in both the Pre-AFST and LASSO periods contained only a single episode (Panel A) for GPS screen-out, GPS screen-in, and CPS episodes; 2) The number of multi-referral episodes (Panel B) was small for all types of episodes and the distribution of these types did not substantially change between the policy periods. Of note, almost all multi-referral episodes (>95%) had 2 referrals in them.

FIGURE A2. **Counts of episodes for month in Pre-AFST and LASSO periods for single-referral and multiple-referral episodes**

Panel A
Episodes/Month by Episode Type for Single Referral Episodes

Panel B
Episodes/Month by Episode Type for 2+ Referral Episodes

GPS Screen-out   GPS Screen-in   GPS Screen-in

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 42
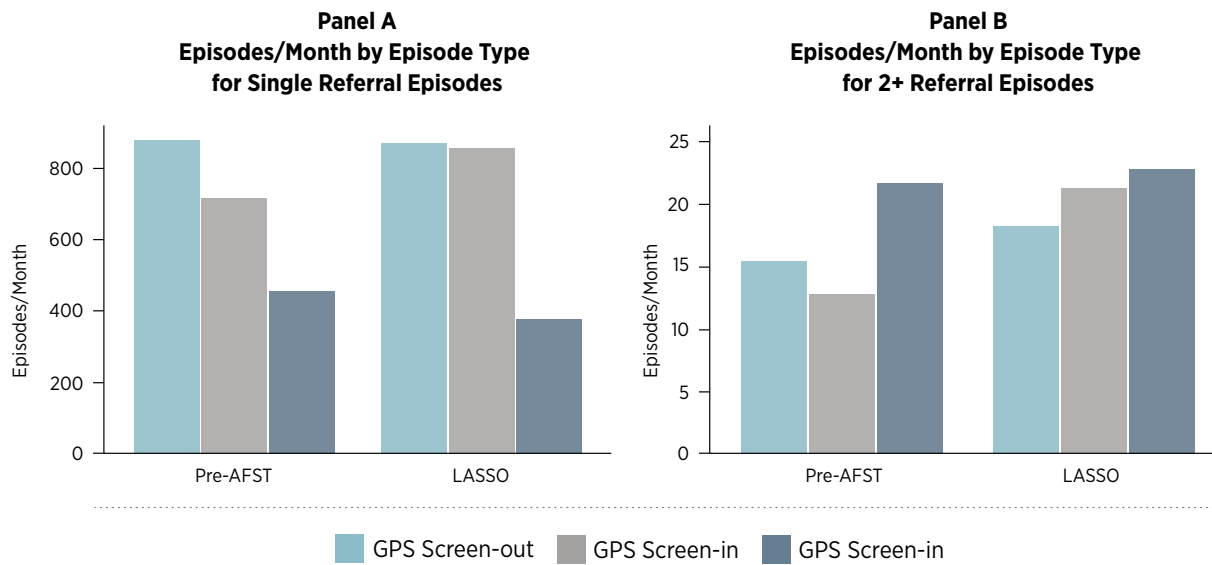
APPENDIX A

The mix of referrals within multi-referral episodes stayed very stable across the Pre-AFST and LASSO periods (**Appendix Figure A3**). The great majority of multi-referral CPS episodes were a mix of GPS referrals and CPS referrals. Similarly, the great majority of multi-referral GPS episodes were a mix of GPS screen-in and GPS screen-out referrals. By definition, multi-referral GPS screen-out episodes had only GPS screen-out referrals.

FIGURE A3. **Referral mix in multiple-referral episodes for Pre-AFST and LASSO periods**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *43*
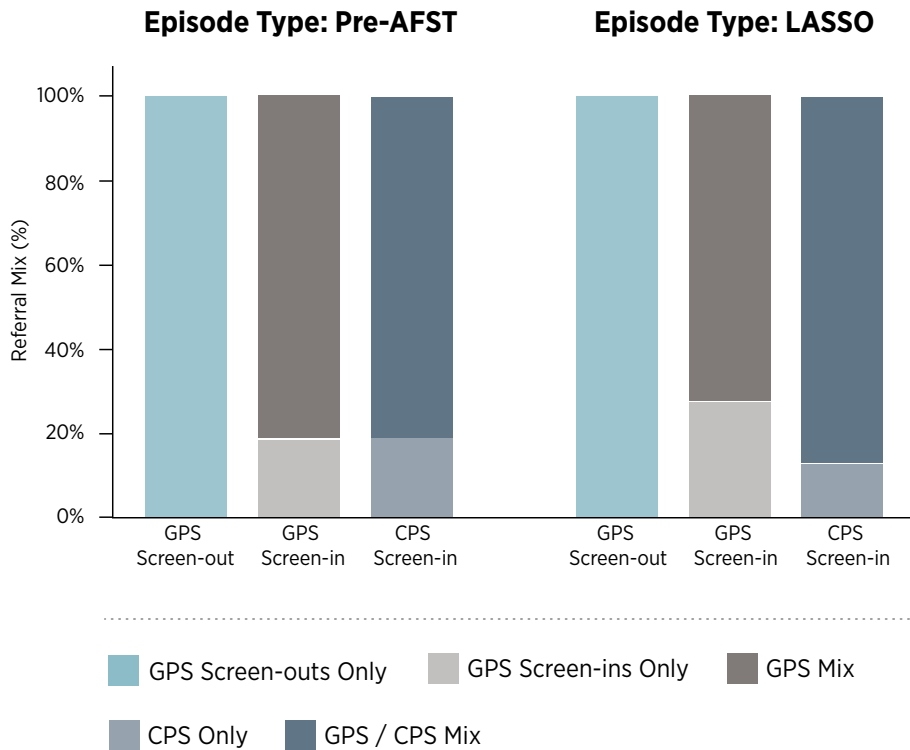
APPENDIX B

APPENDIX SECTION B. REGRESSION RESULTS FOR SHORT-TERM OUTCOMES

TABLE B1A. Adjusted Analysis: probability of a screen-in in a GPS episode

	PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
Pre-AFST	46.6%	45.9%	47.2%
LASSO	48.3%	47.5%	49.2%
Difference	1.8%	0.5%	3.0%

TABLE B1B. Adjusted Analysis: probability of a screen-in in a GPS episode by risk score

	PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST			
LR	33.0%	31.4%	34.6%
LOW	33.6%	32.2%	35.0%
MED	44.6%	43.5%	45.8%
HIGH	52.8%	51.6%	53.9%
HR	55.8%	54.5%	57.1%
LASSO			
LR	25.5%	23.3%	27.7%
LOW	29.0%	27.0%	31.0%
MED	40.8%	39.3%	42.2%
HIGH	51.6%	50.2%	53.0%
HR	77.2%	75.8%	78.6%
DIFFERENCE (LASSO—PRE-AFST)			
LR	-7.4%	-10.2%	-4.7%
LOW	-4.6%	-7.1%	-2.1%
MED	-3.9%	-5.9%	-1.9%
HIGH	-1.2%	-3.1%	0.8%
HR	21.4%	19.4%	23.3%

TABLE B2A. Adjusted Analysis: probability of accept for service in a GPS screen-in episode

	PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
Pre-AFST	40.0%	38.9%	41.0%
LASSO	32.0%	30.8%	33.1%
Difference	-8.0%	-9.8%	-6.2%

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *44*

APPENDIX B

TABLE B2B. Adjusted Analysis: probability of accept for service in a GPS screen-in episode by risk score

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
LR	24.9%	22.2%	27.5%
LOW	29.7%	27.4%	32.1%
MED	34.2%	32.4%	35.9%
HIGH	41.5%	39.8%	43.1%
HR	48.5%	46.7%	50.3%
LASSO			
LR	20.5%	16.2%	24.8%
LOW	20.7%	17.1%	24.3%
MED	25.6%	23.5%	27.7%
HIGH	35.0%	33.1%	36.9%
HR	38.7%	36.9%	40.5%
DIFFERENCE POST — PRE			
LR	-4.4%	-9.4%	0.7%
LOW	-9.0%	-13.3%	-4.7%
MED	-8.5%	-11.4%	-5.7%
HIGH	-6.5%	-9.2%	-3.8%
HR	-9.7%	-12.5%	-7.0%

TABLE B3A. Adjusted Analysis: probability of accept for service in a CPS episode

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
Pre-AFST	20.4%	19.5%	21.3%
LASSO	16.6%	15.6%	17.7%
Difference	-3.8%	-5.4%	-2.1%

TABLE B3B. **Adjusted Analysis: probability of accept for service in a CPS episode by risk score**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
LR	7.7%	6.0%	9.5%
LOW	9.1%	7.5%	10.8%
MED	17.4%	15.8%	19.0%
HIGH	20.8%	19.1%	22.6%
HR	32.9%	30.6%	35.1%
LASSO			
LR	2.7%	0.5%	4.8%
LOW	5.5%	3.2%	7.7%
MED	17.0%	14.8%	19.3%
HIGH	15.3%	13.3%	17.3%
HR	25.2%	22.6%	27.8%
DIFFERENCE POST — PRE			
LR	-5.1%	-7.8%	-2.3%
LOW	-3.7%	-6.5%	-0.9%
MED	-0.4%	-3.3%	2.5%
HIGH	-5.5%	-8.3%	-2.7%
HR	-7.7%	-11.2%	-4.1%

APPENDIX SECTION C. REGRESSION RESULTS FOR LONGER-TERM OUTCOMES

TABLE C1A. Adjusted Analysis: probability of another episode within 180 days of a GPS screen-in episode

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
Pre-AFST	28.2	24.9	31.5
LASSO	26.9	23.5	30.4
Difference	-1.2	-2.3	-0.2

TABLE C1B. Adjusted Analysis: probability of another episode within 180 days of a GPS screen-in episode by risk score

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
LR	21.7	18.2	25.4
LOW	24.4	20.9	28.1
MED	25.9	22.8	29.2
HIGH	29.0	25.8	32.3
HR	31.1	27.7	34.6
LASSO			
LR	21.4	16.7	26.5
LOW	22.9	18.7	27.4
MED	22.8	19.6	26.2
HIGH	27.0	23.8	30.4
HR	31.5	28.2	35.0
DIFFERENCE			
LR	-0.3	-1.4	0.8
LOW	-1.5	-2.6	-0.4
MED	-3.1	-4.2	-2.0
HIGH	-1.9	-3.0	-0.9
HR	0.4	-0.6	1.5

TABLE C2A. Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-in episode

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
Pre-AFST	18.7	15.8	21.7
LASSO	22.5	19.2	26.0
Difference	3.8	2.7	4.9

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *47*

APPENDIX C

TABLE C2B. Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-in episode by risk score

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST			
LR	12.3	9.7	15.4
LOW	15.8	12.8	19.1
MED	16.6	13.9	19.5
HIGH	18.9	16.0	21.9
HR	21.5	18.4	24.8
LASSO			
LR	15.0	10.8	19.9
LOW	17.0	13.2	21.3
MED	17.7	14.7	20.9
HIGH	22.5	19.3	25.8
HR	28.2	24.6	31.8
DIFFERENCE LASSO — PRE-AFST			
LR	2.6	1.6	3.7
LOW	1.2	0.1	2.3
MED	1.1	0.0	2.2
HIGH	3.6	2.5	4.7
HR	6.7	5.6	7.7

TABLE C3A. Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-in episode

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
Pre-AFST	8.1	6.1	10.4
LASSO	10.9	8.5	13.6
Difference	2.8	1.7	3.9

TABLE C3B. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-in episode by risk score**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
LR	4.3	2.8	6.3
LOW	6.5	4.6	8.9
MED	6.6	4.9	8.6
HIGH	7.8	5.9	10.1
HR	10.2	7.9	12.8
LASSO			
LR	7.1	4.3	10.9
LOW	6.0	3.8	8.9
MED	7.7	5.7	10.2
HIGH	10.7	8.4	13.3
HR	14.1	11.4	17.2
DIFFERENCE LASSO — PRE-AFST			
LR	2.8	1.7	3.9
LOW	-0.5	-1.6	0.6
MED	1.2	0.1	2.3
HIGH	2.8	1.7	3.9
HR	4.0	2.9	5.0

TABLE C4A. **Adjusted Analysis: probability of another episode within 180 days of a GPS screen-out episode**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
Pre-AFST	26.5	23.8	29.2
LASSO	25.0	21.9	28.3
Difference	-1.4	-2.5	-0.4

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *49*

APPENDIX C

TABLE C4B. Adjusted Analysis: probability of another episode within 180 days of a GPS screen-out episode by risk score

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
LR	20.4	17.8	23.2
LOW	20.6	17.9	23.4
MED	27.5	24.8	30.2
HIGH	29.8	27.0	32.6
HR	31.3	28.3	34.2
LASSO			
LR	19.7	16.5	23.2
LOW	20.5	17.4	23.8
MED	24.4	21.4	27.5
HIGH	29.1	26.0	32.4
HR	29.2	25.7	32.9
DIFFERENCE LASSO — PRE-AFST			
LR	-0.7	-1.7	0.4
LOW	-0.1	-1.1	0.9
MED	-3.1	-4.1	-2.0
HIGH	-0.7	-1.7	0.4
HR	-2.0	-3.1	-1.0

TABLE C5A. Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-out episode

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
Pre-AFST	17.7	15.2	20.4
LASSO	14.9	12.2	17.8
Difference	-2.8	-3.8	-1.8

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 50

APPENDIX C

TABLE C5B. Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-out episode by risk score

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST			
LR	12.0	9.8	14.5
LOW	12.5	10.3	15.1
MED	18.1	15.6	20.8
HIGH	19.9	17.2	22.8
HR	23.7	20.6	26.9
LASSO			
LR	9.4	7.2	12.0
LOW	9.6	7.3	12.3
MED	13.6	11.1	16.3
HIGH	19.1	16.2	22.3
HR	20.2	16.8	23.9
DIFFERENCE LASSO — PRE-AFST			
LR	-2.6	-3.7	-1.6
LOW	-2.9	-4.0	-1.9
MED	-4.5	-5.6	-3.5
HIGH	-0.7	-1.8	0.3
HR	-3.5	-4.5	-2.5

TABLE C6A. Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-out episode

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
Pre-AFST	6.5	4.9	8.5
LASSO	3.8	2.5	5.6
Difference	-2.7	-3.7	-1.6

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *51*

APPENDIX C

TABLE C6B. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-out episode by risk score**

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST			
LR	3.0	1.9	4.5
LOW	3.9	2.6	5.5
MED	6.9	5.1	9.0
HIGH	7.1	5.3	9.2
HR	10.1	7.8	12.9
LASSO			
LR	2.1	1.1	3.6
LOW	1.8	0.9	3.2
MED	3.3	2.1	4.9
HIGH	4.6	3.1	6.6
HR	6.1	4.0	8.6
DIFFERENCE LASSO — PRE-AFST			
LR	-0.9	-2.0	0.1
LOW	-2.1	-3.1	-1.0
MED	-3.6	-4.6	-2.5
HIGH	-2.4	-3.5	-1.4
HR	-4.1	-5.1	-3.1

TABLE C7A. **Adjusted Analysis: probability of another episode within 180 days of a CPS episode**

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
Pre-AFST	24.5	20.6	28.6
LASSO	25.5	20.9	30.4
Difference	1.0	-0.5	2.5

TABLE C7B. Adjusted Analysis: probability of another episode within 180 days of a CPS episode by risk score

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST			
LR	17.0	13.5	20.9
LOW	15.1	11.8	18.7
MED	25.5	21.4	29.7
HIGH	26.3	22.2	30.7
HR	33.4	28.7	38.1
LASSO			
LR	14.9	10.5	20.0
LOW	19.8	14.8	25.4
MED	25.9	21.3	30.8
HIGH	28.8	24.1	33.7
HR	32.5	27.4	37.8
DIFFERENCE LASSO — PRE-AFST			
LR	-2.1	-3.6	-0.6
LOW	4.7	3.3	6.2
MED	0.4	-1.0	1.9
HIGH	2.5	1.0	4.0
HR	-0.8	-2.3	0.6

TABLE C8A. Adjusted Analysis: probability of another episode that screens-in within 180 days of a CPS episode

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
Pre-AFST	15.8	12.5	19.4
LASSO	19.9	15.7	24.5
Difference	4.1	2.6	5.5

TABLE C8B. **Adjusted Analysis: probability of another episode that screens-in within 180 days of a CPS episode by risk score**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
LR	10.1	7.4	13.4
LOW	9.1	6.6	12.2
MED	16.5	13.1	20.4
HIGH	14.8	11.5	18.5
HR	24.1	19.7	28.9
LASSO			
LR	7.2	4.4	10.9
LOW	13.8	9.6	18.8
MED	18.7	14.4	23.3
HIGH	24.3	19.7	29.2
HR	28.1	22.9	33.6
DIFFERENCE LASSO — PRE-AFST			
LR	-2.9	-4.4	-1.5
LOW	4.7	3.2	6.1
MED	2.1	0.6	3.6
HIGH	9.5	8.1	11.0
HR	4.0	2.5	5.5

TABLE C9A. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a CPS episode**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
Pre-AFST	6.3	4.3	8.9
LASSO	6.4	4.3	9.1
Difference	0.1	-1.4	1.5

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 54

APPENDIX C

TABLE C9B. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a CPS episode by risk score**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
LR	2.2	1.1	3.9
LOW	2.5	1.3	4.4
MED	5.4	3.5	8.1
HIGH	5.2	3.3	7.7
HR	12.3	8.8	16.5
LASSO			
LR	0.7	0.1	2.4
LOW	1.1	0.4	2.6
MED	5.3	3.2	8.1
HIGH	7.6	4.9	11.0
HR	9.9	6.8	13.8
DIFFERENCE LASSO — PRE-AFST			
LR	-1.5	-3.0	-0.1
LOW	-1.4	-2.9	0.1
MED	-0.2	-1.6	1.3
HIGH	2.4	0.9	3.9
HR	-2.4	-3.9	-0.9

TABLE C10A. **Adjusted Analysis: probability of removal within 180 days of a GPS screen-in episode**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
Pre-AFST	8.0	5.9	10.5
LASSO	5.3	3.7	7.4
Difference	-2.7	-3.7	-1.6

TABLE C10B. **Adjusted Analysis: probability of removal within 180 days of a GPS screen-in episode by risk score**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
LR	4.2	2.6	6.3
LOW	3.7	2.4	5.4
MED	4.8	3.4	6.6
HIGH	8.5	6.3	11.0
HR	11.2	8.5	14.4
LASSO			
LR	3.5	1.7	6.3
LOW	2.7	1.3	4.9
MED	3.4	2.2	5.0
HIGH	5.9	4.1	8.0
HR	6.8	4.9	9.1
DIFFERENCE LASSO — PRE-AFST			
LR	-0.7	-1.8	0.4
LOW	-1.0	-2.1	0.1
MED	-1.4	-2.5	-0.3
HIGH	-2.6	-3.7	-1.5
HR	-4.4	-5.5	-3.4

TABLE C11A. **Adjusted Analysis: probability of removal within 180 days of a CPS episode**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
Pre-AFST	4.4	2.6	6.8
LASSO	2.9	1.5	5.1
Difference	-1.5	-2.9	0.0

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *56*

APPENDIX C

Table C11b. Adjusted Analysis: probability of removal within 180 days of a CPS episode by risk score

	PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST			
LR	1.5	0.6	3.0
LOW	3.5	1.9	5.9
MED	3.9	2.2	6.2
HIGH	4.5	2.7	7.1
HR	5.6	3.4	8.6
LASSO			
LR	0.0	0.0	0.0
LOW	1.8	0.6	4.1
MED	4.7	2.5	7.8
HIGH	2.3	1.1	4.3
HR	3.3	1.7	5.8
DIFFERENCE LASSO — PRE-AFST			
LR	-1.5	-2.9	0.0
LOW	-1.7	-3.2	-0.3
MED	0.8	-0.7	2.3
HIGH	-2.2	-3.7	-0.7
HR	-2.3	-3.8	-0.8

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 57

APPENDIX D

APPENDIX SECTION D. RESULTS FOR RACE GROUP DIFFERENCES IN OUTCOMES

Short-Term Outcomes

FIGURE D1: **Adjusted Analysis: probability of a screen-in in a GPS episode by risk score and race**
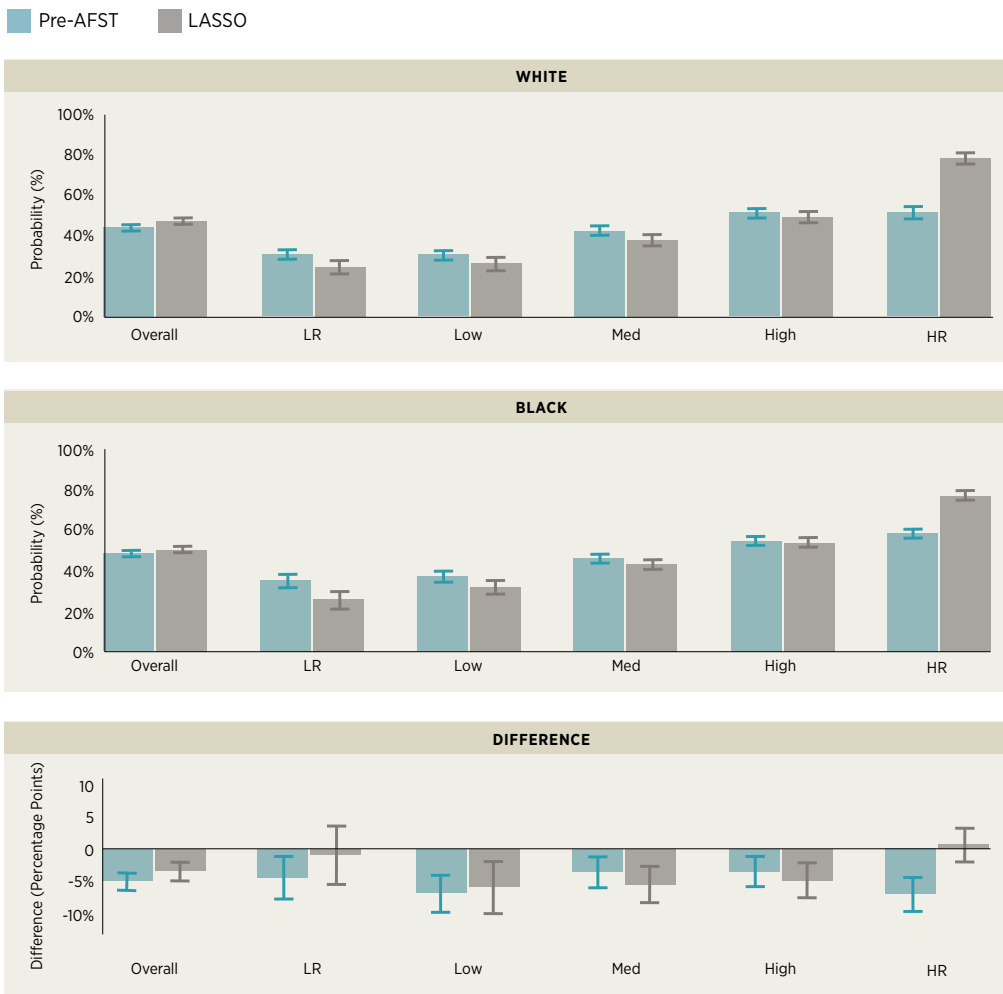
TABLE D1A. **Adjusted Analysis: probability of a screen-in in a GPS episode by race**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
White	44.1%	43.1%	45.1%
Black	49.1%	48.2%	50.0%
Difference	-5.0%	-6.2%	-3.8%
LASSO			
White	47.1%	46.0%	48.3%
Black	50.5%	49.5%	51.6%
Difference	-3.4%	-4.8%	-2.1%
Differential Change	**1.6%**	**-0.2%**	**3.3%**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   58

APPENDIX D

TABLE D1B. Adjusted Analysis: probability of a screen-in in a GPS episode by risk score and race

		PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST				
LR	White	30.9	28.9	32.9
	Black	35.3	32.5	38.1
	Difference	-4.4	-7.8	-1.1
LOW	White	30.4	28.6	32.3
	Black	37.4	35.1	39.7
	Difference	-6.9	-9.8	-4.1
MED	White	42.6	40.9	44.3
	Black	46.1	44.5	47.8
	Difference	-3.6	-5.8	-1.3
HIGH	White	51.3	49.5	53.1
	Black	54.7	53.1	56.3
	Difference	-3.4	-5.7	-1.1
HR	White	51.6	49.4	53.7
	Black	58.7	57.1	60.3
	Difference	-7.1	-9.7	-4.5
LASSO				
LR	White	24.8	22.0	27.5
	Black	25.7	21.9	29.5
	Difference	-0.9	-5.5	3.6
LOW	White	26.1	23.5	28.8
	Black	32.1	29.1	35.2
	Difference	-6.0	-9.9	-2.1
MED	White	37.9	35.7	40.1
	Black	43.4	41.5	45.3
	Difference	-5.5	-8.3	-2.7
HIGH	White	49.3	47.2	51.5
	Black	54.2	52.4	56.0
	Difference	-4.9	-7.5	-2.2
HR	White	78.3	76.1	80.4
	Black	77.6	75.9	79.3
	Difference	0.7	-1.8	3.3
DIFFERENTIAL CHANGE				
LR		-3.5	-9.2	2.3
LOW		-0.9	-5.8	4.1
MED		1.9	-1.7	5.5
HIGH		1.4	-2.1	4.8
HR		-7.9	-11.4	-4.3

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 59

APPENDIX D

FIGURE D2: **Adjusted Analysis: probability of accept for service in a GPS screen-in episode by risk score and race**
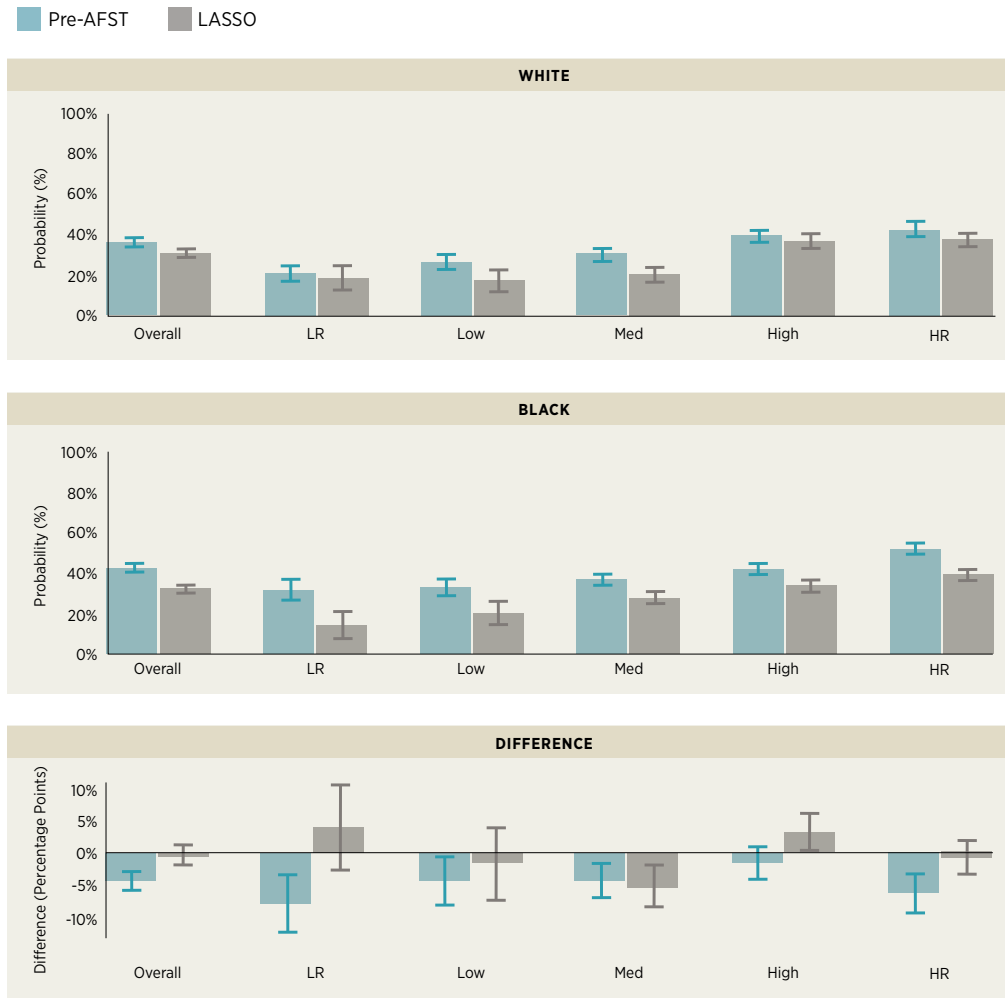
TABLE D2A. **Adjusted Analysis: probability of accept for service in a GPS screen-in episode by race**

	PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST			
White	37.1%	35.5%	38.6%
Black	42.7%	41.4%	44.0%
Difference	-5.6%	3.9%	7.4%
LASSO			
White	31.7%	30.0%	33.3%
Black	32.1%	30.7%	33.5%
Difference	0.4%	-1.5%	2.3%
Differential Change	**-5.2%**	**-7.8%**	**-2.6%**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 60

APPENDIX D

TABLE D2B. Adjusted Analysis: probability of accept for service in a GPS screen-in episode by risk score and race

		PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST				
LR	White	21.5	18.2	24.9
	Black	31.7	27.0	36.5
	Difference	-10.2	-15.8	-4.6
LOW	White	27.4	24.2	30.7
	Black	33.2	29.5	36.8
	Difference	-5.7	-10.4	-1.0
MED	White	31.1	28.6	33.6
	Black	36.7	34.3	39.1
	Difference	-5.6	-8.9	-2.4
HIGH	White	40.1	37.7	42.6
	Black	42.2	40.0	44.4
	Difference	-2.1	-5.2	1.1
HR	White	43.7	40.5	46.9
	Black	51.9	49.7	54.1
	Difference	-8.2	-11.9	-4.5
LASSO				
LR	White	19.6	14.0	25.2
	Black	14.5	8.1	20.9
	Difference	5.1	-3.4	13.6
LOW	White	18.1	13.2	23.1
	Black	20.4	15.1	25.6
	Difference	-2.2	-9.4	4.9
MED	White	21.3	18.1	24.4
	Black	27.9	25.2	30.6
	Difference	-6.6	-10.7	-2.6
HIGH	White	38.0	34.9	41.1
	Black	33.8	31.6	36.1
	Difference	4.2	0.5	7.9
HR	White	38.3	35.5	41.1
	Black	39.3	37.1	41.4
	Difference	-1.0	-4.2	2.2
DIFFERENTIAL CHANGE				
LR		-15.7	-26.8	-4.5
LOW		-3.1	-12.3	6.2
MED		1.6	-3.8	7.1
HIGH		-6.3	-11.2	-1.4
HR		-7.2	-12.1	-2.3

FIGURE D3: Adjusted Analysis: probability of accept for service in a CPS episode by risk score and race
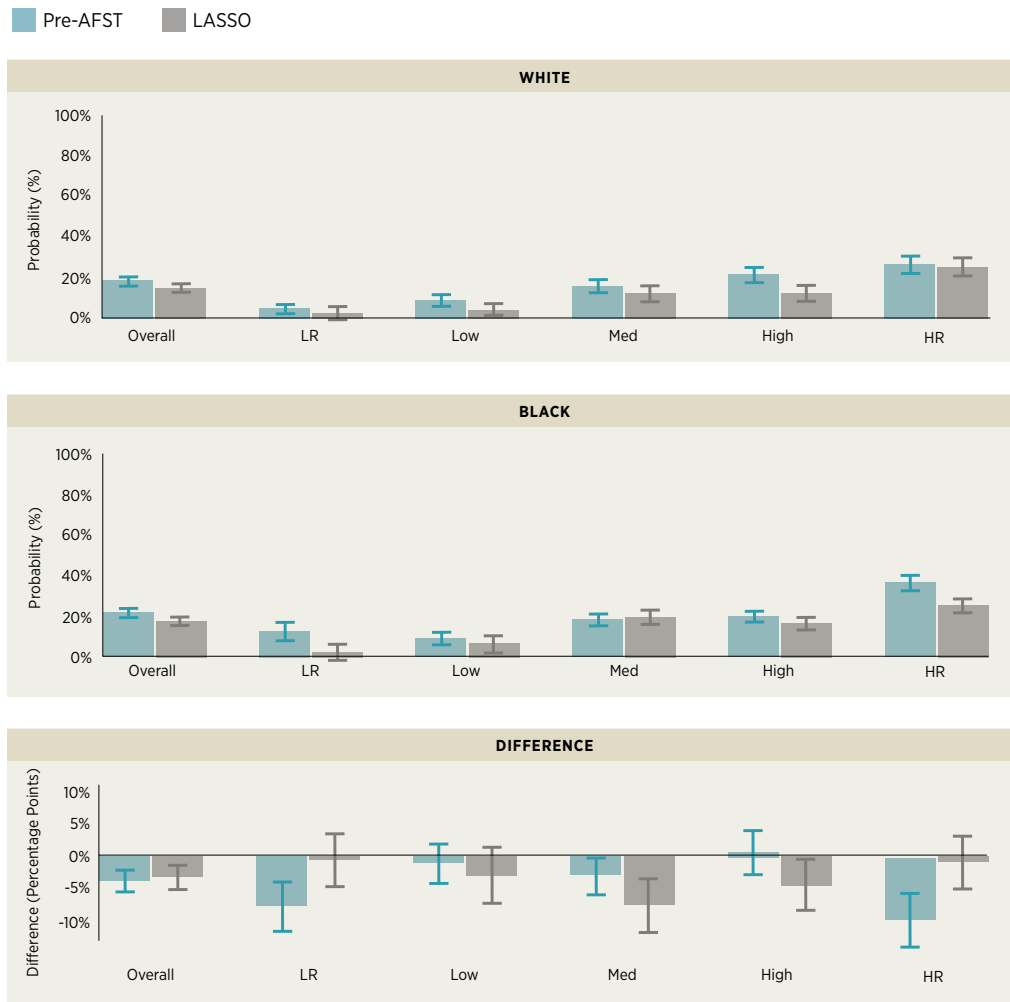
TABLE D3A. Adjusted Analysis: probability of accept for service in a CPS screen-in by race

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST			
White	18.4	17.0	19.8
Black	22.4	21.2	23.7
Difference	-4.1	-5.8	-2.4
LASSO			
White	14.9	13.2	16.5
Black	18.3	16.9	19.7
Difference	-3.4	-5.3	-1.5
Differential Change	**-0.4**	**-3.0**	**2.3**

TABLE D3B. **Adjusted Analysis: probability of accept for service in a CPS episode by risk score and race**

| | | PREDICTED PROBABILITY | [95% C.I.] | |
			LOWER	UPPER
		PRE-AFST		
LR	White	4.8	3.0	6.5
	Black	13.2	9.4	17.0
	Difference	-8.4	-12.6	-4.3
LOW	White	8.8	6.6	11.0
	Black	9.7	7.1	12.3
	Difference	-0.9	-4.2	2.4
MED	White	16.0	13.6	18.4
	Black	19.1	16.9	21.3
	Difference	-3.1	-6.2	0.0
HIGH	White	21.6	18.6	24.6
	Black	20.7	18.5	22.9
	Difference	0.9	-2.8	4.5
HR	White	26.6	23.0	30.2
	Black	37.4	34.5	40.3
	Difference	-10.8	-15.3	-6.3
		LASSO		
LR	White	2.5	-0.2	5.2
	Black	2.9	-0.6	6.4
	Difference	-0.4	-4.8	4.0
LOW	White	4.0	1.0	7.0
	Black	7.1	3.4	10.7
	Difference	-3.0	-7.7	1.7
MED	White	12.2	8.6	15.9
	Black	20.5	17.6	23.4
	Difference	-8.3	-12.8	-3.8
HIGH	White	12.4	9.0	15.7
	Black	17.2	14.6	19.8
	Difference	-4.8	-8.9	-0.7
HR	White	25.4	21.3	29.5
	Black	26.2	23.2	29.1
	Difference	-0.8	-5.2	3.6
		DIFFERENTIAL CHANGE		
LR		-7.6	-14.1	-1.2
LOW		2.6	-3.6	8.7
MED		5.8	0.1	11.4
HIGH		6.0	0.4	11.6
HR		-9.9	-16.3	-3.6

Longer-Term Outcomes

FIGURE D4: **Adjusted Analysis: probability of another episode within 180 days of a GPS screen-in episode by risk score and race**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *64*
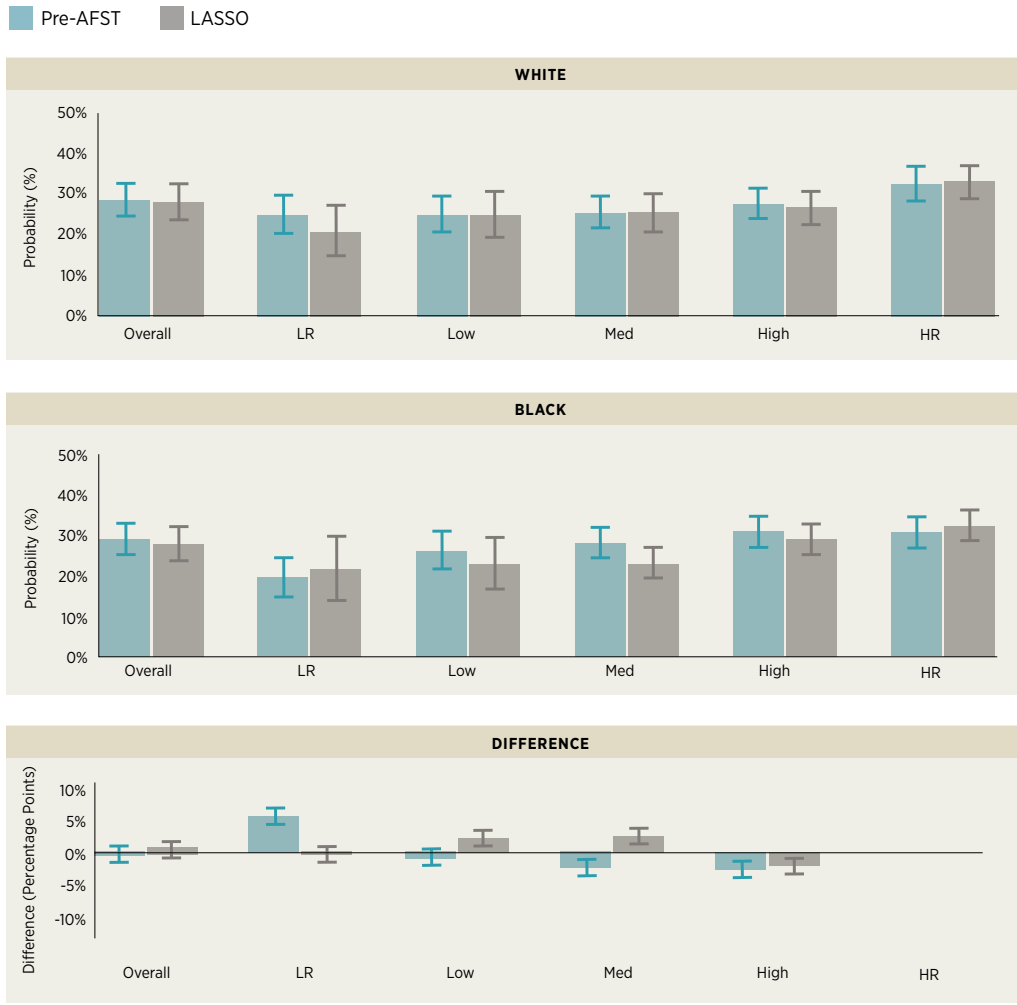
APPENDIX D

FIGURE D5: **Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-in episode by risk score and race**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 65
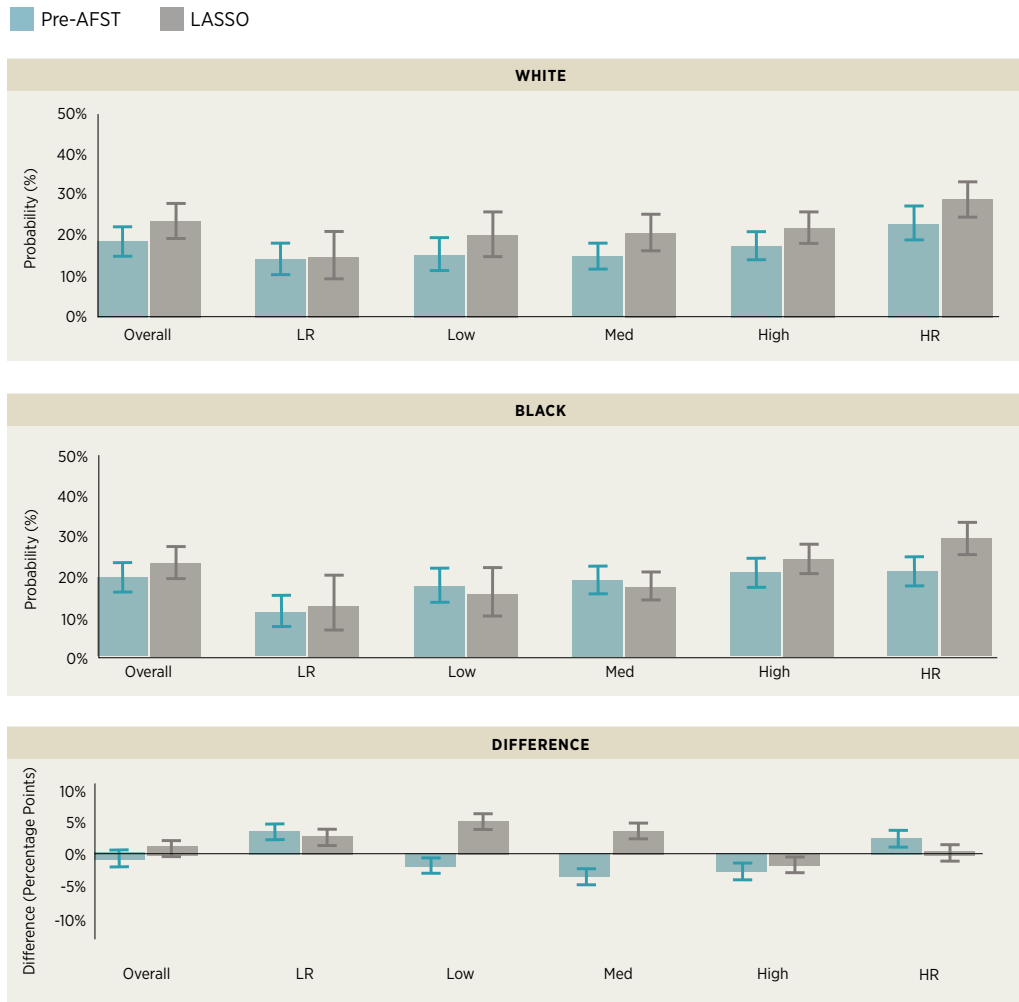
APPENDIX D

FIGURE D6: **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-in episode by risk score and race**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 66

APPENDIX D

FIGURE D7: **Adjusted Analysis: probability of another episode within 180 days of a CPS episode by risk score and race**

FIGURE D8: **Adjusted Analysis: probability of another episode that screens-in within 180 days of a CPS episode by risk score and race**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   68

APPENDIX D

FIGURE D9: **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a CPS episode by risk score and race**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   69
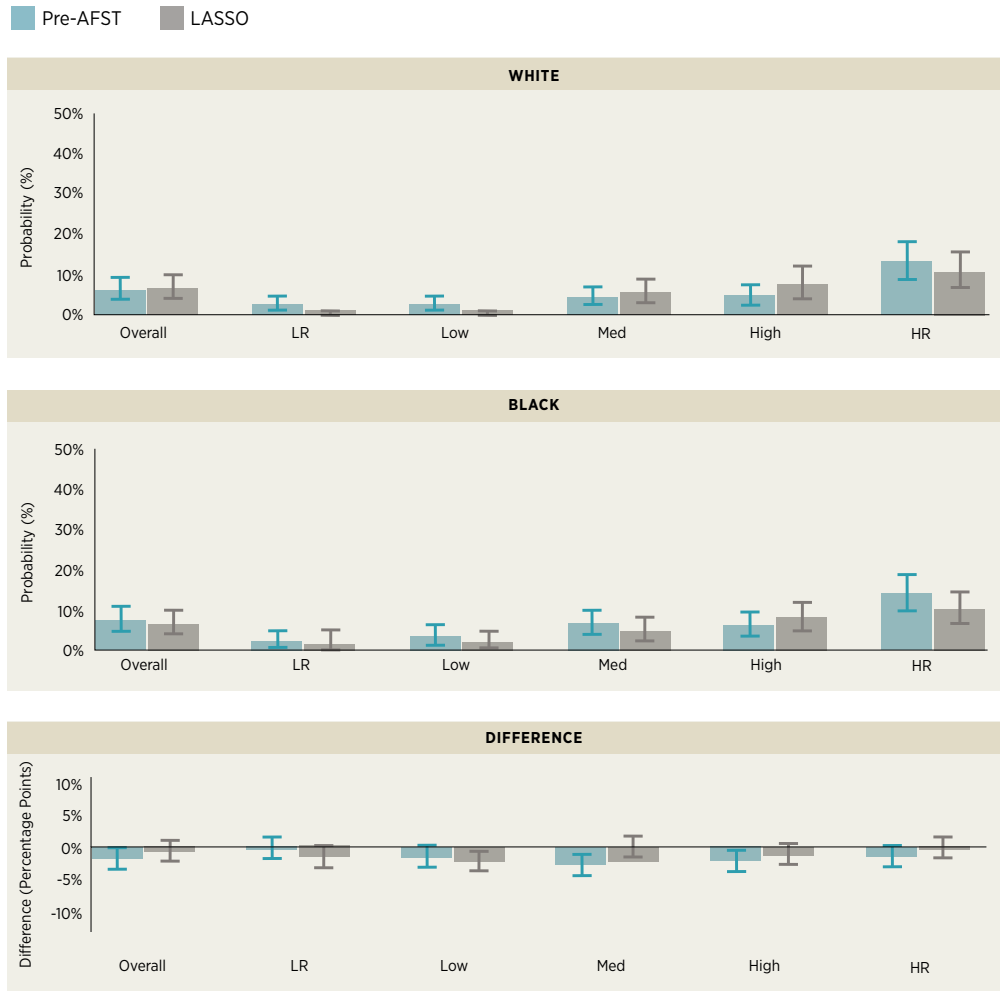
APPENDIX D

FIGURE D10: **Adjusted Analysis: probability of another episode within 180 days of a GPS screen-out episode by risk score and race**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *70*

APPENDIX D

FIGURE D11: **Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-out episode by risk score and race**
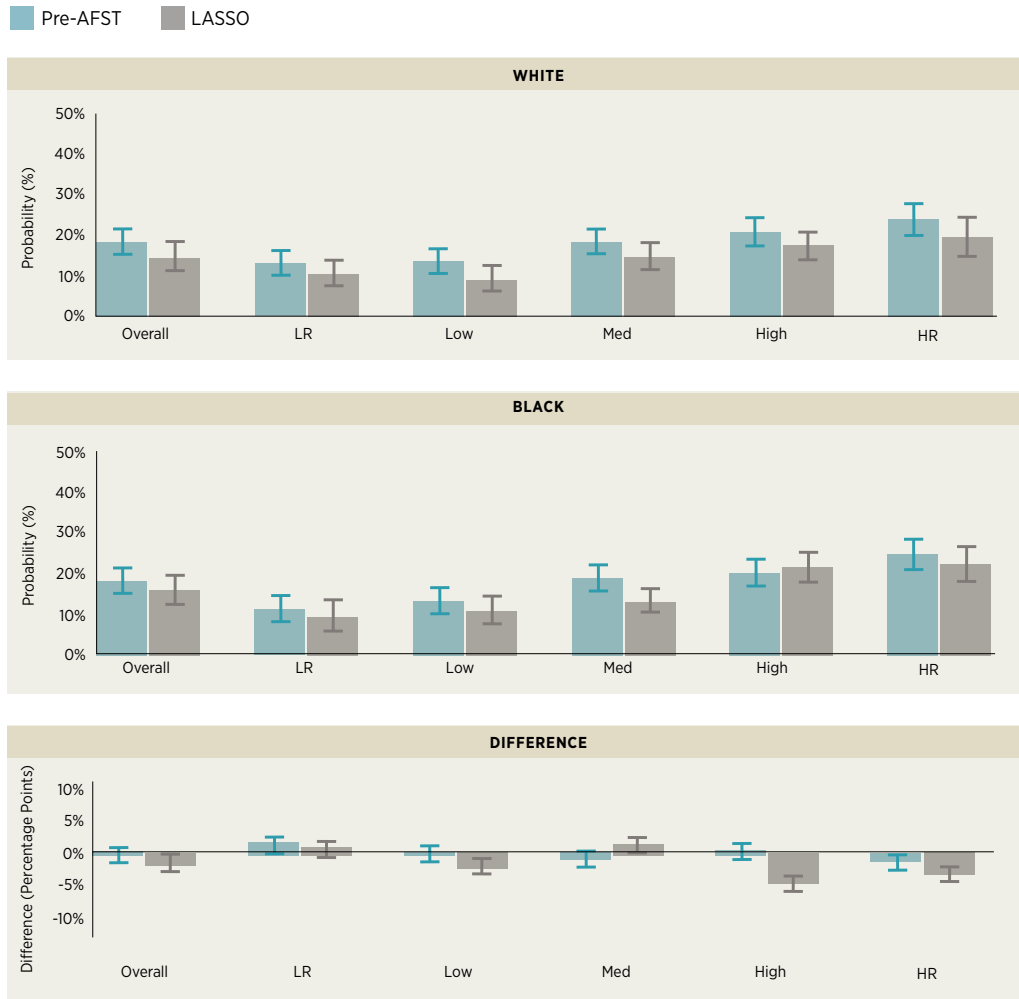
FIGURE D12. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-out episode by risk score and race**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 72
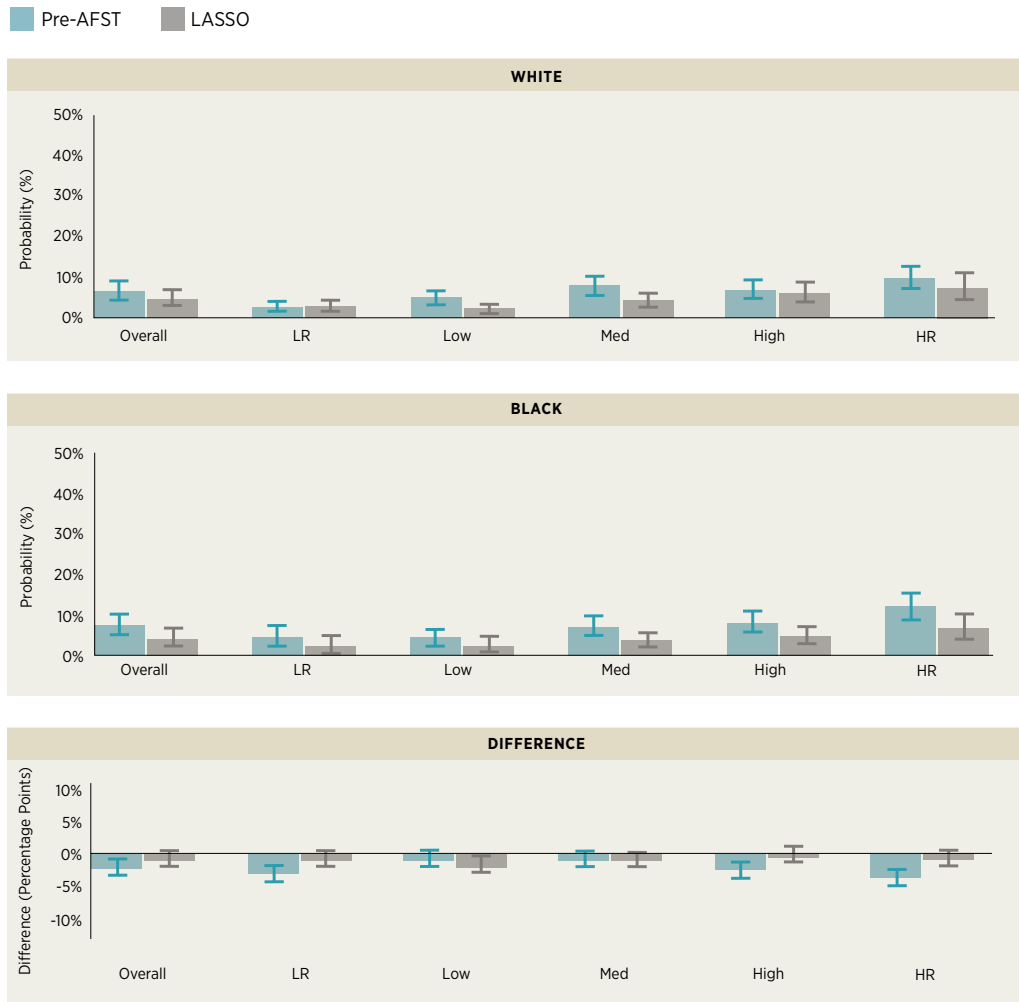
APPENDIX D

FIGURE D13: **Adjusted Analysis: probability of removal within 180 days of a GPS screen-in episode by risk score and race**
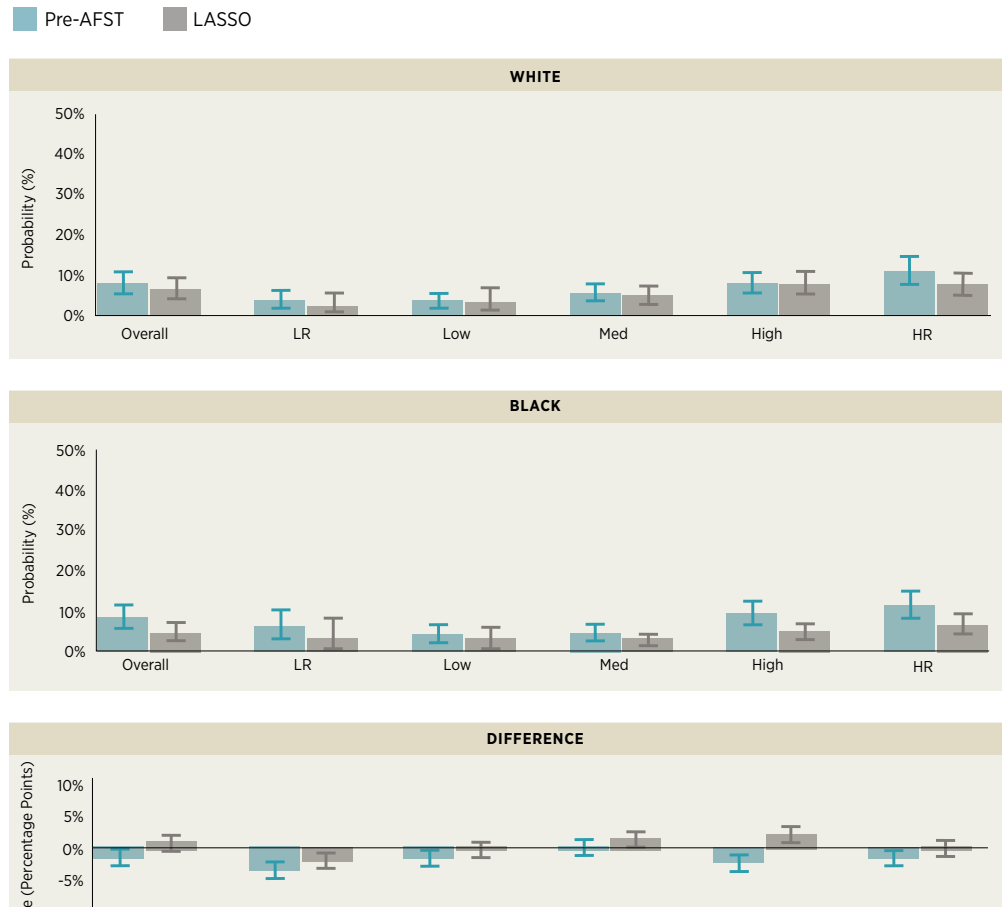
TABLE D4A. **Adjusted Analysis: probability of another episode within 180 days of a GPS screen-in episode by race**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
White	28.7	25.0	32.5
Black	28.9	25.3	32.6
Difference	-0.2	-1.3	0.9
LASSO			
White	28.2	24.1	32.4
Black	27.6	23.8	31.6
Difference	0.5	-0.6	1.7
Differential Change	**0.7**	**-0.9**	**2.3**

TABLE D4B. **Adjusted Analysis: probability of another episode within 180 days of a GPS screen-in episode by risk score and race**

		PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
		PRE-AFST		
LR	White	25.1	20.8	29.6
	Black	19.3	14.9	24.0
	Difference	5.8	4.7	6.9
LOW	White	25.3	21.2	29.7
	Black	26.0	21.7	30.5
	Difference	-0.7	-1.8	0.5
MED	White	25.6	22.2	29.2
	Black	27.9	24.5	31.5
	Difference	-2.3	-3.5	-1.2
HIGH	White	27.9	24.5	31.5
	Black	30.6	27.1	34.1
	Difference	-2.6	-3.7	-1.5
HR	White	32.8	28.7	36.9
	Black	30.4	26.9	34.0
	Difference	2.4	1.2	3.5
		LASSO		
LR	White	21.0	15.4	27.2
	Black	21.2	14.1	29.3
	Difference	-0.2	-1.3	0.9
LOW	White	25.1	19.9	30.6
	Black	22.7	16.8	29.2
	Difference	2.4	1.3	3.5
MED	White	25.5	21.2	30.1
	Black	22.9	19.5	26.5
	Difference	2.7	1.5	3.8
HIGH	White	26.7	23.1	30.6
	Black	28.7	25.3	32.2
	Difference	-2.0	-3.1	-0.8
HR	White	33.2	29.4	37.0
	Black	32.1	28.6	35.7
	Difference	1.0	-0.1	2.1
		DIFFERENTIAL CHANGE		
LR		-6.0	-7.6	-4.4
LOW		3.0	1.4	4.6
MED		5.0	3.4	6.6
HIGH		0.6	-1.0	2.2
HR		-1.3	-2.9	0.3

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 74

APPENDIX D

TABLE D5A. **Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-in episode by race**

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST			
White	18.7	15.4	22.3
Black	19.4	16.2	22.8
Difference	-0.6	-1.8	0.5
LASSO			
White	23.7	19.8	27.9
Black	22.8	19.2	26.7
Difference	0.9	-0.2	2.1
Differential Change	**1.6**	**0.0**	**3.2**

TABLE D5B. **Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-in episode by risk score and race**

		PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST				
LR	White	14.3	10.9	18.2
	Black	10.7	7.4	14.6
	Difference	3.6	2.5	4.7
LOW	White	15.6	12.0	19.6
	Black	17.3	13.5	21.4
	Difference	-1.7	-2.8	-0.6
MED	White	15.1	12.2	18.3
	Black	18.7	15.6	22.0
	Difference	-3.5	-4.7	-2.4
HIGH	White	17.7	14.7	21.0
	Black	20.4	17.2	23.8
	Difference	-2.7	-3.8	-1.6
HR	White	23.2	19.3	27.3
	Black	20.8	17.6	24.2
	Difference	2.4	1.3	3.5
LASSO				
LR	White	14.9	9.9	20.9
	Black	12.2	6.6	19.7
	Difference	2.7	1.5	3.8
LOW	White	20.4	15.5	25.9
	Black	15.3	10.1	21.5
	Difference	5.1	4.0	6.3
MED	White	20.7	16.7	25.2
	Black	17.1	14.0	20.4
	Difference	3.7	2.5	4.8
HIGH	White	22.2	18.6	26.0
	Black	23.8	20.4	27.4
	Difference	-1.6	-2.7	-0.5
HR	White	29.1	25.2	33.1
	Black	28.9	25.2	32.6
	Difference	0.2	-0.9	1.4
DIFFERENTIAL CHANGE				
LR		-1.0	-2.5	0.6
LOW		6.8	5.2	8.4
MED		7.2	5.6	8.8
HIGH		1.1	-0.5	2.7
HR		-2.2	-3.8	-0.6

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 76

APPENDIX D

TABLE D6A. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-in episode by race**

	PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST			
White	7.5	5.3	10.1
Black	8.8	6.6	11.5
Difference	-1.4	-2.5	-0.2
LASSO			
White	11.4	8.6	14.8
Black	10.7	8.3	13.6
Difference	0.7	-0.4	1.8
Differential Change	**2.1**	**0.5**	**3.7**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 77
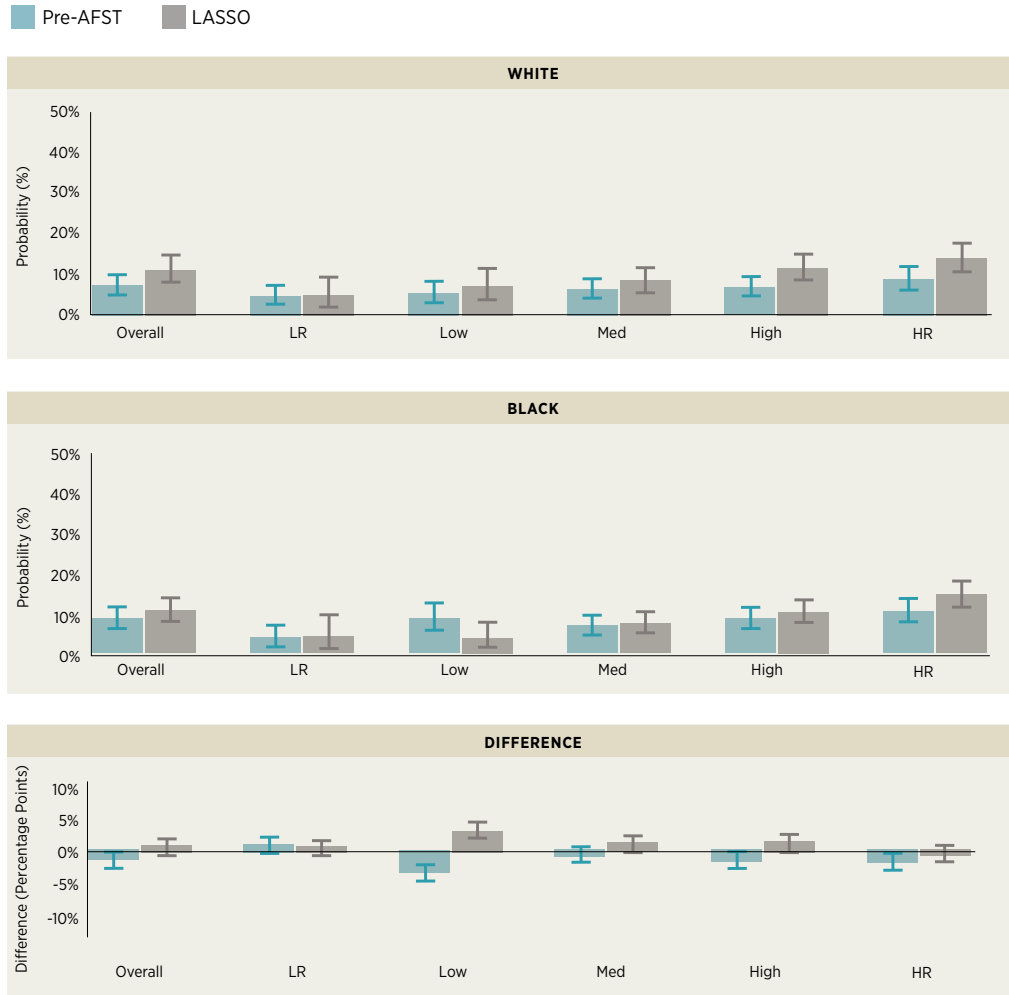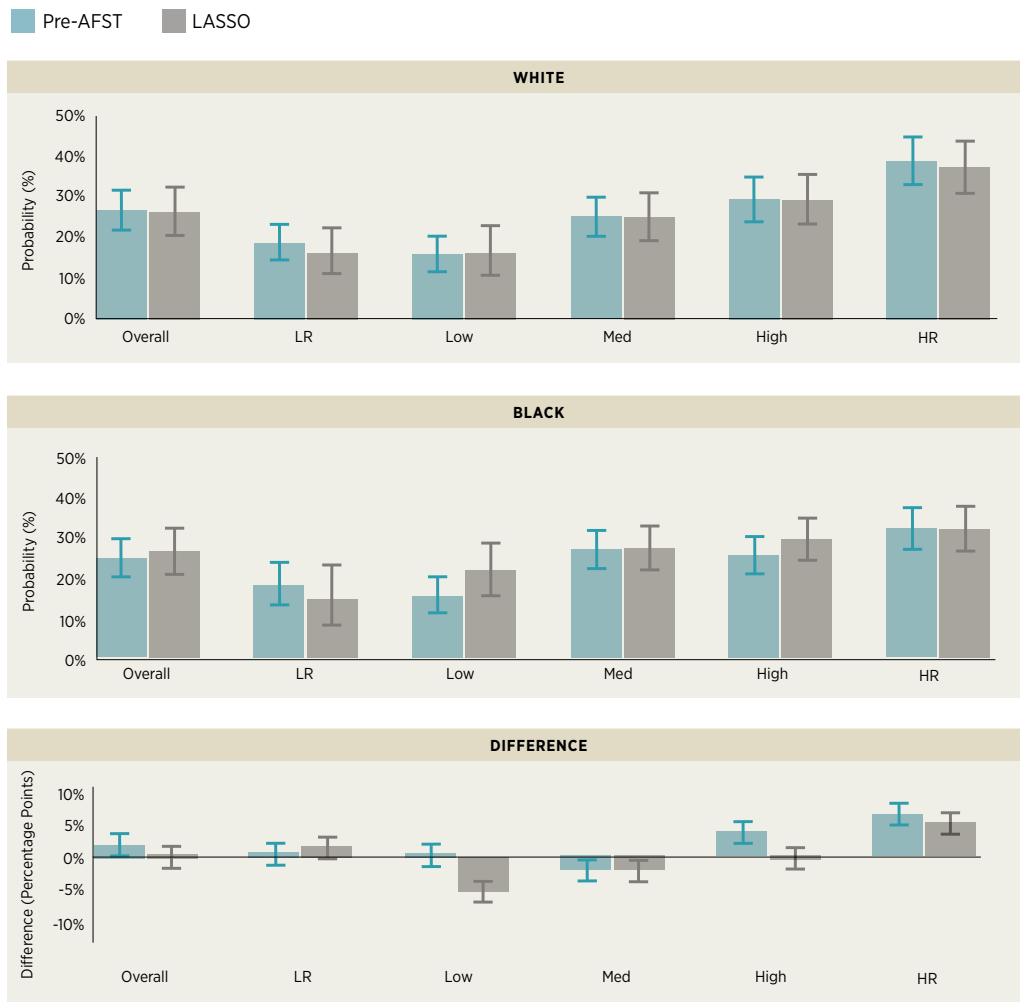
APPENDIX D

TABLE D6B. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-in episode by risk score and race**
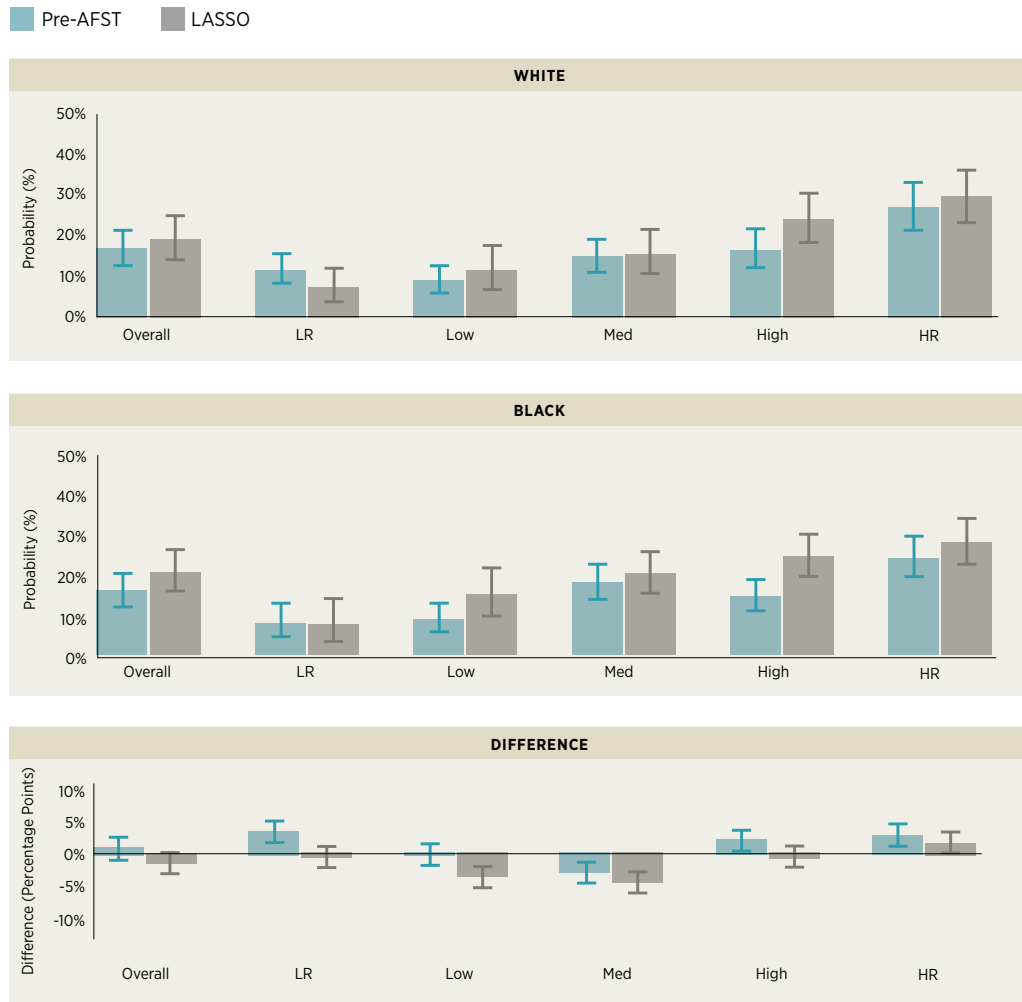
		PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST				
LR	White	5.0	3.1	7.4
	Black	4.0	2.2	6.8
	Difference	0.9	-0.2	2.0
LOW	White	5.6	3.5	8.4
	Black	9.0	6.2	12.4
	Difference	-3.4	-4.5	-2.3
MED	White	6.6	4.7	8.9
	Black	7.0	5.1	9.4
	Difference	-0.5	-1.6	0.6
HIGH	White	7.2	5.2	9.5
	Black	8.7	6.5	11.3
	Difference	-1.5	-2.7	-0.4
HR	White	9.0	6.5	12.0
	Black	10.6	8.2	13.5
	Difference	-1.6	-2.8	-0.5
LASSO				
LR	White	5.0	2.3	9.3
	Black	4.5	1.6	9.6
	Difference	0.5	-0.6	1.6
LOW	White	7.4	4.3	11.6
	Black	4.1	1.9	7.7
	Difference	3.3	2.2	4.4
MED	White	8.6	5.9	11.9
	Black	7.5	5.4	10.1
	Difference	1.1	-0.1	2.2
HIGH	White	11.8	9.1	14.9
	Black	10.4	8.0	13.1
	Difference	1.4	0.3	2.5
HR	White	14.2	11.1	17.7
	Black	14.6	11.7	17.8
	Difference	-0.4	-1.5	0.8
DIFFERENTIAL CHANGE				
LR		-0.4	-2.0	1.2
LOW		6.7	5.1	8.3
MED		1.6	0.0	3.2
HIGH		2.9	1.3	4.5
HR		1.3	-0.3	2.9

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 78

APPENDIX D

TABLE D7A. **Adjusted Analysis: probability of another episode within 180 days of a CPS episode by race**

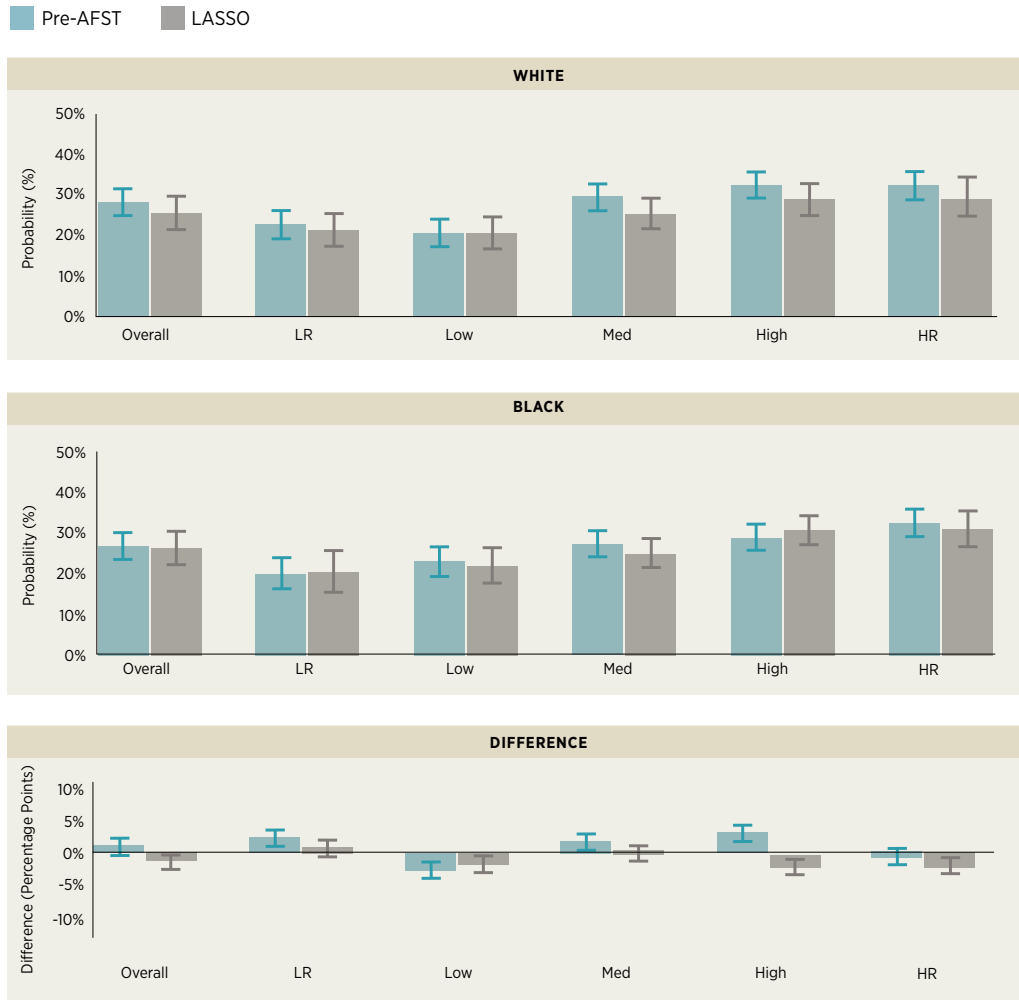|	PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST			
White	26.6	22.1	31.3
Black	24.6	20.4	29.2
Difference	1.9	0.4	3.5
LASSO			
White	26.0	20.6	31.9
Black	26.1	20.9	31.8
Difference	-0.1	-1.7	1.4
Differential Change	**-2.1**	**-4.2**	**0.1**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   79

APPENDIX D

TABLE D7B. **Adjusted Analysis: probability of another episode within 180 days of a CPS episode by risk score and race**

		PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST				
LR	White	18.6	14.6	22.9
	Black	18.2	13.4	23.5
	Difference	0.4	-1.1	2.0
LOW	White	15.7	12.0	19.8
	Black	15.4	11.5	19.8
	Difference	0.3	-1.3	1.8
MED	White	24.7	20.3	29.4
	Black	26.8	22.5	31.3
	Difference	-2.1	-3.7	-0.6
HIGH	White	29.2	24.1	34.6
	Black	25.4	21.1	29.8
	Difference	3.9	2.3	5.4
HR	White	38.6	33.0	44.2
	Black	31.9	27.2	36.7
	Difference	6.7	5.2	8.3
LASSO				
LR	White	16.1	11.1	22.0
	Black	14.7	8.4	22.8
	Difference	1.4	-0.1	3.0
LOW	White	16.1	10.8	22.5
	Black	21.6	15.7	28.2
	Difference	-5.5	-7.0	-4.0
MED	White	24.8	19.2	30.7
	Black	27.0	22.0	32.3
	Difference	-2.2	-3.8	-0.7
HIGH	White	29.1	23.4	35.1
	Black	29.4	24.4	34.5
	Difference	-0.2	-1.8	1.3
HR	White	37.2	30.9	43.5
	Black	31.9	26.6	37.3
	Difference	5.3	3.7	6.8
DIFFERENTIAL CHANGE				
LR		1.0	-1.2	3.2
LOW		-5.8	-7.9	-3.6
MED		-0.1	-2.3	2.1
HIGH		-4.1	-6.3	-1.9
HR		-1.4	-3.6	0.7

TABLE D8A. **Adjusted Analysis: probability of another episode that screens-in within 180 days of a CPS episode by race**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
White	17.0	13.1	21.3
Black	16.1	12.5	20.1
Difference	0.9	-0.6	2.5
LASSO			
White	19.4	14.5	25.0
Black	20.7	16.0	26.1
Difference	-1.3	-2.9	0.2
Differential Change	**-2.3**	**-4.4**	**-0.1**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *81*

APPENDIX D

TABLE D8B. **Adjusted Analysis: probability of another episode that screens-in within 180 days of a CPS episode by risk score and race**

		PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
		PRE-AFST		
LR	White	12.0	8.8	15.7
	Black	8.4	5.1	12.6
	Difference	3.6	2.1	5.2
LOW	White	9.2	6.4	12.7
	Black	9.2	6.3	12.9
	Difference	0.0	-1.6	1.5
MED	White	15.2	11.5	19.4
	Black	18.1	14.2	22.3
	Difference	-2.9	-4.5	-1.4
HIGH	White	16.8	12.5	21.6
	Black	14.7	11.3	18.5
	Difference	2.1	0.5	3.6
HR	White	27.3	21.9	33.0
	Black	24.3	19.7	29.1
	Difference	3.0	1.5	4.6
		LASSO		
LR	White	7.5	4.1	12.1
	Black	7.8	3.7	14.0
	Difference	-0.4	-1.9	1.2
LOW	White	11.7	7.1	17.6
	Black	15.3	10.0	21.6
	Difference	-3.6	-5.1	-2.0
MED	White	15.8	11.0	21.5
	Black	20.3	15.6	25.5
	Difference	-4.4	-6.0	-2.9
HIGH	White	24.4	18.7	30.4
	Black	24.7	19.8	29.9
	Difference	-0.3	-1.9	1.2
HR	White	29.8	23.7	36.2
	Black	28.1	22.7	33.7
	Difference	1.8	0.2	3.3
		DIFFERENTIAL CHANGE		
LR		-4.0	-6.2	-1.8
LOW		-3.5	-5.7	-1.4
MED		-1.5	-3.7	0.7
HIGH		-2.4	-4.6	-0.2
HR		-1.3	-3.5	0.9

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 82

APPENDIX D

TABLE D9A. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a CPS episode by race**

	PREDICTED PROBABILITY	[95% C.I.]	
		LOWER	UPPER
PRE-AFST			
White	5.8	3.7	8.6
Black	7.4	5.0	10.7
Difference	-1.6	-3.2	-0.1
LASSO			
White	6.2	3.9	9.3
Black	6.6	4.3	9.9
Difference	-0.4	-1.9	1.1
Differential Change	**1.2**	**-1.0**	**3.4**

TABLE D9B. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a CPS episode by risk score and race**

		PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST				
LR	White	2.3	1.1	4.2
	Black	2.2	0.8	4.9
	Difference	0.1	-1.5	1.6
LOW	White	2.0	0.9	4.1
	Black	3.4	1.7	6.3
	Difference	-1.4	-3.0	0.1
MED	White	4.0	2.3	6.5
	Black	6.7	4.3	9.9
	Difference	-2.7	-4.3	-1.2
HIGH	White	4.3	2.4	7.0
	Black	6.3	4.0	9.4
	Difference	-2.1	-3.6	-0.5
HR	White	12.8	8.7	17.6
	Black	14.0	10.1	18.7
	Difference	-1.3	-2.8	0.3
LASSO				
LR	White	0.0	0.0	0.0
	Black	1.4	0.3	5.1
	Difference	-1.4	-2.9	0.1
LOW	White	0.0	0.0	0.0
	Black	2.0	0.7	4.8
	Difference	-2.1	-3.6	-0.5
MED	White	5.1	2.8	8.5
	Black	4.9	2.7	8.1
	Difference	0.2	-1.3	1.8
HIGH	White	7.0	3.8	11.7
	Black	8.1	5.2	11.8
	Difference	-1.0	-2.6	0.5
HR	White	10.3	6.7	15.0
	Black	10.2	6.9	14.3
	Difference	0.1	-1.4	1.7
DIFFERENTIAL CHANGE				
LR		-1.5	-3.6	0.7
LOW		-0.6	-2.8	1.6
MED		2.9	0.7	5.1
HIGH		1.0	-1.2	3.2
HR		1.4	-0.8	3.6

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   84

APPENDIX D

TABLE D10A. **Adjusted Analysis: probability of another episode within 180 days of a GPS screen-out episode by race**

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST			
White	28.1	25.0	31.2
Black	27.0	23.9	30.1
Difference	1.1	0.0	2.2
LASSO			
White	25.4	21.7	29.2
Black	26.4	22.7	30.3
Difference	-1.0	-2.1	0.0
Differential Change	**-2.2**	**-3.7**	**-0.6**

TABLE D10B. Adjusted Analysis: probability of another episode within 180 days of a GPS screen-out episode by risk score and race

		PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST				
LR	White	22.5	19.4	25.7
	Black	20.1	16.8	23.7
	Difference	2.4	1.3	3.5
LOW	White	20.4	17.4	23.6
	Black	23.1	19.8	26.6
	Difference	-2.7	-3.8	-1.6
MED	White	29.3	26.4	32.4
	Black	27.5	24.6	30.5
	Difference	1.9	0.8	3.0
HIGH	White	32.3	29.2	35.5
	Black	29.2	26.2	32.2
	Difference	3.2	2.1	4.3
HR	White	32.1	28.8	35.4
	Black	32.6	29.5	35.8
	Difference	-0.5	-1.6	0.5
LASSO				
LR	White	21.2	17.5	25.1
	Black	20.4	15.8	25.6
	Difference	0.8	-0.3	1.8
LOW	White	20.4	16.8	24.3
	Black	22.2	18.2	26.5
	Difference	-1.8	-2.9	-0.7
MED	White	25.2	21.7	28.7
	Black	25.1	21.8	28.6
	Difference	0.0	-1.1	1.1
HIGH	White	28.6	24.9	32.5
	Black	30.9	27.5	34.3
	Difference	-2.2	-3.3	-1.1
HR	White	29.3	24.8	34.0
	Black	31.2	27.1	35.4
	Difference	-1.9	-3.0	-0.8
DIFFERENTIAL CHANGE				
LR		-1.6	-3.2	-0.1
LOW		1.0	-0.6	2.5
MED		-1.8	-3.4	-0.3
HIGH		-5.4	-6.9	-3.9
HR		-1.4	-2.9	0.2

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 86

APPENDIX D

TABLE D11A. **Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-out episode by race**

	PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
PRE-AFST			
White	18.4	15.5	21.5
Black	18.4	15.5	21.6
Difference	0.0	-1.1	1.1
LASSO			
White	14.7	11.6	18.2
Black	16.2	13.0	19.7
Difference	-1.5	-2.6	-0.4
Differential Change	**-1.5**	**-3.0**	**0.1**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 87

APPENDIX D

TABLE D11B. **Adjusted Analysis: probability of another episode that screens-in within 180 days of a GPS screen-out episode by risk score and race**

		PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
		PRE-AFST		
LR	White	13.1	10.5	16.0
	Black	11.5	8.8	14.8
	Difference	1.6	0.5	2.7
LOW	White	13.5	10.9	16.5
	Black	13.4	10.6	16.6
	Difference	0.1	-1.0	1.2
MED	White	18.4	15.6	21.5
	Black	19.1	16.3	22.1
	Difference	-0.7	-1.7	0.4
HIGH	White	20.9	17.8	24.2
	Black	20.4	17.4	23.6
	Difference	0.5	-0.6	1.5
HR	White	23.9	20.3	27.6
	Black	25.1	21.7	28.5
	Difference	-1.2	-2.3	-0.1
		LASSO		
LR	White	10.5	7.9	13.5
	Black	9.7	6.4	13.7
	Difference	0.8	-0.2	1.9
LOW	White	9.1	6.5	12.3
	Black	11.0	8.1	14.4
	Difference	-1.9	-3.0	-0.8
MED	White	14.8	11.8	18.1
	Black	13.4	10.8	16.4
	Difference	1.4	0.3	2.4
HIGH	White	17.2	14.1	20.7
	Black	21.8	18.4	25.4
	Difference	-4.5	-5.6	-3.4
HR	White	19.4	15.1	24.2
	Black	22.5	18.6	26.8
	Difference	-3.1	-4.2	-2.0
		DIFFERENTIAL CHANGE		
LR		-0.7	-2.3	0.8
LOW		-2.0	-3.5	-0.5
MED		2.0	0.5	3.5
HIGH		-5.0	-6.5	-3.5
HR		-1.9	-3.5	-0.4

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 88

APPENDIX D

TABLE D12A. **Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-out episode by race**

	PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST			
White	5.9	4.2	8.0
Black	7.7	5.6	10.1
Difference	-1.8	-2.8	-0.7
LASSO			
White	4.0	2.4	6.1
Black	4.4	2.7	6.7
Difference	-0.4	-1.5	0.7
Differential Change	**1.4**	**-0.2**	**2.9**

TABLE D12B. Adjusted Analysis: probability of another episode that accepts for service within 180 days of a GPS screen-out episode by risk score and race

		PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
		PRE-AFST		
LR	White	1.8	0.9	3.1
	Black	4.5	2.8	7.0
	Difference	-2.8	-3.9	-1.7
LOW	White	3.9	2.5	5.8
	Black	4.4	2.8	6.5
	Difference	-0.5	-1.5	0.6
MED	White	7.0	5.0	9.3
	Black	7.4	5.5	9.8
	Difference	-0.5	-1.5	0.6
HIGH	White	6.2	4.4	8.4
	Black	8.4	6.3	11.0
	Difference	-2.3	-3.3	-1.2
HR	White	8.9	6.5	11.8
	Black	12.3	9.5	15.5
	Difference	-3.4	-4.5	-2.3
		LASSO		
LR	White	1.9	0.9	3.5
	Black	2.4	0.9	5.2
	Difference	-0.4	-1.5	0.6
LOW	White	1.2	0.5	2.4
	Black	2.6	1.3	4.8
	Difference	-1.5	-2.5	-0.4
MED	White	3.3	1.9	5.2
	Black	3.8	2.4	5.7
	Difference	-0.6	-1.6	0.5
HIGH	White	5.2	3.3	7.6
	Black	4.9	3.2	7.2
	Difference	0.2	-0.8	1.3
HR	White	6.7	4.0	10.2
	Black	7.0	4.6	10.2
	Difference	-0.4	-1.4	0.7
		DIFFERENTIAL CHANGE		
LR		2.3	0.8	3.9
LOW		-1.0	-2.5	0.5
MED		-0.1	-1.6	1.4
HIGH		2.5	1.0	4.0
HR		3.1	1.5	4.6

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023* 90

APPENDIX D

TABLE D13A. **Adjusted Analysis: probability of removal within 180 days of a GPS screen-in episode by race**

	PREDICTED PROBABILITY	[95% C.I.] LOWER	UPPER
PRE-AFST			
White	7.3	5.1	10.1
Black	8.7	6.3	11.6
Difference	-1.4	-2.5	-0.2
LASSO			
White	5.9	3.9	8.7
Black	5.0	3.4	7.3
Difference	0.9	-0.2	2.0
Differential Change	**2.3**	**0.7**	**3.9**

Report: *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)]* | *September 2023*   *91*

APPENDIX D

TABLE D13B. **Adjusted Analysis: probability of removal within 180 days of a GPS screen-in episode by risk score and race**

		PREDICTED PROBABILITY	[95% C.I.] LOWER	[95% C.I.] UPPER
		PRE-AFST		
LR	White	3.1	1.6	5.6
	Black	6.5	3.7	10.4
	Difference	-3.4	-4.5	-2.3
LOW	White	2.9	1.6	4.8
	Black	4.4	2.6	6.8
	Difference	-1.4	-2.6	-0.3
MED	White	5.0	3.4	7.1
	Black	4.8	3.3	6.7
	Difference	0.3	-0.9	1.4
HIGH	White	7.4	5.3	10.0
	Black	9.6	7.1	12.6
	Difference	-2.3	-3.4	-1.1
HR	White	10.3	7.4	13.8
	Black	11.7	8.8	15.0
	Difference	-1.3	-2.5	-0.2
		LASSO		
LR	White	1.5	0.3	4.8
	Black	3.4	1.0	8.4
	Difference	-1.9	-3.0	-0.8
LOW	White	2.8	1.0	6.2
	Black	3.0	1.2	6.3
	Difference	-0.2	-1.3	0.9
MED	White	4.3	2.5	6.8
	Black	2.8	1.7	4.4
	Difference	1.4	0.3	2.6
HIGH	White	7.3	5.0	10.2
	Black	5.0	3.4	7.1
	Difference	2.3	1.2	3.4
HR	White	7.0	4.8	9.6
	Black	6.9	4.9	9.3
	Difference	0.1	-1.0	1.2
		DIFFERENTIAL CHANGE		
LR		1.5	-0.1	3.1
LOW		1.3	-0.3	2.8
MED		1.2	-0.4	2.8
HIGH		4.5	3.0	6.1
HR		1.4	-0.2	3.0