Stanford University Response to RFP: Principal Investigator, Dr. Jeremy D. Goldhaber-Fiebert: Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County's Office of Children, Youth and Families

EXECUTIVE SUMMARY

Goals: The proposed impact evaluation will assess how the implementation of the PRM tool within Allegheny County's Office of Children, Youth, and Families improves a set of well-defined metrics. These metrics focus on how use of the PRM tool:

1) *Improves accuracy* of referrals by call screeners (decreasing unnecessary screen-ins and bad screen-outs) to ensure that high-risk children receive further investigation
2) *Reduces unwarranted variation* in whether similar children are referred for investigation
3) *Reduces disparities* in referral rates for similar children from minority groups
4) *Maintains reasonable workload* in terms of the overall rate of referrals

Additionally, the proposed study evaluates different graphical presentations of the risk predictions from the PRM tool to best improve these metrics.

Approach: The proposal uses a methodologically strong and low-cost approach to achieve the impact evaluation's goals. The study consists of a pragmatic experimental design in which performance for 1 year post-implementation of the PRM tool is compared to performance during the 3 years prior to implementation (i.e., an interrupted time-series analysis). It combines this with the randomization of each call screener to 1 of 2 possible graphical presentations of the PRM tool's risk predictions to assess which better achieves the goals of the PRM tool's use (improved accuracy, reduced variation and disparities, reasonable workload).

Experience: The proposed evaluation will be led by Dr. Jeremy Goldhaber-Fiebert, a professor at Stanford University who is a decision scientist and health policy expert and whose group has experience with decision-making under risk, quasi-experimental methods in evaluation of field trials, and child welfare and human services evaluations. Dr. Goldhaber-Fiebert and his group have successfully completed impact evaluations of field studies through multi-institution collaborations conducted in the United States as well as Costa Rica, China, and India. The Stanford team undertakes to work with the existing research team and the Allegheny County team to complete a rigorous impact evaluation that can improve operational performance for the Office of Children, Youth, and Families.

Deliverables: The findings from the proposed impact evaluation will be presented to the existing research team, the Allegheny County team and stakeholders in 3 forms:

1) *Executive Summary:* A short, non-technical document highlighting what was done in the impact evaluation, its key findings, and recommended next steps
2) *Study Report:* A longer document covering the full detail on the design, technical methods, all analyses and their results along with robustness checks and validation, and recommendations for next steps including PRM tool risk prediction presentation
3) *Final Presentation:* A PowerPoint slide deck highlighting elements contained in the executive summary, presented by the evaluation team, and usable by Allegheny County for subsequent presentation and discussion

We also anticipate producing an academic paper in collaboration with the existing research team describing the study and its findings that will provide additional external review and validation of the findings to build confidence for Allegheny County decision makers.

NARRATIVE

1. Organizational Experience

1a. Specific experience in the study of decision-making under risk, quasi-experimental methods in evaluation of field trials, and child welfare and human services evaluations

Dr. Jeremy Goldhaber-Fiebert, who leads the proposed evaluation, has specific experience in all 3 areas listed in the RFP.

Experience with decision-making under risk: Dr. Goldhaber-Fiebert received his PhD in Health Decision Science – the study of decision-making under risk and uncertainty in the context of health – from Harvard University in 2008. After joining the faculty of Stanford University in 2008, he has annually taught 1-2 courses in decision analysis to a range of graduate and undergraduate students. He has published numerous applied and methodological articles addressing decision-making under risk. As a trainee, he was twice awarded the major student prize from the Society for Medical Decision Making and has subsequently served as Trustee of the society. He is currently an Associate Editor of the journal *Medical Decision Making* and has participated in establishing and updating major guidelines and best practices for conducting decision-analytic studies.

Experience with quasi-experimental methods in evaluation of field trials: Dr. Goldhaber-Fiebert has led or participated in numerous field trials in both domestic and international settings. These studies – many published or currently under review – have used a variety of techniques to make valid causal inference in the presence of randomization or quasi-randomization (e.g., randomization, instrumental variables, difference-in-difference, interrupted time-series).

His studies have examined a broad range of topics, many relating to child health and well-being outcomes, and include interventions aimed at:

- Improving exercise habits of adults
- Improving physical activity and nutrition practices for type 2 diabetic patients
- Improving outcomes for individuals with clinical depression and obesity
- Improving and reducing unwarranted variation in the pediatric care provided by less qualified practitioners through the use of telehealth technologies
- Reducing pediatric anemia in school children
- Improving nutritional outcomes in poor children with community and school programs

This work has been published and presented in major journals and has been funded by the National Institutes of Health, Fulbright, the Bill and Melinda Gates Foundation, the International Initiative for Impact Evaluation (3ie), and the UK's Department of International Development.

Dr. Lea Prince, the research analyst for the proposed study, received her PhD in economics from the University of California, Davis in 2014. Her dissertation work included evaluations of caregiver decision pathways regarding sick children, neighbor effects on caregiver behavior and

decisions relating to treating sick children. She used quasi-expermental econometric methods including instrumental variables to conduct this work. She has worked at Stanford with Dr. Goldhaber-Fiebert since 2013, where she has participated in his studies, particularly those focusing on evaluations of interventions to improve child health and well-being.

Experience with child welfare and human services evaluations: Since joining Stanford's faculty in 2008, Dr. Goldhaber-Fiebert has developed active collaborations with child welfare and child health researchers. Supported by Advanced Center grants from the NIH, these collaborations include researchers at UC Berkeley (Lonnie Snowden), USC (Larry Palinkas), Oregon Social Learning Center (John Landsverk, Patricia Chamberlain, and Lisa Saldana), University of Chicago/Chapin Hall (Fred Wulczyn), NYU (Sarah Horwitz). Through these collaborations, Dr. Goldhaber-Fiebert has published 6 papers focused on decision and policy analyses for child welfare, including:

- Assessments of how child welfare agencies explore and adopt evidence-based practices
- Measuring the service use and outcome consequences that may stem from decisions about which policies and practices to use
- Reviews detailing the general lack of decision-analytic research in support of such child welfare decision making
- Decision-analytic modeling to support the evaluation of evidence-based child welfare policies

Dr. Goldhaber-Fiebert has extended his collaboration with these investigators with a major NIH proposal, currently being readied for submission, to evaluate decision-analytic modeling as a decision-support tool for child welfare administrators in all counties in the State of New York.

1b. Key personnel involved in the evaluation

There are two key personnel whose full CVs are also included as attachments:

Dr. Jeremy Goldhaber-Fiebert is a tenured Associate Professor of Medicine and of Health Research and Policy at Stanford University. He is a PhD decision scientist with experience both in program evaluations using both randomized and quasi-experimental designs and in a range of analyses specifically focusing on the use of decision science in the child welfare and human services. He will be responsible for the overall design and conduct of the evaluation as well as the coordination, communication and meetings with the existing research team and the Allegheny County team, and will draft and present interim and final reports, presentations, and papers.

Dr. Lea Prince is a Research Analyst at Stanford University. She is a PhD economist with experience in econometric evaluations that consider randomized and quasi-experimental designs and data. Her prior work has focused on interventions and services that have the potential to improve child health and well-being. Dr. Prince will be responsible for data management, analysis, and coding and running interim and final analyses of the evaluation whose results will be included in the reports, presentations, and papers.

2. Approach to Impact Evaluation

2a. Overall approach

The goal of the impact evaluation is to determine how the implementation of the PRM tool improves the accuracy of referrals by call screeners, reduces unwarranted variation in the likelihood of referring similar children, reduces disparities in referral rates for minority groups, and alters overall referral rates and workload.

We propose a pragmatic experimental design that uses 3 years of call screeners' historical performance as the control against which to compare their performance in the year after the PRM tool is implemented (commonly termed "interrupted time-series analyses"). Additionally, the experimental design has each call screener randomized to 1 of 2 possible graphical presentations of the PRM tool's risk predictions. These presentations of the risk predictions vary in terms of how granularly risk is presented, motivated by the idea that on the one hand summarized information is often easier to act upon and leads to standardization of practice but on the other hand that finer grained information allows for more accurate risk classification and targeting of call screener effort.

Questions addressed by the study's analyses: Analyses of the impact evaluation will comment both on how the PRM tool has altered accuracy, variation, disparities, and workload and which graphical presentation of the PRM tool's risk predictions performs best.

2b. Management of the evaluation, including working with self-selected collaborators, the existing research team, and the Allegheny County team

The evaluation team, as described above in section 1, has extensive experience working with and coordinating between multiple institutions and teams in various geographic locations. We will conduct the present evaluation using best practices from this prior work, managing the evaluation in three phases described below.

Management in Phase I: Initial Operational Planning
Initial operational planning is central to the success of a coordinated impact evaluation because it is important to build trust, ensure agreement on definitions, standardize and practice procedures, and plan for key moments in a study. We will hold on a series of video conference meetings, phone calls, and email exchanges intended to:

1) Understand the data elements extracted from Allegheny County's data warehouse and how the existing research team have processed and cleaned these data elements. These would be used to refine exact definitions in the analytic plan below
2) Confer on the exact user interface design for the presentation of the PRM tool's risk prediction per the description of 2 such presentations listed in the analytic plan below
3) Confirm the processes for maintaining randomized assignment of each call screener
4) Plan intermediate data extracts (historical data, end of the first month of the study and at month 6) used for preliminary analyses

5) Confirm expectations for final presentation of the results to the Allegheny County team

We expect that coordination and discussion prior to the study during this planning phase would be the most intensive involvement between our team, the existing research team, and the Allegheny County team. We expect to visit the County team in person during this period as well.

Management in Phase II: Conduct of the Study

We believe short weekly video conference meetings conducted between Dr. Goldhaber-Fiebert and designees from the existing research team and from Allegheny County is an efficient way to exchange status, raise any issues and make plans to resolve them.

Management in Phase III: Reporting of the Study Results

After analyzing the data per the protocol described in the analytic methods section, we will draft versions of a final presentation, executive summary of findings, and full report including motivation, data, methods, results, and conclusions. We will circulate these for comments and questions by the existing research team and by Allegheny County, incorporating feedback into a finalized version of each. We will present these to Allegheny County stakeholders in person.

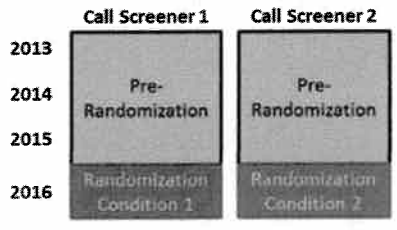
2c. Analytic methods to assess impact of the PRM Tool in practice

Outcome Goals for PRM: The RFP describes a set of outcomes for the impact evaluation that encompass the goals of implementing the PRM tool itself in terms of:

1) Accuracy: Increase the fraction of high risk children referred to investigation and reduce the fraction of low risk children referred to investigation.
2) Consistency: Reduce variation in referral rates for children with similar characteristic and risks.
3) Fairness: Reduce disparities in referral rates between children from various disadvantaged groups with otherwise similar characteristics and risks.
4) Workload: Keep overall work volume at a manageable level by maintaining overall referral rates at a reasonable level.

Design Overview: We have designed an impact evaluation of the PRM tool to measure its effects on these outcomes over a 1-year period. Our impact evaluation measures unbiased causal effects using a randomized design that, in keeping with the requirements of the RFP, does not employ a traditional control group. Instead, it uses each call screener's referral patterns over the 3 years prior to randomization as a historic control (2013-2015). Notably, although not available to call screeners in this period, it will be important to compute PRM risk scores for all reports in this 3-year period for use in analyses described below. Effects of the PRM tool are measured relative to the 3-year pre-trends (a form of "interrupted time-series analysis"). At the start of the evaluation, each call screener is randomized to 1 of 2 different presentations of the PRM tool's risk prediction information (referred to as a "Randomization Condition"). Figure 1 illustrates the pre-randomization and randomization periods for 2 call

Figure 1



screeners who are randomized to different Randomization Conditions.

During the year-long study, each call screener only sees 1 of the Randomization Conditions on his/her computer for all children screened for referral (Figure 2, left). Randomization Condition 1 provides a fine-grained graphical display of the predicted risk level for a given child in (on a scale of 1-20, where 1 is the Lowest risk and 20 is the Highest). Randomization Condition 2 provides a grouped graphical display of risk level categorized as high, medium, or low level of risk. Cut-offs shown for Randomization Condition 2 in Figure 2 are illustrative only, based on prior work from the existing research team who have implemented PRM tool's in settings including New Zealand (Figure 3, right). In Figure 3, high risk involves the top 2 deciles whose risk is >25% on average, whereas low risk involves the bottom 3 deciles whose risk is <5% on average. For the proposed evaluation, the exact shape of Allegheny County's risk curve in the pre-period would determine the cutoffs to be used for Randomization Condition 2.

Figure 3

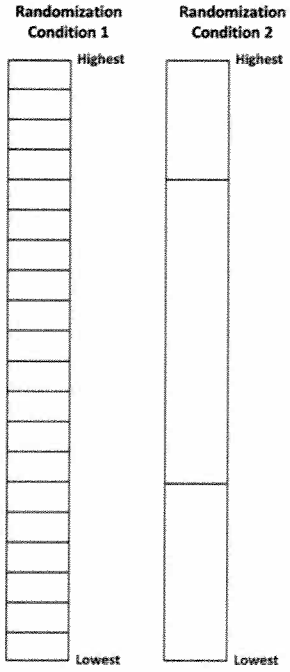
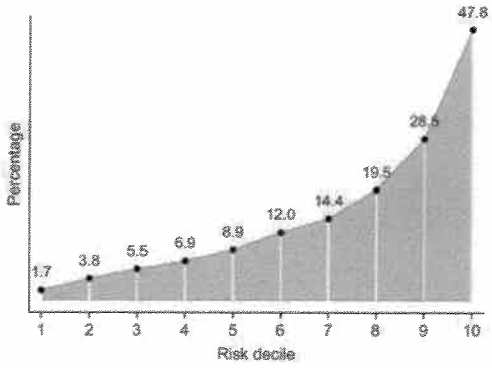


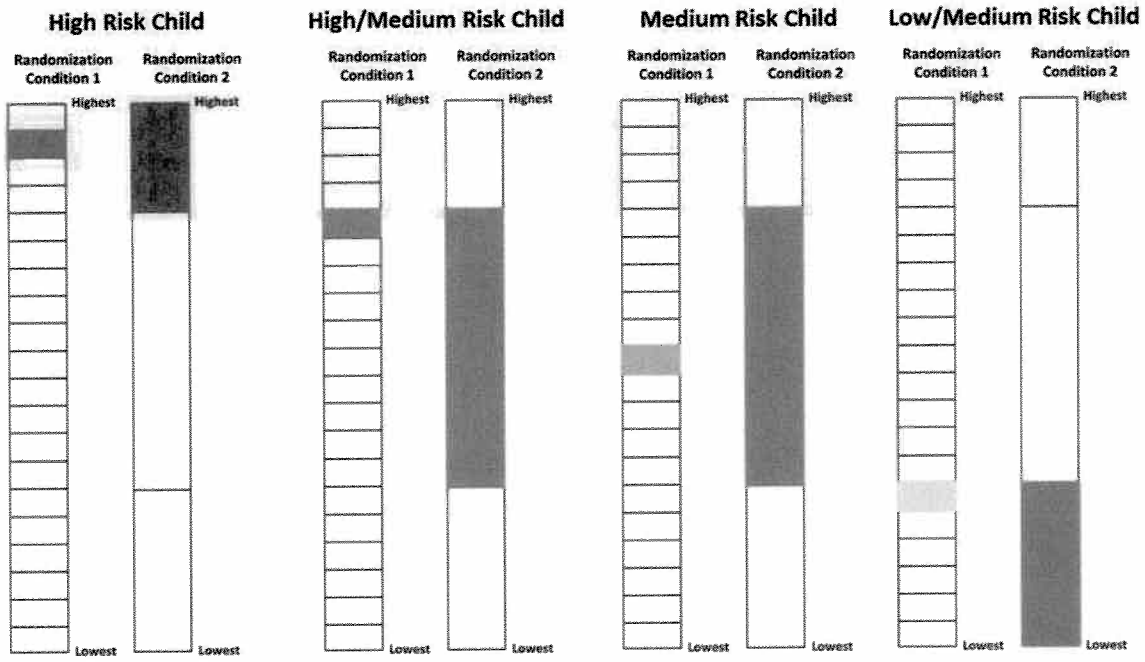
Figure 2



Rationale for Randomization of PRM Risk Presentation: Evaluating multiple displays of the PRM tool's information is important to determine how best achieve the goals of the PRM tool itself. Evaluation is necessary because based on theory alone it is unclear whether Randomization Condition 1 or Randomization Condition 2 (show in Figures 2 and 3 above) will be more effective. On the one hand, because Randomization Condition 1 provides a graphical display of each child's risk that is finely grained, it should lead to higher Accuracy since only the most similar children are grouped together when determining whether or not to refer for investigation. It may also be best for controlling overall Workload, since each call screener can focus on children that, in his/her best judgment, most need further assessment (e.g., those children whose risk level falls somewhere in the middle of the range) prior to making a referral determination. On the other hand, when information is too finely presented, it can be overwhelming and lead to "information overload" so that the best effects on Accuracy and Workload are by no means certain. Randomization Condition 2 addresses "information overload" by presenting children in 1 of 3 simple risk categories (high, medium, and low). This simplification also could help to better promote Consistency and Fairness (e.g., each call screener does not need to establish his/her own threshold for calling a child high risk because common pre-specified thresholds are used).

Figure 4 below illustrates the trade-offs between the two Randomization Conditions – two types of risk presentation. On the left, high risk children are shown similarly for the two Randomization Conditions, though how high a high risk child's level is in Randomization Condition 2 will be less clear to the call screener. However, it may be sufficient to know that risk is >25%, with further refinements not relevant to the decision to investigate (i.e., all children at this risk level should be referred). The latter 3 panels show that children at various points in the medium risk range will be more finely differentiated in Randomization Condition 1 than Randomization Condition 2. Again, lack of differentiation may be more efficient at focusing the call screeners' attention on the middle risk group in general and in eliminating unwarranted variation in the threshold for investigation that each call screener might otherwise use.

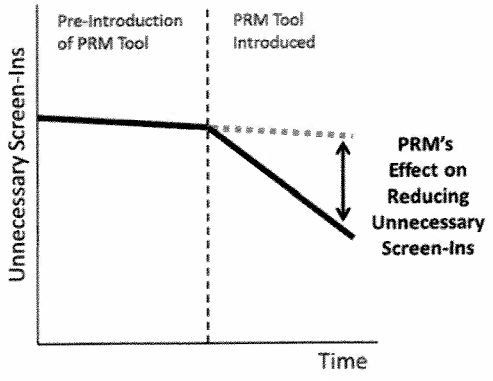
Figure 4



Outcomes and Their Analysis: The analysis will examine multiple outcomes. These include:

Outcome 1: Overall rates of calls "screening-in": We define the rate of "screening-in" as equal to the number of children in reports referred by the call screener for further investigation divided by the total number of children in reports, computed for 1-month windows of time.

Figure 5



- We will conduct two analyses of Outcome 1 (also illustrated graphically in Figure 5):
  1. *Does the introduction of the PRM tool alter the overall rate of calls screening in*? This is an interrupted time series analysis where the

rate of calls screening in and its trend are established in the 3-year pre-randomization and their change during the 1-year post-randomization period is then measured. The interrupted time-series analysis interacts an indicator for the introduction of the PRM tool with time before/after the introduction of the PRM tool (a significant interaction coefficient would show how the PRM tool had altered the overall screen-in rate). As this is a measure of workload, we believe the goal is less to change this overall rate than to reduce unnecessary screen-ins and replace them with additional good screen-ins by reducing the likelihood of bad screen-outs (see other outcomes below).

2. *Does one presentation of the PRM tool's risk data (Randomization Condition 1) differentially alter the overall rate of calls screening compared to how another presentation does so (Randomization Condition 2)?* This is an interrupted time series analysis similar to the one described above except that there are two separate indicators for which randomization condition each call screener is assigned to that are interacted with time before/after the introduction of the PRM tool (a significant test for equality for the randomization condition-specific interactions would show that the two presentations have differential effects).

Outcome 2: Likelihood of "bad" screen-outs: We define the rate of "bad" screen-outs as the number of children in reports who were not referred by the call screener for further investigation but subsequently either died, nearly died, or were re-referred within a given number of months divided by the total number of children in reports who were not referred. For the re-referral period, we plan to use 1 and 2 months because using longer lengths for this period will cause too many reports coming in towards the end of the study to be excluded for not having complete follow-up prior to study end.

- We will conduct two analyses for each duration (1-month and 2-month follow-up) used for Outcome 2 that use the same analytic approaches as those described above for Outcome 1.

Outcomes 3: Likelihood of unnecessary screen-ins: We define the rate of unnecessary screen-ins in two ways:

1. *Investigator-defined unnecessary:* The number of children in reports referred by the call screener for further investigation for whom the investigation concluded that no further action was necessary divided by the total number of children in reports referred for further investigation.
2. *Investigator AND outcome-defined unnecessary*: The number of children in reports referred by the call screener for further investigation for whom the investigation concluded that no further action was necessary and who did not die, nearly die, or get referred within a given number of months (1 or 2) divided by the total number of children in reports referred for further investigation.

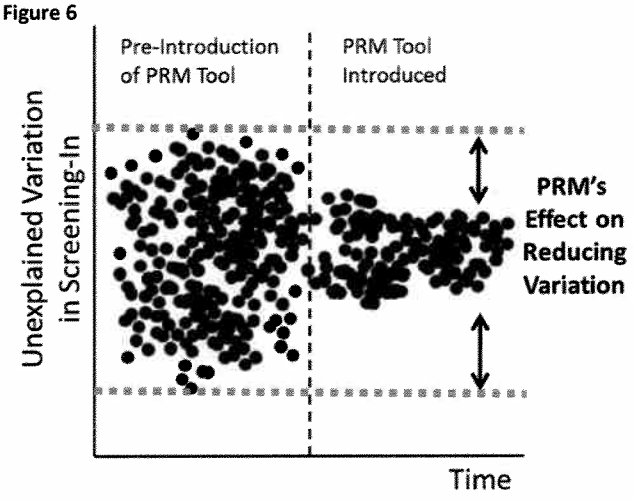
- We will conduct two analyses for each definition of unnecessary (including 2 analyses with 1-month follow-up and 2-month follow-up for this second definition of unnecessary) used for Outcome 3 that use the same analytic approaches as those described above for Outcome 1.

Outcome 4: Unwarranted variation in screening decisions: We define unwarranted variation in screening decisions in two ways:

1. *Individual call screeners:* unwarranted variation of an individual call screener occurs if a given call screener is not self-consistent in how he/she refers children with highly similar characteristics (variance within each call screener in the probability of "screening-in" similar reports computed for 3-month time windows).
2. *Call screeners as a group:* even if call screener are self-consistent in how they handle reports with children with highly similar characteristics, different call screeners may have different thresholds for "screening-in" (variance between call screeners in their average probability of "screening-in" similar reports computed over 1- an d 3-month time windows).

- We will conduct two analyses for each type of variation listed in Outcome 4 (also illustrated graphically in Figure 6).

Figure 6



1. *Analysis of individual call screeners*: For each call screener, we will model the likelihood of screening-in each child they handled given the child's observable characteristics as well as the time before/after the introduction of the PRM tool and an interaction of an indicator for the introduction of the PRM tool with time before/after the introduction of the PRM tool. After these adjustments, the variation that remains (the residual in the regression) is the unexplained/unwarranted variation. If we plot the residuals with respect to time before/after the introduction of the PRM tool, we should see a decrease in the range of the residuals after the PRM tool was introduced if the PRM tool reduces unwarranted variation, which we can test for statistically (Figure 6).
2. *Analysis of call screeners as a group*: We will conduct a similar analysis to the one described for individual call screeners except we will analyze the data of all call screeners for each Randomization Condition as a group simultaneously and include call screener fixed effects and their interactions with time before/after the introduction of the PRM tool, and with the interaction of time and an indicator for the introduction of the PRM tool. We will again examine the residuals as described above (Figure 6).

Outcome 5: Disparities in screening decisions: This outcome and its analysis are closely related to the unwarranted variation in screening decisions outcome (Outcome 4). We will consider four types of disparities:

1. *Gender disparities:* exist if call screeners have different probabilities of "screening-in" otherwise similar reports for male and female children.

2. *Racial disparities:* exist if call screeners have different probabilities of "screening-in" otherwise similar reports for children of different racial backgrounds.
3. *Economic disparities:* exist if call screeners have different probabilities of "screening-in" otherwise similar reports for children of different economic backgrounds. Here, we would use a child's Medicaid or public assistance eligibility as a proxy for being economically disadvantaged.
4. *Composite disparities:* Racial minorities are often more economically disadvantaged on average than the majority. Therefore, we will define 4 composite categories of group membership based on whether a child is a member of a racial minority and/or economically disadvantaged and examine whether call screeners have different probabilities of "screening-in" otherwise similar reports for children in each of these composite categories.

As noted by the Institute of Medicine, disparities exist when there are differences in how groups are treated that cannot be explained for appropriate reasons. Hence, in the setting of call screeners and "screening-in" reports, an important question is if otherwise similar children are considered similar based on their predicted risk. If using PRM compresses unwarranted variation in "screening-in" similar reports, it may also reduce various types of disparities in the probability of "screening-in" similar reports. However, it may still not compress such variation evenly across these groups, and hence may still leave residual albeit smaller disparities.

- The analysis of Outcome 5 is very similar to the analysis of unwarranted variation in call screeners as a group for Outcome 4 described above. For each type of potential disparity, we use a similar regression model except that we interact the child's characteristic that defines the potential disparity (e.g., female gender) with the time before/after the introduction of the PRM tool and an interaction of an indicator for the introduction of the PRM tool with time before/after the introduction of the PRM tool in the regression model of the probability of "screening-in" we will interact. If the three-way interaction between disparity-category/time/PRM introduction is significant, then there is a differential effect of PRM introduction on otherwise similar individuals based on their disparity category. We then similarly examine the residuals for this type of regression to determine whether the PRM introduction-caused change in variation is differential between disparity-categories. Findings that either the expected change or change in unexplained variation differs systematically by a disparity category should be interpreted as reason for additional investigation on the part of Allegheny County as to the reasons for this and consideration of ways to further address these differences.

Data Requirements: All analyses requires de-identified data on individual children including race, gender, age and other sociodemographic and family characteristics including poverty status as well as the timing and types of experiences with the call screeners and related referrals and investigations as well as with child welfare and other child-serving systems from 2013 through study end. Importantly, the PRM tool's risk prediction for all children including those in the pre-randomization period is needed, despite the fact that screeners would not have had this information. This is to adjust for predicted risk so that the effects of other factors (the PRM tool introduction or a disparity category) can be examined for "otherwise similar" children. Finally, these analyses require de-identified data on individual call screeners including

their sociodemographic characteristics, length of tenure in their current job, and the Randomization Condition they were assigned to. The proposal assumes that these data can be extracted and cleaned by the existing research team in a manner similar to that used by the team to develop the PRM algorithm in order to provide a clean analytic dataset for the proposed analysis plan described above.

2d. Perceived requirements/limitations/challenges; plans for assessing result validity

The main challenge to an impact evaluation that does not have a control group is that the effects of other changes that are contemporaneous with the implementation of the PRM tool are attributed to the impact of the PRM tool itself. For example, simply training the call screeners and letting them know that they will be evaluated may lead to differences in their behavior. We will address this in two ways. First, we will run a statistical placebo test which essentially involves randomly assuming different "placebo" start dates for the PRM intervention and running the analyses described above – if differences we observe are due to periodic training or other activities, then we would see spurious effects for start dates of the PRM intervention other than the actual date. Second, comparison of the effect estimates from the two Randomization Conditions will help since it is unlikely that they will not differ at all – their difference will be evidence of a PRM effect since other contemporaneous changes will not be perfectly correlated with the PRM Randomization Condition.

A second challenge is that there a relatively small number of call screeners (~11) which means that each randomization condition will have ~6 call screeners assigned to it and could differ by chance alone. This is why we conduct overall PRM intervention analyses as well as separate Randomization Condition analyses. Despite having ~10,000 children screened per year, the study may still be underpowered to detect smaller differences between Randomization Conditions. However, selecting one way to present the PRM risk data over another is less important if the differences between the two in terms of their effect are not large.

Finally, not all possible ways to summarize the PRM risk information can be tested in a randomized design and so it is possible that some other risk thresholds and categorization could produce an even bigger effect. Examination of the effect of Randomization Condition 1 on outcomes stratified by PRM risk score would enable the identification of threshold effects (i.e., PRM risk scores above which the intervention has a greater effect or below which it has a greater effect) which could then be used to design a variant of Randomization Condition 2 where risk scores are categorized at these empirically derived thresholds. This can be accomplished by repeating the analyses above with a non-linear spline for PRM risk score and then examining the coefficients of the spline's interaction with PRM tool introduction to identify PRM predicted risk thresholds. Connected to this is the idea that call screeners may spend more effort on medium risk children, and hence findings from the process evaluation about how call screeners change their time allocation for different risk children could be very helpful in continuing to enhance the effectiveness of the PRM tool. We would be very open to working with the evaluation team responsible for the process evaluation to accomplish such tasks.

BUDGET NARRATIVE
TITLE: Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County's Office of Children, Youth and Families?
PI: Jeremy Goldhaber-Fiebert

This budget was constructed for the period 3/1/2016-02/28/2017. Per Final Agreement dated 5/20/2015 between Stanford University and the Office of Naval Research, effective 9/01/2015 the Facilities and Administrative cost for this proposal is the on-campus organized research rate of 58%. The Facilities and Administration rate is 58.0% for this proposal. The fringe benefits rates are those negotiated between the U.S. Office of Naval Research and Stanford University. 30.6% for faculty and regular staff, 24.3% for postdoctoral affiliates, 5.2% for graduate students and 8.8% for contingent workers. Stanford's agreement with the Office of Naval Research provides for 8.9% vacation accrual/disability sick leave (DSL) for exempt employees and 7.7% for non-exempt employees. The vacation accrual/DSL rates will be charged at the time of the salary expenditure. No salary will be charged to the award when the employee is on vacation.

The majority of the budget is for salary support of two key personnel:
Dr. Jeremy Goldhaber-Fiebert (10% FTE, 1.2 calendar months): Dr. Goldhaber-Fiebert is a tenured Associate Professor of Medicine and of Health Research and Policy at Stanford University. He is a PhD decision scientist with experience both in program evaluations using both randomized and quasi-experimental designs and in a range of analyses specifically focusing on the use of decision science in the child welfare and human services. He will be responsible for the overall design and conduct of the evaluation as well as the coordination, communication and meetings with the existing research team and the Allegheny County team, and will draft and present interim and final reports, presentations, and papers.

Dr. Lea Prince (60% FTE, 7.2 calendar months): Dr. Prince is a Research Analyst at Stanford University. She is a PhD economist with experience in econometric evaluations that consider randomized and quasi-experimental designs and data. Her prior work has focused on interventions and services that have the potential to improve child health and well-being. Dr. Prince will be responsible for data management, analysis, and coding and running interim and final analyses of the evaluation whose results will be included in the reports, presentations, and papers.

Travel (2 trips totaling $3,000): A member of the evaluation team will make two trips to Allegheny County for face-to-face coordination of the evaluation during Phase I and for final presentation of results in Phase III.

BUDGET
TITLE: Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County's Office of Childre
PI: Jeremy Goldhaber-Fiebert
Period of Performance: 03/01/2016-02/28/2017

Personnel

	Role	Effort
Jeremy Goldhaber-Fiebert	Principal Investigator	10%
Lea Prince	Research Analyst	60%

Personnel Subtotal

Travel

2 trips to Allegheny County

Travel Subtotal

Total Direct
Facilities and Administration (58%)

Total Requested

:n, Youth and Families?

	Amount
$	27,175
$	57,734
$	**84,909**
$	3,000
$	**3,000**
$	**87,909**
$	**50,987**
$	**138,896**

REFERENCES

John Landsverk, PhD
Senior Research Scientist, Oregon Social Learning Center
Senior Scholar, Brown School of Social Work, Washington University in St. Louis
Senior Research Professor, School of Social Work, University of Southern California
10 Shelton McMurphey Blvd
Eugene, OR 97401
[redacted]

Sarah McCue Horwitz, PhD,
Professor, New York University School of Medicine,
Department of Child and Adolescent Psychiatry
1 Park Avenue, 7th Floor
New York, NY 10016
[redacted]

Patricia Chamberlain, PhD
Science Director, Oregon Social Learning Center
10 Shelton McMurphey Blvd
Eugene, OR 97401
[redacted]