

# Use of a community advisory board to build equitable algorithms for participation in clinical trials: a protocol paper for HoPeNET

Nicole Farmer,<sup>1</sup> Foster Osei Baah,<sup>2</sup> Faustine Williams,<sup>3</sup> Erika Ortiz-Chapparo,<sup>2</sup> Valerie M Mitchell,<sup>2</sup> Latifa Jackson,<sup>4</sup> Billy Collins,<sup>2</sup> Lennox Graham,<sup>5</sup> Gwenyth R Wallen,<sup>1</sup> Tiffany M Powell-Wiley,<sup>2,3</sup> Allan Johnson<sup>6</sup>

**To cite:** Farmer N, Osei Baah F, Williams F, *et al.* Use of a community advisory board to build equitable algorithms for participation in clinical trials: a protocol paper for HoPeNET. *BMJ Health Care Inform* 2022;**29**:e100453. doi:10.1136/bmjhci-2021-100453

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjhci-2021-100453>).

Received 31 July 2021  
Accepted 07 January 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr Tiffany M Powell-Wiley; [tiffany.powell-wiley@nih.gov](mailto:tiffany.powell-wiley@nih.gov)

Dr Allan Johnson; [ajohnson@howard.edu](mailto:ajohnson@howard.edu)

## ABSTRACT

**Introduction** Participation from racial and ethnic minorities in clinical trials has been burdened by issues surrounding mistrust and access to healthcare. There is emerging use of machine learning (ML) in clinical trial recruitment and evaluation. However, for individuals from groups who are recipients of societal biases, utilisation of ML can lead to the creation and use of biased algorithms. To minimise bias, the design of equitable ML tools that advance health equity could be guided by community engagement processes. The Howard University Partnership with the National Institutes of Health for Equitable Clinical Trial Participation for Racial/Ethnic Communities Underrepresented in Research (HoPeNET) seeks to create an ML-based infrastructure from community advisory board (CAB) experiences to enhance participation of African-Americans/Blacks in clinical trials.

**Methods and analysis** This triphased cross-sectional study (24 months, n=56) will create a CAB of community members and research investigators. The three phases of the study include: (1) identification of perceived barriers/facilitators to clinical trial engagement through qualitative/quantitative methods and systems-based model building participation; (2) operation of CAB meetings and (3) development of a predictive ML tool and outcome evaluation. Identified predictors from the participant-derived systems-based map will be used for the ML tool development.

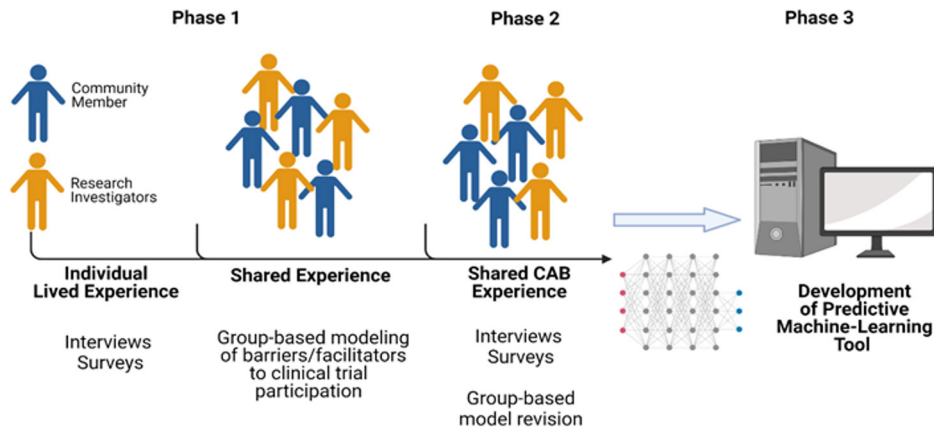
**Ethics and dissemination** We anticipate minimum risk for participants. Institutional review board approval and informed consent has been obtained and patient confidentiality ensured.

## INTRODUCTION

Machine learning (ML) can identify statistical patterns from generated data to train computers to perform tasks intended to aid in human decision making.<sup>1</sup> Emerging use of ML is occurring in clinical trial evaluation and clinical trial recruitment,<sup>2-4</sup> a field in which improvement in reaching and recruiting racial and ethnic minorities is increasingly essential. For individuals from groups who

are recipients of societal biases, utilisation of ML can lead to the creation and use of biased algorithms.<sup>5,6</sup> The design of equitable ML tools that advance health equity could be guided by community engagement processes which leverage collective knowledge and experience to inform clinical trial development and design.

Participation from racial and ethnic minorities in clinical trials has been burdened by issues surrounding mistrust and access to healthcare, both of which ultimately impact referral to clinical trials.<sup>7,8</sup> Furthermore, participation barriers may extend beyond these recognised factors. To address barriers, community-based participatory research (CBPR) has emerged to involve communities at all stages of a research study life cycle from study design through dissemination of results. A component of CBPR research is the formation of a community advisory board (CAB) to advise and direct research questions, recruitment plans and evaluate disseminated results of the study.<sup>9,10</sup> Although considered central in securing participation from under-represented communities, utilisation of the CAB's experience more broadly in therapeutic clinical trials has been limited. To date, the utilisation of the CAB experience to generate data has not been used to develop ML algorithms. We, therefore, seek to conduct a study in which CAB input is utilised for ML development through capturing 'lived experience-based knowledge' generated during a 12-month CAB participation study. We will achieve this goal by addressing three specific aims: (1) measure perceived barriers/facilitators to clinical trial engagement among African-Americans using qualitative and quantitative approaches; (2) use group-based model building as a systems



**Figure 1** Graphical abstract of HoPeNET protocol: a community advisory board (CAB)-based protocol to evaluate lived experiences from multiple stakeholders, to create systems-based understanding of barriers and facilitators to clinical trial participation. HoPeNET will aid in creating a predictive algorithmic tool to help increase African-American clinical trial participation. Figure created by coauthors (NF, FOB and EO-C).

science approach to identify key activities to improve community trust and engagement and (3) develop an ML-based tool for predicting community engagement in clinical trials using data from the group-based model building. **Figure 1** is a graphical representation of the three phases of the study.

**METHODS**

**Recruitment and characterisation of study participants**

This is a multisite 24-month, triphased study of 56 participants. The two sites for the study are Howard University (HU) and the Intramural Research Programme (IRP) of the National Institutes of Health (NIH). Both sites are located within the metropolitan Washington, D.C. area.

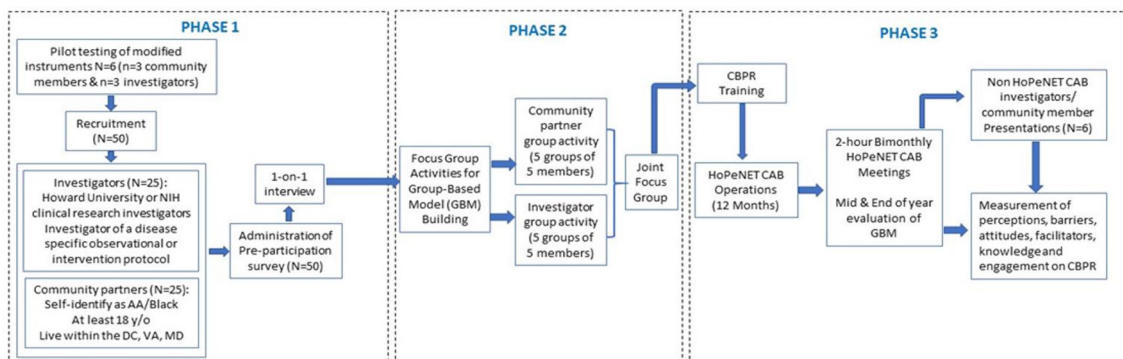
During phase 1 (**figure 2**), formation of a 50-member HoPeNET CAB will occur through the recruitment of two groups: 25 community partners and 25 investigators. Recruitment of participants will occur through multiple channels including emails to our current CAB on cardiovascular disease and obesity, the Washington D.C. Cardiovascular Health and Obesity Collaborative (D.C. CHOC)<sup>11</sup>; flyer distribution targeting members of community-serving non-profit organisations, through institutional communications and snowball recruitment

for HU and the NIH IRP investigators. Inclusion and exclusion criteria for each group of HoPeNET CAB members are provided in **figure 2**.

A preparticipation survey and one-on-one interviews will be used to capture demographics, knowledge of CBPR principles, perceptions and beliefs surrounding clinical trial participation, and the role of social determinants and implicit bias on research outcomes (**table 1**). The preparticipation survey and interview guide will be pilot tested prior to administration.

**Focus groups/workshops: group model building**

Three focus groups/workshops with (1) community members, (2) investigators and (3) both groups combined will be conducted using a group model building (GMB) activity methodology<sup>12</sup> (**figure 2**). GMB is a powerful participatory method for actively engaging stakeholders or communities to provide perspective on a complex problem, structure, or dynamic process as well as the results and solutions.<sup>12 13</sup> To facilitate knowledge sharing, discussion and consensus on the issue, participants will be engaged in a number of activities (scripts). The objective of the first two sessions is to elicit discussion on factors contributing to a lack of African-American clinical trial participation by creating a systems-based map or causal



**Figure 2** Study procedures. Figure created by coauthors (NF, FOB and EO-C). CAB, community advisory board; CBPR, community-based participatory research.

**Table 1** Assessment data and measurement tools

Assessment of recruited participants Phase 1	Assessment of CAB members Phase 2	CBPR measurements Phase 3
<ul style="list-style-type: none"> <li>▶ Engagement</li> <li>▶ Self-efficacy scales</li> <li>▶ Resilience measures</li> <li>▶ Transcribed attitudes interviews</li> </ul>	<ul style="list-style-type: none"> <li>▶ Transcribed identified barriers/motivators for CAB participation</li> <li>▶ Identification of latent factors affecting CAB members:               <ul style="list-style-type: none"> <li>▶ RACE scale</li> <li>▶ Discrimination</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>▶ Perception/barriers</li> <li>▶ Attitudes</li> <li>▶ Knowledge/engagement</li> <li>▶ Mid and end of year evaluation of group-based modelling</li> </ul>

CAB, community advisory board ; CBPR, community-based participatory research; RACE, Race Attributes in Clinical Evaluation.

loop diagrams (CLDs). Our rationale for conducting separate workshops prior to the combined sessions is to develop a richer understanding of participants' unbiased perceptions surrounding clinical trial participation. Additionally, this strategy improves stakeholders' engagement and participation in group activities that may suffer due to power imbalances. The research team will work offline between sessions to refine and synthesise the CLDs or informal causal maps, created during sessions 1 and 2 by participants. During the third combined session, participants will evaluate the synthesised map. This is necessary to ensure that the model reflects the insights and stories shared by participants. In addition, this allows for participants to identify potential areas in the system where they believe change is needed. Following the iterative process, all participants will be provided copies of the combined systems map to provide final feedback to the research team.

### CAB intervention

In phase 2 (figure 2), the HoPeNET CAB experience will commence. Prior to the start of bimonthly meetings, the CAB participants will receive asynchronous training on CBPR principles, code of practices, and confidentiality procedures. Training will be provided by NIH site team members (TP-W, GRW, NF and VMM) and current members of the D.C. CHOC CAB. To ensure equitable discussions during meetings, 'Ambassadors' from each stakeholder group will be trained to lead CAB meetings.

The CAB will meet bimonthly for 2 hours over a 12-month period. During the meetings, non-CAB members (n=6) consisting of investigators and community members will be invited to present on specific disease areas from ongoing protocols or community health projects. Inclusion criteria for the HoPeNET presenters are based on criteria for the HoPeNET CAB members (see figure 2), except presenters do not have to self-identify as African-American. Exclusion criteria are that presenters cannot participate as CAB members or focus group facilitators. After the presentations, the HoPeNET CAB will evaluate the studies and provide feedback on ways to engage African-Americans in their protocols/programmes or to engage researchers in the community projects. All HoPeNET CAB meetings will be audio recorded for anonymised transcription. To continue the iterative process of the system-based map, at the mid-year

time point, the HoPeNET CAB will reevaluate the systems-based model of facilitators and barriers to clinical trial participation. All revisions of the model will be provided to research staff for analysis. Participant engagement at the mid-year and end-of-year time points will also occur using a validated standardised metric.<sup>13</sup>

### ML tool development

Phase 3 will occur over a 6-month period and will focus on the implementation of data results from phases 1 and 2 for the development of the ML-based tool and outcome evaluation. The primary input data for the ML algorithm will be the results of the facilitators and barriers model (table 2) created from the group-based model activity and analysis of collected qualitative and survey data. Initial survey responses will be aggregated thematically across multiple responses and Likert scales to create ordinal scales to use supervised ML approaches. This will include regression models and decision trees to identify patterns in the data. Supervised ML approaches are useful in identifying patterns where we have labelled and structured data. Meanwhile, coded responses from free form assessments such as focus groups will be analysed using unsupervised ML approaches such as hierarchical and non-hierarchical clustering of responses/participants. Unsupervised approaches can highlight and detect previously undetected patterns in data that provide insights into the perspectives of study participants. For example, each model will have an identified central latent variable that is directly and indirectly related to measurable or observable factors (second level or third level variables) for a clinical trial.

### Outcomes and evaluation

Aligned with our goal to develop an ML predictive tool based on the lived experience of stakeholders, evaluation of the HoPeNET CAB study will be guided by an adaptation of the conceptual logic model of CBPR<sup>14 15</sup> (online supplemental figure 1). The evaluation approach includes careful consideration of community context and understanding group dynamics to build an equitable partnership that explicitly values reciprocal learning as illustrated in the model.

In addition to developing an ML tool, we anticipate that the HoPeNET CAB experience will influence investigators' behaviours and perceptions. To assess this, we will

**Table 2** Analytical approaches used in HoPeNET machine learning algorithm development

Product/goal: **creation of data-trained predictive tool for examining barriers in future trials that will inform changes in recruitment, screening and enrollment of AA/Black participants**

Analysis	Data/tools
Correlation analysis (Corrplot)	Self-reported self-efficacy/engagement between initial/terminal time points with CAB outcome assessments
t-distributed stochastic neighbour embedding	Initial and final participant self-efficacy, knowledge, engagement, barriers and attitudes
Natural language processing	Transcribed community and investigator focus group data to identify within and between group differences and similarities in attitudes, knowledge, perceptions of bias, etc.
Structural equation modelling (SEM-LAVAAN) and group-based modelling (GBM-CrimCV)	Phase 2 focus group attitudes to clinical trial participation among community and investigator group models.
Path analysis	Path analysis of phase 3 changes in perceptions on group-based model.

AA, African-American; CAB, community advisory board.

conduct post-CAB surveys and interviews. Engagement and reach in the community will be assessed using the following metrics from studies presented at CAB meetings: (1) number of participants screened who are directly from the community and (2) volume of requested recruitment materials from members of the community.

### Ethics and dissemination

Participant confidentiality and privacy will be strictly maintained and held in trust by the participating investigators, and their staff. No information concerning the study, or the data will be released to any unauthorised third party without prior written approval of the Principal Investigator. The study data entry and study management systems used by research staff will be secured and password protected. At the end of the study, all records will continue to be kept in a secure location for as long a period as dictated by the reviewing institutional review board, institutional policies or sponsor requirements. We anticipate minimal risk for this study. However, we will ask participants to express their perceptions surrounding barriers and facilitators of clinical trial participation during the one-on-one interviews and focus group activities. We recognise that this activity may elicit emotional distress. Study participation will be voluntary and interviews can be stopped at any time. To be consistent with CBPR principles and to the stated programme evaluation, study findings will be disseminated to the HoPeNET CAB, presented at departmental and institutional levels at HU and the NIH IRP. We will also present our findings at national and international conferences, and peer-reviewed manuscripts from our project will also be submitted for publication.

### Author affiliations

<sup>1</sup>Translational Biobehavioral and Health Disparities Branch, NIH Clinical Center, Bethesda, Maryland, USA

<sup>2</sup>Social Determinants of Obesity and Cardiovascular Risk Laboratory, Cardiovascular Branch, Division of Intramural Research, NHLBI, Bethesda, Maryland, USA

<sup>3</sup>Intramural Research Program, NIMHD, Bethesda, Maryland, USA

<sup>4</sup>Department of Pediatrics, Howard University, Washington, DC, USA

<sup>5</sup>Department of Health Sciences and Management, College of Nursing and Allied Health Sciences, Howard University, Washington, DC, USA

<sup>6</sup>Department of Nutritional Sciences, College of Nursing and Allied Health Sciences, Howard University, Washington, DC, USA

**Acknowledgements** The authors would like to acknowledge members of the D.C. Cardiovascular Health and Obesity Collaborative (CHOC) Community Advisory Board for their commitment to community-based research.

**Contributors** NF, FOB, EO-C, FW, LJ, VMM, BC, LG, GRW, TP-W, AJ conceptualised the protocol and manuscript; NF, FOB, EO-C, FW, LJ, VMM, BC, LG, GRW, TP-W and AJ developed the methodology for the protocol; TP-W and AJ are the protocol and manuscript supervisors; NF, FOB and EO-C were involved in writing the original draft; NF, FOB, EO-C, FW, LJ, VMM, BC, LG, GRW, TP-W and AJ were involved in revising and editing the manuscript.

**Funding** This work was supported by Genentech Health Equity Fund Award, Grant number G-89258. The coauthors, NF, FOB, FW, EO-C, VMM, BC, GRW, TP-W, are supported by intramural funding from the National Institutes of Health Clinical Centre, the National Heart, Lung and Blood Institute and the National Institute on Minority Health and Health Disparities. This research was made possible through the NIH Medical Research Scholars Program, a public-private partnership supported jointly by the NIH and generous contributions to the Foundation for the NIH from the Doris Duke Charitable Foundation, Genentech, the American Association for Dental Research, the Colgate-Palmolive Company, Elsevier, alumni of student research programs, and other individual supporters via contributions to the Foundation for the National Institutes of Health.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by the Institutional Review Board, Howard University, Washington D.C., IRB approval number IRB-2021-0071, approval received on 6 August 2021. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; internally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

**Author note** The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institute on Minority Health and Health Disparities; the National Institutes of Health; or the U.S. Department of Health and Human Services.

## REFERENCES

- 1 Toh TS, Dondelinger F, Wang D. Looking beyond the hype: applied AI and machine learning in translational medicine. *EBioMedicine* 2019;47:607–15.
- 2 Zeng K, Pan Z, Xu Y, *et al.* An ensemble learning strategy for eligibility criteria text classification for clinical trial recruitment: algorithm development and validation. *JMIR Med Inform* 2020;8:e17832.
- 3 Ni Y, Beck AF, Taylor R, *et al.* Will they participate? Predicting patients' response to clinical trial invitations in a pediatric emergency department. *J Am Med Inform Assoc* 2016;23:671–80.
- 4 Ni Y, Kennebeck S, Dexheimer JW, *et al.* Automated clinical trial eligibility prescreening: increasing the efficiency of patient identification for clinical trials in the emergency department. *J Am Med Inform Assoc* 2015;22:166–78.
- 5 Rajkomar A, Hardt M, Howell MD, *et al.* Ensuring fairness in machine learning to advance health equity. *Ann Intern Med* 2018;169:866–72.
- 6 Pfohl SR, Foryciarz A, Shah NH. An empirical characterization of fair machine learning for clinical risk prediction. *J Biomed Inform* 2021;113:103621.
- 7 Claudel SE, Ceasar JN, Andrews MR, *et al.* Time to listen: a mixed-method study examining community-based views of mobile technology for interventions to promote physical activity. *BMJ Health Care Inform* 2020;27:e100140.
- 8 Trauth JM, Musa D, Siminoff L, *et al.* Public attitudes regarding willingness to participate in medical research studies. *J Health Soc Policy* 2000;12:23–43.
- 9 Schulz AJ, Parker EA, Israel BA, *et al.* Addressing social determinants of health through community-based participatory research: the East side village health worker partnership. *Health Educ Behav* 2002;29:326–41.
- 10 Robinson JM, Trochim WMK. An examination of community members', researchers' and health professionals' perceptions of barriers to minority participation in medical research: an application of concept mapping. *Ethn Health* 2007;12:521–39.
- 11 Yingling LR, Mitchell V, Ayers CR, *et al.* Adherence with physical activity monitoring wearable devices in a community-based population: observations from the Washington, D.C., cardiovascular health and needs assessment. *Transl Behav Med* 2017;7:719–30.
- 12 Williams F, Colditz GA, Hovmand P, *et al.* Combining community-engaged research with group model building to address racial disparities in breast cancer mortality and treatment. *J Health Dispar Res Pract* 2018;11:160–78. Article 11.
- 13 Vennix JAM. Group model-building: tackling messy problems. *Syst Dyn Rev* 1999;15:379–401.
- 14 Goodman MS, Sanders Thompson VL, Johnson CA, *et al.* Evaluating community engagement in research: quantitative measure development. *J Community Psychol* 2017;45:17–32.
- 15 Sandoval JA, Lucero J, Oetzel J, *et al.* Process and outcome constructs for evaluating community-based participatory research projects: a matrix of existing measures. *Health Educ Res* 2012;27:680–90.