IRB STUDY PROTOCOL

Developing a Childhood Asthma Risk Passive Digital Marker

Protocol Version: 1.0
Protocol Date: August 1, 2022
IRB Number: 15873
ClinicalTrials.gov Identifier: NCT05826561

Principal Investigator	Arthur H. Owora, MPH, PhD
Department/Unit	Division of Pediatric Pulmonology, Allergy/Immunology and Sleep Medicine, Department of Pediatrics, Indiana University School of Medicine
Institution	Indiana University School of Medicine / Riley Hospital for Children
Funding	NIH K01 career development award, K01HL166436
Study Type	Pilot randomized clinical trial using a Solomon 4-group design
Population	Practicing pediatric clinicians; deidentified pediatric EHR-derived vignettes

1. Study Summary / Abstract

This pilot randomized clinical trial evaluates the perceived acceptability, perceived usability, and feasibility of an electronic health record (EHR)-integrated machine learning-based Passive Digital Marker (PDM) for identifying preschool-aged children at high or low risk for persistent school-age asthma using routinely collected clinical data [1]. Clinicians will evaluate standardized pediatric patient vignettes derived from longitudinal EHR data from birth through age 3 years.

The study uses a Solomon 4-group randomized design to assess study procedures and PDM implementation outcomes while evaluating potential pretest exposure effects. The protocol is informed by reporting and protocol guidance for artificial intelligence-based clinical interventions [2,3]. The primary outcomes are perceived PDM acceptability, perceived PDM usability, and study feasibility. The secondary outcome is clinician prognostic accuracy.

2. Background and Rationale

Early identification of children at risk for persistent asthma is clinically challenging because preschool respiratory symptoms are heterogeneous and frequently overlap with transient viral wheezing illnesses. Existing clinical prediction tools provide structured risk stratification but have had modest predictive performance and limited workflow integration [5]. Longitudinal EHR data provide an opportunity to synthesize early-life diagnoses, symptoms, medication prescriptions, allergic comorbidities, respiratory infections, and family history into a decision-support marker [1].

The PDM is an EHR-integrated machine learning risk marker designed to classify preschool-aged children as high or low risk for persistent school-age asthma using routinely collected structured clinical data [1]. This protocol evaluates the clinical decision-making impact, usability, acceptability, and feasibility of presenting the PDM risk classification to pediatric clinicians in a controlled vignette-based setting, consistent with recommendations that AI-based clinical interventions evaluate both performance and implementation-related outcomes [2,3].

3. Study Objectives and Endpoints

Primary Objective

To evaluate perceived PDM acceptability, perceived PDM usability, and study feasibility among pediatric clinicians participating in a controlled vignette-based assessment.

Secondary Objectives

1. To estimate clinician prognostic accuracy when using the PDM compared with standard assessment without decision support.

2. To estimate accuracy for case and control vignettes and explore whether PDM-supported assessment improves identification of children who later develop asthma.

3. To estimate effect sizes and variance components to inform a future fully powered pragmatic randomized trial.

Primary Endpoints

1. Perceived PDM acceptability, measured using a behavioral intention/acceptability scale assessing clinicians' likelihood of using the tool if integrated into routine EHR workflows [6].
2. Perceived PDM usability, measured using a modified clinical decision support usability instrument adapted from the Healthcare Systems Usability Scale [4].
3. Study feasibility, measured using recruitment completion within the planned accrual period, completion rate for vignette assessments, completion rate for post-assessment questionnaires, and other predefined completion metrics.

Secondary Endpoint

Clinician prognostic accuracy, measured as the proportion of correctly classified vignettes out of total vignettes evaluated. Descriptive analyses may also examine accuracy separately for case and control vignettes, reflecting the exploratory decision-support evaluation described for the PDM [1].

4. Study Design

This is a pilot randomized clinical trial using a Solomon 4-group design. Eligible clinicians will be randomized to one of four groups: (1) standard assessment with pretest and posttest, (2) standard assessment posttest only, (3) PDM-supported assessment with pretest and posttest, and (4) PDM-supported assessment posttest only.

Pre-post PDM intervention: Clinicians randomized to this arm complete an initial set of asthma prognostic vignette assessments without PDM support, then receive access to the Passive Digital Marker risk classification, and subsequently complete posttest vignette assessments with PDM support. This arm permits evaluation of within-clinician change in prognostic accuracy before versus after exposure to the PDM.

Pre-post SAU: Clinicians randomized to this arm complete an initial set of asthma prognostic vignette assessments without PDM support and then complete posttest vignette assessments, again without PDM support. This arm serves as the standard-assessment comparison condition and permits evaluation of changes in prognostic accuracy over time that may occur because of repeated testing, familiarity with the vignette format, or other non-PDM effects.

The Solomon design allows estimation of the PDM intervention effect while assessing whether pretest exposure influences subsequent performance. The protocol is aligned with Artificial Intelligence (AI) clinical trial reporting and protocol guidance [2,3].

Clinicians cannot be blinded to whether the PDM risk classification is provided. Clinicians will be blinded to the true vignette case/control outcome status, namely whether the child represented in each vignette later had documented asthma diagnosis between ages 6 and 11 years.

5. Study Setting

The study will recruit practicing pediatric clinicians from outpatient pediatric care settings in Indiana. Vignette assessments and study questionnaires may be completed electronically or in a controlled study environment, according to approved study procedures.

6. Study Population

6.1 Clinician Participants

Participants will be practicing pediatricians or pediatric clinicians who provide outpatient pediatric care. The planned enrollment is 100 pediatricians, with 25 participants assigned to each of the four Solomon design study arms. Each clinician will complete 10 vignette assessments. No restrictions are applied based on years of practice, practice setting, sex, race, ethnicity, or clinician demographic characteristics unless required by local credentialing or eligibility verification.

6.2 Pediatric Vignette Data

The pediatric cases are represented only as standardized vignettes derived from authentic longitudinal EHR data spanning birth through age 3 years. Five vignettes represent children with documented asthma diagnoses between ages 6 and 11 years, and five represent children without such diagnoses. Use of longitudinal EHR-derived data for passive childhood asthma risk stratification is based on prior PDM development and validation work [1].

7. Eligibility Criteria

Inclusion Criteria for Clinician Participants

1. Board-certified, board-eligible, or otherwise credentialed pediatric clinician providing outpatient pediatric care.

2. Willing and able to complete informed consent.

3. Willing and able to complete vignette assessments and study questionnaires.

4. Able to complete study procedures in English.

Exclusion Criteria for Clinician Participants

1. Unable or unwilling to provide informed consent.

2. Unable to complete vignette assessments or study questionnaires.

3. Prior involvement that would compromise blinding to vignette case/control status or substantially bias study assessments, as determined by the study team.

8. Recruitment and Consent Procedures

Potential clinician participants will be recruited through professional networks, pediatric clinics, department communications, email invitations, and/or direct outreach approved by the IRB. Recruitment materials will describe the study purpose, procedures, time commitment, voluntary nature of participation, confidentiality protections, and contact information for study staff.

Clinicians who express interest will be screened for eligibility. Eligible clinicians will complete written or electronic informed consent before any study-specific procedures. The consent process will emphasize that participation is voluntary, refusal will not affect employment or professional standing, participants may withdraw at any time, and responses will be analyzed in aggregate whenever possible.

9. Randomization and Allocation

Participants will be randomized using a computer-generated allocation sequence within the Solomon 4-group framework, with planned allocation of 25 pediatricians per study arm: (1) standard assessment with pretest and posttest, (2) standard assessment posttest only, (3) PDM-supported assessment with pretest and posttest, and (4) PDM-supported assessment posttest only. Allocation records will be maintained securely by the study team. Because the intervention involves presenting decision-support output, participants cannot be blinded to intervention exposure; however, the true future asthma outcome status of each vignette will not be disclosed during assessment.

10. Intervention and Comparator

10.1 Intervention: Passive Digital Marker-Supported Assessment

The PDM is an EHR-integrated machine learning algorithm that generates a binary asthma risk classification, high risk or low risk for persistent school-age asthma, using routinely collected early-life EHR variables [1]. Clinicians assigned to the PDM condition will review the same standardized vignette information as standard-assessment clinicians and will also see the PDM risk classification when making prognostic judgments.

10.2 Comparator: Standard Assessment Without Decision Support

Clinicians assigned to standard assessment will review standardized pediatric vignettes and make prognostic judgments based only on the clinical information presented in the vignette, without PDM output.

11. Study Procedures

1. Confirm eligibility and obtain informed consent.

2. Collect baseline clinician demographic and professional characteristics as approved by the IRB.

3. Randomize participants to one of the Solomon design groups.

4. For groups assigned to pretest, have clinicians complete vignette prognostic assessments before the intervention/posttest phase.

5. Present 10 standardized pediatric vignettes in randomized order. Each vignette summarizes early-life clinical information, including respiratory symptoms, allergic comorbidities, medication use, respiratory infections, and family history, consistent with variables used in prior EHR-based asthma risk marker development [1].

6. Ask clinicians to estimate whether each child is likely to develop asthma by school age.

7. For PDM groups, provide the PDM high/low risk classification during the assigned PDM-supported assessment phase.

8. Collect post-assessment usability and behavioral intention measures from clinicians exposed to the PDM, including usability items adapted from the Healthcare Systems Usability Scale for clinical decision support systems [4].

9. Record feasibility metrics, including recruitment duration, completion of vignette assessments, and completion of questionnaires.

12. Data Elements Collected

Clinician-level data may include age category, sex, race, ethnicity, years in practice, specialty, practice setting, randomization group, vignette responses, usability ratings, behavioral intention ratings, and completion indicators. Vignette-level data include deidentified early-life clinical features summarized from EHR records, PDM risk classification, true case/control outcome status, clinician prediction, and whether the prediction was correct.

13. Risks and Discomforts

This study is expected to pose minimal risk to clinician participants. Potential risks include loss of confidentiality regarding clinician responses or demographic information, discomfort related to performance evaluation, and time burden. The study is not designed to evaluate individual clinician competency, and results will be reported in aggregate. Clinician identifiers will be stored separately from analytic data whenever feasible.

Use of deidentified pediatric vignettes minimizes risk to pediatric patients. The study does not involve direct patient contact, clinical intervention, change in patient care, or disclosure of patient identity to clinician participants.

14. Potential Benefits

Clinician participants may not receive direct personal benefit. Potential societal and scientific benefits include improved understanding of the perceived usability, acceptability, and feasibility of machine learning-enabled EHR decision support, along with exploratory evidence on clinician prognostic accuracy. Findings may inform future pragmatic trials and eventual improvements in early identification of children at risk for persistent asthma.

15. Privacy, Confidentiality, and HIPAA Considerations

All study data will be maintained according to institutional policies and IRB requirements. Clinician identifiers will be limited to information necessary for recruitment, consent, scheduling, and study administration. Analytic datasets will use study IDs rather than direct identifiers whenever feasible.

Pediatric EHR-derived vignettes will be deidentified or coded before presentation to clinician participants. Vignettes will not include names, medical record numbers, exact dates of birth, addresses, telephone numbers, or other direct identifiers. Any dates or ages included in vignettes will be generalized to the extent compatible with the study aims. Access to source EHR data and linkage files, if any, will be restricted to authorized study personnel.

If the local IRB determines that protected health information is used to create vignettes, the study team will obtain required HIPAA authorization, waiver of authorization, or preparatory-to-research/data-use approvals as applicable. Data handling will comply with institutional HIPAA and information-security standards.

16. Data Management and Security

Study data will be stored on secure, access-controlled institutional systems. Access will be limited to approved study personnel. Data transfers will use institutionally approved secure methods. Study datasets exported for analysis will be coded and will exclude direct identifiers to the extent possible. Consent documents and linkage files, if any, will be stored separately from analytic datasets. Data will be retained according to institutional policy, sponsor requirements, and IRB approval, then destroyed or archived securely.

17. Sample Size and Feasibility Rationale

This pilot study is designed to estimate perceived acceptability, perceived usability, feasibility metrics, intervention effect sizes, and variance components rather than provide definitive efficacy evidence. The planned sample is 100 pediatricians, allocated evenly across the four Solomon design study arms at 25 participants per arm. With each clinician completing 10 vignette assessments, the study is expected to yield up to 1,000 vignette assessments to support feasibility assessment and pilot estimation for future pragmatic evaluation of the PDM [1].

18. Statistical Analysis Plan

Descriptive statistics will summarize clinician characteristics, primary outcomes (perceived PDM acceptability, perceived PDM usability, and study feasibility), and the secondary outcome of prognostic accuracy. Continuous measures will be summarized using means, standard deviations, medians, interquartile ranges, and/or 95% confidence intervals, as appropriate. Categorical measures will be summarized using counts and percentages.

Primary analyses will focus on perceived PDM acceptability [6], perceived PDM usability, and study feasibility metrics, including recruitment completion, vignette assessment completion, and questionnaire completion. Secondary analyses will evaluate prognostic accuracy using the Solomon design factorial framework, including fixed effects for intervention group, assessment timing, and the interaction between intervention and timing. Because clinicians contribute multiple vignette assessments, mixed-effects models will include random intercepts for clinician to account for clustering. Regression coefficients with 95% confidence intervals will be reported. Accuracy for case and control vignettes may be examined separately using appropriate group comparisons, such as Welch-adjusted t tests or model-based contrasts. Two-sided P values less than .05 will be considered statistically significant for pilot inference, with emphasis on estimation and confidence intervals. This analysis approach is consistent with pilot evaluation of an Artificial intelligence (AI)-supported clinical decision intervention and related AI trial reporting considerations [2,3].

19. Data and Safety Monitoring

Given the minimal-risk nature of the study, a formal Data and Safety Monitoring Board is not anticipated unless required by the IRB or sponsor. The principal investigator will oversee study conduct, protocol adherence, data quality, confidentiality protections, and reporting of unanticipated problems. Any unanticipated problems involving risks to participants or others will be reported according to institutional IRB policy.

20. Compensation

Clinician participants will receive a $50 gift card as compensation for their time and effort after completing the study procedures, including vignette assessments and required study questionnaires. Compensation will not be prorated unless otherwise required by institutional policy or the IRB-approved consent form.

21. Vulnerable Populations

The direct research participants are adult clinicians. Pediatric individuals are represented only through deidentified or coded EHR-derived vignettes. No children are directly recruited, consented, contacted, randomized, or assigned to clinical interventions. No prisoners, pregnant individuals as a target population, or cognitively impaired adults are specifically recruited.

22. Dissemination of Results

Results may be disseminated through scientific conferences, peer-reviewed publications, reports to the funder, clinical informatics and pediatric research forums, and future grant applications. Publications and presentations will not identify individual clinician participants or pediatric patients. Deidentified study data and analytic code may be made available upon reasonable request to the corresponding investigator, subject to IRB, institutional, and data-use restrictions.

23. References

1. Owora AH, Jiang B, Shah Y, Gaston B, Boustani M. External validation and update of the pediatric asthma risk score as a passive digital marker for childhood asthma using integrated electronic health records. EClinicalMedicine. 2025;84:103254.

2. Liu X, Rivera SC, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-1374.

3. Rivera SC, Liu X, Chan AW, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. BMJ. 2020;370:m3210.

4. Ghorayeb A, Darbyshire JL, Wronikowska MW, Watkinson PJ. Design and validation of a new Healthcare Systems Usability Scale for clinical decision support systems. BMJ Open. 2023;13(1):e065323.

5. Kothalawala DM, Kadalayil L, Weiss VBN, et al. Prediction models for childhood asthma: a systematic review. Pediatr Allergy Immunol. 2020;31(6):616-627.

6. Kaczmarek LD, Kashdan TB, Drazkowski D, et al. Why do people prefer gratitude journaling over gratitude letters? The influence of individual differences in motivation and personality on web-based interventions. Pers Individ Dif. 2015;75:1-6.

Appendix A. Schedule of Study Activities

Activity	Screening/Consent	Pretest if assigned	Assessment/Posttest	Post-assessment
Eligibility confirmation	X			
Informed consent	X			
Clinician demographics	X			
Randomization	X			
Vignette assessment without PDM		X	X for SAU groups	
Vignette assessment with PDM			X for PDM groups	
Usability questionnaire				X for PDM-exposed clinicians
Behavioral intention questionnaire				X for PDM-exposed clinicians
Feasibility tracking	X	X	X	X