Advancing the Collection and Use of Patient-Reported Outcomes Through Health Information Technology

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
AHRQ Task Order: HHSP23337004T
Contract Number: HHSP233201500022I
Contracting Officer's Representative: Chun-Ju (Janey) Hsiao, PhD, MHS

Prepared by:

MedStar Health
Deliya B. Wesley, MPH PhD (PI)
Raj M. Ratwani, PhD (Co-PI)
Kathryn M. Kellogg, MD, MPH (Clinical Advisor)
Robin Littlejohn, MS
Zoe Pruitt, MS
Ram Dixit, MS
Laura C. Schubel
Joseph Blumenthal
Shrenikkumar Shah
Shrey Mathur

AHRQ Publication No. 20-0044
June 2020

This project was funded under contract number**:** HHSP233201500022I from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services. The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of AHRQ. Readers should not interpret any statement in this product as an official position of AHRQ or of the U.S. Department of Health and Human Services. None of the authors has any affiliation or financial involvement that conflicts with the material presented in this product.

Public Availability Notice. This product is made publicly available by AHRQ and may be used and reprinted without permission in the United States for noncommercial purposes, unless materials are clearly noted as copyrighted in the document. No one may reproduce copyrighted materials without the permission of the copyright holders. Users outside the United States must get permission from AHRQ to reprint or translate this product. Anyone wanting to reproduce this product for sale must contact AHRQ for permission.

Suggested Citation

Advancing the Collection and Use of Patient-Reported Outcomes Through Health Information Technology. Prepared under Contract No.**:** HHSP233201500022I. Rockville, MD: Agency for Healthcare Research and Quality; Pub. No. 20-0044. June 2020.

Table of Contents

Advancing the Collection and Use of Patient-Reported Outcomes Through Health Information Technology i

Executive Summary Overview 1

Stakeholder Input 1

App Development and Pilot Tests 2

Findings and Recommendations 2

Key Lessons Learned and Summary Recommendations from Pilot Test 1 (OBERD App): 2

Key Lessons Learned and Summary Recommendations from Pilot Test 2 (PRISM App): 3

Conclusion 5

Introduction to the Work 6

Current State and Challenges with Collecting Patient-Reported Outcomes 7

Contract Goals and Objectives 7

Our Human Factors and Sociotechnical Systems Approach 7

Document Organization 8

Chapter I: Stakeholder Feedback 9

Introduction 10

Methods 10

Key Findings and Summary Recommendations 11

Chapter II: Pilot 1: Modification and Implementation of an Existing Healthcare PRO App 13

Overview 14

Section I: Modification of Existing App 14

Original OBERD App: Technical Integration 14

Modification of Existing OBERD App 14

Communication Standards of Patient PRO Measures 15

Implementation Modification of PROMIS Survey Distributor 15

Security Model Modifications 16

OBERD App Modification: Challenges and Lessons learned 17

Section II: Integration with Different Electronic Health Records 18

Integration of the OBERD App 19

Iterative Development of Provider Visualization 20

Integration of Provider-Facing Visualization: Challenges and Lessons Learned 20

Technical Implementation: Perspective and High-Level Recommendations 21

Section III: Implementation 22

Development of Training Materials 23

Clinical Pilot Site Selection and Recruitment 23
Evaluation and Metrics 23
Data Analysis 25
Section IV: Clinical Implementation Findings 25
Practice Site Assessments 25
Patient-Level Assessments 27
Provider-Level Assessments 30
Health IT Staff-Level Assessments 32
Section V: Overall Lessons Learned: Clinical Implementation and App Usage 33
Recruitment and Patient Enrollment Methods Matter 33
Workflow and Staff Communication Can Be Just as Important as the Technology 34
Certain Practice Characteristics May Be More Predictive of Success Than Others 34
The Content of the Survey Questions May Need More Careful Consideration Than the App 35
Chapter III: Integration and Implementation of a New App 36
Overview 37
Section I: The PRISM App 37
Background for Challenge Competition and the Winning App 37
PRISM App Technical Integration 37
PRISM Usability Testing and Development Process 38
PRISM App Architecture 39
FHIR Standards and EHR Integration 39
Section II: PRISM App Integration with Different Electronic Health Records 40
Scalable Architecture for EHR Integration 40
Integration of the PRISM App 41
Initial MedStar Health Systems-Level Approvals and Processes 41
Section III: Pilot Site Implementation Processes 42
Recruitment and Selection of Clinical Pilot Sites 42
Assignment of Participating Practices: Tiers of Implementation Support 42
Development of Patient, Staff, and Provider Training Materials 43
Section IV: Data Collection Procedures and Tools 44
Evaluation Metrics and Data Analysis 44
Section V: Technical Implementation Findings 44
PRISM App Implementation: Challenges and Lessons Learned 44
PRISM Integration with the Hub 45
Integration of the Provider-Facing Visualization: Challenges and Lessons Learned 47

Technical Implementation: Developer Perspective and High-Level Recommendations 47
Technical Challenges for Patients Using the PRISM App: Bugs and Issues with the Mobile Platforms 48
Section VI: Clinical Implementation Findings 49
Practice Site Assessments 49
Site Characteristics Within Tiers 52
Patient-Level Assessments 53
Provider-Level Assessments 55
Health IT Staff-Level Assessments 56
Section VII: Summary of Key Findings by Stakeholder Group: Clinical Implementation and App Usage 56
Patients Using the PRO App 56
Primary and Specialty Care Providers Using the PRO Data 57
Developers and Health IT Staff 58
Primary or Specialty Care Practices 58
Institutions or Health Systems Implementing PRO Technology 59
Chapter IV: Discussion, Crosscutting Themes and Conclusions 60
Discussion 61
Crosscutting Themes Across Pilot Test 1 (OBERD) and Pilot Test 2 (PRISM) 63
Conclusion 65
Limitations 66
References 67

Executive Summary Overview

Problem Statement: Despite the development of patient-reported outcome (PRO) measures, their widespread clinical adoption has not been attained. Technical tools, including software applications (apps), can serve to improve the collection and analysis of PRO data; however, this technology must be usable by patients, providers, and researchers.[1] Further, there must be clear standards driving the development of PRO apps, and the integration of these apps with electronic health records (EHRs), so that they can be consistently implemented in a seamless fashion. However, currently there are no such widely used standards for PRO apps, resulting in challenges for providers, patients, and technical teams.

Goal: This project was driven by three overarching goals. First, to gather and incorporate input from key stakeholders to guide development, implementation, and use of a PRO app. Second, to assess the Office of the National Coordinator for Health Information Technology's (ONC's) PRO Fast Healthcare Interoperability Resources® (FHIR) technical specifications for PRO app development and implementation, by modifying an existing app. Third, to integrate and pilot test the modified app and a new app developed using the FHIR technical specifications.

Stakeholder Input

Methods: Interviews included demographic questions followed by focused knowledge elicitation questions to inform the development and use of a general PRO app from different perspectives. Patients were asked questions related to the relevance of a PRO app for personal use, the usability and functionality of an app, and potential concerns about privacy. Providers were asked questions about the relevance of PRO data to inform clinical practice, and the design and integration needs for PRO data to fit into their workflow to ensure data were useful and accessible. Health information technology (IT) stakeholders were asked about prior integration and implementation experience with PRO apps and standards, as well as specific needs for implementation in different healthcare settings. **Participants** included nine patients, nine providers, and nine health IT professionals from MedStar Health and our partner sites. **Recruitment:** Participants were recruited from MedStar Health practices and from our partner sites in the Capital Area Primary Care Research Network (CAPRICORN), a practice-based research network (PBRN). Participants were identified using purposive recruitment from a convenience sample with an effort to enhance diversity in perspectives on PRO apps.

Findings by stakeholder group:

Patients: Patients were generally only interested in using a physical function (PF) PRO app if they had an active health issue to track. Patients desired optimal usability and bidirectional communication in an app with transparency about privacy. **Providers:** PRO data would be most useful, relevant, and clinically significant for providers if key patient populations were targeted (e.g., Medicare patients). Providers expressed a need for accurate data that are seamlessly integrated into the EHR and intuitive to access, underscoring that these data should be easy to visualize and interpret. **Health IT Professionals:** Participants confirmed that ideal PRO data integration within the EHR would conform to open standards, enabling easy sharing of data across different institutions. Critical to success were leadership buy-in, coordination across different IT teams, and support from EHR vendors.

App Development and Pilot Tests

Methods: Guided by a sociotechnical systems (STS) approach, we used rigorous mixed methods to evaluate the facilitators and barriers to app implementation and use in diverse ambulatory settings. For **Pilot Test 1** the focus was the Outcomes Based Electronic Research Database (OBERD) app, a web-based app that was used in several MedStar Health orthopedic practices to collect PRO PF data. We applied the ONC PRO FHIR Implementation Guide (IG) to modify the OBERD app, enabling data to be integrated with the EHR. In **Pilot Test 2** the focus was the **P**ROMIS (Patient-Reporting Outcomes Measurement Information System) **R**eporting and **I**nsight **S**ystem from **M**innesota (PRISM™) app, the winner from AHRQ's Step Up App Challenge, aimed to enhance the quality of clinical discussion between healthcare providers and patients by allowing for continued patient engagement outside of the clinical setting. The app supports any PRO instrument and allows seamless data integration with EHRs via FHIR.

Each app was pilot tested in nine primary and specialty care practices across MedStar and CAPRICORN networks, with 10 patients targeted to enroll and use the app at each site. Participating sites were of various size, geographic locations, with varied workflows, and using three different EHRs. For Pilot Test 2, sites were further allocated to three different levels of technical assistance by the research team. Three sites functioned independently following training, three sites functioned with partial technical assistance, and the three remaining sites functioned with full technical assistance. While both pilots collected patient-reported PF data, Pilot Test 1 patients entered PF data into an app on a tablet provided by the research team. Pilot Test 2 patients used their personal smartphones (iOS or Android platform) to enter the PF data. Both pilots used the same evaluation methods. Practice-level assessments included readiness for change and clinic workflow observations. Patient-level assessments focused on usability, functionality, and general preferences using a structured survey and semi-structured interviews. Provider and health IT staff assessments focused on app usability, functionality, and general lessons learned from the integration and implementation. All qualitative data were analyzed using a grounded theory approach, whereby two researchers coded and analyzed the interview data. Survey data were analyzed using standard statistical software.

Findings and Recommendations

Key Lessons Learned and Summary Recommendations from Pilot Test 1 (OBERD App):

App modification and integration: We successfully applied the FHIR IG to adapt the OBERD app and integrate it with three different EHRs. Key drivers of this success included the technical guidance from the IG that clearly translated technical specifications for developers. In addition, the ability to scale Substitutable Medical Applications, Reusable Technologies (SMART) on FHIR across EHR systems (including the External Assessment Center [EAC], which made the PRO technical solution more modular). Several critical lessons were learned during the app modification and integration process. Key recommendations for future developers implementing the standards include:

1. Building an EAC to provide adaptive content is technically challenging and resource intensive, thus implementing an existing EAC is recommended.

2. Achieving "light integration" via external hosting (e.g., hosting a web app in the cloud) avoids many integration challenges including EHR variability.

3. Even when integrating with the same EHR at different sites, engaging local IT support as early as possible in the process is critical for identifying and mitigating implementation challenges and delays; local context is critical.

Practices: Based on our experience with nine primary and specialty care practices, we identified several barriers and facilitators for ambulatory sites wishing to implement similar FHIR based PRO apps to capture PF data. Highly engaged practices with providers and staff who were open to the implementation of new health IT to improve care delivery generally experienced the fewest challenges. Uptake and use of the modified app was most successful in practices with: (1) dedicated staff members to facilitate the implementation of the new technology (i.e., these sites were better able to integrate the data collection into their clinical workflow without it seeming an overly burdensome process), 2) team cohesion and consistent staffing (i.e., where the provider had a dedicated medical assistant and there was little/minimal rotation in front desk staff from day to day), and 3) efficient methods for patient screening and communication with eligible patients to ensure successful patient enrollment and data collection.

Patients: Overall the app was successfully used by 90 patients to collect their PF data, which was available in the EHR in real time during their provider encounter. Patients generally reported that the app was easy to use and that the survey could be completed quickly. We noted a few challenges: 1) touch-responsiveness of the tablet did not always function as intended for the patient population under study, 2) many patients struggled with the relevance and context of questions in both the PF survey and the usability survey, 3) many patients did not perceive an impact on communication with their provider based on availability of their PF scores.

Providers: Nine providers interviewed generally found the visualization of the scores easy to understand and appreciated the layout and design of the PRO score as presented in the EHR. The most notable challenges expressed included: 1) the number of clicks required, or workaround to access the score within the EHR; 2) inconsistent communication and/or unreliable notification when a patient's PRO score was available in the EHR; 3) provider's uncertainty about the validity of the scores for certain patients (e.g., when scores indicated much lower or greater PF than what provider knew was their patient's reality). Most notably, however, providers stated despite believing the PF data were important, they did not always find clinical utility in having the data available, except for the specialty providers who noted in at least two instances that they were alerted to an issue due to the PF score and patient's response, which resulted in a critical referral for a patient.

Key Lessons Learned and Summary Recommendations from Pilot Test 2 (PRISM App):

Key Findings by Stakeholder:

App Developers and Technical Integration Teams: The FHIR standards allowed for the PRISM app to quickly and easily switch from one FHIR server to another and enabled a successful abstraction layer to bridge the mobile app and two different EHRs. Key considerations for future developers for this type of technical implementation include:

1. Setting up a new Amazon Web Services (AWS) cloud instance of the app can be time consuming, and support for the app under AWS can be time and resource intensive.

2. Regardless of the size of the implementation, dedicated technical staff are needed to customize and troubleshoot app integration to ensure successful adoption for different clinic workflows.

3. Any apps that use SMART on FHIR technology need to also support access from alternative devices that do not rely on app downloads (e.g., web-based apps or embedded apps on clinic tablets).

Practices: Effective teamwork is one of the most important facilitators of implementation. Efficient and consistent communication between the different levels of clinic staff is critical for successful adoption of this technology and integration into each clinic's workflow. Key considerations with implications for scaling this include:

1. Practices looking to adopt and implement this PRO technology would benefit from having dedicated staff to facilitate patient use and adoption, and a clear delineation of roles and responsibilities for support staff.

2. Practices with a high patient volume may not have enough resources or staff to engage and provide patients with technical assistance.

3. The availability of Wi-Fi and cellular service are critical to adoption and sustained use of this technology. This can be particularly challenging for clinics in rural areas with poor cellular service, or individual practices with specific architectural or physical constraints (e.g., lead walls installed for protection from nuclear medicine).

Primary and Specialty Care Patients: In general, patients found the PRISM app to be usable and easy to navigate for completion of the PROMIS survey. The app facilitated efficient survey completion (74 seconds on average), highlighting the potential for this technology to be scaled for a large battery of clinically relevant PROs. Key findings with implications for scale include:

1. Successful engagement with the app in the pilot test setting was often dependent on the patient generally being in good health, having the appropriate smartphone, and having the infrastructure necessary to make the technology work (i.e., a reliable cellular network or strong Wi-Fi connection).

2. Password issues when registering an account or downloading from the App Store or Google Play were a common challenge for many patients attempting to engage with the app.

3. Security and privacy concerns remain paramount for many patients and would likely impact their willingness to continuously use this type of technology to collect PROs.

Primary and Specialty Care Providers: Providers play a key role in patients' willingness to use PRO apps to collect PF data, as patients were more willing to use the app if they knew the provider was looking at their data. We found that surfacing PRO data via a dynamic template in the EHR proved to be a viable solution, allowing providers to easily access and consume the PRO data in the EHR. However, some providers still expressed hesitation about interpreting the data because it ultimately represented a patient's perception, rather than an objective physiological measure. Key findings with implications for scale include:

1. Many primary and specialty care provider workflows are not conducive to accessing the patient's PRO data in the EHR prior to the patient encounter (e.g., providers who do not review patient charts until after the encounter, or where patient volume limits this option). Alternate workflows need to be considered.

2. Provider notification or an EHR alert of a patient's survey completion is needed so as not to hinder workflow and to prevent providers from missing the data.

3. Providers need flexibility incorporating the PRO data into their EHR workflows. For example, the option to copy and paste the data into notes if they want to review the data at another time after the patient encounter would be helpful.

Health Systems and Institutions: While SMART on FHIR technology is generally relatively easy to implement, many health systems do not currently have the tools or technology resources on hand to support it. Further, several EHRs do not currently support the latest SMART on FHIR standards, and some clinic sites may not be willing to invest in these apps. Key recommendations with implications for scale include:

1. Systems should anticipate significant challenges with using cloud-based services, including the need for security assessments, HIPAA, business associate agreements and related internal processes, such as gaining approval from IT, and internal app privacy and security assessments.

2. Institutions and health systems with complex and uniform vetting processes for adopting any new technology may benefit from standardized, expedited approval processes for integrating novel patient-facing health IT.

Involving all system stakeholders, including legal and compliance, as early in the implementation process as possible is critical to successful and timely adoption of this PRO technology given the multiple policies and regulatory considerations.

Conclusion

This pilot study was an important step toward testing the application of ONC's PRO FHIR Implementation Guide, and a demonstration of the various factors critical to the successful adoption, potential scaling, and sustained use of this technology in ambulatory settings. Our pilot tests successfully implemented a patient-facing PRO app by leveraging the modern standards of SMART on FHIR**.** Importantly, through this testing process we were able to demonstrate the types of settings and system factors that are most conducive to this type of technical implementation, and settings where this technology is most likely to be successfully adopted. Our pilot findings all highlight critical points at which the coordination of human and technical processes using a systems approach is crucial to ensure the successful use of PRO data. It is important to note that even a successful implementation of a PRO data collection system does not necessarily guarantee or imply long-term adoption or meaningful use of the data. Successful use of PROs is complex, tending to be context-dependent and strongly coupled to the existing relationships between patients and their providers.

Introduction to the Work

Current State and Challenges with Collecting Patient-Reported Outcomes

Patient-reported outcomes (PROs) are assessments that directly reflect the "voice of the patient," capturing the medical symptoms of the patient based on their own perception. They also provide an increasingly important avenue for patient engagement. These include measures such as physical function, pain, and health-related quality of life, without interpretation of the patient's responses by a clinician.[2,3] PROs are critical to the effective care of patients, with PRO data informing provider practice as well as supporting important research on patient outcomes.[1]

Despite the development of PRO measures, their widespread clinical adoption has not been attained.[4–6] Technical tools, including software applications (apps), can serve to improve the collection and analysis of PRO data; however, this technology must be usable by patients, providers, and researchers.[1] Further, there must be clear standards driving the development of PRO apps, and the integration of these apps with electronic health records (EHRs), so that they can be consistently implemented in a seamless fashion, and the resulting data can be easily combined for large-scale analyses. Each implementation can be unique and require extensive resources from information technology (IT) teams at each clinical site, making the current implementation of PRO apps burdensome, with technology having suboptimal usability. Few existing examples demonstrate how to best optimize the collection of PRO data from the perspective of key stakeholders including patients, providers, and health IT professionals.

Contributing to this knowledge base, the Agency for Healthcare Research and Quality (AHRQ) funded this contract, and concurrently sponsored the AHRQ Step Up App Challenge competition in 2018. The goal of this nationwide challenge contest was to develop a mobile app that leveraged digital technologies to empower patients, improve patient outcomes, and increase value in the healthcare system.

Contract Goals and Objectives

The overarching goal of this contract was threefold. First, to seek and incorporate feedback from key stakeholders to inform the optimal development of a PRO app and the approach for the pilot tests. Second, to assess the Office of the National Coordinator for Health Information Technology (ONC's) Fast Healthcare Interoperability Resources® (FHIR) technical specifications for PRO app development and implementation, and third, to assess the usability of two apps and the generated PRO data by providers and patients through two distinct pilot tests. In the first pilot test (Pilot 1), we assessed the use of the PRO FHIR Implementation Guide (IG) to modify an existing PRO app used in the MedStar Health system. We integrated and pilot tested the modified app in nine diverse primary and specialty care clinics. In the second pilot test (Pilot 2) we sought to integrate and implement a new app developed using the PRO FHIR IG within the MedStar Health system. We worked in collaboration with the AHRQ Step Up App Challenge winning team to integrate and pilot test the winning PRISM app in nine primary and specialty care settings across the MedStar Health system.

Our Human Factors and Sociotechnical Systems Approach

Recognizing that PRO apps will be part of the complex healthcare system, successful implementation and use requires seamless integration. To address these system factors, the project plan was informed by the **Sociotechnical Systems (STS)** approach for the successful implementation and use of health IT.[7] The STS model provides a multidimensional framework within which any health IT implemented within a complex adaptive healthcare system can be studied. This model integrates specific technological and measurement dimensions of health IT with other STS dimensions, including 1) hardware and software computing

infrastructure; 2) clinical content; 3) human computer interface; 4) people; 5) workflow and communication; 6) internal organizational policies, procedures, and culture; 7) external rules, regulations, and pressures; and 8) system measurement and monitoring. When applicable, the STS approach was considered to ensure a smooth, effective, and considerate approach to implementation for each pilot test.

Our research team utilized a human factors and sociotechnical systems approach to development of the stakeholder feedback instruments for the first phase of the project, and to development of the app, integration, rollout, and data collection processes for both pilot tests. This ensured continuous and rigorous evaluation through each phase.

Document Organization

This report includes an overview of the stakeholder feedback (Chapter I), and a detailed overview of each pilot test. We document the goals, methods and findings, and recommendations for each pilot separately (Chapters II and III). We then present an overall discussion, drawing on crosscutting findings from both pilot tests before presenting general conclusions (Chapter IV).

Chapter I: Stakeholder Feedback

Introduction

With a focus on a PF PRO measure, we sought to identify perceived barriers and facilitators to the successful implementation and use of PRO apps through stakeholder interviews. Further, to inform the focus and development of the planned pilot tests of two PRO apps, we interviewed participants from three stakeholder groups instrumental to successful implementation and use of these PROs. Participants were: *1) patients from whom PRO data are collected, 2) providers as the intended primary users of PRO data to inform clinical decisions, and 3) health IT professionals responsible for implementing and maintaining the technology in a healthcare setting*. We sought to understand the usability and functional requirements from the perspective of each of these stakeholders.

Interview questions were informed by the **STS model**[1] integrating specific technological and measurement dimensions of health IT with other STS dimensions. Questions were developed to obtain feedback on the following dimensions of the STS model which were relevant to each stakeholder group: *1) hardware and software computing infrastructure*; *2) clinical content; 3) human computer interface; 4) people; 5) workflow and communication; 6) internal organizational policies, procedures, and culture; 7) external rules, regulations, and pressures; and 8) system measurement and monitoring.* Not all these factors are relevant to each stakeholder group.

Methods

Table 1 outlines the purpose of the interviews for each stakeholder group.

Table 1. Stakeholder interview goals

Stakeholder	Interview Domain	Purpose
Patient	Relevance	Understand if and when patients would use PRO apps
Patient	Usability and Functionality	Understand design needs for patients entering PRO data
Patient	Privacy	Understand privacy concerns for patients using a PRO app
Provider	Relevance	Determine if and how providers would use PRO data to inform clinical practice
Provider	Functionality and Workflow	Understand design and integration needs for PRO data to fit into workflow and ensure data were useful and accessible
Health IT	Relevant Experience	Understand lessons learned and prior integration and implementation experience with PRO apps and standards
Health IT	Needs for Implementation and Integration	Identify the resources required for a successful integration
Health IT	Perceived Barriers to PRO Use	Understand potential challenges to implementation and use of PRO data at all levels

Interviews included demographic questions followed by focused knowledge elicitation questions to inform the goals described in Table 1 above. **Participants** were nine patients, nine providers, and nine health IT professionals from MedStar Health and our partner sites. **Recruitment:** Participants were recruited from MedStar Health sites and from our partner sites in the Capital Area Primary Care Research Network (CAPRICORN), a practice-based research network (PBRN). Participants were identified using purposive recruitment from a convenience sample. They were purposefully recruited to enhance diversity in perspectives on PRO apps.

Analysis: The interview responses were first transcribed and then segmented into discrete statements by question. The segments for each stakeholder group were analyzed using a grounded theory approach.[8,9] A preliminary code book was iteratively developed by having two researchers review the segments from the first three participants of each stakeholder group to identify codes that best described the segment through discussion. Once the code book was agreed upon, the two researchers applied the coding scheme to the remaining data using a consensus process. Throughout the coding process if new themes emerged, they were added to the code book, and previous responses were re-examined to determine if that new code was more appropriate. More than one code may have been applied to each segment. If the two researchers could not reach consensus during coding, a third researcher was consulted and ultimately the majority determined the category label.

Key Findings and Summary Recommendations

Patients: Patients were generally only interested in using a PF PRO app if they had an active health issue to track. Patients desired optimal usability and bidirectional communication in an app with transparency about privacy.

- Design an app that can be tailored to different user groups by adapting content for and testing with different patient populations; focus on functionality, customizable settings, and interpretation of physical function scores.

- Provide assurances and include clear messaging related to data security and data usage.

Providers: PRO data would be most useful, relevant, and clinically significant for providers if key patient populations were targeted (e.g., Medicare patients). Providers expressed a need for accurate data that are seamlessly integrated into the EHR and intuitive to access, underscoring that these data should be easy to visualize and interpret.

- Enable flexibility for providers to choose and load preferred PRO measures and enable data validation steps between patient entry and integration of data into the patient's medical record.

- Integrate data with the EHR in a way that is considerate of workflow and provider needs (e.g., generation of customized visualizations within the chart), and provide clinical decision support.

- Address potential provider concerns about the impact of PRO measures on assessment of quality metrics for physicians and determine institution-level response and implications prior to implementation.

Health IT Professionals: Participants confirmed that ideal PRO data integration within the EHR would conform to open standards, enabling easy sharing of data across different institutions. Critical to success were leadership buy-in, coordination across different IT teams, and support from EHR vendors.

- Design the app using open and flexible industry standards, such as Substitutable Medical Applications, Reusable Technologies (SMART) on FHIR, and apply the same industry security standards for protecting patient-generated data to encourage adoption and ease of cross-platform implementation.

- Consider a lighter integration to address team and time constraints; use a custom data view to surface data in the EHR that has access to the PRO data source, allowing flexible presentations of the data.

These key takeaways informed the functionality and user requirements for the development and implementation of an app to collect PRO data and also served to guide implementation plans for both Pilots 1 and 2.

Chapter II: Pilot 1: Modification and Implementation of an Existing Healthcare PRO App

Overview

Using the PROMIS PF PRO measure, the pilot involved multiple phases over an 11-month period. Modification of an existing app used in our health system, technical integration of the modified app, and the pilot implementation of this app at nine primary and specialty care practices in Washington DC, Maryland, and Virginia. During each phase we identified facilitators and barriers from the perspective of key stakeholder groups.

Section I: Modification of Existing App

Original OBERD App: Technical Integration

The Outcomes Based Electronic Research Database (OBERD) is a cloud-based app hosted by HIPAA-compliant Amazon Web Services (AWS). The OBERD app currently facilitates the collection of PRO measures by implementing logic to adaptively serve content as validated by researchers at Northwestern University. The app is accessed by clinicians and patients using a personal login created by registering an account for each user. The technical goal of Phase 1 of this pilot was to modify the existing OBERD app and provider visualization app currently in use by MedStar Health. The OBERD app's ecosystem relied on proprietary standards and technologies. The patient's registration and authorization to be enrolled in the OBERD system relies on extensive and costly integration between the clinical and vendor systems. The original app was not integrated with the EHR and required providers to navigate to an external, vendor-hosted site with separate credentials to retrieve and view the PRO data. **The technical challenge addressed by the MedStar Health team was to modify the existing OBERD backend and the provider-facing visualization to bring them up to the modern industry standard of HL7 SMART on FHIR, which allows for a more interoperable, scalable, and workflow-centered solution.** In this section, we describe the original app, detail the modification process using the PRO FHIR IG, and outline key lessons learned and recommendations.

At the time of this report, the original OBERD app for collection of PRO data is in use at MedStar Health in the orthopedic setting. The workflow is such that when a patient has an appointment scheduled at MedStar Health, the institution's interface engine sends an HL7 Admit, Discharge, Transfer message to the OBERD backend for inclusion in their data store. The data feed to provide this connection is an effective system once in place but requires deep integration, which is cumbersome, costly, and burdensome to the health system.

Modification of Existing OBERD App

The MedStar team rebuilt the entire app's ecosystem to leverage the powerful, scalable, secure, and modern standards of SMART on FHIR. To facilitate the rebuild, the communication standards, and backend functionality, the MedStar team used the PRO FHIR IG which can be found here.

This IG provides specific, concise, and clear guidance on what FHIR resources are to be implemented as well as examples, diagrams, and overviews. Figure 1 shows a high-level architecture diagram of the elements modified during this pilot. The OBERD data repository to store these PRO data remains the same while the app programming interface (API), patient-facing OBERD app, and the provider-facing app were modified. The sections below outline specific elements of the OBERD app that were modified to achieve the pilot project.

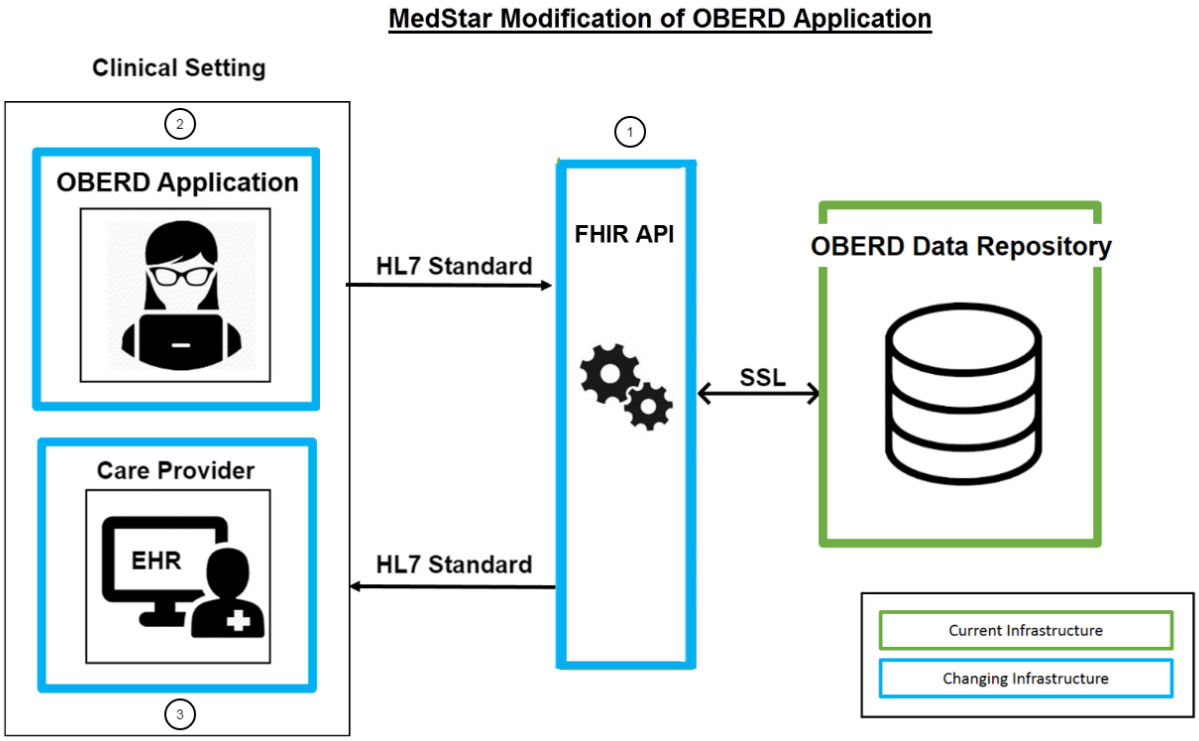

Figure 1. Elements of the OBERD app modified in the phase 1 pilot

Communication Standards of Patient PRO Measures

The proprietary structure used in the original OBERD app was reimplemented using the SMART on FHIR[10] and Questionnaire Response resources as per HL7 FHIR version 3.5.0.[7] These changes allowed the data flow to function identically between the user-facing mobile device for PRO collection, the OBERD database and FHIR API (Number 1 in Figure 1), and the provider-facing visualization (Number 3 in Figure 1) inside the EHR. The FHIR API implemented individual resources to facilitate this communication, including patient demographics, the adaptive questionnaire content, and the patient's responses to be stored for display in the provider-facing visualization. This standardization of bi-directional data flow allowed the MedStar team's app to provide a streamlined solution relying on open standards.

Implementation Modification of PROMIS Survey Distributor

The most challenging architectural component we modified was the engine that serves the PROMIS instrument's content to the patient-facing survey app (Patient-Facing Administration App in Figure 2). The OBERD distributor or EAC (External Assessment Center in Figure 2) is currently implemented in a database with proprietary APIs. The logic for the adaptive content is implemented to mirror the logic of the algorithms validated by Northwestern researchers. In the pilot phase, the PRO FHIR IG for the new distributor was given to OBERD to implement using the architecture shown in Figure 2 and is described in this section.

The existing OBERD app's backend distributor was tightly coupled with the infrastructure of their database and FHIR API. The PRO FHIR IG details a FHIR-based communication standard for administration of adaptive questionnaires. The Questionnaire Next Question Operation is shown in Figure 2.

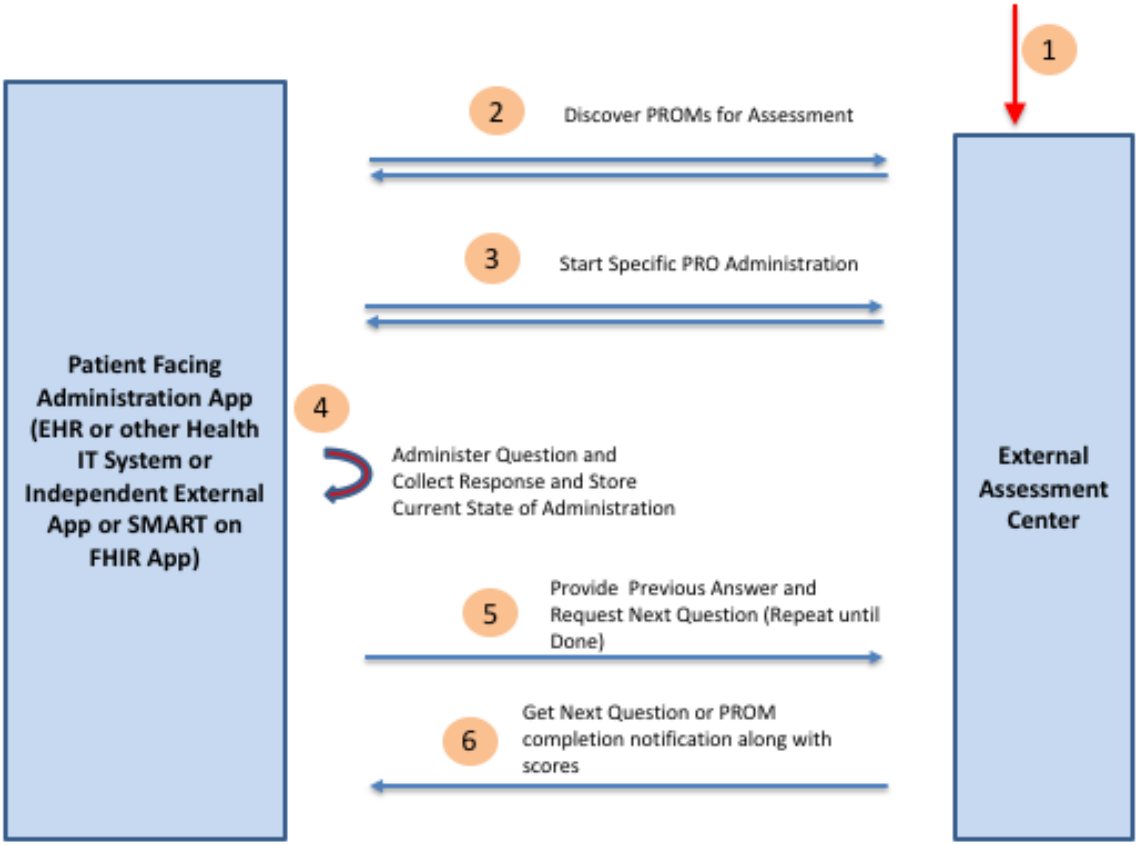

Figure 2. Architecture for administration of CAT using an EAC

Rather than relying on the proprietary structure of OBERD's implementation of the PROMIS measures, the FHIR specifications define an EAC. This substantive modification of the OBERD app allows for a more generalizable, scalable, and secure solution to building a distributor. The numbered data flow outlined in Figure 2 illustrates how a computer adaptive test (CAT) is administered using the EAC, questionnaire, and next question operation. An advantage of implementing an EAC is to decouple this component from the main data store to allow greater flexibility to integrate the patient-facing app. Another key feature of an EAC is that it never requires the use or communication of protected health information (PHI). The only relevant data for its algorithm is to know the response to the last question algorithm to determine and provide the next question for the patient-facing app.

Security Model Modifications

To conform to the HL7 SMART security model, we modified the method of authenticating the user. The original OBERD app authorizes the user to access the provider-facing visualization using an OBERD-specific set of credentials. The required modification needed to be both SMART on FHIR compliant and provide single sign on (SSO). This avoids requiring the user to enter another set of credentials after logging into the EHR. To implement this desired functionality, we modified the OBERD app to leverage the power of the cutting-edge technology of the SMART backend authentication (https://hl7.org/fhir/uv/bulkdata/authorization/index.html). Figure 3 shows the data flow for the SMART authentication of a client.

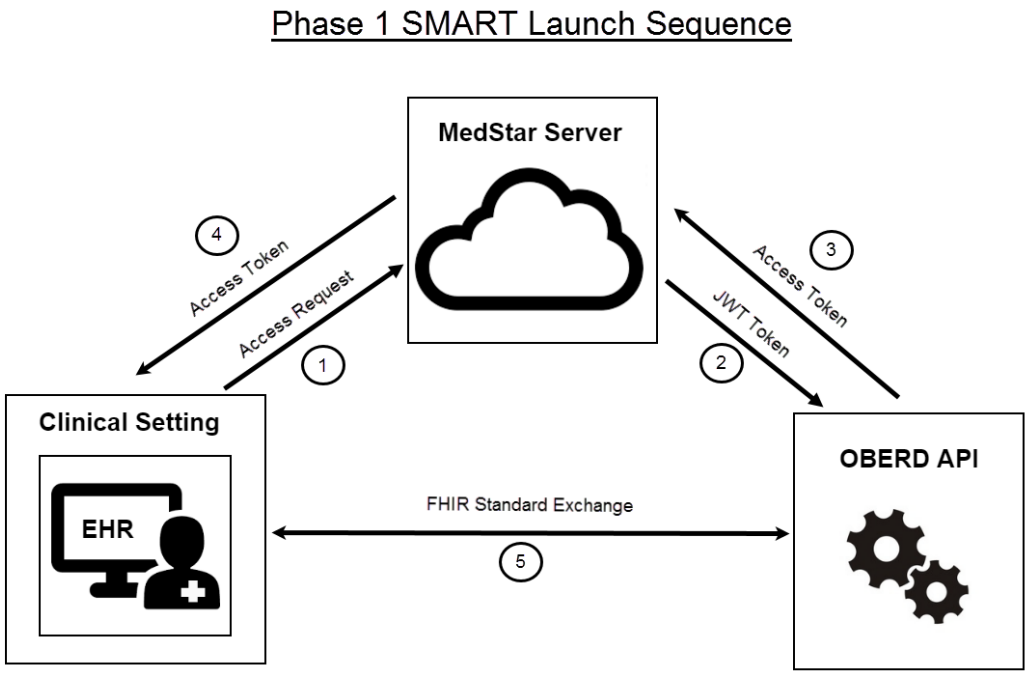

Figure 3. Phase 1 SMART secure authentication for SSO and light integration

The process of authorizing and authenticating the user begins when the provider logs into the EHR with their username and password. This allows for provisioning at the EHR system level to allow greater flexibility using group or role-based security. When the provider clicks to load the visualization, the EHR requests access from a secure cloud server (Number 1 in Figure 3), which in turn brokers a server-to-server bulk SMART authentication mechanism (Number 2 in Figure 3). Once the encrypted signed JSON Web Token (JWT) token is verified, the server returns an access token (Numbers 3 and 4 in Figure 3), which the EHR can use to fetch content from the OBERD backend dynamically (Number 5 in Figure 3). This model of access is a data-on-demand model, which provides flexibility, scalability, enhanced security, and avoids deep integration with the EHR.

OBERD App Modification: Challenges and Lessons learned

As expected with most technical implementations, several unforeseen challenges arose during the design, planning, and validation criteria. We demonstrated that the key to successful integration is understanding, documenting, and iteratively updating implementation plans from lessons learned.

Implementation of PRO FHIR IG

- **Key challenge:** Driving parallel development of both the provider-facing visualization (client) and (OBERD) server-side APIs.
  (*Related STS Dimensions: (1) hardware and software computing infrastructure; (4) people; (5) workflow and communication)*

- **Lessons learned:** Planning the parallel phases to test integration as early as possible is extremely important. OBERD's modification plan was to provide a development environment for MedStar Health to begin implementation and testing as soon as possible.

Implementation of EAC

- **Key challenge**: Re-implementing the distributor to use the EAC guidelines as defined by ONC. The greatest challenge was implementing the Questionnaire Next Question functionality. *(Related STS Dimensions: (1) hardware and software computing infrastructure)*

- **Lessons learned:** Although re-implementing the method by which the OBERD app served the questionnaire content required a significant effort, the new EAC as defined by the ONC IG offered significant advantages, providing a scalable solution that enabled the success of subsequent implementation of the patient-facing app and the EHR integration.

Implementing the New SMART Security Model

- **Key challenge:** Clearly defining the technical approach and security protocols (particularly around PHI) to non-technical stakeholders and decision makers at each site. (*Related STS Dimensions: (6) internal organizational policies, procedures, and culture; (7) external rules, regulations, and pressures; (4) people; (5) workflow and communication)*

- **Lessons learned:** Engage decision makers and stakeholders as early as possible. The key to successful recruitment and implementation was the ability to provide concise reference materials that were created and tailored to site-specific needs or concerns.

Section II: Integration with Different Electronic Health Records

Once the app was successfully modified, our team integrated the app with different EHRs. Of the participating pilot sites, six sites were on Cerner's EHR and three were on NextGen. Two of the three NextGen sites were running the same version of the software, with the third site running a different version, making this effectively two instances of the EHR since the same solution did not work for both versions. In order to generalize our approach for implementation, we chose the app architecture described earlier. It was designed to be scalable, modular, and web-based, thereby reducing the level of effort to deploy the provider-facing visualization in the EHR. The app's architecture was guided by the principle of **"data on demand"** rather than storing these PRO data natively inside the EHR. The principle of "data on demand" allows for the EHR to dynamically retrieve the PRO data from an external FHIR-enabled server when the provider-facing visualization is launched. **There are several clear advantages to this approach, which proved critical to the successful of this integration given the timeline and multiple contexts:**

1. **Reduces points of integration:** These PRO data are securely hosted outside the EHR, which eliminates the need for resource-intensive and time-consuming integration work through interface engines or feeds. With this architecture, when the provider launches the visualization at the point of care, these data are dynamically retrieved from the OBERD FHIR-enabled server.

2. **Reduces variance between EHR systems:** Every EHR's clinical data store is implemented in proprietary data models. If patient-generated data are stored natively in the her-specific data model, the functionality to write data into the EHR would be unique not only to the EHR vendor but to each implementation. The creation of a modular solution that dynamically retrieves content from an external FHIR server allows for a reduction in the variability between each site's EHR integration process.

3. **Reduces effort of IT support:** Externally hosting the PRO data on OBERD's FHIR servers reduces the level of effort required by the site's IT support. As the solution and data are externally hosted, the EHR's IT support teams only need to help deploy the visualization layer and not host the PRO data or code base inside the EHR.

Integration of the OBERD App

This pilot included the modification of the OBERD app to feature an in-situ provider-facing visualization inside the EHR. The original OBERD provider-facing app was a vendor-hosted website the provider was able to access using a separate username and password. This was modified to be available inside the EHR (Number 3 in Figure 1). Both NextGen and Cerner systems have pathways to surface custom visualizations using modern web programming languages, which allow visualizations to natively communicate with APIs upon which the FHIR standard is modeled. The modification of the app is detailed below for each system.

- **Cerner**: Cerner provides customized visualizations known as Cerner Millennium Pages (mPages). The Cerner mPages have access to both the Document Object Model (DOM) in which the custom code resides and the launch context of the chart in which the visualization is surfaced.

- **NextGen**: The NextGen system provides a native functionality to allow system developers to create composable interfaces called Dynamic Templates. These templates can include custom web components (DOMs), which are able to dynamically load client-side HTML, CSS, and JavaScript. These dynamic templates can dynamically retrieve content by communicating with external hosts.

Figure 4 shows the PRO visualization inside the Cerner PowerChart app as an mPage on a test patient. This visualization is accessible inside the patient's chart by clicking on the tab on the bottom left marked "MedStar Patient Mpages."

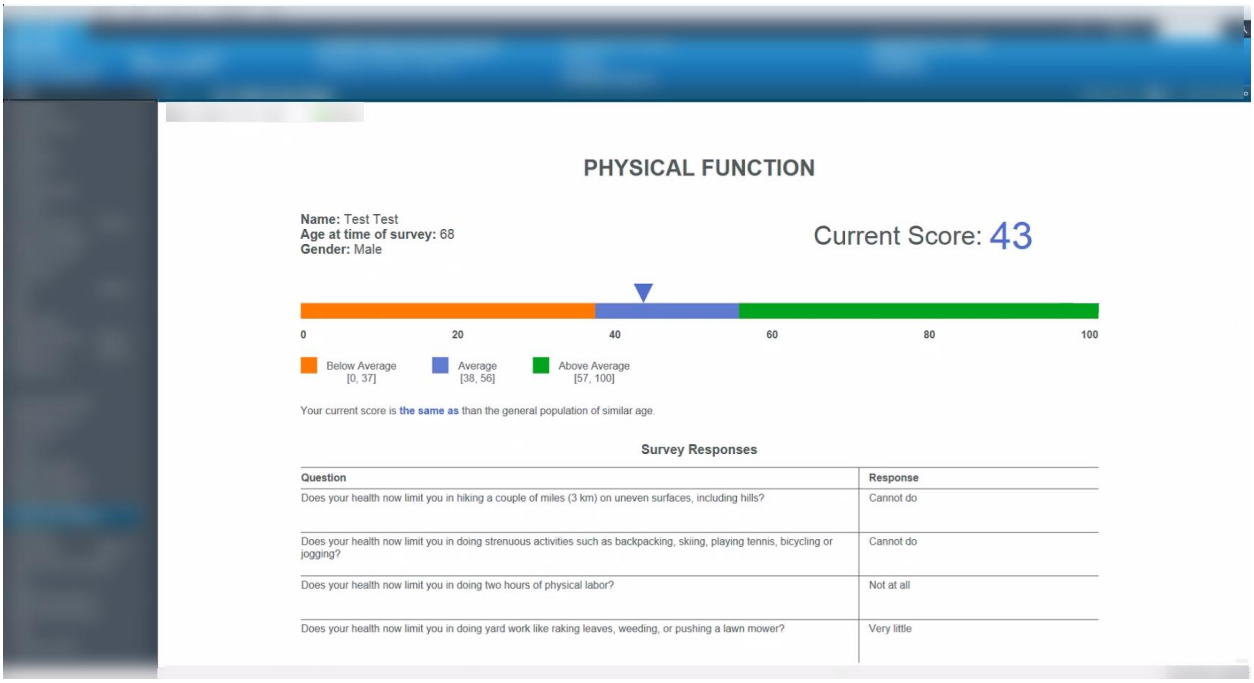

Figure 4. EHR integrated provider-facing visualization shown as a Cerner mPage on a test patient

Iterative Development of Provider Visualization

To develop the actual design to represent the visualization of the patient's score in the EHR, we utilized an iterative development process to generate a usable final visualization from the initial mockup we conceptualized based on examples in the literature and from input from stakeholders in the first phase of the project. We incorporated elements from other examples of PRO visualizations to develop an initial provider visualization. The visualization progressed through several iterations, which included modifications to layout and color usage to increase readability and understanding. Color scheme changes were also made to better align with users' mental models related to color interpretation. We then conducted a round of expert evaluation and subject matter feedback sessions to further iterate on these designs through 30-minute interview sessions with three representative primary care physicians. All feedback was incorporated to produce the final iteration of the visualization used for the pilot (shown in Figure 4 above).

Integration of Provider-Facing Visualization: Challenges and Lessons Learned

Throughout the process of integration, we encountered a series of both anticipated and unanticipated challenges associated with the EHR.

Cerner:

- **Key challenges:**
    - *(Related STS Dimensions: (1) hardware and software computing infrastructure)*
    - Given our familiarity with Cerner's architecture and building SMART on FHIR apps, the only challenge was the (anticipated) need to develop a sophisticated, modern, web-based app to be deployed inside an EHR. As Cerner uses an embedded browser window to surface web pages no modern tools are available for debugging.
- **Lessons learned:**
    - Begin iterative testing inside the EHR as soon as possible in the integration process. This strategy creates more work during the development phase avoids many downstream consequences that could arise during external development of a provider-facing app in EHR.

NextGen:

- **Key challenges:** We note two significant unexpected challenges from integrating with two NextGen systems:
    1. Seemingly unrelated and unknown system upgrades have serious consequences. The team was unaware of a major system upgrade which was only discovered when attempting to launch the app at one site that was ready to enroll patients. The system upgrade had caused the existing implementation of the NextGen provider-facing visualization to be deleted, requiring us to re-build.
        - *(Related STS Dimensions: (1) hardware and software computing infrastructure; (4) people)*
    2. Different instances of the same software did not function the same way. NextGen provides functionality to export a provider-facing visualization in Extensible Markup Language (XML). When attempting to port an existing solution from one NextGen environment to another NextGen instance, we discovered it would not work due to different EHR versions and had to implement a separate solution for each version.
        - *(Related STS Dimensions: (1) hardware and software computing infrastructure; (4) people; (5) workflow and communication)*

- **Lessons Learned:** EHR integration into two separate NextGen instances provided the insight into how to anticipate related challenges.
    1. The importance of communicating with system administrators of the EHR to understand the exact schedule of any system code changes and their downstream consequences on integrated components.
    2. The second lesson learned is that the same EHR vendor's product may or may not support a wide range of functionality depending on the versions.

Technical Implementation: Perspective and High-Level Recommendations

Recommendations for Optimizing the User Experience and Implementation

1. **Design the solution to have the shortest pathway or number of actions necessary to reach the provider visualization app within the patient context**
   EHRs can launch external apps using a hyperlink or button from a patient list, from a menu item, or from a user interface (UI) element visible only when you are in a single patient chart. We determined the optimal way to launch it was from a patient-detail-level window. For both Cerner and NextGen, the provider visualization app opened with a single click once a provider was in a specific patient's detail screen. As there was no need for any input or further authentication, we did not need any additional dialog or UI elements in a workflow that opened the provider visualization app.

2. **Develop a presentation layer that does not require further scrolling or any other visual manipulation to see the data**
   Once the provider visualization app is visible within the EHR window, avoid further need for the provider to scroll or zoom to see the main data points. This requirement is highly dependent on the EHR's capabilities and sometimes is not technically possible. In this case, design so that data presented only require vertical scrolling. Avoid putting any zoom/enhancing UI elements for presenting data, as this feature makes it harder to visualize the overall context quickly at one glance.

3. **Keep the presentation layer locked down to the patient context**
   It is possible for web apps to show up in a floating (non-modal) window; however, this design incurs a risk of displaying data not bound to the current patient, creating a potential hazard and safety risk. We worked with integration experts from both EHRs to find a way to display the provider visualization app as an embedded UI control within the EHR. This causes the provider visualization app to close as soon as the patient is changed.

Anticipating Challenges and Mitigation Strategies

Throughout the technical implementation, we faced some unexpected and unforeseen challenges. The following recommendations and key mitigation strategies could potentially prevent significant delays to a similar implementation, if scaled:

Finish EHR "plumbing" (technical integration) work early in the process

Different ambulatory practices will have different ways to manage their EHRs, including those who rely entirely on the EHR vendors, others with internal teams who can directly facilitate customizations, and others with a mix of both. Ensuring that the basic steps—such as launching the UI with the click of a button

or passing patient context—work as early as possible can help determine the level of effort or vendor support required to make EHR specific changes, if needed.

Anticipate that EHR version differences add to client variability

The provider visualization app that worked in the EHR at one facility did not work at another facility, despite the same EHR software. Troubleshooting revealed that differences in EHR versions was causing rendering issues, as one facility was running an older version of the EHR, which used Internet Explorer 7-based rendering control.

Incorporate testing from the beginning of the solution design

We tested each part of the PRO provider app alone first, before testing the whole app, and this allowed for a significant time saving since it enabled early identification of potential issues before building on. The PRO app had multiple security features that included client context passing, which relied on client-side rendering (CSS, Javascript, HTML) for UI rendering, and it worked within the confines of the EHR. We developed a testing tool that helped test each component separately and used these in initial phases of implementation. This helped significantly when we integrated all the pieces together for a final test.

Implement security from the beginning – it should not be an afterthought

From the beginning, our discussions with the vendor included major portions dedicated to security implementation. All our functional design and implementation had security pieces stubbed out from the beginning which was helpful when we started layering in security pieces. This also allowed troubleshooting when the app did not function due to security-related issues.

Include change management plans for both technical and project planning, as well as timelines

Depending on the institution, change management processes can range from being very informal to highly structured, but some level of approvals and reviews by the app owners will always be required. At one of our pilot institutions, this only required a simple explanation and demo in a test system. In another institution, it was a much more complex processes requiring completion and submission of multiple forms and participating in technical as well as architectural change management calls with institutional officials to obtain approval and implement our app.

Anticipate that system and software updates will add variability

Though EHR systems in this implementation only displayed provider-facing data using a web app, we encountered issues when there was an unanticipated system update that wiped out changes we had made in its template, which stored the provider-facing web app's URL and other parameter configuration. We had a backup of the original template from that EHR system, which helped us get it back online with very little downtime. Although the outside system and its configuration changes are not under our control, it helps to get documentation or exports of any changes as a backup so that it is easy to recover from downtime in the event of data loss.

Section III: Implementation

This section presents findings from the implementation of the app at nine pilot sites in the Washington, DC, Maryland, and Virginia regions. We describe the process of launching the pilot at each ambulatory

site, including the development of appropriate training materials and methods for data collection and evaluation. We then present the results findings and lessons learned relative to the practices, patients, providers, and the health IT staff.

Development of Training Materials

To facilitate implementation at each pilot site, we provided each practice with training materials for each user group for reference once the app was launched in each clinic. Materials were developed for patients, providers, and practice staff, using a rigorous evaluation process. Our team worked directly with OBERD to develop the initial prototype training materials once the app was fully developed. We then conducted user testing to iteratively develop these materials.

Clinical Pilot Site Selection and Recruitment

Participating sites were identified through convenience and snowball sampling. First, we identified potentially interested providers through our direct contacts within our research network. We initiated face-to-face meetings with potentially interested providers to establish rapport, explain the study, and establish a relationship. Eligible patients included individuals who were: 1) at least 65 years of age or actively rehabilitating from injury, 2) able to communicate in English, and 3) able to provide a verbal consent to participate in the study. Once their interest was confirmed, we enrolled them as a site and proceeded with subsequent interactions via phone or email as appropriate to keep the sites engaged until we were ready to initiate the pilot study in clinic. For sites within the MedStar system, we obtained leadership buy-in by way of the head of MedStar Ambulatory Group. We leveraged this to facilitate introductions to potentially interested providers throughout our system, and this resulted in successfully reaching our recruitment goal for MedStar sites. For the CAPRICORN sites, we leveraged the network's coordinator, who had dedicated time to assist with reaching out to providers and identifying those who were interested in participating and serving as a pilot site. Finally, in two cases, we asked one participating provider to recommend another practice to us, and we were able to successfully recruit the final site in this fashion.

Seven sites were MedStar Health sites, and two sites were from the CAPRICORN PBRN. Practices ranged from small practices with two providers, to larger practices with 18 providers. None of the sites had used PROs prior to this effort with the exception of one site using them to report on quality metrics. Staffing ratios varied, with some sites having two medical assistants and two administrative staff, while other sites had up to eight administrative staff and eight medical assistants rotating through the day. The sites were in different geographic areas ranging from urban, to suburban, to rural, and serving patients of different racial/ethnic and socioeconomic background accordingly. All but two sites were primary care, and the other two were a specialty cardiology clinic and a rheumatology clinic. Each site received a $2000 participation incentive upon completion of the pilot.

Evaluation and Metrics

Our team conducted a rigorous evaluation of the app implementation and use. Select metrics identified facilitators and barriers to implementation and use at each site. Our evaluation and assessments were conducted on three levels: 1) general practice level, 2) patient level, and 3) provider level.

1. **General practice-level assessment:** These were conducted to understand, at baseline, the factors that may impact implementation and use of the app and PRO data at an organizational level.

a. **Baseline organizational *Readiness for Change* survey**[11]: The structured survey assessed each pilot clinic's readiness to implement this new app and provided an overall indication of the likelihood of success prior to implementation.

b. **Clinic baseline observation visit:** Our research team conducted an in-person observation visit prior to launching the pilot. The purpose was to observe standard clinic workflow and to understand any logistical considerations and nuances (e.g., lack of Wi-Fi or presence of kiosks for check-in).

c. **Clinic site kickoff:** Following integration testing, the kickoff meeting consisted of the research team presenting an overview of the pilot study, including proposed workflow, patient data collection procedures, and a training on both the app interface and provider visualization. Staff were also provided with the training materials at this time.

2. **Patient-level assessment:** We conducted multiple assessments with patients to understand 1) usability, 2) usage, and 3) satisfaction/preferences/barriers.

   a. **Systems usability scale (SUS):** This was administered to each patient after they submitted their responses in the app. The SUS is a simple, brief scale that is considered the industry standard for evaluating usability of hardware, software, mobile devices, websites, and apps.[12] The PRO app was programmed to include the SUS, which was integrated seamlessly into the interface.

   b. **Semi-structured interviews:** We conducted a total of nine semi-structured interviews across all sites, with the goal of interviewing one patient from each site. Interviews were structured to elicit feedback on general use of the PRO app, including ease of use, preferences, perceived usefulness, and experience with the app.

3. **Provider assessment:** From providers, we sought to understand 1) usability, 2) usage and usefulness, and 3) satisfaction/preferences/barriers.

   a. **Semi-structured interviews:** One provider was interviewed from each pilot site. Interviews addressed topics including: 1) how providers actually used the PRO data to inform clinical decision making; 2) reasons for use or non-use of the PRO data; 3) how the process could be optimized; and 4) general satisfaction, preferences, and barriers.

4. **Health IT staff assessment:** The goal of these interviews was to understand the experience with integration of the technology in each pilot site, and to identify barriers and promoters of successful integration.

   a. **Semi-structured interviews:** We conducted nine semi-structured interviews with health IT staff at various levels with individuals involved in various aspects of the technical integration. Interviews also focused on satisfaction with the implementation process, preferences, and barriers.

Data Analysis

Semi-Structured Interviews Data

Data were analyzed using the same qualitative data coding approach employed for the analysis of the stakeholder feedback (*Chapter I Stakeholder Feedback- Methods*). If the researchers coding the data encountered technical terms that they were unfamiliar with, such as information technology terms mentioned by the health IT stakeholders, the researchers consulted experts on the research team to clarify and code appropriately. Data were completed to determine the most prominent themes in the context of the interview questions and goals of the questions. The project team identified where particular themes were more frequent than others by participant. These frequently occurring themes are presented in the results section below.

Readiness for Change Data Analysis

The point of contact at each site was asked to respond to the survey based on their practice's attitudes and perceptions toward health IT. In most instances (n=7), the point of contact was the practice manager. In two instances, the point of contact was a site coordinator. Differences in the role of the point of contact should be considered when reviewing the data.

Overall, perception of e-health technology could score a minimum of 24 points and a maximum of 120; organizational readiness for change could score a minimum of 21 points and a maximum of 105; and e-health experiences within the organization could score a minimum of 12 points and a maximum of 60. Responses were summed across each section, with lower numbers signifying a higher readiness. The data were then analyzed by identifying common themes across all sites in the context of each question.

Systems Usability Survey Data Analysis

Each participant who completed the PRO survey also completed the system usability scale (SUS), a standardized measure of usability.[12] The SUS is scored between 0 and 100, which is a score specific to SUS and not a percentile. This score is calculated by applying a formula that subtracts one from the score of all even questions on the survey, subtracts the scores of the odd questions from five, adds the total, and then multiplies the resulting score by 2.5. This calculation was completed on 10 SUS evaluations per site, for a total of 90 evaluations.

Section IV: Clinical Implementation Findings

Practice Site Assessments

Readiness for Change

Based on our review of the readiness for change data, we made several general observations that helped us characterize our participating sites according to the three core categories on this assessment.

1. **Overall acceptance of health IT across pilot sites:** Overall, results show that all nine sites were generally accepting of and had positive attitudes toward new health IT. Only one site (Site D) stated they were not satisfied with currently available e-health technology at their practice, a question within the domain, "Overall perception of e-health." This same site also agreed with the statement "I may lose some of my autonomy using e-health technology," as well as, "I feel a lot of pressure to be more effective by using e-health technology." Similarly, Site I agreed with the

item, “I feel a lot of pressure to be more efficient because of e-health technology.” Site H stated they preferred to use e-health technologies that would not change, and they found e-health technology generally complex. Almost all sites strongly agreed that e-health technology can improve patient outcomes, improve patient care, and enhance a team’s approach to care.

2. **Pilot site perceptions of organizational readiness:** All sites generally had positive perceptions regarding their respective organizational readiness; however, there was some variability in the extent to which the sites felt ready to implement and use e-health technology. Notably, Sites E and A both reported having limited success with implementing e-health technology in the past, and Sites D and A both disagreed with the statements “My organization has an adequate number of IT staff for technical support,” and “My organization provides timely and flexible support to users of e-health technology.” Site D also strongly disagreed with the statement, “My organization offers staff opportunities to provide input into the early phases of e-health implementation and change.” All sites agreed or strongly agreed with the statement, “My organization is committed to making e-health projects successful.”

3. **Pilot site prior experiences with Health IT:** While almost all sites reported positive experiences, Sites D and E both disagreed with the statement, “E-Health technology generally performs at an adequate speed,” and Site D disagreed with the statement, “E-Health technology is user-friendly.” Overall, Site A scored as least ready, responding as “Uncertain” to most questions. Notably, most sites, which opted to participate in this pilot due to a stated interest in health IT and research, scored relatively high on overall readiness. As one practice manager expressed to the research team, “Our practice loves doing [research studies]. We’re always open to opportunities to be involved.”

Readiness for Change Scores

All practices scored in the lower end of the range with a total score of 49 or below (with the lowest possible score being 24; highest possible score 120). For perceived organizational readiness, one practice scored almost in the median of the range (score of 59 with a lowest possible score of 21 and highest possible score of 105), while the rest scored below this. Finally, scores across sites suggested a wide variability in terms of prior experiences with e-health as an organization. Approximately half of the sites recorded scores in the mid-range suggesting some prior experience with e-health. Three practices had relatively low scores suggestive of minimal prior experience with e-health technologies in general.

Summary of Observations During Practice-Site Data Collection

We conducted basic observations of each clinic both at baseline and during the process of data collection. Assessing each site individually allowed our team to tailor each implementation plan to meet the needs of each site. These observations ensured workflow, roles, and expectations were thoroughly discussed for all staff. In general, we noted observations in four key areas that directly impacted how the pilot was implemented at each site.

Variability in workflow: The workflow for each site varied greatly from the check-in process to the patient-provider encounter. We noted critical differences related to 3 key factors: (1) whether the site used kiosks or in-person check-in, (2) patient volume and practice size, and (3) how medical staff were alerted to the patient’s arrival.

Variability in communication among practice staff: At some sites, communication about the pilot study involved the research team and all practice staff, while at others it only involved one or two staff members (e.g., just the participating provider and the practice manager). As a result, the level of penetration and reach of communication from the research team about project goals and training varied between sites. Similarly, depending on workflow, practice size, and staff resources, communication between the participating providers, front desk staff, practice manager, and/or other participating staff members was variable. In these practices, there were no issues with notifying or alerting the provider when the patient they were seeing had a PRO measure to be reviewed in the chart.

Overall, based on observation alone, the greatest barrier to successful implementation across clinics was related to communication breakdowns. This included the practice alerting the research team to arriving eligible patients, the research team having an adequate opportunity to explain the study to the patients, and the practice staff alerting the provider when patients participated.

Variability in engagement and involvement with the project by patients and practice staff: Based on comparison and observations of the nine practices, we noted that another determinant of efficient data collection was the level to which the practices and their staff were invested in the project, and the level to which patients themselves were engaged. This was highly variable across sites. For instance, we found that practices that were invested and fully engaged performed extra activities ensuring their success. This included willingness to put additional time into pre-screening and identifying patients who would be ideal participants (based on their prior knowledge of patient characteristics or existing issues with physical function).

Patient experiences during data collection: Based on observations alone, patients appeared to experience challenges or had questions related to survey progression, technological challenges, comprehension of the PROMIS and SUS questions, and general questions about security of their data. Across all nine sites, patients asked clarifying questions about the purpose of the questions. One patient wanted to ensure that the data would only be for his doctor's use.

Summary of Enrollment and Recruitment by Site

Each site met its target accrual of 10 patients over different time periods. Once each site had reached their target of 10 patients enrolled with complete surveys, one patient at each site was contacted for a follow-up interview. At one site, none of the patients could be reached for a follow-up interview despite indicating interest in participating. To ensure the target sample size was still met, two interviews were conducted at Site H.

Patient-Level Assessments

Below we describe the findings from the patient-level assessments, including the usability survey and the patient semi-structured interviews. We first describe the standard ranges and relative interpretation of the SUS score, and then present the median and interquartile ranges of the scores from the subset of patients at each site.

Systems Usability Survey (SUS)

The SUS is scored between 0 and 100, which is a score specific to SUS and not a percentile. Bangor et al. (2008) ascribed the acceptability of SUS scores compared to the average.[13] SUS scores ranging from 1-50 are deemed 'acceptable,' 51-70 are deemed 'marginal,' while 71-100 are deemed 'acceptable.'

Patient Observations: Usability

Based on observations of participants interacting with the tablet to capture their PRO data, we identified several usability issues. These included complications with tablet functionality, survey progression, and comprehension of survey questions.

1. **Tablet functionality**: At all nine sites, we observed that some patients experienced issues with manual interaction with the tablet. Common issues included: tap duration (too short or too long), tap force (too light), or point of contact (fingernail versus finger pad).
    a. *(Related STS Dimensions: 1) hardware and software computing infrastructure; 3) human computer interface; 4) people)*

2. **Survey Progression**: Patients across all nine sites experienced challenges with the survey progression. Many patients were confused about the auto-progression of the SUS questions. Whereas the PROMIS measures required patients to select their intended response and then click "continue," the SUS auto-progressed onto the next question following an entered response.
    a. *(Related STS Dimensions: 1) hardware and software computing infrastructure; 3) human computer interface; 4) people)*

3. **Comprehension and applicability of the survey items:** The most commonly observed usability issues for most patients across all sites were challenges with comprehension and readability of the PF survey questions. <u>Patients often expressed that the survey content was not directly applicable to their lifestyle or context, and thus had trouble conceptualizing the question to provide the most accurate response</u>.

    a. In addition, while filling out the survey, patients commented that the wording of the questions and the formatting of the SUS was confusing. This confusion introduces an important caveat and potential limitation to interpreting the SUS data.

    b. *(Related STS Dimension: 2) clinical content, Semi-Structured Interviews (n=9))*

Two overarching themes related to ease of use were identified from patient interviews: interaction with the app and general comprehension.

1. **Patients generally found the app was usable and functional for completing the PF survey.**

    - Most patients (n=6) found survey completion using the app was quick (<5 minutes).
    - All patients (n=9) found the survey easy to complete.
    - One patient (with a disability) needed assistance during the data collection process. The remaining eight patients did not need any assistance.
    - Exemplar quotes:
        - *"The [PRO] process worked well. It's a good time to catch patients because they have to wait before their appointment anyway."*

- *"[The questionnaire] was quick and easy. I'm not really a computer or tablet person, I don't care for them. But this was pretty easy."*

2. **Most patients did not perceive that completing the PRO measure via tablet had much utility nor did they believe it impacted or changed the encounter with their provider.**

    - When asked whether the survey had any impact on their conversation with their doctor, all patients (n=9) answered no, it did not.

    - When given the option between completing a PRO survey on a phone, tablet, or computer, most patients selected a tablet (n=5), followed by a computer (n=3). The remaining patient did not have a preference.

    - Exemplar quotes:

        - *[When asked whether the patient would like to complete the PRO questionnaire on her phone]: "It's too small of a screen, I can't see it. Plus, it's a flip phone so it wouldn't work."*

        - *"It's hard to say if questions were relevant. Most were, but they were very general. It would be better to tailor questions to patient based on visit."*

        - *"My doctor didn't talk about survey during the appointment, but my appointment was just a follow-up."*

Summary of Barriers and Facilitators for Implementation from the Patient Perspective

Based on the usability findings and follow-up interviews with patients, we gleaned some insight into potential facilitators of and barriers to successful, scalable use of PRO apps in primary and specialty care. These have potential implications for adoption and sustained use from the patient perspective. Potential facilitators included:

1. **Ease of interaction with the app:** Most patients did not report significant challenges in terms of the usability of the app itself. Usability scores ranged from marginal to acceptable across sites, but across all 90 patients, they were in the marginal range. This suggests there was room for improvement in the app itself, and we highlight some of these areas related to manual interaction in the previous section.

2. **Speed of completion of the survey enabled by computer adaptive test (CAT) format on a tablet:** Almost all the patients noted that the app facilitated quick completion of the survey, allowing them to finish in around 5 minutes. The computer adaptive test ensured that the patients on the average completed an average of 6-8 items and patients appreciated the efficiency of not having to fill out a lengthier survey, yet still generating the needed results.

Potential barriers for implementation of PRO apps included:

1. **Lack of perceived utility of the PRO score by the patient:** All nine patients interviewed noted the PF score they entered did not impact communication with their provider. Patients suggested that since their provider did not address it, it must not be important. While providers may not have seen a need to discuss an unremarkable score during that patient's appointment, the apparent lack of validation from the patient's perspective may in fact dissuade their subsequent use or collection of these data due to the lack of perceived utility.

2. **Lack of directly relevant or contextually appropriate survey content for all patient populations:** The most common complaint from patients was about the wording of the questions (reading level too high or hard to comprehend), or the context of the questions (not relevant for certain patients). During observations some patients appeared visibly frustrated when this was the case, or they communicated this to the research team.

Provider-Level Assessments

Semi-Structured Interviews

In general, we found that providers were pleased with the visualization, though some believed it could be even easier to access. Several providers also mentioned that the survey changed the conversations they had with patients to be more focused on their physical function. More general responses suggested that in some instances, providers did not think the survey items applied to the provider's patient population, or responses did not correspond with the reason for their patient's visit. Finally, providers alluded to workflow issues that emerged during the data collection, the most egregious and challenging being the communication breakdowns that resulted in the provider not being notified when surveys were complete and needing to be accessed.

In general, providers identified several factors that served as <u>key facilitators</u> of this implementation:

1. **Most providers found it easy to access the data within the EHR and use the reference materials when needed.**
    a. *(Related STS dimension: 3) human computer interface)*

2. **The presentation of the data in the visualization was clear.**
    a. *(Related STS dimension: 3) human computer interface)*

3. **Having a member of the research team available onsite during data collection made implementation much easier.**
    a. *(Related STS dimension: 4) people; 5) workflow and communication; 6) internal organizational policies, procedures, and culture)*

4. **Having the PRO PF data positively changed patient/provider communication on occasion for primary care providers.**
    a. *(Related STS dimension: 2) clinical content; 4) people; 5) workflow and communication)*

5. **Data clinically useful for cardiologist**.
    a. *(Related STS dimension: 2) clinical content; 4) people; 5) workflow and communication)*

b. One of the two specialty providers, a cardiologist, reported that upon review of a patient's physical function data PF score during an encounter, the score appeared worse than expected. This triggered further inquiry and influenced clinical decision making. In this instance, the provider referred the patient to another specialist (pulmonologist). Data were clinically relevant and impacted decision making for the specialty provider.

Providers collectively noted several barriers to the full success of this implementation, if scaled:

1. **Provider notification/alert to survey completion.**
   - *(Related STS dimensions: 3) human computer interface; 4) people; 5) workflow and communication)*

   - Because the app did not notify providers through the EHR that patients had completed the survey, notification of the provider was dependent on reliable and consistent communication between the relevant staff and provider within each practice.
     - Five of the seven sites struggled to consistently notify providers of survey completion.

     - The two sites where the front desk consistently communicated with the providers used intra-clinic communication methods (i.e., secure messaging systems, headset communication) to notify providers.

     - At another two sites, physician notification was the sole responsibility of the site coordinator, as requested by the practice. This proved completely effective.

     - Because communication within the clinics could be unreliable, providers suggested that having the EHR flag patients who completed the survey would result in more consistent use of the survey data.

2. **Lack of applicability or clinical utility of the data to the patient population**.
   - *(Related STS dimensions: 2) clinical content; 4) people; 5) workflow and communication)*

   - Eight providers said that PF PRO data were important to know for their patients; however, the survey results were not clinically useful. Two providers noted they were already capturing these data in a more useful format, while three said these data were not relevant for that visit.

     - Five providers explained that the survey had changed their conversation with patients; however, all but one of these providers expressed that the conversation did not change clinical decisions.

     - Four providers noted that ideally they needed customizable questions for the visit.

3. **Number of clicks to access the data**.
    - *(Related STS dimensions: 3) human computer interface; 4) people; 5) workflow and communication)*

    - Although providers were generally pleased with the ease of accessing the data, two providers mentioned that the link to the visualization was outside of their typical workflow and took too many clicks.

Health IT Staff-Level Assessments

Semi-Structured Interviews

Overall, technical experts conveyed the value of our multidisciplinary team in an implementation of this magnitude, given that successful implementation required a competent technical team as well as several other disciplines, and open and consistent communication between them all.

Overall health IT staff identified three key facilitators:

1. **Having technical experts on the team along with expertise in other areas of implementation.**
    - *(Related STS dimensions: 4) people; 5) workflow and communication)*

    - Participants noted it was helpful that the research team included experts in the technology and standards. The information systems (IS) processes were made simpler by having members of the team who could directly participate in conversations necessary to obtain institutional clearances. Institutional IS staff otherwise had limited knowledge of the project, and these factors helped to facilitate conversations, troubleshoot, and ensure project objectives were met.

2. **Consistent communication/help from research team**.
    - *(Related STS dimensions: 4) people; 5) workflow and communication)*

    - Having a designated point person who is easily accessible and keeps an open line of communication for questions and concerns from the IS team.

3. **Transparent and accessible FHIR technical standards worked well**.
    - *(Related STS dimensions: 5) workflow and communication; 7) external rules, regulations, and pressures)*

    - Two technical experts reported that the FHIR standards were relatively straightforward to implement and generally worked well for the technical teams involved.

Interviews revealed four key barriers to a successful implementation of this nature, and particularly scaling:

1. **Miscommunication about the distribution of labor and assignment of roles when working with outside technical teams and vendors.** This level of integration will often require work with technical teams outside of the core team. For this level of implementation, it required that the expectations for each party be revisited and refined.

2. **Delays in involving IT teams in all aspects, despite how straightforward the task.** Several participants noted that to successfully scale the implementation of this technology, technical experts and institutional IT teams needed to be involved early. Specifically, including these key members early could improve inefficiencies in the discovery process, and many of the major hurdles to implementation would not be a surprise or could be avoided altogether.

3. **Failure to anticipate the full range of institutional hurdles.** Technical experts argued that the project team did not anticipate the amount of time it would take to obtain institutional clearances (e.g., from the hospital system's IS department).

4. **Issues with EHR version compatibility and upgrades**. Two of the technical experts directly involved with integration highlighted the initial confusion caused by compatibility issues due to upgrades. One of these experts said that their site was multiple versions behind on their EHR upgrades; therefore, the template provided by our technical team was not compatible with the site's EHR.

Section V: Overall Lessons Learned: Clinical Implementation and App Usage

In addition to the facilitators and barriers that have been described in each section of the preceding, we also identified several lessons learned that span the entire project process from site recruitment to final data collection and site closeout. These lessons learned have direct implications for scalability of the standards and PRO technology. We provide details of the four key lessons learned below.

Recruitment and Patient Enrollment Methods Matter

The process of patient recruitment and engagement varied widely between clinics, and success largely relied on clinic communication and engagement throughout. Clinics used four methods to recruit patients, with varying levels of success:

1. **Full pre-screening and advance patient selection:** This required the most involvement on the part of the practice staff but yielded the fastest and most efficient patient recruitment. In some clinics, this also involved provider participation, allowing the clinic to curate which patients would participate and, in theory, provide more relevant data. This required high motivation and involvement by the practice staff and required a higher workload for the entire project team, call to screen, recruit, and communicate with patients throughout the process.

2. **Partial pre-screening with pre-identified patient list:** This method was most commonly used by mid-sized practices. At the beginning of the week, the practice manager communicated to the research team which days and time periods had the highest number of eligible patients, considering the patient's reason for visit. A member from the research team would then show up at the designated time, communicate with the front desk to identify the eligible patients, and

recruit on the spot with the patient having little to no prior awareness of the research study. This method had moderate success.

3. **Recruitment by intercept (no pre-screening):** This was used by two clinics with exceedingly high patient volumes and involved little to no screening on the practice's part. As the practice saw many eligible patients on any given day, the research team was instructed to show up in the morning on any day most convenient to them. From there, the patient schedule was printed and given to the researcher, who communicated with the front desk about which patients were ideal participants. The front desk would then inform the researcher when the designated patients arrived, and recruitment would commence, with the patient having no prior knowledge of the research study. This method had the least success, with the lowest participation rate of recruited patients.

4. **Involvement of a dedicated practice site coordinator:** This involved the participation of a site coordinator. Used in two practices, the site coordinator worked closely with the provider to identify and recruit eligible patients. The site coordinator called the patients prior to their visit, recruited and consented the patients when they arrived for the visit, and communicated directly with the provider upon completion. This method was extremely successful, as the process, from screening to provider communication, was efficiently streamlined and did not involve extraneous coordination or communication with the practice.

Workflow and Staff Communication Can Be Just as Important as the Technology

The greatest barrier to successful project execution and data collection at all sites was related to communication among practice staff—in particular, the coordinated communication between the front desk staff or medical assistant, and the provider. Communication with the front desk was the cornerstone to the project's success, as it was involved at every point of the recruitment process. First, whether patients had been screened prior to their arrival or not, the front desk was expected to be aware of and inform the researcher of patient arrival. After successful recruitment and patient participation, the researcher would inform the front desk, who would then be responsible for informing the provider. If a communication breakdown occurred at either of these points, the project would be compromised for that patient.

Front desk engagement varied from clinic to clinic as well. One practice's front desk staff went as far as to wait to inform the medical assistant (MA)/nurse of the patient's arrival until after recruitment occurred. This permitted a more relaxed recruitment process, giving the patient ample time to progress through the consent and survey before being called back for their appointment. Conversely, front desk staff that did not engage thoroughly in the study resulted in researchers not being informed upon patient arrival, recruitment and consent being rushed (oftentimes, the researcher had to ask the MA/nurse for additional time prior to the patient being called back, thus delaying workflow), and providers not being consistently informed in patient participation (thus not knowing the visualization could be accessed for that patient in the EHR).

Certain Practice Characteristics May Be More Predictive of Success Than Others

Based on our learnings across the nine pilot sites, we suggest that to successfully scale PRO data collection in ambulatory settings using this model, several key factors need to be carefully considered relative to the practice. Specifically, the size of the practice, the general acceptance of technology, and the practice's general patient characteristics.

1. **Practice size and staffing:** In our pilot, we noted that the size of the clinic directly corresponded with key factors including eligible patient volume and the number of clinical and administrative staff. We found that larger practices had great variations in workflow, including less streamlined check-in processes, scattered communication between clinical and administrative staff, and patient coordination. For reasons such as these, data collection for PROs need to be effectively tailored to each practice and should not rely on a one-size-fits-all approach.

2. **Patient demographics and characteristics:** In this pilot, we observed that patient demographics were another key determining factor of successful execution and would be critical for determining scalability. Depending on the targeted patient population, practices may see many eligible patients within the same timeframe and will thus need to ensure they have the necessary resources to accommodate this.

3. **Acceptance and uptake of health IT**: We found that in general, the acceptance of health IT by the practice and staff was another key determinant of successful PRO implementation. In general, sites that were most open to new technologies also experienced the fewest issues in executing this pilot specifically related to workflow, patient recruitment, and clinical comprehension of the data.

The Content of the Survey Questions May Need More Careful Consideration Than the App

While the general SUS scores for all participants at all sites suggested average usability, we noted three key usability challenges with the app: 1) touch-responsiveness on the tablet, 2) skip patterns, and 3) readability and relevance of question items (which may have led to the responses that providers felt were not clinically meaningful). While the first two usability challenges related to the human-computer interaction (HCI), patient and provider feedback alike suggested that the content of the questions and, subsequently, the responses may have the most direct impact on the clinical utility on the PRO PF scores.

Chapter III: Integration and Implementation of a New App

PRISM Team Contributors
Steve Johnson, PhD (PI)
Pinar Karaca-Mandic, PhD (Co-PI)
Zach McGill
Doug DeBold
Elisha Friesema, MBA
Michael Borofsky, MD
Andrew Portis, MD
De Liu, PhD
Carla Pavone, PhD
Sue Neises

Overview

Using the PROMIS PF PRO measure, the pilot involved two distinct phases over a 14-month period: 1) the technical integration of a new PRO app at MedStar Health, and 2) utilization of the app at nine primary and specialty care practices in Washington DC, Maryland, and Virginia. The goal was to identify facilitators and barriers from the perspective of key stakeholders.

Section I: The PRISM App

Background for Challenge Competition and the Winning App

In 2018, AHRQ hosted the Step Up App challenge to develop new apps to collect standardized PRO data by using the PRO FHIR IG. In March 2019, AHRQ announced that a multi-disciplinary team from the University of Minnesota, Fairview/Health East Kidney Stone Institute (KSI) as well as a Minnesota-based technology company, Perk Health, won first prize and advanced to the final phase of the competition—the pilot stage—in collaboration with MedStar to pilot the app in nine clinics across three different EHR systems (two instances of NextGen and one in Cerner).

The primary goal of the winning app, PRISM, is to address the current deficiencies in collecting and using validated PROs as part of routine clinical care by developing a mobile app and saving these data in the EHR. For the development of the PRISM mobile app, the PRISM team conducted extensive user experience design and feedback sessions with patients from diverse populations (across different age ranges, clinical conditions, and technology comfort levels). Features of the app include score trending, peer group comparisons, relevant recommendations for follow-up actions, and provision of educational materials aimed at further engaging patients in their care.

PRISM App Technical Integration

Development of the PRISM App

Based on the goals of the PRISM app for collecting PRO data, the app has the following key features and functionalities:

1. Provide patients multiple ways to answer the PROMIS questions, before or during a visit, including having clinic staff help fill it out. Be accessible via multiple devices (including mobile phone, iPad, patient desktops, and clinic desktops), and be supported via a secure mobile app.

2. Display a patient's historical trends of their PROMIS measure scores over time, a comparison of the patient's score relative to other patient populations, and personalized recommendations of educational materials. These comparisons add value for patients in understanding PROMIS measures, which promotes patient engagement with the app and allows them to see the value in filling out the surveys.

3. Make it easy and quick to answer PROMIS questions, including using the CAT API from the Assessment Center to reduce the number of questions a patient must answer while still maintaining validity of the results.

4. Enable seamless integration with the EHR using the HL7 FHIR standard. If providers cannot easily find and review the PROs, then their value in the care process is lost.

5. Make it easy for organizations to deploy the app into their clinical operations. The Amazon AWS cloud was chosen to deploy the app, giving organizations a standard and secure way to create their own instance of the PRISM app and have it integrated with their EHR and comply with HIPAA.

A high-level overview of the PRISM architecture is shown in Figure 5.

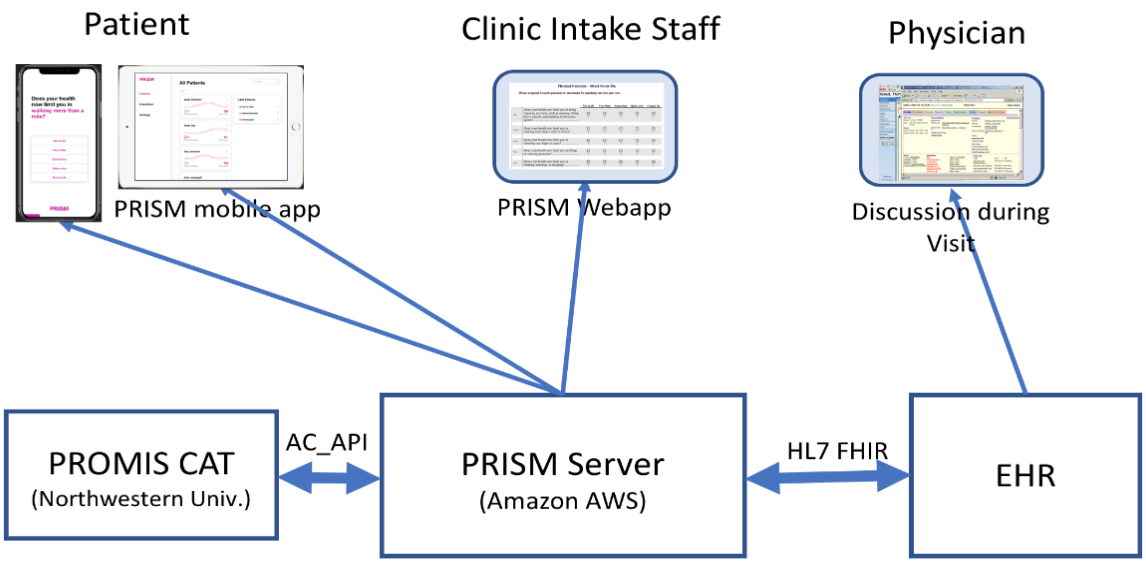

Figure 5. High-level app architecture

The resulting mobile app is easy and intuitive for patients to use. The app uses short message services (SMS) messaging to remind the patient to complete a survey and then walks them through the survey questions and finishes by calculating and displaying a score. The app also maintains a history of all the patient's scores and can display graphs showing trends and recommend educational material for follow up.

PRISM Usability Testing and Development Process

PRISM's development path took a multi-step approach. During usability testing, the team reviewed current clinic workflows to assess whether and how clinics currently collected PROs. Based on the learnings, the prototype was further developed into a clickable wireframe that was tested using a structured interview protocol with additional patients. Three cohorts of patients were interviewed for this phase of usability testing. After each cohort of patients, the team reviewed observations and made iterative changes to prototype to offer to the next patient cohort.

In general, patients found the app easy to use and found the text message and push notification features to be highly favorable; however, they asked that the scoring and the social comparison be explained better. It was determined that those who lack access to cell phones with data plans or email would not be target users for PRISM. The critical functionality to adoption was determined to be the app's ease of registration and use, text/push reminder notifications, long-term results tracking, educational materials, and seamless integration into EHR systems. Based on these results, a final iteration of the prototype was developed for use in the MedStar pilot.

PRISM App Architecture

The PRISM app was designed and built using modern open web standards. The other guiding principle in designing this app was not to reinvent the wheel and to use pre-built, industry-standard resources for hosting, security, and data persistence. This principle led to the development of a serverless architecture where the chosen cloud provider hosts and scales the code. Serverless architecture provides several benefits including speed of development, security, and scalability (Figure 6).

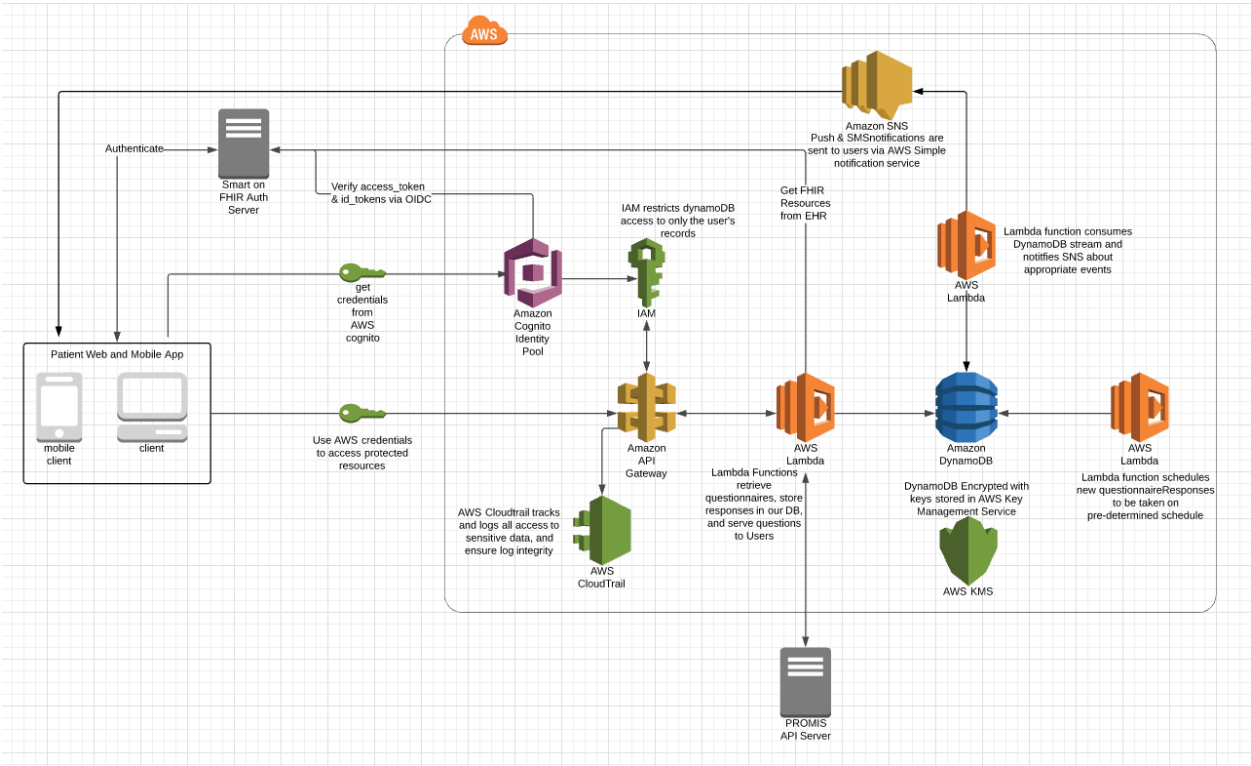

Figure 6. PRISM AWS app architecture

HIPAA Compliance

The PRISM app complies with HIPAA and is architected to allow PHI and the PROMIS measures collected by the app to be stored and maintained both within the EHR and outside of the EHR. The app is hosted on secure HIPAA-compliant servers from AWS and uses industry best practices for healthcare software development and delivery to ensure HIPAA compliance.

FHIR Standards and EHR Integration

The PRISM app uses the latest FHIR protocol (R4 v4.0.0) of the FHIR spec released December 27, 2018, to implement the following:

Authentication and Authorization: When a user launches the app by opening the native mobile app, the first step in the SMART Standalone App Launch[14] (http://www.hl7.org/fhir/smart-app-launch/) is to authenticate with their EHR's FHIR authorization server and authorize the SMART app using the standard SMART Patient Standalone Launch Sequence.

Login Redirect Screen Sequence: Once the user completes authorization and logs into the app, they are redirected to the app to complete the survey.

Retrieving and Displaying PROMIS Surveys: After completing authentication, the user's device begins communicating with the server using the FHIR API. First, the client retrieves the stored patient record. The device then searches for all completed PROMIS measure questionnaires for the patient. The app brings the user to the Patient Home screen which displays the history of these scores.

During the PROMIS survey administration, the client interacts with the PROMIS API to administer the CAT. As the user completes each question, their answer is sent to the CAT server (using FHIR to interact with the server hosted at Northwestern University) and gets back the next question. Each progressive answer is saved in the items and contained questionnaire per the FHIR adaptive structured data capture (SDC) specification.

Upon survey completion, the questionnaire and the responses are saved and returned to the client. The results are calculated such as how long it has been since the last completed questionnaire and what the t-score difference was. At this point the questionnaire is returned to the FHIR EHR endpoint for long-term storage.

Section II: PRISM App Integration with Different Electronic Health Records

Of the participating pilot sites, six sites were on Cerner's EHR and three were on NextGen. The technical teams purposefully started the integration phase with MedStar Health's Cerner instance. This allowed for more rapid and iterative development due to our degree of familiarity with the platform. Lessons learned in the first EHR integration were used to refine the process for integration into the NextGen EHR. The primary focus for this technical integration was using the PRISM app on different user devices (bring your own device – BYOD) and EHR integration.

Scalable Architecture for EHR Integration

The solution's architecture was designed to be scalable, modular, and web-based, thereby reducing the level of effort to deploy the provider-facing visualization in the EHR. In this implementation, OBERD's backend solution was replaced by Perk Health's backend solution, referred to as "the Hub." Because we implemented the provider visualization to communicate with the original OBERD backend with SMART on FHIR specifications, and as the Hub was implemented to follow the same outbound specifications for FHIR APIs, we did not expect many changes, nor did we anticipate having any issues during implementation.

Initial Challenges:

Though we did not need to make many changes to the architecture for connectivity and data transfer, we did undergo significant challenges making the Hub work with the NextGen EHR-based provider visualization app. This was due to an inability to connect and transfer data when the connection was moved from OBERD to the Hub. Most of these issues were related to very old Internet Explorer control that NextGen was using to render the provider visualization app. Even after multiple joint debugging sessions, this issue was difficult to resolve, as the NextGen EHR did not offer descriptive error logs or mechanisms to probe deeper into the underlying connectivity issues.

Solution:

We ultimately resolved this issue by creating another SMART on FHIR wrapper in a different technology stack, which was able to communicate successfully with NextGen, which in turn communicated with the Hub. EHR vendor support for SMART and FHIR should continue to improve over the next couple of years, so issues like this are not as likely to present such a challenge in the future.

We developed a "data on demand" architecture, which reduced points of integration, helped with variances between EHR systems, and reduced the effort of local IT support. We found all those benefits not only remained true during this integration, and in a way got validated based on how little additional work was required to make the Cerner integration work. However, the NextGen integration presented some new challenges, which we detail further in *Section V: Technical Implementation Findings - PRISM Integration with the Hub*.

Integration of the PRISM App

This pilot included replacing the app from the previous pilot test with the BYOD-based PRISM app running on personal mobile devices. Data captured by the PRISM app was surfaced in the provider-facing visualization inside the EHR. The original OBERD app was a vendor-hosted web app that the patient was able to access while on clinical site. This was a different format compared with the PRISM app which the patient could download on his/her personal device using standard app installation procedures. The PRISM app allowed the user to login using his/her own credentials (user ID and password) that the patient created during the first launch of the app.

In order to surface the provider-facing visualization inside the EHRs, the MedStar team worked within the EHR's unique framework. Both NextGen and Cerner systems have pathways to surface custom visualizations using modern web programming languages, which are summarized below.

- **Cerner**: Cerner provides customized visualizations known as Cerner Millennium Pages (mPages). The Cerner mPages have access to both the Document Object Model (DOM) in which the custom code resides and the launch context of the chart in which the visualization is surfaced.

- **NextGen**: The NextGen system provides a native functionality to allow system developers to create composable interfaces called Dynamic Templates. These templates can include custom web components (DOMs) which are able to dynamically load client-side HTML, CSS, and JavaScript. These dynamic templates can dynamically retrieve content by communicating with external hosts.

Initial MedStar Health Systems-Level Approvals and Processes

There were several levels of institution-level approvals required to implement the PRISM app. Introducing a new non-native app into the MedStar IS environment required us to apply for security vetting through the institutional "Demand Management" process, the primary goal of which is to ensure any new app does not pose a security threat to existing systems or introduce any new vulnerabilities. While the approval processes in the first pilot test were largely the same, the level of scrutiny differed due to specific details of the solution for the PRISM app, regardless of architectural similarities between the two solutions. Further, the OBERD app used in the Phase I pilot had been previously cleared through the health system IS processes, since it was already in use.

The following decision points posed the greatest complexity in the approvals processes:

- **Use of cloud platforms**: Cloud platforms are largely preferred in solution architecture for their flexibility, scalability, and functionality. They are easier to instantiate compared to physical servers, and they provide pre-built functionality specific to SMART on FHIR apps. Although generally preferred in most industries, these environments lack adoption and familiarity in healthcare. While the use of cloud platforms has the ability to expedite development, deployment, and cost, their additional scrutiny from security may cause undue delays as was the case in this pilot.

- **Patient-facing language and disclaimers**: The health system's legal and compliance teams required the inclusion of additional language, which added an additional step to our approval process. When a patient opts in to use their own device and downloads the app, compliance mandated there be a message related to what the app and any data collected can be used for by the health system.

The health system does not currently have an expedited means to review and clear new technology being used for a limited time within a smaller scope and as such is subject to the same lengthy process. This institutional barrier caused delays in "spinning up" solutions for time-limited pilots.

Section III: Pilot Site Implementation Processes

Recruitment and Selection of Clinical Pilot Sites

The processes used in Pilot Test 1 were also used to recruit pilot sites and patients in Pilot Test 2. This included the same eligibility criteria for patients, and the same 'top-down' approach to identifying potential practice sites first through leadership buy-in and referral, and then snowball sampling. Three sites from the first pilot study were also recruited based on their level of engagement and provider interest in participating in the second pilot. An additional facilitator of recruitment was the engagement of a research coordinator internal to the MedStar network who was able to recommend engaged providers based on prior research studies. This resulted in the successful recruitment of the final three ambulatory practices.

Participating sites for this pilot were all MedStar sites ranging in practice size from four providers, to one site with up to 24 providers. Patient volume among these sites varied significantly with some sites seeing 30 patients a day to some closer to 90 patients a day. All sites were primary care practices and were a mix of suburban, urban, and rural sites. Each site received a $2,000 participation incentive upon completion of the pilot.

Assignment of Participating Practices: Tiers of Implementation Support

A key objective of this study was to determine the feasibility of practices successfully adopting and integrating the PRISM app into their native system and implementing it in their workflow (outside the context of a research study). We developed a three-tier system based on the level of assistance provided. The protocol was broadly divided in two parts:

1. Research tasks such as recruiting, consenting, and compensating patients who chose to participate in the study.

2. App-related tasks such as registering participants in the PRISM admin hub, sending participants the app invitation, troubleshooting issues with the PRISM app, and notifying physicians of survey completion.

The three tiers of implementation support are briefly described below, and tier assignment is discussed in *Section IV: Data Collection Procedures and Tools - Evaluation Metrics and Data Analysis*.

- **Fully assisted.** Fully assisted sites received the most support from the research team during data collection. In this tier, the research team was responsible for both research and app tasks. The clinic staff's involvement was limited primarily to informing researchers when eligible patients arrived and, at some sites, notifying physicians of survey completion.

- **Partially assisted.** At partially assisted sites, the research team was responsible for research tasks, and clinic staff were responsible for app tasks. However, research staff were on site for ongoing training and as a troubleshooting resource for the clinic staff in case any issues arose during data collection.

- **Independent tier.** Independent sites received no support from the research team. Instead, they relied on a site-specific research coordinator to perform all research and app tasks, from identifying eligible patients to compensating participants who completed the study.

Development of Patient, Staff, and Provider Training Materials

To facilitate the implementation of the PRISM app at clinical sites, the MedStar and PRISM teams collaborated to develop instructional materials for patients, practice staff, and providers. The design process for all materials was based on a workflow analysis of the PRISM admin hub and patient-facing app, human factors principles for the design of instructional materials,[15] and multiple rounds of iterative design and feedback from team members with expertise in design, technology development, and field research. The overarching goal of this process was to provide standalone guidance for all users of the PRISM technology in the absence of support by the research team, as would be the case for this technology outside the context of this study. The main design features of the patient, staff, and provider materials are described below:

- **Patient materials** were designed to provide guidance on the registration and use of the PRISM app by older adults. The patient user guide included instructions for patients to create their PRISM account, download the PRISM app, take the PF survey, and finally complete the SUS survey. These instructions were tailored for the needs of older patients by establishing the context of this app, concrete and active language to describe the steps needed to use the app, large font sizes, and a combination of screenshots and text to support comprehension.

- **Clinic staff materials** were designed to provide guidance on the use of the PRISM admin hub and troubleshooting for potential issues patients may have with the PRISM app flow. The instructions for the admin hub were designed using similar principles as described above and included specific instructions about which patient identifier to use from the her, which varied across EHRs and is necessary for the proper functioning of the app. Additionally, a troubleshooting section was created from a failure mode analysis of the PRISM system that highlighted potential issues such as the need to validate or edit patient information, reset passwords, or assist the patient in downloading the app. Finally, contact information for the research staff was included in case the clinic staff was unable to troubleshoot an issue with the system.

- **Provider materials** were designed to direct the provider to the PRO visualization in the EHR. As the processes for this step were unchanged from the first pilot test, these materials were reused for this study as well.

Final AHRQ-approved versions of the training materials were distributed to practices during the site kickoffs to be used as guides or job aids throughout the process of data collection.

Section IV: Data Collection Procedures and Tools

Evaluation Metrics and Data Analysis

The same evaluation and metrics used for the Pilot Test 1 were used for this pilot. These metrics and evaluation methods are described fully in Chapter II, *Section III: Implementation - Evaluation and Metrics*. The only new evaluation method for this phase of the pilot test was the use of an approach log for further patient-level assessment, capturing metrics including the number of patients approached and reasons for refusal. Quantitative data from the approach log were used to produce recruitment tables for each site, tier, and across all sites. Qualitative data include the reason for refusal and patient observations. These qualitative data were then tabulated to provide reasons for decline and workflow breakdowns. These data were further reconciled with clinic observation notes to validate specific issues and to highlight key breakdown areas in the app workflow.

Section V: Technical Implementation Findings

In this section, we outline the findings from the technical team based on experiences during integration of the PRISM app with two different EHRs in the MedStar system.

PRISM App Implementation: Challenges and Lessons Learned

As designed, the PRISM app is intended to be very easy to setup and maintain. When a health system wishes to start using the PRISM app, they will create a new secure instance of the PRISM server environment and then configure the environment to interact with their specific EHR. The section below lists detailed steps taken when the app's ecosystem was deployed.

1. Deployed a new instance of the app within the health system's AWS environment, which was accomplished by using the provided scripts to ease deployment.

2. Configured the public-facing web services belonging to the health system.

3. Added the PRO measure. The health system comes pre-populated with the PROMIS PF measure. The health system can add any additional PRO profiles to the app.

4. Configured the username/password for the health system's credentials to the CAT server.

5. Configured the file transfer protocol FHIR endpoint where the completed PROMIS surveys would be sent.

6. Configured the cloud service leveraging the health system's authentication pathway (i.e., patient portal credentials verification).

Challenges

Setting up a new AWS instance can be time consuming: Medstar fortunately already had a relationship with Amazon and, therefore, had a business associate agreement (BAA) agreement already in place that covered the work proposed for this pilot. For organizations that do not have an existing account with Amazon, the process of agreeing to a BAA may take several weeks and is dependent on the organization accepting the standard AWS BAA.

Support for the app under AWS can be time and resource intensive: It was relatively straightforward to work with the MedStar technical operations staff to setup a new PRISM instance. They were then able to delegate administrative control to the PRISM team to support the app during the pilot. However, if MedStar had to support the PRISM app under AWS on its own, more training would have been required of the MedStar staff on how to monitor the health of the app and address issues that arose.

PRISM Integration with the Hub

As part of the deployment of the PRISM app at MedStar, the FHIR server that the app was using had to be switched from the Cerner Sandbox to the Hub. The Hub was meant to be a bridge between the patient-facing PRISM app and the different EHRs that were in use at MedStar. The Hub would be used to bridge the version differences and adjust for other issues. In general, this process went quite smoothly; however, there were some minor sticking points due to differences in versions causing slight incompatibilities.

1. The PRISM app was built using FHIR R4, but for expediency, the Hub was built using STU3. This led to some communication issues when archiving the completed PROMIS measure responses back to the Hub.

2. SMART App Launch supports authentication, an optional extension to the SMART App Launch Sequence. The PRISM App leverages this to federate identity from the EHR identity provider to the app. Due to the configuration of the testing environment, parameters had to be changed to when integrating with the Hub.

Overall, having well laid out standards allowed the PRISM app to quickly and easily switch from one FHIR server to another—a process that would be far more complicated without the standards.

Cerner vs. NextGen Specific Differences in Provider Visualization App Implementation

A finding from the Pilot Test 2 was the importance of having direct access to a testing environment. For a developer, it is critical to have access to test independently to prevent bottlenecks in development. As the MedStar technical team was unable to secure direct access to even a NextGen development environment, each testing session had to be led by a member of the NextGen's IT team. This finding underscores the impact of delays in development and the ability to rapidly iterate on the visualization.

Key Findings and Lessons Learned from the Technical Integration of the PRISM App with Two EHRs

During this pilot test, the technical team encountered several challenges yielding important lessons learned for other developers or technical teams attempting to achieve a similar implementation. These unanticipated EHR integration challenges were instructive regarding the variability in EHR implementations and the variability that can occur even while implementing two nearly identical in-situ provider-facing solutions. Between the first pilot test with the OBERD app and this pilot with the PRISM

app, only the patient-facing devices and hosting of the data hub changed. The architecture and deployment of the provider-facing visualization otherwise remained the same, yet the technical team encountered the following challenges. We highlight each one in turn and relate each set of challenges to the related dimension within the sociotechnical systems model where appropriate.

Cerner:

- **Key challenges:**
    - *(Related STS Dimensions: 1) hardware and software computing infrastructure)*

    - One unanticipated challenge was switching the data hub from the instance built by OBERD to one built by PRISM and Perk Health Teams. In both phases, the data hub was responsible for communicating the PROMIS measures recorded by the patient's completion of the survey. Surprisingly, while functionality and architecture remained the same between the two phases as well as their conformance to the IG specifications, some modifications were required to make the PRISM app work.

- **Lessons learned:**
    - The lesson learned from this challenge is to include as much detail as possible up front, when delivering functional specifications to vendors who are tasked with building a data store for PROMIS measure questions, responses, and scores. Although vendors in Pilot Tests 1 and 2 were able to independently conform their solution to the PRO FHIR IG, it is critical to anticipate variance because the specification allows elements that are required vs optional, and apps need to be able to handle the differences in implementations.

NextGen:

- **Key challenges:**
    - *(Related STS Dimensions: 1) hardware and software computing infrastructure)*

    - The most significant challenges in implementing the EHR provider-facing visualization surfaced during implementation of the PRISM app in NextGen. Although the functionality and protocol for communication with the data hub remained the same between Pilot Test 1 and Pilot Test 2, there were additional unexpected consequences of changing their respective internet protocol (IP) addresses. Each health system's EHR has specific firewall rules governing which IP addresses or sets are open for communication. While testing the solution using our integration testing app, built to detect and diagnose communication issues in Pilot Test 1, the teams established that the provider-facing visualization in the EHR was unable to communicate with the data hub through the pathway in Pilot Test 1. Though repeated attempts were made by both the Perk Health and MedStar team to diagnose the problem, the teams were unable to troubleshoot due to the lack of modern debugging tools in the EHR. The problem was finally resolved by communicating with a FHIR API surfaced on a different IP as part of the same server that contained the data hub.

- **Lessons Learned:**
  - Simply identifying a potential issue in advance does not ensure it will be resolved expeditiously. Despite building and deploying a digital communications testing suite to detect communication and extensive server logging, the lack of modern debugging tools in the EHR made this process not only cumbersome but ultimately futile. This underscores the importance of not only anticipating challenges but also the need for contingency planning in the event they materialize.

Integration of the Provider-Facing Visualization: Challenges and Lessons Learned

As the initial architecture integrated the provider visualization app using loosely coupled architecture, we anticipated very minimal changes in the base code while integrating this app with Perk Health's Hub from the original modified app's backend.

For the Cerner integration, we changed security artifacts including security keys, secrets, and certificates on the backend. This enabled full communication. We were able to get FHIR data using the same function signatures and were able to render without any issues.

For the NextGen integration, the team faced significant and unexpected issues. The primary issue was because the NextGen EHR being operated during this pilot was older than the instance used during the first pilot test. The version of IE control being used to surface the provider-facing app during the initial pilot was antiquated and did not support modern web functionality. This posed significant challenges as soon as a security layer (HTML headers) was added, resulting in communication being stopped.

This challenge was exacerbated by the fact that the EHR did not have any tooling for debugging or error trapping. Despite several attempts to fix this, communication between the Hub and EHR did not work. The final solution required us to develop a wrapper around the Hub's FHIR API, which enabled communication and an ability to surface the visualization in NextGen, which mirrored the solution enabled in Cerner.

Early access to and knowledge of the exact EHR setup for testing communication and rendering capabilities can help with anticipating and designing mitigation strategies for any possible communication errors. Successful mitigation strategies depend on access to a development environment which mirrors the targeted platform upon which the application will be deployed. Without access to such a testing environment, creating alternative implementation pathways and mitigation strategies included rollback procedures is exceedingly difficult.

Technical Implementation: Developer Perspective and High-Level Recommendations

EHR Integration – General Context

Based on our prior experience designing provider-facing apps that cater to a single patient, we decided to use the EHR as a gateway for user authentication. This alleviated the need for users (providers in this case) to authenticate again and reduced the need for IT to manage another set of login information. This approach provided additional benefits for scalability and usability of the solution for provider-facing apps. Throughout the Pilot Test 2 technical implementation, we faced some unexpected and unforeseen challenges. While any technical integration is likely to have some unanticipated outcomes, our experience provided us with some general mitigation strategies that may assist future developers with a similar implementation. The two key challenges related to implementing security from the beginning, and

factoring in change management plans. These two factors both posed challenges over and above the level experienced in Pilot Test 1 because this was a completely new app to the MedStar system.

Technical Challenges for Patients Using the PRISM App: Bugs and Issues with the Mobile Platforms

The PRISM App encountered several bugs and issues during the pilot. The team was able to troubleshoot these issues in real time and, in every instance, resolve the issue to facilitate continued data collection. The identified bugs and their resolution are detailed below.

Passwords

- **Challenge:** The workflow for a patient was to receive an SMS message with a unique link, follow that link to a page where they needed to set their password, download the mobile app, then log in using the password that was just created. A significant number of patients ran into difficulty remembering the password that they had just created.

- **Lessons Learned:** There are two ways to mitigate this issue. First, use a mobile web app instead of a mobile app so that the patients do not have to log in a second time. Second, use different OAuth web containers for iOS (ASWebAuthenticationSession) and Android (AppAuth) to leverage the existing session in the mobile browser to log in.

Font Sizing

- **Challenge:** Some users, especially older patients, had set their device font sizes to be much larger than standard to help them see the screen. This would lead to text flowing off the screen and not being readable while they were attempting to use the PRISM app.

- **Lesson Learned:** Testing on many different devices and testing each device with various font size settings is important to ensure that apps display correctly on all devices.

Google Play Store Links

- **Challenge:** Some Android users did not have the Google Play store installed on their device. This made the process of downloading an app slow and difficult, as it added an extra layer of activity.

- **Lessons Learned:** Downloading an app can be difficult for some patients. A simpler alternative to downloading an app is using a web browser to access the FHIR app. This can be helpful in accelerating the adoption of a mobile app.

Lack of Cell Phone Service and Connectivity

- **Challenge:** Two of the nine sites did not have adequate cell reception for mobile data based on their geographic location and did not have public Wi-Fi available. In these situations, the mobile app was rendered completely unable to function.

- **Lessons Learned:** To be inclusive and accessible to patients even in areas with limited cell service or internet connectivity, the app should be amenable to function via a web-based platform that can be accessed via an amplified Wi-Fi signal if clinics are able to use a hotspot or other means to enable internet connectivity that does not solely depended on cellphone service.

Observed Challenges

The team observed additional challenges to inform important findings, such as the importance of end-to-end testing of the system. When building an ecosystem architecture to facilitate the capture of PROMIS measures initiated in the clinical setting, completed by the patient, then visualized by the provider, a walkthrough is key to successful launch. Although the MedStar, PRISM, and Perk Health teams performed rigorous module testing of each component, there was less integration testing to ensure each component of the patient- and provider-facing architectural components functioned properly together. One such example was the importance of practicing and implementing a system for provisioning providers to access the visualization. The MedStar team codified a process by which upcoming sites go-live by closely tracking user provisioning requirements, as well as incorporating previous lessons learned. This key finding speaks to the importance of establishing all systems prior to go-live aside, from the app's functionality itself.

Section VI: Clinical Implementation Findings

Practice Site Assessments

Readiness for Change Themes

Based on our review of the readiness for change data, we made several general observations that helped us characterize our participating sites according to the three core categories on this assessment.

1. **Overall acceptance of health IT across pilot sites:** Most sites were generally accepting of and had positive attitudes toward new health IT. Within the "Overall perception of e-health" domain, all sites either strongly agreed or agreed with the statements: "e-health technology can improve patient outcomes," "e-health technology can enhance a team to approach care," and "I usually try hard to learn how to use new e-health technology." However, individual responses revealed an underlying feeling of dissatisfaction with the state of current health IT systems.

2. **Pilot site perceptions of organizational readiness:** These items assessed the extent to which sites felt ready to implement and use e-health technology in their practice. There was wide variability among sites regarding their perception of organizational readiness.

3. **Pilot site prior experiences with health IT:** While almost all sites reported positive experiences with prior and existing e-health technology, two sites notably remarked on having negative experiences in the past.

Readiness for Change Scores

Scores across sites suggested a wide variability in terms of prior experiences with e-health as an organization with no trends in tiers. It is also worth noting that the practice manager (or corresponding individual) filled out the survey measure for each site. Results should therefore be interpreted cautiously due to the subjective nature of the questions.

Summary of Recruitment and Clinic Observations

The following results describe the overall recruitment and clinic observations from the study across all sites and by tier. First, the workflow observations describe important variables at each site, such as age and socioeconomic status of the population. The section also details facilitators and barriers to successful

data collection at the site. The data below also show the most common reasons that patients did not complete the study, including reasons that patients declined to participate, and problems encountered while attempting to download and use the app.

Site Characteristics Across All Tiers

Recruitment Across All Sites

Of the 173 total patients approached to participate in this study, 84 patients consented to participate, and 58 patients successfully completed the survey process. The most common reason patients declined to participate in the study was due to cellphone issues (49%), or patient-specific issues (22%). The most common workflow breakdown occurrences were a result of technical issues (42%) and patient complications (39%).

Not all sites reached the target accrual of 10 patients due to insurmountable barriers to data collection (i.e., no cell service or lack of eligible patients). These data are explored further in the sections below, which highlight barriers and facilitators to recruitment, enrollment, and completion by practice tier groups.

A series of breakdowns occurred during data collection, which prevented the consented patients (n=84) from completing the study. Potential points of failure resulting incomplete survey completion include: registration, receive the link to the registration page, create a password, access the App Store, download the app, log into the app, and complete the survey. Workflow breakdowns include 12 patients who would not have been able to complete the survey without assistance (i.e., could not download the app and used the researcher's phone instead). Looking at all sites, the most common reasons that the patients failed to complete the study were problems receiving the link to the registration page (n=9), creating a password (n=8), and inability to download the app (n=7).

Barriers and Facilitators Across All Tiers

Based on observations and approach log data across sites, major facilitators and barriers to data collection were identified. A summary of the facilitators and barriers across tiers is presented below.

Facilitators Across All Tiers

1. **Dedicated Research Staff.** Having dedicated support for data collection facilitated the process because this person could focus their time on the study tasks and the patient's needs for the study. Another advantage of having a dedicated staff member was limiting the need for communication between the researcher and clinic staff at multiple steps of the process, which could either delay data collection or result in a missed opportunity to enroll a patient.

2. **Staff Preparation and Workflow Accommodation.** Staff at multiple sites were able to contribute to data collection by performing preparatory tasks and adapting workflows to streamline the process (e.g., flagging patients in the EHR on mornings of data collection, pre-registering patients in the admin hub before data collection, modifying their workflow to allow patients extra time in the waiting room or exam room to complete the study).

3. **Clear Roles and Responsibilities.** At most sites, researchers and clinic staff were able to clearly define roles and responsibilities. These clear goals and responsibilities made sure that steps in the data collection process were completed in a timely fashion.

4. **Pre-existing Cell Communication Processes**. Sites that used text message reminders for appointments typically had patients' cell numbers on file. The patient's cell number is a required input when registering patients in the admin hub, and it is essential for sending the patient a link to the password creation screen. Sites that had patients' cell numbers on file were able to preregister patients before they arrived, streamlining the process.

5. **Staff Rapport with Patients and Providers**. Clinic staff at some sites recommended who the researchers should target throughout the day based on inclusion criteria (English speaking and ability to consent) as well as general knowledge of the patient's personality, technology usage, and relationship with the clinic. This resulted in more efficient data collection efforts.

6. **Provider Interest/Clinical Champion.** Sites with highly engaged providers and staff were more accommodating to the research study and input additional time and effort into completing tasks (e.g., prescreening and preregistering patients) that led to successful data collection. Engaged providers screened patients for whom they thought the study would be useful and encouraged patients to talk with the research staff. Provider buy-in also encouraged practice staff to accommodate workflow or allow more time for the study.

Barriers Across All Tiers

1. **Staff Workload.** In fully assisted and partially assisted tiers, app tasks typically fell to the front desk staff, who would often already have extremely high workload demands answering phones, checking in patients, and supporting the practice otherwise. Because of their high workload and low priority of the research study compared to core clinic functions, the front desk staff were not always able to complete study-related tasks such as inputting patients into the admin hub, sending the text message, or communicating survey completion with the provider.

2. **Tight Clinic Schedules.** There was significant variability in the timeliness of the different clinics. Some clinics were always on schedule and MAs began the intake process as soon as the patient arrived. Other clinics ran behind, which would give time for patients to complete the study in the waiting room. It was more difficult to recruit patients at clinics with timely schedules because the patients had less time to complete the study before their arrival and their appointment. Workarounds included approaching patients in the exam room instead of the waiting room or not approaching patients in the morning when the clinics were typically on schedule.

3. **Unclear Roles/Responsibilities.** There was some confusion about who was responsible for completing tasks involved in the study. At one site, there was confusion about who would input patients into the admin hub; initially it was the responsibility of the practice manager, but when it became clear that she did not have time to complete the tasks it was pushed to multiple other parties, none of whom truly knew who was responsible for the task. Consequently, preregistration was not always completed at this site before data collection.

4. **Lack of Pre-existing Cell Communication Processes.** Some sites did not have patients' cell numbers on file. This required the research team to get the patient's phone number from them after consent, relaying the number to the front desk staff, and waiting for staff to update the cell

number in the admin hub before finally sending the text message. This process took between 2 and 10 minutes, depending on the workload of the front desk staff, lengthening the time of the patient's participation and making it more likely that the patient would be called back for their appointment before completing the study.

5. **Poor Cell and Internet Service.** Sites that experienced poor cellular connection struggled to complete data collection. Due to poor cell service, patients were unable to receive the text message link to begin the registration process, nor were they able to download the app from the App Store. This was a critical hindrance to patient accrual and data collection in two of the independent sites, which were all located in rural areas with limited cellular signal.

6. **Clinical Workflow Delay.** At one site, the provider noted after the kick-off that they would not be able to look at the data before or during the appointment as they were concerned about the study delaying their appointments. Another provider allowed the research team to approach patients if they were early enough for their appointment but made it clear not to approach patients if it would delay their scheduled visit time at all.

Site Characteristics Within Tiers

Below is a description of the different patient population characteristics of the clinics. These characteristics are not inherently barriers or facilitators; however, they did appear to influence the success of data collection to varying degrees.

1. **Socioeconomic Status/Technology Restrictions.** There was significant variability in the socioeconomic status of patients approached during the study. Anecdotally, the patients recruited in typically higher socioeconomic status areas were more likely to have iPhones rather than Android phones. The opposite was true for patients recruited in clinics that were in areas with lower socioeconomic status. Here, Android phones were more common. When asked if they had a smartphone, a handful of patients in the lower socioeconomic status areas shared that they could not afford one. Consequently, it appears that socioeconomic status is a key factor determining whether patients have the necessary technology to access the PRISM app.

2. **High Patient Volume.** Pilot sites where there was a high patient volume had more eligible patients; however, staff members at these sites also experienced higher workloads and therefore had competing demands trying to ensure the passage of patients through the clinic, while also incorporating research activities. Staff with higher workloads struggled to complete tasks associated with the study in addition to their normal duties.

3. **High Acuity Patients.** In some cases, the patient's illness or level of function due to illness or age prevented them from participating in the study. Some patients struggled to talk or move. Others had serious medical conditions that put them in severe pain or gave them trouble breathing. In some cases, the front desk recommended that the research team not approach a particularly unwell patient. In another case, an eligible patient was rushed to the emergency room.

4. **Cancellation/No-Show Rate**. Clinics had many no shows or cancellations. This greatly reduced the number of patients that researchers could approach for the study and minimized the chances of reaching goal recruitment.

5. **Age.** In general, the younger patients encountered during this pilot tended to be more tech savvy and need less assistance with typical smartphone tasks, such as clicking the registration link and downloading the app. Interestingly, these patients still faced trouble creating passwords and remembering their App Store ID.

Patient-Level Assessments

Below we describe the findings from the patient-level assessments, including decline reasons from the approach log, usability survey, patient observations, and the patient semi-structured interviews.

Decline Reasons

Phone issue declines included incompatible devices (e.g., a non-smartphone or internet-enabled device or device left at home) or issues with phone battery. This was the most common reason for decline among patients. As one patient said, "I'm old school, my phone is just a regular flip top."

Patient issues included concerns about the app download process, issues with appropriateness of the timing (e.g., patient preoccupied with another task or focus), poor health (e.g., patient not feeling well enough to participate or physical function preventing technology use), privacy concerns, and concerns with the process being too complicated. Quotes highlighting some of these reasons are included below:

- App Download: "[I already receive] so many appointment notifications and messages. It takes me too long to clear everything off of it."
- Privacy concern: "I don't put apps on my phone. Everything on my phone is confidential."
- Too complicated: "Do I have to put something on my phone? I don't know how to use it."

Patients who heard the study description before choosing to not participate were categorized as informed declines. In these cases, the exact reason for declining was unspecified by the patient. Most commonly, informed declines were accompanied by a general lack of interest in participating in research.

Immediate declines were categorized as patients declining to participate before hearing out the description of the study. Patients who immediately declined were simply not interested in anything outside of their scheduled visit and had no interest in conversation with the research staff.

The final reason for decline was an inability to consent. Only one patient was recorded as such, as they were called back to the exam room after the researcher introduced themselves but before the patient was given the opportunity to consent.

Systems Usability Survey

The SUS is scored between 0 and 100, which is a score specific to SUS and not a percentile. Bangor et al. (2008) ascribed the acceptability of SUS scores compared to the average.[13] A SUS score of 0-50 is deemed "unacceptable," 51-70 is considered "marginal," and 71-100 is deemed "acceptable." Data analysis was performed on seven sites that completed the SUS evaluations, resulting in a total of 51 SUS evaluations. Sites that reached 10 patients did not necessarily report 10 SUS scores. Some scores were not captured, as patients failed to complete the entire survey. Scores from two sites were not captured, as PRO data

were not captured for these sites due to lack of accrual. The overall median SUS score across all 90 patients was 47.50 (range=[21.25, 82.50]).

Patient Observations: Usability

Based on observations of participants interacting with the app and recordings from the approach log, we identified several usability issues. These included password creation, App Store ID, and comprehension and applicability of survey items.

1. **Password Creation**
    - *(Related STS Dimensions: 3) human computer interface; 4) people)*

    - During data collection, participants were required to create a password for their PRISM account. This password needed to have a number and capital letter. Identical passwords also needed to be input twice. Many users struggled to input a password for three main reasons. One, the users did not understand the password requirements before creating their first attempt at a password. Two, users did not accurately input two identical passwords. Three, older users and those with mobility issues struggled to use the keyboard on the small phone screen. In total, approximately 24 of the 84 patients (28.6%) who consented to participate in the survey struggled to create their password. This was a critical failure point for five users who did not complete the study after consenting because of password issues.

2. **App Store ID**
    - *(Related STS Dimensions: 3) human computer interface; 4) people)*

    - A less common, but highly disruptive, occurrence was patients not having or not knowing their Apple/Google Play ID or password. To download an app in either store, the user is required to have an account with the company. Some patients did not have an App Store account. Other patients had clearly created an account; however, they did not know the password associated with the account. In five cases, this led to the study being stopped prematurely as without downloading the app, the survey could not be completed. In two cases, the researcher provided their own phone for the patient to use to complete the study.

3. **Comprehension and Relevance of Survey Items**
    - (Related STS Dimensions: 2) clinical context)

    - Commonly observed usability issues for most patients across sites were challenges with comprehension and readability of the PF survey questions. Patients often expressed that the survey content was not directly applicable to their lifestyle or context, and thus had trouble conceptualizing the questions to provide the most accurate response. For example, when one patient was asked whether he could row a boat vigorously for 30 minutes, he remarked that the question was ridiculous as he had never rowed a boat before. Another patient asked what "two hours of hard labor" meant in the context of the question.

- In addition, while filling out the survey, patients commented that the wording of questions and the formatting of the SUS was confusing. As mentioned above, this confusion introduces an important caveat and potential limitation to interpreting the SUS data.

Semi-Structured Interviews

A total of five patients opted-in to complete a follow-up interview to provide additional context and general feedback on their experience using the app. We were unable to retain representative patients from four sites for interview. Patients were interviewed and transcripts for these interviews were analyzed to extract commonalities and identify emergent facilitators and barriers. Patients ranged in age from 61 to 70 years.

Summary of Barriers and Facilitators for Implementation from the Patients' Perspectives

Most eligible patients who were approached declined to participate in the study. The primary barrier to consent was the lack of a compatible phone, which made it impossible for patients to download the PRISM app. Another significant barrier for patients was concern about downloading an app on their personal device or the privacy of their data. Many participants simply had no interest in participating in the study at all.

In general, patients found the PRISM app itself easy to use. Patients did appear to struggle with the required process of creating a password, perhaps due to the size of the phone keyboard and small text of the password requirements. Some patients also struggled to remember their App Store ID. However, overall feedback from the interviews and SUS data suggests that the survey function of the app itself was straightforward.

Unfortunately, the context and language of some of the PROMIS survey questions presented a challenge for some participants. Three of the four patients who completed the interviews felt like they were the right audience for the questions; however, researcher's observations suggest that many people found the questions confusing or in some instances 'absurd.'

Lastly, it appears that patients had mixed feelings about the impact of the survey on their visit. Three patients interviewed expressed that they found the survey meaningless, while others (n=2) said that it helped them better participate in their own care. This suggests that the survey has merit when targeting applicable patients and used by providers in an engaging way.

Provider-Level Assessments

Semi-Structured Interviews

After the completion of data collection at each site, semi-structured interviews were completed with a participating provider to understand their experience. Except for four sites, one provider from each site was interviewed. The exceptions included three sites with unavailable providers, and one site in which the provider did not have any success recruiting patients. To account for these discrepancies, we interviewed two providers from one site and the research coordinator of the independent tier who had a particularly unique role in the project. In total, seven interviews were conducted. Providers' primary specialty was family medicine or internal medicine, and they ranged in years of practice from 6 years to

24 years. The site coordinator was interviewed as a replacement to the provider. The site coordinator is not a practicing clinician with a specialty.

Interview questions for providers focused on usability and functionality of the EHR visualization, the utility of PRO measures and their impact on workflow, and patient/provider communication. Interviews were coded for consistent themes and analyzed to extract facilitators and barriers.

Summary of Provider Interviews

Overall, providers were satisfied with the PRO workflow and data collection process. While no providers reported the output having an impact on patient care, all providers reported value in collecting and reporting on PRO measures. From these interviews, providers emphasized the importance of having support on site to facilitate data collection, ease of access within the EHR and existing workflow, and guidance for PRO interpretation and action. Future projects should attempt to anticipate these needs when implementing PRO measures within ambulatory care.

Health IT Staff-Level Assessments

Semi-Structured Interviews

Upon completing of the technical implementation, we interviewed five health IT staff who were involved at various stages of this implementation. Questions related to FHIR compatibility with different EHRs, institutional requirements and policies, and the general success and pitfalls of the project.

Summary of Health IT Interviews

Overall, technical experts were satisfied with the interaction between project teams. They believed that the close working relationship between the teams helped to facilitate the sharing of essential information that moved the project forward. One exception was communication about institutional hurdles, which led to choices about cloud-based services that complicated institutional barriers even further. Future projects should attempt to anticipate and understand the scalability and regulatory needs before selecting technical solutions. Additionally, knowledge of these hurdles must be disseminated early and to all key team members to facilitate effective decision making.

Section VII: Summary of Key Findings by Stakeholder Group: Clinical Implementation and App Usage

In addition to the facilitators and barriers that have been detailed in the previous sections, we identified key lessons learned from the entire implementation process. These lessons learned have direct implications for scalability of the standards and PRO technology. We outline these by stakeholder group.

Patients Using the PRO App

- In general, the PRISM app was found by patients to be usable and easy to navigate for completion of the PROMIS survey.

- Downloading from the App Store or Google Play and password issues were common challenges for many patients attempting to engage with the app.

- Successful engagement with the PRISM app was highly dependent on patients having the appropriate technology (smartphone) and having the tools to make the technology work (i.e., a reliable cellular network or strong Wi-Fi connection).

- The PRISM app facilitated efficient survey completion (74 seconds on average), highlighting the potential for this technology to be scaled for a large battery of clinically relevant PROs.

- Successful engagement with the PRISM app in the pilot test setting was often dependent on the patient generally being in good health.'

- The value proposition for patients engaging with a PRO app is critical—patients have to perceive value in how their data will be used by their provider or impact their care. Many patients were skeptical about whether or not their provider was actually going to use the data they were reporting.

- Security and privacy concerns remain paramount for many patients and would likely impact their willingness to continuously use the PRISM app or similar technology.

- For many patients, the language and context of PROMIS survey items were potentially confusing (e.g., "row a boat vigorously for 30 minutes").

- The ideal timing for patients to complete their PRO survey can be highly variable and is not always optimal immediately before their appointment or in clinic.

Primary and Specialty Care Providers Using the PRO Data

- Providers play a key role in patients' willingness to use PRO apps to collect PF data, as patients were more willing to use the app if they knew the provider was looking at their data.

- Surfacing PRO data via a dynamic template in the EHR proved to be a viable solution allowing providers to easily access and consume the PRO data in the EHR.

- Many provider workflows are not conducive to accessing the patient's PRO data in the EHR prior to the patient encounter; alternate workflows need to be considered (e.g., for providers who do not review patient charts until after the encounter, or where patient volume limits this option).

- Providers expressed the need for flexibility in terms of what they could do with the PRO data (e.g., paste it in notes, save it to the chart).

- Provider notification or an EHR alert of a patient's survey completion is needed so as not to hamper workflow and to prevent providers from missing the data.

- Despite increasing use of PROs, some providers still expressed hesitation about their interpretation because it represented a patient's perception rather than an objective physiological measure.

- Providers generally still noted they were unsure of what to do with the PF PRO data and which scores were actionable.

- Several providers in this pilot suggested limited clinical utility of PF PRO data for primary care in this context.

- Certain primary and specialty care providers may benefit more from patients using this technology than others (e.g., providers in certain specialties who routinely track and interpret these data as part of a standard periodic evaluation, or who can pair it with a targeted intervention).

Developers and Health IT Staff

- Having well laid-out standards facilitated the PRISM app being able to quickly and easily switch from one FHIR server to another—a process that would have been far more complicated in the absence of such standards.

- The FHIR standards enable a successful abstraction layer to bridge the mobile app and the multiple EHRs—assuming that it is supported by the vendor and site.

- While SMART on FHIR technology is generally relatively easy to implement, many health systems do not currently have the tools or technology resources on hand to support it.

- Several EHRs do not currently support the latest SMART on FHIR standards, and some clinic sites may not be willing to purchase the app.

- SMART on FHIR solutions for PRO apps are only part of the solution of scaling to diverse populations where cell connectivity may be unreliable. Apps need to also support alternative architecture that does not rely on app downloads (e.g., web-based apps or embedded apps on clinic tablets).

Primary or Specialty Care Practices

- Practices most likely to adopt and implement this PRO technology would have dedicated staff to facilitate patient use and adoption, and a clear delineation of roles and responsibilities for staff supporting patients and the data collection.

- Practices with efficient communication between the different levels of clinic staff are most likely to successfully adopt this technology and integrate it with their workflow.

- Practices that see a high patient volume may experience challenges with engaging patients with this technology in clinic, given the lack of resources to also provide technical assistance to patients.

- Practices are likely to adopt and sustain this technology if they also have supportive technical infrastructure (i.e., Wi-Fi or connectivity).

- Providing technical assistance to practices implementing this technology may not necessarily be the key driver of adoption (fully assisted sites still had practice-specific challenges to adoption and use).

- The availability of Wi-Fi of and cellular service are critical to adopting and sustaining use of this type of technology. This can be especially challenging in hospitals in rural areas with poor cellular

service, or individual practices with specific architectural or physical constraints (e.g., lead walls installed for protection from nuclear medicine).

Institutions or Health Systems Implementing PRO Technology

- Involving all system stakeholders, including legal and compliance, as early in the implementation process as possible is critical to successful and timely adoption of this PRO technology, given the multiple policies and regulatory considerations.

- Systems should anticipate significant challenges of using cloud-based services, including security assessments, HIPAA, BAAs, and related internal processes.

- Institutions and health systems with complex and uniform vetting processes for adopting any new technology may benefit from standardized, expedited approval processes for integrating novel patient-facing health IT.

- Dedicated technical staff are needed regardless of the size of the implementation to ensure successful integration and adoption for different clinic workflows.

Chapter IV: Discussion, Crosscutting Themes and Conclusions

Discussion

Broadly, the findings from both pilot studies highlight important barriers and facilitators for PRO implementation, some of which align with the exiting literature, and several that generate important new considerations for the potential scalability of technology-enabled PRO data collection. Notably, these pilots (particularly Pilot Test 2 with the PRISM app) demonstrated that **for a technical implementation of this scale, the PRO app itself was rarely the most significant challenge; rather, the interplay of multiple sociotechnical factors impacted the implementation.** In general, key barriers arose in three areas of clinical implementation: technical and workflow integration, patient engagement, and perceived clinical utility. Across both pilot tests, a variety of communication breakdowns impacted the efficiency of patient enrollment, completion of the PF survey, and in some instances the providers' review and use of the PF data. While some of these issues, such as patient recruitment and consent, were particular to the execution of a research study, others, such as the usability of the survey app and the notification/presentation of the survey data in the her, highlight critical points at which the coordination of human and technical processes is crucial to ensure the successful use of PRO data.

Across each pilot site several broad themes held true:

The mobile form factor is particularly amenable for patients taking surveys. In Pilot Test 2, patients clearly demonstrated that the use of mobile phones is a great way to take PRO surveys in a way that is convenient and familiar to most patients. Once able to download and access an app, patients could efficiently complete a PRO survey with their own technology. In Pilot Test 1, patients were able to easily access and complete the survey already loaded on the tablet in clinic. While both methods worked well, the PRO literature does not stress the clear opportunity of using a patient's own technology to gain widespread collection of PROs from patients in primary and specialty care, in addition to the more commonly used methods in clinics using tablets and kiosks.

The apps themselves were usable and presented limited challenges for patients once they were able to access the survey. In Pilot Test 1, there were minimal challenges for patients accessing the PRO survey, as it was pre-loaded once the patient signed onto the app. For Pilot Test 2, patients found the app highly usable once they were able to access and download it. In general, it appeared to be mostly factors outside the app that hindered survey completion. For users of a range of technical abilities, the apps presented a good user experience; however, patient frustration, perceived utility, and value in the app were mostly related to factors external to the app itself.

Technical assistance or additional staffing is a critical factor to ensure adoption for many practices. In Pilot Test 1, where our research team was fully present and managed all the logistics of the implementation and patient enrollment, we achieved 100% completion of all the surveys for patients who successfully enrolled. In Pilot Test 2, despite that we obtained only a 50% average completion rate for the PRO survey across all sites, there were significant differences between the different tiers in terms of how many potentially eligible patients could be reached. The completion and consent rates for each tier, as discussed, do not accurately reflect or clearly illustrate the number of patients missed because the front desk staff were overextended tending to patients, or the days the research team were unable to go to clinic because the clinics were short staffed and therefore unable to facilitate the study on those days. To that end, one of the main facilitators at the fully assisted and independent sites was the availability of a full-time study operator. This individual had the sole task of collecting data and was not overburdened by additional tasks of the clinic, such as intaking patients or answering phones. At the partially assisted tiers, where clinic staff assumed some tasks for study completion (i.e., entering information into the admin hub, sending the text, notifying the physician), they struggled to keep up with the demands of the study. This

resulted in researchers being unable to approach eligible patients or eligible patients being taken back to the exam room before their appointment. **In future implementations, it will be essential for staff workload to be appropriately addressed so that clinic staff have the appropriate resources to perform survey collection tasks.**

No amount of planning can replace testing in the 'real world'—which is critical to successful local implementation. Despite detailed and rigorous planning, technical teams need to test this type of technology in the real world to determine what it takes for this to work "in the wild." Given the complexity of healthcare contexts, all the planning and preparation still rarely prepares an organization for the realities of local implementations. Many realities are simply impossible to anticipate. For example, one interesting aspect in Pilot Test 2 is that during the initial PRISM usability testing, most of the issues surfacing were due to patients' lack of familiarity with their own mobile phones. Even in the prototype testing, patients often did not know how to download an app or find an SMS message. Similar issues were found when PRISM was used at the clinics, where most of the issues were not due to the PRISM app but were rather general technology and mobile issues (forgetting passwords, not understanding app downloading, etc.).

An abstraction layer in front of the EHR was critical to our success. Using the Hub as an abstraction layer was critical in being an adapter and bridge between all the technical and version issues that we encountered integrating with the various EHRs. A standards-based, loosely coupled architecture enables rapid reuse of existing apps, and allows developers to focus on building valuable tools and not integrating them. We saw this particularly in Pilot Test 2, when we transitioned PRISM from the Cerner Sandbox to the Hub, there were only a couple of issues relating to message formatting and versions, which were quickly and easily solved. Unfortunately, the world isn't quite ready for this future, and so having the Hub in the middle to align with the standards and abstract away EHR idiosyncrasies made development much faster.

Impact on clinical decision-making is still unclear. While assessing clinical impact and utility was not a primary goal tested in this pilot, we were able to obtain feedback and insight from participating clinicians from both pilot tests related to how, if at all, their clinical decision-making was in any way impacted. It is likely this was more a function of the study design and the use of the PF PRO measure, which several clinicians noted was not the PRO they would have most liked to collect for their patients. The study was designed to assess the issues with patients providing PROs. It was not designed to measure clinical impact. Physicians did look at the PRO scores and responses from patients, but anecdotal feedback from participants suggested the PF PRO was not necessarily appropriate or useful for a specific patient's condition.

Institutional policies can make all the difference. One of the most significant barriers encountered, especially in Pilot Test 2, was an unanticipated delay in implementation due to unforeseen complex institutional policies and regulations that we had to navigate to secure approval and clearance to integrate the PRO technology with the health institution's native information systems architecture. These policies and regulations will likely be highly variable by institution and health systems but will need to be anticipated in any implementation timeline. Technical teams should allow significant time to navigate these processes and to meet institutional requirements, particularly when new vendors are working with a system for the first time. These issues all highlight critical points at which the coordination of human and technical processes is crucial to ensure the successful use of PRO data.

Crosscutting Themes Across Pilot Test 1 (OBERD) and Pilot Test 2 (PRISM)

The primary goal of the two pilot tests and implementation phases was to identify facilitators and barriers of implementation of PRO apps (modified OBERD and new PRISM apps, respectively) across diverse ambulatory practices. The Pilot Test 2 of the PRISM app was built off the pilot test of the OBERD app, with the research team leveraging key lessons learned. Naturally, we found that the integration and implementation of Pilot Test 2 was relatively more seamless. This suggests true utility in our findings and lessons learned for the benefit of other practices in the future. In addition, Pilot Test 2, having the three different tiers for implementation, generated an additional layer of rich contextual findings, which triangulated our findings from the first pilot test in terms of what ambulatory practices really need to do in order to implement SMART on FHIR app-based PRO data collection, integration, and use. In Table 2 and Table 3 below, we present a final summary of the cross-cutting learnings across the two pilots, highlighting some of the key similarities and differences, respectively.

Table 2. Summary of key learnings for similarities across pilot phases

Phase	**Similarity**
Implementation	• **Integration testing** should begin as early as possible.
Implementation	• **Engage all stakeholders** as early as possible.
Implementation	• **Communicate with system administrators** to best understand schedules of system code changes and their subsequent consequences.
Implementation	• Ensure a thorough **understanding of both the vendor product** (and its enabled functionality) **and any site-specific EHR builds** (including web browsers and other technology used in the workflow).
Technical Implementation of the App /Technical Integration	• Design the provider solution to have the **shortest pathway or number of actions** necessary to reach the app within the patient context.
Technical Implementation of the App /Technical Integration	• **Develop a presentation layer that does not require further scrolling** or any other visual manipulation to see the data.
Technical Implementation of the App /Technical Integration	• **Keep the presentation layer locked down** to the patient context.
Technical Implementation of the App /Technical Integration	• Design for the **least amount of client variability**.
Technical Implementation of the App /Technical Integration	• **Avoid user re-authentication** and use the EHR as the trusted app.
Technical Implementation of the App /Technical Integration	• **Finish EHR "plumbing"** (technical integration) work early in the process.
Technical Implementation of the App /Technical Integration	• Anticipate that **EHR version differences add to client variability**.
Technical Implementation of the App /Technical Integration	• **Incorporate testing from the beginning** of the solution design.
Technical Implementation of the App /Technical Integration	• **Implement security from the beginning** – it should not be an afterthought.

Phase	Similarity
Technical Implementation of the App /Technical Integration	• **Use source control** and backup for all technical artifacts.
Technical Implementation of the App /Technical Integration	• Anticipate that system and **software updates will add variability**.
Patient Perspective	• Ensure devices and applications are **designed for the targeted user population** (e.g., usability testing with users).
Patient Perspective	• Ensure measures used are **applicable and painless** (e.g., quick and easy and relevant to the user).
Patient Perspective	• Ensure there is **value in the questions**. Survey content should be relevant, appropriate, and have the potential to influence care.
Provider and Practice Perspective	• **Data should be easily accessible** and, when possible, integrated into the existing workflow.
Provider and Practice Perspective	• **Dedicated personnel need to be present** for seamless data collection (e.g., research team on site).
Provider and Practice Perspective	• Survey output should be accompanied with an **alert/notification** if it is not embedded within the existing workflow.
Provider and Practice Perspective	• **Survey measure should be relevant**, actionable, and have potential to influence care.
Health IT Observations	• **Consistent and clear communication** is paramount.
Health IT Observations	• Technology **standards (FHIR) should be transparent** and accessible.
Health IT Observations	• **Distribution of labor** and the assignment of roles needs to be top priority (OBERD vs PRISM).
Health IT Observations	• **Multidisciplinary teams** need to engage as early as possible, technology implementation should start as early as possible.
Health IT Observations	• **Anticipate institutional hurdles** and a lengthy approval process.
Health IT Observations	• **Anticipate EHR compatibility issues**, including issues in EHR updates that may interfere with integration.

Table 3. Summary of key learnings for differences across pilot phases

Phase	Difference
Implementation	• **Identifying implementation issues (Pilot Test 1) does not ensure quick solutions (Pilot Test 2)**. Anticipate challenges, but also be prepared for contingency planning in case challenges arise.
Implementation	• If building for the first time, **use scalable solutions** for implementation (e.g., EAC defined in the PRO FHIR IG).
Technical Implementation of the App /Technical Integration	• No differences in recommendations between the two pilot tests.
Patient Perspective	• **Devices will differ** in their challenges—iPads were troublesome for different reasons than phones (e.g., issues with touch screen responsiveness and screen dimming vs. issues with the App Store passwords).
Health IT Observations	• Use architecture that **does not rely solely on having a steady and reliable Wi-Fi** network (Pilot Test 2).
Overall Observations	• Sites that operate independently **need the necessary infrastructure,** including technical assistance, Wi-Fi/hotspots, and cellular service.

Conclusion

This study was an important step toward testing the application of the PRO FHIR IG, and the identification of the various system factors critical to the successful implementation, adoption, scalability, and sustained use of this technology to collect PF PRO data in primary and specialty care sites. We demonstrated the ability to successfully adapt the architecture of an existing PRO app by leveraging the modern standards of SMART on FHIR, and to integrate a new patient-facing app and an EHR-based provider visualization app. Key drivers of this success included the ability to design, deploy, and maintain a loosely coupled architecture to centralize communication and decrease variability between sites and EHR implementations. Another key driver of success was the ability to engage with sites and their technical teams as early as possible to begin early testing of data connectivity, as well as any limitations of the EHR's underlying capabilities to render advanced custom visualizations. Each of these drivers may vary based on the team building and deploying the solution, as well as the level of support from local IT teams or their EHR vendor.

We characterized the types of settings, system factors and circumstances that are most conducive to this type of technical implementation. Further, we highlighted the settings where this technology is most likely to be successfully adopted, while identifying the potential challenges faced by others. Our pilots present a model of a large healthcare system integrating a SMART on FHIR PRO app and the various facilitators and barriers faced at this level. For smaller organizations or standalone practices, our findings highlight the infrastructure necessary to implement and support this technology. The multiple-tier design of our implementation for Pilot Test 2 allows practices to better assess and determine the optimal way to resource their implementation based on their context.

Many of our findings were consistent with input from stakeholder feedback in the first part of this project. The pilot tests, however, allowed us to validate the perceptions suggested, and to more definitively

articulate how these factors interplayed with other system factors to either result in successful execution, or presented challenges to the execution of the pilot. Importantly, they demonstrated that developing optimal technology is only part of the story and, as suggested by the Sociotechnical Systems framework guiding this study, there are critical factors related to the providers, practice settings and policies, and the patients themselves. With respect to technology adoption, it is important to note that even a successful implementation of a PRO data collection system does not necessarily guarantee or imply long-term or meaningful use of the data. Successful use of PROs is complex—tending to be context-dependent and strongly coupled to the existing relationships between patients and providers.

Limitations

We note some limitations specific to our pilot. First, implementation and patient and provider enrollment for this study was relatively controlled, even at the independent sites in Pilot Test 2. This experience is, therefore, likely not entirely representative of how the clinical implementation of such a PRO app would function "in the wild" and in clinics with a more typical workflow that does not include some level of intervention by a research coordinator or other facilitator of this alternate workflow. Our pilots rather serve as a proof of concept. Second, there is likely selection bias in terms of the clinics, providers, and even patients who agreed to participate in this pilot. Actual adoption by providers and practices is likely highly variable. Third, the SUS survey results from the patient's usability assessment of the app should be interpreted with caution. Patients were often confused by the wording of some of the SUS items relative to the app. For example, the SUS items reference a "website," and patients were sometimes confused what the term referred to. In addition, patients were asked for feedback on the usability of the app but were likely providing feedback on the entire process, including having to download the app and create passwords. Fourth, we only used one type of PRO (physical function). We expect that patients would find other surveys as easy to use but did not assess that aspect. Finally, this study only asked patients to take an initial survey in a clinic setting. It did not test how well patients could take the survey at home or as a periodically scheduled survey between office visits.

References

1. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Qual Saf Heal Care*. 2010;19(Suppl 3):i68-i74. doi:10.1136/qshc.2010.042085

2. Snyder CF, Jensen RE, Segal JB, Wu AW. Patient-reported outcomes (PROs): Putting the patient perspective in patient-centered outcomes research. *Med Care*. 2013;51(8 SUPPL.3). doi:10.1097/MLR.0b013e31829b1d84

3. What Is Human Factors/Ergonomics? - The Human Factors and Ergonomics Society. https://www.hfes.org/about-hfes/what-is-human-factorsergonomics. Accessed August 30, 2019.

4. Lizzio VA, Blanchett J, Borowsky P, et al. Feasibility of PROMIS CAT Administration in the Ambulatory Sports Medicine Clinic With Respect to Cost and Patient Compliance: A Single-Surgeon Experience. *Orthop J Sport Med*. 2019;7(1). doi:10.1177/2325967118821875

5. Seneviratne MG, Bozkurt S, Patel MI, et al. Distribution of global health measures from routinely collected PROMIS surveys in patients with breast cancer or prostate cancer. *Cancer*. 2019;125(6):943-951. doi:10.1002/cncr.31895

6. Bennett A V., Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. *CA Cancer J Clin*. 2012;62(5):336-347. doi:10.3322/caac.21150

7. Index - FHIR v3.5.0. http://hl7.org/fhir/2018Sep/. Accessed August 30, 2019.

8. Lingard L, Albert M, Levinson W. Qualitative research: Grounded theory, mixed methods, and action research. *BMJ*. 2008;337(7667):459-461. doi:10.1136/bmj.39602.690162.47

9. Pope C. Qualitative research in health care: Analysing qualitative data. *BMJ*. 2000;320(7227):114-116. doi:10.1136/bmj.320.7227.114

10. Patient_Reported_Outcomes. http://hl7.org/fhir/us/patient-reported-outcomes/2018Sep/index.html. Accessed August 30, 2019.

11. Touré M, Poissant L, Swaine BR. Assessment of organizational readiness for e-health in a rehabilitation centre. *Disabil Rehabil*. 2012;34(2):167-173. doi:10.3109/09638288.2011.591885

12. Brooke J. SUS: A quickdirty usability scale. In: Jordan P, Thomas B, Weerdmeester B, eds. *Usability Evaluation In Industry*. London, UK: Taylor & Francis; 1996:252. https://books.google.com/books?hl=en&lr=&id=IfUsRmzAqvEC&pgis=1. Accessed August 30, 2019.

13. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Int J Hum Comput Interact*. 2008;24(6):574-594. doi:10.1080/10447310802205776

14. SMART App Launch Framework. http://www.hl7.org/fhir/smart-app-launch/. Accessed March 21,

2020.

15. Wickens CD, Hollands JG, Banbury S, Parasuraman R. *Engineering Psychology and Human Performance (4th Edition)*. Vol 27. Psychology Press; 2012.