AO Patient Outcomes Center (AOPOC)—Design, Implementation, and Evaluation of a Software Application for Collection of Patient-Reported Outcome Measures in Orthopedic Outpatient Clinics

Nan Rothrock, PhD¹, Michael Bass, MS¹, Andrea Blumenthal, PhD², Richard C. Gershon, PhD¹, Beate Hanson, MD², Alexander Joeris, MD², Aaron Kaat, PhD¹, Suzanne Morrison, MPH³, Robert V. O'Toole, MD⁴, Shalini Patel, BS¹, Michael Stover, MD⁵, Michael J. Weaver, MD³, Raymond White, MD⁶, Maria D. Varela, MS¹, and Mark S. Vrahas, MD⁷

¹ Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA
² Clinical Investigation and Documentation, AO Foundation, Dübendorf, Switzerland
³ Harvard Medical School Orthopaedic Trauma Initiative, Harvard Medical School, Boston, MA, USA
⁴ Department of Orthopaedics, University of Maryland School of Medicine, Baltimore, MD, USA
⁵ Department of Orthopaedic Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA
⁶ Orthopaedic Surgery, Maine Medical Center, Portland, ME, USA
⁷ Orthopaedics, Cedars Sinai Health System, Los Angeles, CA, USA

Corresponding Author:

Nan Rothrock, PhD
Department of Medical Social Sciences
Feinberg School of Medicine, Northwestern University
625 N. Michigan Ave, Suite 2700
Chicago, IL 60660
Office: (312) 503-3514
Fax: (312) 503-4800
n-rothrock@northwestern.edu

Paper Type: Original Paper
Keywords: Orthopedics, Patient Reported Outcome Measures, Tablet Computers
Abstract

Background: Patient-reported outcomes (PROs) are increasingly utilized in routine orthopedic clinical care. Computer adaptive tests (CATs) from the Patient-Reported Outcomes Measurement Information System (PROMIS) offer brief and precise assessment that is well suited for collection within busy clinical environments. However, software applications that support the administration and scoring of CATs, immediate access to PRO scores, and minimize clinician burden are not widely available.

Objective: Our objective was to design, implement, and test the feasibility and usability of a web-based system for collecting CATs in orthopedic clinics.

Methods: AO Patient Outcomes Center (AOPOC) was subjected to two rounds of testing. Alpha testing was conducted in 3 orthopedic clinics to evaluate ease of use and feasibility of integration in clinics. Patients completed an assessment of PROMIS CATs and a usability survey. Clinicians participated in a brief semi-structured interview. Beta phase testing evaluated system performance through load testing and usability of the updated version of AOPOC. In both rounds of testing, user satisfaction, bugs, change requests, and performance of PROMIS CATs were captured.

Results: Patient feedback supported the ease of use in completing an assessment in AOPOC. Across both phases of testing, clinicians rated AOPOC as easy to use, but noted difficulties in integrating a web-based software application within their clinics. PROMIS CATs performed well; the default assessment of 2 CATs was completed quickly (mean=9.5 items) with a satisfactory range of measurement.

Conclusion: AOPOC was demonstrated to be an easy to learn and easy to use software application for patients and clinicians that can be integrated into orthopedic clinical care. The workflow disruption in integrating any type of PRO collection must be addressed if patients’ voices are to be better integrated in clinical care.
Introduction

There is an increasing demand to utilize patient-reported outcomes (PROs) in clinical care for a variety of aims. PROs can capture important outcomes of care such as reduction of disability or depressive symptoms (e.g., [1, 2]). This is particularly true within orthopedics, as treatment is often initiated to improve a patient’s physical function and reduce pain. Quantification of the patient’s perspective can be utilized in treatment decision making such as when the patient’s pain and disability has progressed enough to consider joint replacement and to help judge the success of treatment [3]. Payors have also introduced PROs as a method for assessing healthcare quality (PRO Performance Measures, PRO-PMs) rather than only utilizing measures that evaluate the process of care [4-6]. For example, the Centers for Medicaid and Medicare Services now includes submission of PRO data for total hip and total knee replacement reimbursement as part of their Comprehensive Care for Joint Replacement Payment Model [7]. It is hoped that using PRO-PMs will alter the definition of healthcare quality to include the function and symptom burden of patients [8].

In an effort to improve measurement precision, efficiency, applicability, and interpretability, the National Institutes of Health (NIH) invested in the development and validation of the Patient-Reported Outcomes Measurement Information System® (PROMIS®), a collection of PRO measures that assess important domains of self-reported health. PROMIS measures include computer adaptive tests (CATs)—a tailored administration in which questions are selected dynamically based upon past responses using the most informative question for that specific level of function or symptom severity. This offers rapid assessment with high measurement precision across a wide range of functional abilities and symptom severities. CATs are particularly suited for integration in clinical care as there is little time in the clinic workflow for assessment and a need for highly precise measures when evaluating individual patient data [9].
Our aim was to develop a software application that enables the administration and scoring of PROMIS CATs and other relevant PROs in the clinical routine across diverse orthopedic clinics. Specifically we hypothesized that a) we could develop a software application that is easy to learn and easy to use for clinicians and patients, b) AOPOC is feasible for integration into orthopedic clinics, and c) PROMIS CATs provide rapid quantification of symptoms and function across a wide range of patients.

Methods

AOPOC Design

Design Principles

AO Patient Outcomes Center (AOPOC) followed five design principles: 1) a primary focus on in-clinic data collection for use in the clinical encounter including display of longitudinal PRO data for an individual patient, 2) a common assessment battery of PROMIS CATs for all patients, 3) the capability to add other patient- and clinician-reported measures to the assessment, 4) an easy to learn and easy to use interface that imposes minimal burden on surgeons, and 5) the ability to export data for a group of patients for research analyses.

Figure 1: AO Patient Outcomes Center Homepage
Description of AOPOC

AOPOC is a web-based software application to collect PROs in orthopedic clinics. A clinician is able to establish one or more “Patient Groups” (Figure 1). Assessment content can be tailored for each Patient Group. For example, clinicians may utilize different PROs for individuals with traumatic injuries than individuals receiving total joint replacement. A library of high quality PROs used in orthopedic care is available (Figure 2) as are numerous clinician-provided variables such as fracture classification code, mechanism of injury, and body mass index. All Patient Groups include the PROMIS CATs for Pain Interference and Physical Function. A patient is registered in AOPOC with full name, date of birth, and is assigned to one or more Patient Groups. The interface for the patient to complete the assessment is a simple design maximized for viewing on an iPad (Figure 3). Each CAT is scored in real time and displayed in a longitudinal graph along with the questions and responses from the most recent assessment (Figure 4). This report is available as a PDF to share with a patient or manually add to an electronic health record (EHR). Data from a Patient Group can be exported. It includes raw response data, PROMIS measure scores, time and date of a patient’s responses, and response time. Optionally, multiple consent forms can be included to facilitate use of AOPOC in research data collection.

Figure 2: Portion of AOPOC PRO Library
Figure 3: AOPOC Patient Interface

Figure 4: Portion of AOPOC Patient PRO Report
A clinician can give access to a Patient Group to other clinical users who belong to the same HIPAA organization. For example, a surgeon can establish a Trauma Patient Group and give access to fellows, front desk staff, and a physician assistant. Each member of the team has a unique login and password and is assigned a level of access to determine whether or not he or she can register patients, modify the assessment content, export data, or carry out other actions.

**Default PRO Battery**

The default assessment includes 2 PROMIS CATs. A minimum of 4 and a maximum of 12 questions are administered per CAT. PROMIS T-scores have a mean of 50 (SD=10) in the U.S. General Population.

*PROMIS v1.0 Pain Interference CAT:* Consequences of pain on relevant aspects of one’s life are measured, including the extent to which pain hinders engagement with social, cognitive, emotional, physical, and recreational activities. Higher scores indicate more difficulties from pain.

*PROMIS v1.2 Physical Function CAT:* Self-reported capability of physical activities including upper extremities, lower extremities, and central regions (neck, back), as well as instrumental
activities of daily living, such as running errands. The initial version of AOPOC utilized PROMIS CATs for Mobility (v1.2) and Upper Extremity Function (v1.2) instead of the single PROMIS Physical Function CAT.

**AOPOC Development**

To develop the application, we utilized a modified Agile methodology [10-12] that enabled iterative development with continuous feedback. A multidisciplinary team of orthopedic trauma surgeons, clinic staff, and PRO scientists participated in Joint Application Design sessions with a business analyst to set a vision and scope. The team then compiled requirements and constructed use cases with a business analyst and informatics project manager. Each feature’s requirements including functionality, terminology, navigation, and user interface was defined in addition to nonfunctional requirements such as security and scalability. The team reviewed and edited screen mockups. Once use cases were approved, software development and quality assurance testing (QA) were completed. User Acceptance Testing (UAT) was conducted by demonstrating each use case in the development environment to provide early feedback and ensure the feature functioned as desired. Requested application modifications were implemented followed by QA and UAT iteratively until a use case was approved. Upon completion of all requirements, AOPOC was moved to a staging environment in which all functionality was again tested with errors corrected as needed. When all test cases passed QA, AOPOC was made available in an online production environment.

AOPOC is a web application that administers PROMIS CATs, short forms and other custom instruments that have been loaded and tested in Assessment Center [13]. It follows a Model View Controller (MVC [14]) pattern built upon Microsoft technologies in which the controllers are coded in C# and compiled into Microsoft .NET 3.5 dlls. The models are represented as XML documents [15] that are dynamically generated by the controllers and rendered to the views through XSL transformations [16]. Microsoft SQL
Server is used for data persistence where data transfer takes place over Hyper Text Transfer Protocol Secure (HTTPS) protocol. All user-entered data (i.e. demographic and PRO responses) are encrypted at the field level using AES encryption and user passwords are one way hashed. AOPOC is hosted on a third party server in a secure environment only accessible through VPN connections.

Alpha-phase Testing

Alpha phase testing was conducted to test ease of use and feasibility in integrating AOPOC in orthopedic clinics. Data were collected from both patients and clinic staff at 3 orthopedic trauma clinics in the U.S. associated with a surgeon in the project team. Each site’s IRB determined the project was not considered human subjects research.

Patients:

The site PI identified the patient population for alpha testing (e.g., single surgeon’s patients, specific clinic’s patients). Adult, English-speaking patients were eligible to participate. Participants were asked to complete a battery of PROs selected by his or her surgeon. Following the PROs, an 11-item usability survey consisting of questions related to past computer use experience, comfort using the data collection device (e.g., tablet computer), and satisfaction with the user interface was administered. Ease of use questions had four response options (0=not at all, 1=a little bit, 2=somewhat, 3=quite a bit). User interface questions utilized five response options (4=excellent, 3=very good, 2=good, 1=fair, 0=poor).

Clinic Staff

After 3 to 9 weeks of AOPOC experience, the site PI identified staff including surgeons, other clinicians, and administrative personnel (e.g., front desk staff) who interacted with AOPOC on more than one
occasion. All were invited via email to participate in a 20 minute semi-structured interview by phone. The interview included open ended questions targeting specific features of AOPOC including 1) completing the application for implementing AOPOC at his or her site, 2) establishing the assessment content for specific patients (Patient Group set-up), 3) enrolling clinical users, 4) registering patients, 5) having patients complete the assessment, and 6) accessing patient data. Multiple choice questions assessing the ease of use of specific features, system usefulness, and degree of disruption to clinic workflow were also administered. Questions used a 5-point scale ranging from 1 (Not at all) to 5 (Very). The interview script was modular so that interviewees were only asked about those portions of AOPOC they utilized. All issues, requested modifications, and responses to multiple choice items were recorded in a database. Additionally, communication with the AOPOC support desk was used to identify areas of confusion or errors as well as requested system modifications. This feedback was added to the database. Throughout alpha testing, bugs (instances when AOPOC was not functioning as it should) were distinguished from change requests (e.g., modification to improve usability or expand system capabilities). Bugs were resolved by a software developer immediately. Following prioritization by the project team, high priority change requests were implemented using the same software development protocol (e.g., QA, UAT).

**Beta-phase Testing**

The aim of Beta-phase testing was to prepare for public release of AOPOC. It included load testing to quantify and improve system performance and usability testing to again evaluate ease of learning and using an updated version of AOPOC.

**Load Testing**

In order to determine AOPOC performance capabilities, a test harness was developed that ran scripts which simulated common use cases identified by the business analyst and informatics project manager.
With this test harness, load testing was conducted by generating a simulated load for a typical AOPOC use case of registering a patient and administration of the default PRO assessment battery. The load testing gradually increased the number of simultaneous patients, starting with 100 and increasing until the system started to return timeout errors caused by server requests exceeding the defined default response time (90 sec). Rounds of testing were conducted to identify server settings that maximized performance, identified bottlenecks (area within the software application that slowed overall performance due to concurrency), and re-evaluate performance after modifications were implemented. A benchmark to double simulated patients was established.

Usability
Surgeon members of AO Trauma, a non-profit international organization of clinicians and researchers aimed at fostering and improving medical care for musculoskeletal trauma, were surveyed about their interest in serving as an AOPOC beta test site. The project team reviewed 66 interested orthopedic clinics to identify representatives from specific types of clinics (e.g., academic medical center, community hospital practice), and invited a diverse group of 36 sites to participate with a goal of enrolling 20. After 6 months, 16 sites had completed enrollment. Similar to alpha testing, usability feedback was collected in multiple ways. First, data were extracted from AOPOC to evaluate clinicians’ ability to tailor assessment content and patient response burden with PROMIS CATs. Second, interactions with clinical users including emails and calls to the AOPOC support desk and questions, comments, or difficulties during demonstrations were used to identify bugs as well as areas of confusion and errors. Third, beta site leaders identified active AOPOC users at their sites to participate in a 20-minute semi-structured interview by phone. Interviews utilized the same modular guide as alpha testing. All issues, requested modifications, and responses to multiple choice items were recorded in a database. Again, bugs were identified throughout beta testing and resolved immediately. Following prioritization by the project
team, all high priority change requests that fit within the available resources were implemented using the same software development protocol (e.g., QA, UAT).

Analytic Plan

**Alpha-phase Patient-level Data**

All patient-level data was exported from AOPOC by a software developer and de-identified by a data manager following a standardized protocol. Score distributions for each PROMIS CAT were constructed. Frequency distributions for the number of items administered and mean PROMIS score were calculated. Time to complete each CAT was calculated. Frequency distributions and means for individual usability survey items were calculated.

**Alpha-phase Clinician-level Data**

The database of user feedback was reviewed and redundant entries combined noting the number of users reporting the same issue. If required, additional clarification was sought from the source. All change requests were prioritized for implementation using a MoSCoW approach. This approach assigns each change request a Mo=Must (a mandatory change for the intended functionality of AOPOC), S=Should (a high priority and desirable change that is not mandatory), Co=Could (a change that would improve AOPOC, but is not critical to its success), or W=Would (a change that may be considered, but is not critical or appropriate now) rating. Ratings are based upon 1) alignment with the intended scope of AOPOC, 2) number of affected users, 3) frequency of request, and 4) whether or not a user could circumvent the issue. The project PI made an initial rating assignment. This was provided to the project team who reviewed and revised ratings individually, then met for a consensus meeting to finalize ratings.
The available development effort was used to implement change requests in order of priority until it was expended.

**Beta-phase Usability Data**

All beta phase patient-level data was exported by a software developer and de-identified by a data manager following a standardized protocol. All available data were included in analysis, including sites which did not participate in usability interviews. Descriptive statistics including the number of Patient Groups set up within each beta test site, number of measures in addition to the AOPOC battery that were administered per Patient Group, and number of unique patients providing data were calculated. Frequency distributions were constructed for the use of other PROMIS CATs and number of items needed to complete the default and supplemental PROMIS CATs.

“Should” change requests that were identified but not implemented during alpha testing were merged with issues identified in beta testing. All were reviewed and redundant entries combined. Potential system modifications to address user issues were identified. All change requests were prioritized using new MoSCoW ratings. The project PI provided the initial prioritization, distributed it to the project team, and a meeting was held to reach consensus on ratings. The available development effort was used to implement change requests in order of priority until it was expended.
**Results**

**Alpha-phase Testing**

**Patients**

Of 2446 patients who were registered in AOPOC, 935 (38.2%) completed an initial assessment. Of these, 125 completed a second assessment at a follow-up clinic visit. The 61.8% of registered patients who did not complete an assessment were primarily from one site that registered patients the day prior to his or her clinic visit.

**PRO Scores:** In alpha testing, separate Mobility and Upper Extremity Function CATs were used instead of a single Physical Function CAT. Half of the subjects completed the entire assessment battery (3 CATs) in under 3.7 minutes (mean=4.7, SD=5.2, 95% CI 4.5 – 5.0). Long completion times were due in part to interruptions from clinic staff (e.g., moving patient to visit room). About 85% of patients answered only the minimum of 4 questions for the Mobility and Pain Interference CATs compared to only 22.1% of patients for the Upper Extremity CAT. For the Upper Extremity CAT, 42.3% of patients answered the maximum number of questions which was 12 (Table 1). Mean T-scores were in the moderate impairment range for Mobility (mean=38.7, SD=8.7, 95% CI=38.1-39.2) and at the border between mild symptoms and within normal limits for Pain Interference (mean=59.2, SD=9.4, 95% CI=58.6-59.9) with fairly well distributed scores (Figures 5 and 6). Upper Extremity Function was also in the moderate impairment range (mean=39.8, SD=10.7, 95% CI=39.1-40.5). The best possible score (T=56.4) was received by 15.7% (n=142) of patients, indicating the measure may be unable to distinguish between patients with excellent function (“ceiling effect”).

**Table 1: Performance of Instruments in the AOPOC Battery**

<table>
<thead>
<tr>
<th></th>
<th>PROMIS Mobility CAT</th>
<th>PROMIS Upper Extremity CAT</th>
<th>PROMIS Pain Interference CAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completion time (sec)</td>
<td>Mean (SD)</td>
<td>Median (min, max)</td>
<td></td>
</tr>
</tbody>
</table>
|----------------------|-----------------|-------------------|---
|                      | 131.5 (133.7)   | 85.0 (5.0, 1108.0)|
|                      | 96.6 (109.3)    | 70.0 (4.0, 1010.0)|
|                      | 60.0 (69.7)     | 42.0 (4.0, 790.2) |

| No. of items administered | Mean (SD) | Median (min, max) |   
|---------------------------|-----------|-------------------|---
|                           | 4.7 (2.1) | 4.0 (4, 12)       |
|                           | 8.2 (3.5) | 7.0 (4, 12)       |
| Assessments with min no. | 84.6 (781)| 7.5 (69)          |
| of items: % (n)           |           | 4.0 (4, 12)       |
| Assessments with max no.  | 22.1 (200)| 42.3 (384)        |
| of items: % (n)           |           | 85.4 (766)        |

| T-score                  | Mean (SD)       | Median (min, max) |   
|--------------------------|-----------------|-------------------|---
|                          | 38.7 (8.7)      | 37.2 (18.3, 60.2) |
|                          | 39.8 (10.7)     | 38.7 (14.7, 56.4) |
|                          | 59.2 (9.4)      | 60.1 (38.5, 80.1) |

AOPOC = AO Patient Outcomes Center, CAT = Computer adaptive test, PROMIS = Patient-Reported Outcomes Measurement Information System

Figure 5: Frequency Distribution of PROMIS Mobility and Upper Extremity Function CAT T-scores

Figure 6: Frequency Distribution of PROMIS Pain Interference CAT T-scores
Usability: Overall, most patients were familiar with using computers; 86% (n=736 of 857) used a computer within the past year. A similar number (85%, n=724 of 856) used a touchscreen such as ATM or airline check-in kiosk. A majority (86%, n=721 of 839) owned a device with internet connectivity and reported use 5-7 days a week (74%, n=551 of 741). AOPOC was not difficult for patients to use (Table 2). Most participants (81%, 686 of 842) reported they had no difficulty using the data collection device (tablet or desktop computer). Only 7% (58 of 840) were “somewhat” or “quite a bit” uncomfortable, anxious, or nervous using the data collection device. Most (86%, 721 of 841) had no difficulty answering the PRO questions. A strong majority (91%, 760 of 839) would be willing to complete a similar assessment at a future clinic visit. Furthermore, ratings of the AOPOC interface design were favorable, including data collection screens (88% good, very good, or excellent, 741 of 843) and the response button design (90% good, very good, or excellent, 743 of 825).

Table 2: AOPOC Usability for Patients

<table>
<thead>
<tr>
<th>What is your overall rating of the design of the screens</th>
<th>Excellent % (n)</th>
<th>Very good % (n)</th>
<th>Good % (n)</th>
<th>Fair % (n)</th>
<th>Poor % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 (182)</td>
<td>30 (256)</td>
<td>36 (303)</td>
<td>10 (87)</td>
<td>2 (15)</td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Not at all % (n)</td>
<td>A little bit % (n)</td>
<td>Somewhat % (n)</td>
<td>Quite a bit % (n)</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>What is your overall rating of the buttons on the screens, including their size and shape?</td>
<td>24 (197)</td>
<td>33 (275)</td>
<td>33 (271)</td>
<td>8 (65)</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Did you have any difficulty using this computer?</td>
<td>81 (686)</td>
<td>11 (89)</td>
<td>4 (35)</td>
<td>4 (32)</td>
<td></td>
</tr>
<tr>
<td>Did you ever feel uncomfortable, anxious, or nervous while using the computer?</td>
<td>82 (685)</td>
<td>12 (97)</td>
<td>4 (33)</td>
<td>3 (25)</td>
<td></td>
</tr>
<tr>
<td>How difficult was it to answer the questions shown on this computer?</td>
<td>86 (721)</td>
<td>9 (77)</td>
<td>3 (27)</td>
<td>2 (16)</td>
<td></td>
</tr>
<tr>
<td>If you were asked to complete this survey again at a future appointment, would you?</td>
<td>91 (760)</td>
<td>9 (79)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Clinicians**

Of 13 clinic staff who were invited, 11 participated in a usability interview. Interviewees were from all 3 sites and were providers (surgeon, physician assistant, n=5), research staff (n=3), or in other positions (nurse manager, program director, medical secretary; n=3). One provider did not personally interact with AOPOC and therefore questions concerning application usability were skipped. All 3 sites utilized a single
Patient Group with only the default assessment content. Each site followed a slightly different workflow (e.g., staff register patient in AOPOC when s/he checks-in, staff registers all patients for the following day). Data collection occurred during an existing wait time such as in the clinic waiting room on an iPad or in the visit room on an iPad or desktop computer.

Clinicians described feeling comfortable using AOPOC at the time of the interview (mean=4.7, SD=0.5, n=10). AOPOC was easy to learn (mean=4.7, SD=0.7, n=10) and easy to use (mean=4.5, SD=0.8, n=10). The process for completing the application for access to AOPOC was described as "A little bit" difficult by the two respondents who did this task. Registering as a clinical user was "Not at all" or "A little bit" difficult for the two clinicians who answered this question. Clinicians who set up a Patient Group reported it as not difficult (mean=1.3, SD=0.5, n=4). Clinicians also had little difficulty registering patients (mean=1.4, SD=0.7, n=8), starting a patient’s assessment (mean=1.3, SD=0.8, n=6), and accessing a single patient’s data (mean=1.8, SD=0.8, n=5). Only two attempted to export data from a Patient Group and both reported no difficulty. Areas where clinicians experienced more difficulty included entering clinical data (e.g., fracture classification code, mean=3.5, SD=2.1, n=2) and understanding the patient’s data including PROMIS CAT scores (mean =3.0, SD=1.3, n=6). There was a wider range in responses concerning how disruptive AOPOC was to the clinic workflow (mean=2.3, SD=1.2, n=9).

The qualitative feedback from the usability interviews and queries to help desk identified 17 bugs and 104 change requests. Ten of 17 bugs were resolved during the pilot phase. The remaining bugs could not be re-created (n=4), were re-categorized to change requests (n=2), or required additional investigation time after the alpha test period (n=1). Change requests were clustered in several areas. First, the process to register as a new site, register as a clinical user, and to provide access to other site members was
reported to be challenging. Second, there was technical difficulty in utilizing an iPad as a data collection device. For example, users were not aware of the set-up requirements to enable AOPOC on an iPad (e.g., turn off pop-up blocker). Multiple clinicians commented that staff had more difficulty interacting with an iPad than patients. Although support material including instructions for iPad set-up was available, some users did not know it existed. Finally, the AOPOC Administrator identified multiple areas of difficulty in registering new users including being unable to view those who were sent a registration email, but had not completed the registration process, restrictions in the ability to make edits to user email addresses, HIPAA entity names, and demographic fields (e.g., State). In the consensus meeting, 17 "Must", 31 "Should", 44 "Could", and 12 "Would" change requests were identified. The Must requests were implemented prior to the beginning of beta testing. In addition, based upon the range restriction of the PROMIS Mobility and Upper Extremity Function CATs, these measures were replaced with the PROMIS v1.2 Physical Function CAT.

Beta-phase Testing

Load Testing

Throughout beta testing, no users reported problems with system speed. However, load testing identified two web server settings that contributed to reduced speed: "keep alive" and the maximum number of connections. Tests using simultaneous requests and an initiation rate of 2-4 seconds between requests were conducted. Changing default Microsoft IIS 7 settings and increasing the number of CPUs did not increase the number of requests. After suspecting that requests were being blocked from finishing, the extra CPU used on the staging database was applied to the staging web server that was reaching capacity. This resulted in doubling the performance of requests finishing in multiple rounds of testing. Several recommendations were identified to prepare for wider distribution of AOPOC including 1) updating hardware and increasing CPUs to four, 2) continuous monitoring of performance by server
host, and 3) engage in ongoing evaluation with updated information about volume of simultaneous users.

Usability

Of the 16 enrolled sites, 8 had completed Business Associate Agreements (BAAs) between their institution and AOPOC at the time of usability interviews. The remaining 8 sites completed BAAs after usability interviews were concluded but were able to provide data on the initial set-up of AOPOC at their clinics. Reasons for not enrolling as a test site included no support from other clinic personnel, utilization of other data collection systems (e.g., state mandated system), and difficulties in executing a BAA. BAA challenges included a) difficulty identifying who had authority to sign this agreement, b) institutional staff requiring modifications to the existing BAA which required negotiation (e.g., removing AO Foundation’s right to use a de-identified dataset from AOPOC for research purposes), and c) requiring that the local institution’s own standard BAA be utilized requiring review and discussion with the project team.

Across 16 sites, 71 unique Patient Groups were created. There was a wide range (2 to 24) of patient- and clinician-completed instruments per Patient Group. Approximately 34% (n=24) only collected the default battery. Three to 5 additional instruments were included in 38% (n=27) of Patient Groups. There were 9725 assessments completed by 5088 patients. Most patients completed one or two assessments, and the maximum completed by the same person was 14. The default assessment (2 CATs) required 9.5 items on average. For Pain Interference, 80% of assessments (n=6355) required only the minimum number of items per CAT (see Table 3). Similarly, 78% (n=6178) of Physical Function assessments used the minimum number of items. An additional 8 PROMIS CATs were administered by at least one clinician. Most (n=6) were able to be completed with 4-5 items. Depression (mean=7.2, SD=3.8) and Upper Extremity...
Function (mean=8.0, SD=3.5) required the most items, but in both cases the number of items was bimodally distributed, with a substantial number completing the CAT in 4 or 5 items (for Depression and Upper Extremity Function, respectively) and a minority requiring the maximum of 12 items.

Table 3: Utilization and Average Length of PROMIS CATs in Beta-phase

<table>
<thead>
<tr>
<th>PROMIS CAT</th>
<th>No. of Assessments(a)</th>
<th>CAT Length Mean no. of items (SD)</th>
<th>Assessments with min no. of items: % (n)</th>
<th>Assessments with max no. of items: % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain Interference (default assessment)</td>
<td>7895</td>
<td>5.0 (2.5)</td>
<td>80 (6355)</td>
<td>11 (857)</td>
</tr>
<tr>
<td>Physical Function (default assessment)</td>
<td>7965</td>
<td>4.5 (1.4)</td>
<td>78 (6178)</td>
<td>2 (167)</td>
</tr>
<tr>
<td>Upper Extremity Function</td>
<td>488</td>
<td>8.0 (3.5)</td>
<td>24 (118)</td>
<td>39 (191)</td>
</tr>
<tr>
<td>Mobility</td>
<td>636</td>
<td>5.8 (3.2)</td>
<td>71 (452)</td>
<td>21 (132)</td>
</tr>
<tr>
<td>Pain Behavior</td>
<td>78</td>
<td>4.7 (2.2)</td>
<td>90 (70)</td>
<td>8 (6)</td>
</tr>
<tr>
<td>Depression</td>
<td>39</td>
<td>7.2 (3.8)</td>
<td>51 (20)</td>
<td>36 (14)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>19</td>
<td>4.9 (2.5)</td>
<td>89 (17)</td>
<td>11 (2)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>75</td>
<td>4.3 (1.2)</td>
<td>89 (67)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Sleep Disturbance</td>
<td>37</td>
<td>5.1 (2.4)</td>
<td>51 (19)</td>
<td>5 (2)</td>
</tr>
<tr>
<td>Ability to Participate in Social Roles and Activities</td>
<td>18</td>
<td>4.7 (1.9)</td>
<td>78 (14)</td>
<td>6 (1)</td>
</tr>
</tbody>
</table>

\(a\)Some patients completed more than one assessment.

Sixteen clinicians were invited to participate in a usability interview. Of these, 10 were not yet using AOPOC in clinic. Five clinicians across four sites completed the interview. All interview participants reiterated the ease of learning and using AOPOC. The most significant concern was the impact of PRO collection on clinic workflow. Change requests were related to usability, security, and default assessment content. Following prioritization, 9 change requests related to usability were evaluated as “Musts” and implemented (e.g., improve consistency in use of checkboxes, increase visibility of location within the
application, clarify reminders to clinicians about missing clinical information). Two of these were also reported and rated as “Shoulds” during alpha-phase testing. Six bugs were identified with only 1 being critical – one test site blocked incoming registration emails from AOPOC. All bugs were resolved during the beta testing phase. Additionally, 9 security improvements (e.g., server side validation, improved encryption of URLs, construction of log of all user logins and failed logins) were implemented.

Discussion

AOPOC was demonstrated to be an easy to learn and easy to use software application for patients and clinicians that can be integrated into routine orthopedic care. In alpha-phase testing, a battery of PROMIS CATs was completed in under 5 minutes with usually only the minimum number of questions required. The measures performed well, although Upper Extremity Function demonstrated a limitation in assessing patients with better functioning. Patients were comfortable using the data collection device and answering questions with 91% willing to complete an assessment again. Clinicians also reported AOPOC was easy to learn and its multiple features were easy to use. The most frequently cited areas for improvement were in onboarding as a new site and clinical user and mitigating the impact of PRO collection on the clinic workflow. A small number of bugs were reported, none of which were critical (i.e., prevented data collection). In beta-phase testing, areas for improving system speed were identified and addressed. PROMIS CATs performed well; they were completed quickly and there were no ceiling effects in the revised default assessment. Only one critical bug was identified which was resolved. Clinicians demonstrated the ability to establish more complex assessment plans through multiple assessment types and tailored content. Of the small number of clinicians who completed interviews, concerns focused on executing a BAA and achievable modifications to improve usability.
AOPOC addresses some of the barriers to the collection and use of PROs in routine orthopaedic clinical practice. First, similar to other studies in orthopaedic patient populations (e.g., [17, 18]), the PROMIS Physical Function, Mobility, and Pain Interference CATs demonstrated acceptable assessment at the very high and very low ends of the range. However, the PROMIS v1.2 Upper Extremity Function CAT did demonstrate poor ability to distinguish between levels of excellent function and therefore may not be able to capture the full degree of improvement following intervention. All measures have a limit to the range of function or symptom that they can quantify. One of the advantages of utilizing CATs is that the range of measurement can be extended by adding new questions to the item bank that capture these extremes (e.g., [19]). This allows for ongoing measurement improvement while retaining the ability to compare scores from different versions of a measure. CATs also remove a second assessment barrier – time burden [20]. By tailoring the items that are administered, measures were most frequently completed in 4-5 items. Finally, AOPOC provided measure scores immediately in table and graphical presentations thus removing the need to calculate and display results [21].

This project highlights two issues related to integrating PRO collection in clinical practice that are not specific to AOPOC: 1) managing the impact of adding a PRO assessment to the clinic workflow is critical and 2) the site-specific requirements for utilizing a software application in parallel with an EHR are substantial. In this project, clinic workflow disruption was frequently identified as a concern which is consistent with previous research [5]. Low participation from clinicians at beta-test sites and by registered patients at one alpha-test site may also reflect workflow problems. Collection of PROs requires patients’ time to complete measures and care providers’ time to access and review results [22]. Enabling patients to complete assessments outside of the clinic setting can reduce the time demand on patients at the care setting, though participation has been low [23, 24]. Another strategy has been to improve the usefulness of PRO results as a way to increase patients’ and clinicians’ engagement in PRO
collection. Aligning the assessment content with the clinical purpose (e.g., screening versus monitoring a primary outcome of care) [23, 25] and making PRO results more interpretable and actionable [26-29] have been recommended. Clinician training programs have also been successful in improving interpretation and use of PRO results [30].

Software applications that work in parallel with an EHR require BAAs. Most of the beta-testing sites required more than 4 months for a BAA to be finalized and some sites were not successful in reaching an agreement even after 8 months. As reimbursement is increasingly tied to PRO collection [7, 31], the prioritization by healthcare systems to enable PROs in clinical care is expected to increase. Non-EHR data collection software applications like AOPOC are able to quickly integrate advances in PRO measurement science such as improved measures (e.g., a PROMIS Upper Extremity CAT with a higher measurement ceiling) and graphical displays of results (e.g., integrate newly published normative scores for a particular patient population). Without increasing the efficiency of adopting software applications into clinical care, the breadth of advances in patient-centered care will remain difficult to implement.

Limitations
Several limitations are noted. Most of the patient participants had previous experience interacting with a computer and a touch screen. Consequently, the positive usability feedback may not represent the experience of those patients without past experience. Information on the number of patients who were approached to complete an assessment but declined was not captured. Therefore, it is not known to what extent patient factors (e.g., computer fluency, concerns about privacy, English literacy) and/or clinic factors (e.g., barriers to adding an assessment to clinic workflow, insufficient data collection devices) contributed to patients not completing an assessment. Similarly, there may have been selection bias inherent in getting clinician feedback from those who used the system; clinicians who were more
comfortable or open to computerized testing might have been more likely to use AOPOC. Additionally, both alpha and beta testing occurred in the clinics of orthopedic surgeons who were members of the AO Foundation which may not be representative of all orthopedic surgeons.

Conclusion

AOPOC is a feasible, robust software application that enables collection of CATs within orthopedic clinical care. Barriers to routinely integrating PROs including modification of clinic workflow and execution of BAAs remain. However, addressing those barriers enables the integration of the patient’s perspective in his or her healthcare. This is particularly important in orthopedics as physical function and pain are regularly the targets of interventions and the reason patients seek care.

Acknowledgements

Financial support for this project was provided by AO North America and AO Trauma. AO Clinical Investigation and Documentation (AOCID) staff contributed to software design (BH) and manuscript preparation (AB, AJ). NER, MB, RCG, AK, SP, and MDV also receive grant funding for PROMIS from the National Institutes of Health (U2C CA186878).

Conflicts of Interest

RCG and MSV serve on the AOPOC Inc Board of Directors.
References


14. Reenskaug TMH. MVC XEROX PARC 1978-79. http://heim.ifi.uio.no/~trygver/themes/mvc/mvc-index.html. Archived at: https://urldefense.proofpoint.com/v2/url?u=http-3A__www.webcitation.org_6yd7TxA9L9d=DwIDAw&c=yHIS04HhBraes5B9ueu5zKhE7rtNXt_d012z2PA6ws&r=yyEGD5WKfGbOQtLy9nLOSziJBR-qfZlZDVl-hkJuk&m=brSknfgVmLhWboCYSARy0M55Geu_g30e3d225UVPS18&s=wpOsv0t_oHEz3JMFFTIOwblvdcwpGuVnrnNvFvp5gw&e

15. Quin, L. Extensible Markup Language (XML). http://www.w3.org/XML. Archived at: https://urldefense.proofpoint.com/v2/url?u=http-3A__www.webcitation.org_6ycyZqXrC&d=DwIDAw&c=yHIS04HhBraes5B9ueu5zKhE7rtNXt_d012z2PA6ws


