



# Large-scale clinical implementation of PROMIS computer adaptive testing with direct incorporation into the electronic medical record

M. O. Papuga<sup>a,b,c</sup>, C. Dasilva<sup>a,c,d</sup>, A. McIntyre<sup>a,c</sup>, D. Mitten<sup>c,d</sup>, S. Kates<sup>a,c</sup> and J. F. Baumhauer<sup>a,c,e</sup>

<sup>a</sup>Center for Musculoskeletal Research, University of Rochester, Rochester, NY, USA; <sup>b</sup>Department of Research, New York Chiropractic College, Seneca Falls, NY, USA; <sup>c</sup>Department of Orthopaedics and Rehabilitation, University of Rochester, Rochester, NY, USA; <sup>d</sup>Center for Clinical Innovation, University of Rochester, Rochester, NY, USA; <sup>e</sup>Department of Orthopaedics and Rehabilitation, University of Rochester Medical Center, Rochester, NY, USA

## ABSTRACT

The objective of this research was to assess the implementation of collecting patient-reported outcomes data in the outpatient clinics of a large academic hospital and identify potential barriers and solutions to such an implementation. Three PROMIS computer adaptive test instruments, (1) physical function, (2) pain interference, and (3) depression, were administered at 23,813 patient encounters using a novel software platform on tablet computers. The average time to complete was  $3.50 \pm 3.12$  min, with a median time of 2.60 min. Registration times for new patients did not change significantly,  $6.87 \pm 3.34$  to  $7.19 \pm 2.69$  min. Registration times increased for follow-up ( $p = .007$ ) from  $2.94 \pm 1.57$  ( $p < .01$ ) min to  $3.32 \pm 1.78$  min. This is an effective implementation strategy to collect patient-reported outcomes and directly import the results into the electronic medical record in real time for use during the clinical visit.

## ARTICLE HISTORY

Received 27 November 2015  
Revised 7 October 2016  
Accepted 19 October 2016

## KEYWORDS

orthopaedic outcomes;  
patient-reported outcomes;  
PROMIS; physical function;  
pain; depression

## 1. Background

Most clinicians are time pressured in clinic and collection and analysis of outcome measures during the clinical visit is nearly impossible. However, the use of a standard set of patient-reported outcomes (PROs) may help inform treatment decisions during the visit.

Currently, there is a lack of standardized outcome measures or methods to study the effects of treatment and to effectively monitor and document patient outcomes. (Black, 2013; Chen, Ou, & Hollis, 2013; Reeve et al., 2007) To that end, the National Institutes of Health (NIH) previously funded a consortium from (2004 to 2014) in order to produce and refine the Patient-Reported Outcome Measurement Information System (PROMIS). (Cella, Riley, Stone, et al., 2010; Cella et al., 2007; Fries, Bruce, & Cella, 2005) This system is a collection of freely available evaluation assessments for a variety of physical, mental, and social domains (Cella et al., 2007, 2010) that have been found reliable and valid in a variety of clinical settings. (Bjorner et al., 2014; Broderick, Schneider, Junghaenel, Schwartz, & Stone, 2013; Hung et al., 2014a, 2014b; Khanna, Maranian, Rothrock, et al., 2012; Overbeek, Nota, Jayakumar, Hageman, & Ring, 2015) To date, the primary focus of PROMIS-based research has been focused on construct development (Cella et al.,

2007; Fries et al., 2005), primary construct validation (Cella et al., 2010; Fries et al., 2014), and small convergent validation studies. (Hung, Clegg, Greene, & Saltzman, 2011; Hung et al., 2014; Papuga, Beck, Kates, Schwarz, & Maloney, 2014; Tyser, Beckmann, Franklin, et al., 2014).

The PROMIS computer adaptive testing (CAT) instruments are validated, publicly available, NIH-developed instruments that select questions based on item response theory (IRT) to determine standard scores in a few questions. (Broderick et al., 2013; Fries et al., 2014) In most cases, after answering 4 to 7 questions, the assessment is complete. It can be administered on a tablet computer, which is simple and convenient in a healthcare setting. Specific barriers to implementation of systems that improve efficiency and value in health systems have previously been identified. (Coons et al., 2015; Schick-Makaroff & Molzahn, 2015; Segal, Holve, & Sabharwal, 2013; Selby, Beal, & Frank, 2012; Wu & Snyder, 2011; Yoon, Wilcox, & Bakken, 2013) A report by Segal et al. (2013) details many of the issues surrounding the implementation of PROs in a healthcare setting. The cost of to develop, validate, and license software, platforms, and other tools are of specific concern. (Segal et al., 2013; Selby et al., 2012) Reoccurring themes related to clinical use of PRO include the patient experience, privacy, workflow, transparency, utility, and usage rates. (Segal et al., 2013) Healthcare

systems are complex with dynamic interactions between various stakeholders, including payers, regulators, staff, administrators, clinicians, and patients. Added to this are the physical settings, organizational structures, technologies, and procedural norms within which these systems operate. Identifying common barriers and developing methodologies by which to overcome these barriers is necessary to realize PRO integration.

The implementation of PROs as part of each office visit is an effort to improve healthcare value and its effectiveness, efficiency, and patient-centeredness. Previous work has shown that the routine use of PROs has positive effects on patient – provider communication, monitoring of treatment responses, unrecognized problem detection, and patient satisfaction. (Chen et al., 2013) There is also evidence that there are effects on overall patient management. (Chen et al., 2013) Widespread mandated national use in England has led to large effects of patient management, especially in the realm of elective surgeries. (Black, 2013) Incorporation of PRO data into a patient’s EMR can serve as important evidence for future outcomes research as to the efficacy of commonly used medical treatments. Currently, efficacy of treatment is judged by legacy outcome measures, or process measures such as range of motion, radiographic healing, survivorship, reoperation, and infection rate or hospital readmission. The ability for physicians to have access to powerful, standardized, and efficient tools with results imported into the EMR that can be shared with the patient and family to help them understand their progress in treatment is a significant advance in patient care. This is concordant with the Institute of Medicine recommendation for patient-centered care that is respectful of and responsive to individual patient preferences, needs, and values and ensures that patient values guide all clinical decisions. (Institute of Medicine, 2001) With this in mind we set out to (1) develop a system for collection of PROMIS CAT surveys during episodes of care in the clinic setting, and directly import results into the EMR for immediate access by clinicians and (2) consistently attain a greater than 80% capture rate for patients during their clinic visit, using 3 PROMIS CAT instruments.

## 2. Methods

The present study examined the integration of the PROMIS computer adaptive test (CAT) instruments (1) physical function, (2) pain interference, and (3) depression as part of each and every visit made to the orthopaedic clinic of a large academic hospital system. The PROMIS CAT Instruments were chosen for their ease of use, speed, accuracy, scoring, and domain-specific approach. Outcomes data were available to providers during their interaction with their patient to enhance the understanding of the patient’s health situation. This undertaking is part of a university-wide initiative to

operationalize PRO collection as standard of care in all departments. The goals in doing so are to collect outcomes on all patients across all visits in the health system, to minimize the impact on patient flow, and to provide real-time results in an easy-to-use format for providers in order to have the greatest immediate impact on patient care. The present study was reviewed by the local IRB and was deemed exempt.

The implementation strategy was developed and piloted within the Department of Orthopaedic Surgery and Rehabilitation. The primary orthopaedic clinic sees approximately 180,000 outpatient visits per year. Patients treated in the clinic are from a variety of cultural, racial, and ethnic backgrounds and represent a very diverse patient population seeing over 700 patients per day.

A focus group was organized to include patient, clinician, and administrative stakeholders. This focus group was convened bi-weekly to report on progress and receive feedback from all of the stakeholders. The main themes were collected during these focus group meetings, and actions steps were formulated based on those themes (Table 1). One major element of the implementation was the development of support staff positions devoted to help patients in the waiting areas use the tablet and fill out the assessment instruments. This line of communication for patient feedback has been used to focus our aims and keeping patients concerns at the forefront of our planning. A frequently asked questions (FAQ) document, with responses, was created and distributed to registration staff (Appendix 1).

After consulting with information technology consultants from within the institution, we concluded that

**Table 1.** Focus group themes.

Themes	Action items	Week of implementation
Variability in patient engagement	Script development	Week 5
Clean Hardware	Clean and dirty baskets	Week 7
iPad movement with patient	Techs return iPads	Week 1
Hardware Theft	Determent label	Week 8
Depression scale as a “Diagnosis”	Renamed “mood” indicator	Week 2
Stylus use	Make available for limited use	Week 1
Older patients	Assistance available	Week 1
Help staff in waiting room	Define and limit role	Week 1
Provider use of results	Faculty meeting presentations	Week 2, Week 7, Week 12
Provider support for effort	Faculty meeting presentations	Week 2, Week 7, Week 12
Dissemination of information to staff	Staff meeting presentations	Week 1, Week 3, Week 8
Addressing Frequent Patient Questions	Create FAQ document	Week 5
iPad welcome screen	Include explanation/description	Week 8

Notes: Listed are general themes collected in the 8 meetings of the focus group. Action steps are listed for each theme. The purpose of this group was to identify barriers to the patients, staff, and providers in the use of the software, hardware and collection of the patient responses, and any possible solutions. The week of implementation of action is given; however, no appreciable differences in the tracking metrics were found.

unmounted tablets offered the most flexibility. Collecting the data in the waiting room immediately after registration was found to be the best fit for our patient flow. The iPad Mini was selected for our data collection primarily due to the consistency of the software and hardware. The iPad Minis use a core operating system and hardware form factor that have remained predominately unchanged with only small evolutionary updates. This consistency is important for a process of this scale, to avoid having to upgrade hardware frequently. We utilize a mobile device management (MDM) software called Xenmobile Worx Home (Citrix, Fort Lauderdale, FL, version 10) to remotely configure, inventory, secure, and manage the iPad Minis. As theft was a concern, the use of the iPad Mini allowed us to utilize a less expensive tablet thereby mitigating the risk of cost associated with possible loss. 100 iPads were deployed in the clinic none were lost during the 15 weeks.

Based on our experiences setting up and administering PROMIS CAT questionnaires through the Assessment Center web page interface ([www.assessmentcenter.net](http://www.assessmentcenter.net)), it was clear that this method was not going to be able to satisfy the goals of our current project. Web page access to the data was cumbersome to administer, and was unable to provide real-time data. In order to overcome this roadblock, a new piece of software was developed to address the project's goals. This new software permitted administration and collection of the PROMIS CAT data via tablet, and providing access to the data via the EMR. The software integrates with our scheduling software, Flowcast (GE Healthcare, Wilkes-Barre, PA, version 5.1.1). After a patient is registered, the employee triggers the software on their desktop computer which then receives the patient information from Flowcast. It then displays a Quick Response (QR) code which is unique to that patient's encounter. The QR code is scanned by the employee using the iPad Mini which receives the specific information needed for that particular patient encounter. The iPad Mini is now ready for the patient to use. This process requires approximately 10 seconds per patient.

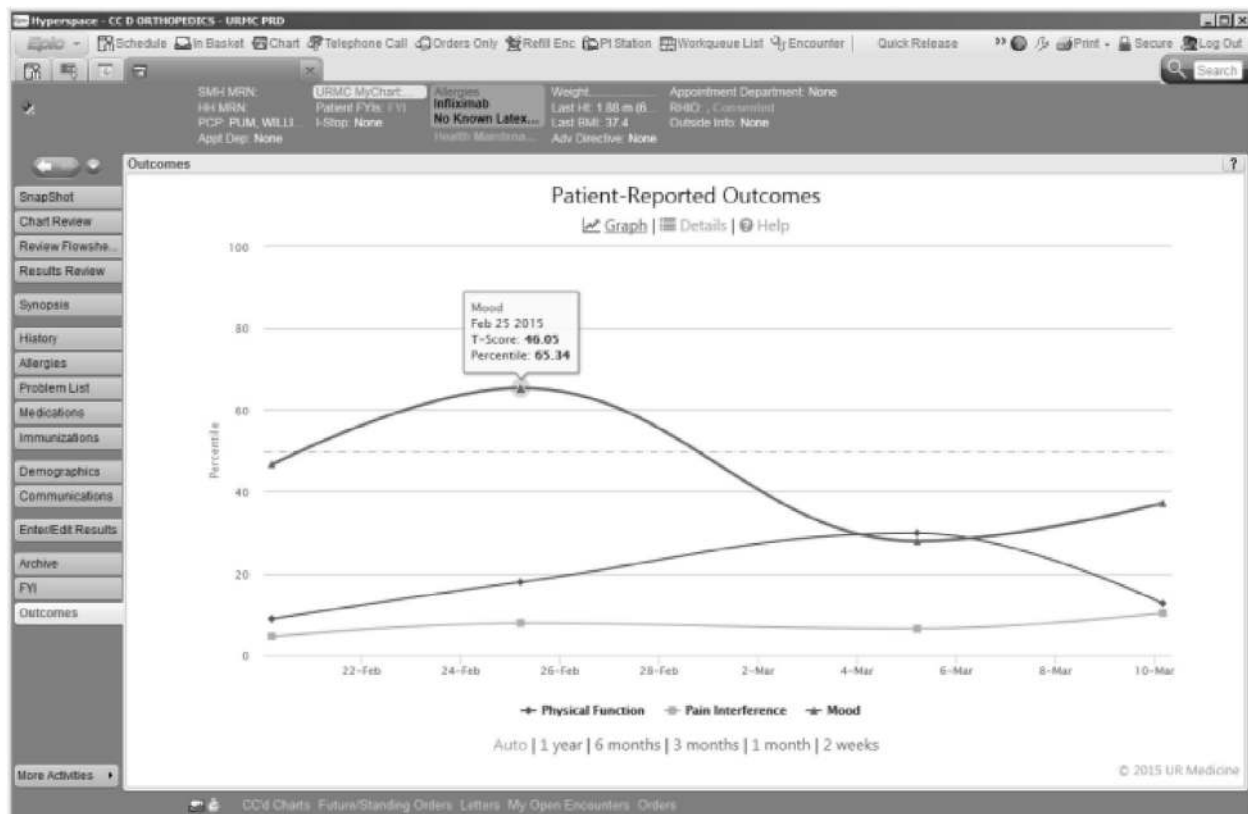
The new software works with the Assessment Center's application program interface which is run on the University servers behind the firewall reducing data privacy concerns. Using information provided by the patient's scheduled encounter, the software selects the PROMIS CAT instruments that are appropriate for that patient. To date we have utilized Physical Function v1.2, Pain Interference v1.1, and Depression v1.0 for all patients. The questions used as part of each PROMIS assessment are copyrighted material and the authors are unable to reproduce them in their entirety. Sample questions can be found at [www.nihpromis.org](http://www.nihpromis.org). The entire sets of questions are freely available to the public through [www.assessmentcenter.org](http://www.assessmentcenter.org), and a demonstration of

the computer adaptive tests is available at <https://www.assessmentcenter.net/ac1/assessments/catdemo>. The physical function assessment asks patients to rate their difficulty in performing activities of daily living such as walking, working, climbing stairs, and getting out of bed. The pain interference assessment asks patients to rate the degree to which pain interferes with aspects of their life such as relationships, leisure activities, household chores, and the ability to take in new information. The depression assessment asks patients to estimate the frequency of feelings such as sadness, loneliness, disappointment, and guilt. Patient participation in data collection is voluntary and they are provided options to quit at any time prior to or during the assessments.

As the PROMIS instruments are completed by patient, the score is immediately available in the EMR system, Epic (Verona, Wisconsin, version 2014). The integration with the EMR was achieved using Epic's Web Activity feature. This feature allows a web page to be displayed within the context of Epic. Epic passes information to the web page about which patient is selected and which user is currently logged in, and the web page then displays the appropriate data. This functionality is available in the "Outcomes" web activity (Figure 1), and is accessible by all providers with access to the EMR. It is active through Epic's Chart Review, Visit Navigator, Patient Station, or Inbasket screen for any particular patient. Selecting the Outcomes activity will display an interactive graph of the patient's scores over time. There are also options to view each question and answer responded to by the patient as well as the option to view helpful information on how PROMIS scoring works. The graph allows the scores to be quickly visualized (Figure 1).

PROMIS data are typically presented as T-scores, a standardized score with a population mean of 50 and a standard deviation of 10 (i.e., where a score of 40 is 1 standard deviation below the population mean). This reporting method was used during the first 3 weeks of implementation. However, based on both Clinician and Patient feedback we chose to display PROMIS scores as a percentile rank which is a more familiar presentation of this type of data. T-scores are converted using a standard conversion table. (Richard Chin, 2008) Percentile rank is plotted on the vertical axis and the visit date is plotted on the horizontal axis. Each instrument is represented on the graph as a different series. Data are presented so that a higher percentile rank represents a better outcome (lower depression, higher physical function, lower pain interference). Hovering over an individual data point shows the underlying T-Score value.

The implementation of the new PRO collection started with one provider, and grew by adding providers and subspecialties weekly. In order to track the progress and efficacy of implementation, several descriptive



**Figure 1.** Longitudinal Presentation of PROMIS CAT Scores in Electronic Medical Record. An example of a patient's PROMIS scores as seen in the patient's Electronic Medical Record (EMR). Scores are graphically represented as a percentile rank, based on the general population. This display is present in the larger context of the EMR under the Web activity "Outcomes" vertical tab. The graph is interactive and can display one or all three of the PROMIS scores, and using the mouse hovering over one of the points will show details of the that particular data point, including the associated *T*-Score.

tracking statistics were collected over time. The specific tracking measures are as follows:

Encounters Administered	Number of patients arrived for clinic appointment.
Declined	Number of patients starting the PRO system.
Quit	Number of patients immediately declining to participate on welcome screen.
Abandoned	Number of patients choosing to quit at some point other than on welcome screen.
Completed	Number of patients who do not complete all assessments but whom do not quit.
	Number of patients answering all questions reaching final thank you screen.

Registration time was tracked by timestamps available through the scheduling software and defined as the time from the opening the patient's schedule at the intake desk until exiting that patient's schedule/record when the patient is sent to the waiting room. This time does not include the completion of the survey instruments, which are completed while the patient is waiting to see the doctor after registration. The iPad is collected from the patient primarily by staff in the waiting room or when directing patients to their exam room.

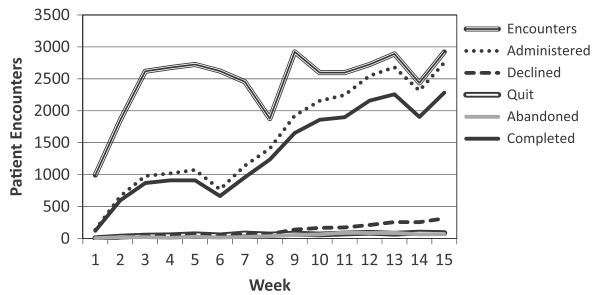
These metrics can be tracked by specific location, individual provider, day, hour, and by any combination of such. This has enabled us to identify problems in the implementation and address them in order to improve our ability to track outcomes. Demographic data were collected from the scheduling software for each patient encounter. Differences based on implementation time

point, patient demographics, and orthopaedic were tested by means of a multivariate ANOVA conducted on raw data. Tukey's Post hoc analysis was performed to address pairwise comparisons between individual group means. All analyses used a significance level of .05.

### 3. Results

Over the first 15 weeks of implementation using our new software interface, we administered the three PROMIS CAT instruments in 26,227 patient encounters out of the 29,227 we intended. This includes 17,892 unique individuals. The average number of questions for the physical function assessment was  $4.57 \pm 1.33$  with a median of 4 questions. The average number of questions for the pain interference assessment was  $4.48 \pm 1.93$  with a median of 4 questions. The average number of questions for the depression assessment was  $6.32 \pm 3.54$  with a median of 4 questions. The average number of questions for the entire set of assessments was  $15.39 \pm 4.75$  with a median of 13 questions. The average time to complete all assessments was  $3.48 \pm 3.10$  min, with a median time of 2.58 min. Registration times for new patient appointments did not change significantly, lasting  $7.19 \pm 2.69$  min compared to  $6.87 \pm 3.34$  min with and without the use of the new software, respectively. However, registration times were statistically significantly higher for

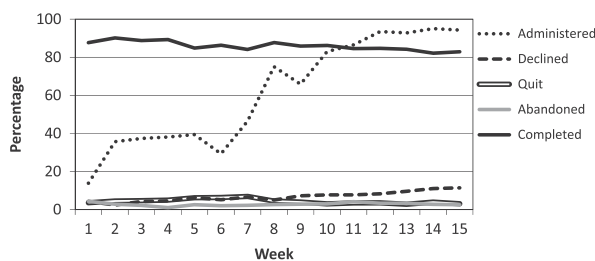




**Figure 2.** Weekly Implementation Tracking Metrics; Volume. The graph shows the weekly progression in administration volume over the first 15 weeks. The increase in volume is directly related to the “rollout” by which additional providers and sub-specialties were added each week. Week-to-week decreases were due to scheduling irregularities whereby providers were not seeing patients.

follow-up appointments ( $p = .007$ ) taking  $3.32 \pm 1.78$  min with the new software compared to  $2.94 \pm 1.57$  min without. Over the 15 weeks reported here, one iPad Mini was lost/stolen.

During the first week, patients seeing one provider were given questionnaires using the new software, this resulted in 138 administrations. Administration expanded in subsequent weeks to additional providers and across sub-specialties and by week 10 the entire orthopaedic outpatient clinic (excluding the Pediatric patients) was collecting PROMIS data. This expansion resulted in 2450 administrations on average from weeks 10–15. Tracking metrics (described above) show the gradual increase in usage (Figure 2) relative to the total number of patient encounters in the outpatient clinic, as well as a breakdown of how each administration was terminated. These indicators show a steady rise in the administration rate with stabilization during weeks 12–15 (Figure 2). The percentage of administrations that result in patients declining to participate had a significant increase ( $p < .001$ ) over time which contributed to a significant decrease in completion rate



**Figure 3.** Weekly Implementation Tracking Metrics; Completion rates. The graph shows the weekly progression in administration and completion rates over the first 15 weeks. The completion rate stays above our goal of 80% throughout the implementation. Increasing decline rates after week 9 may be a result of patient fatigue. Relative low rates of quit and abandoned indicate adequate time to complete and adequate wireless infrastructure.

( $p < .001$ ) over the same period (Figure 3). Pairwise Post hoc analysis based on the time from patients last appointment (by week) failed to show that those with more recent visits were most likely to decline the assessments.

While there were no significant difference based on gender found in the administration, declined, or completion rates, the higher rate of men who quit ( $p = .021$ ) or abandon ( $p = .002$ ) was statically significant (Table 2). Also we found that time to complete all assessments was significantly higher ( $p = .017$ ) for men  $3.53 \pm 3.10$  min than for women  $3.46 \pm 3.09$  min. However, this equates to approximately 6 s difference which is likely not practically meaningful. Calculated PROMIS scores showed statistically significant differences by gender as a whole across all patients with men having higher physical function scores ( $43.3 \pm 10.1$  vs.  $41.0 \pm 9.2$ ,  $p < .001$ ), lower pain interference scores ( $58.6 \pm 8.6$  vs.  $\pm 8.4$ ,  $p < .001$ ), and lower depression scores ( $47.9 \pm 10.3$  vs.  $50.4 \pm 10.1$ ,  $p < .001$ ). However, distribution-based threshold of evaluating the minimal clinically important difference (MCID) of  $1/2$  the standard deviation fail to show these differences as meaningful. (Norman, Sloan, & Wyrwich, 2003; Sloan, Symond, Vargas-Chanes, & Friendly, 2003).

Significant differences in administration rate were found based on age ( $p < .001$ ). Pairwise Post hoc analysis showed that only those 80 + were administered assessments less frequently. However, significant differences were found between age groups ( $p < .001$ ) for all other tracking metrics as well, with consistent increases in declined, quit, and abandoned assessments with increasing age, resulting in gradually decreasing rates of completion (Table 2). Pairwise Post hoc analysis based on age showed that an increase in age difference increased the likelihood of significant differences between groups for each of the tracking metrics measured. Time to complete increased significantly with age ( $p < .001$ ); post hoc analysis showed that those over 50 took significantly longer than younger groups and that subsequent age groups were significantly different from all other age groups (Table 3). PROMIS scores also differed significantly by age ( $p < .001$ ); however, only function seems to clearly be clinically important based on the MCID threshold of  $1/2$  SD (Table 3).

Significant differences in administration rate were found based on specialty ( $p < .001$ ); post hoc analysis showed that for declined, quit, and abandoned metrics a difference of more than 1% between groups was significantly different, while the completion metric differing by more than 5% was found to be significant (Table 4). A significant difference in time to complete was found based on specialty ( $p < .001$ ); however, none of these differences were found to be practically significant. PROMIS scores also differed significantly by specialty

**Table 2.** Completion rates by gender and age.

Gender	Administered (n)	Declined (%)	Quit (%)	Abandoned (%)	Completed (%)
F	13,563	7.96	3.80	2.61	85.63
M	10,250	7.88	4.29*	3.30*	84.53
Age	Administered	Declined (%)	Quit (%)	Abandoned (%)	Completed (%)
18–19	649	4.01	2.93	2.62	90.45
20–29	2161	2.59	2.45	1.57	93.38
30–39	2308	5.20	2.47	1.26	91.07
40–49	3907	6.81	3.25	1.51	88.43
50–59	6050	7.06	3.69	2.26	86.99
60–69	4861	8.66	4.28	3.48	83.58
70–79	2580	10.93*	6.20	5.81	77.05
80+	1297*	22.36*	8.33*	7.48*	61.84*

\*Indicates a significant difference from all other groups based on post hoc pairwise comparisons.

**Table 3.** Assessment duration and T-scores by age.

Age	Time to complete (Min)		Function		Pain		Depression	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
18–19	2.83	1.89	46.53	10.64	55.59	8.78	45.24	9.69
20–29	2.87	3.05	44.74	10.61	57.60	8.56	47.65	10.55
30–39	2.84	2.27	42.60	10.03	60.24	8.37	49.86	10.77
40–49	3.00	2.38	42.04	9.83	60.54	8.52	50.09	10.88
50–59	3.38*	2.97	41.84	9.23	60.04	8.27	50.06	10.23
60–69	3.82*	3.49	41.86	9.14	58.78	8.32	48.93	9.61
70–79	4.65*	3.74	39.92	8.99	58.57	8.64	48.50	9.31
80+	5.12*	3.69	36.32	8.81	58.98	8.27	50.21	9.35

Notes: The table indicates Assessment duration given in minutes. Assessment scores for each of the three domains are given as a T-Score.

\*Indicates a significant difference from all other groups based on post hoc pairwise comparisons.

**Table 4.** Completion rates by specialty.

Specialty	Administered	Declined (%)	Quit (%)	Abandon (%)	Completed (%)
Foot and Ankle	4514	8.26	3.54	1.77	78.25
General	575*	12.87*	4.7	5.57	70.26
Hand and wrist	6786	7.16	4.23	2.42	75.6
Joint replacement	1984	6.60	3.83	4.23	75.76
Spine	4621	7.88	2.12	2.45	67.13
Sports	7102	5.18	3.11	2.04	82.77
Trauma	2587	8.39	3.52	3.32	66.87
MISC	1058*	7.37	4.91	3.69	55.86*

\*Indicates a significant difference from all other groups based on post hoc pairwise comparisons.

**Table 5.** Assessment duration and T-scores by specialty.

Specialty	Time to Complete (Min)		Function		Pain		Depression	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Foot and ankle	3.24	3.1	39.65	9.51	58.67	8.95	49.11	10.02
General	3.77	3.6	43.28	9.5	58.02	8.53	49.16	10.27
Hand and wrist	3.6	3.05	45.42	9.78	57.96	8.63	48.63	10.43
Joint replacement	3.71	2.8	42.53	9.55	59.01	8.22	49.04	9.99
Spine	3.39	3.1	38.73	7.94	63	7.47	52.38	10.27
Sports	3.31	2.96	42.78	8.84	59.03	7.75	48.12	9.86
Trauma	3.79	3.35	37.43	10.32	60.4	9.15	51.42	10.29
Miscellaneous	4.16	3.99	45.71	11.23	58.14	9.26	46.55	9.25

Note: The table indicates assessment duration given in minutes. Assessment scores for each of the three domains are given as a T-Score.

( $p < .001$ ), using clinically important based on the MCID threshold of  $\frac{1}{2}$  SD those differences larger than 5 are most likely to be clinically important (Table 5). Averages and standard deviations for duration, function, pain, and depression are given stratified by both specialty and age (Table 6).

## 4. Discussion

### 4.1. Implementation

The tracking metric results in general indicate a robust collection rate of above 80% for the vast majority of the patient population that we have examined. Those above

**Table 6.** Assessment scores for each of the three domains are given as an average T-Score (+/- stdev).

Age	Duration															
	Foot and ankle		General		Hand and Wrist		Joint replacement		Spine		Sports		Trauma		MISC	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
18–19	2.35	1.18	3.22	1.84	3.41	2.57	2.46	1.16	2.89	1.65	2.72	1.72	3.1	2.62	3.08	1.86
20–29	2.42	1.93	2.61	1.31	3.23	3.83	2.79	2.41	2.65	2.1	2.78	3.34	3.26	2.68	3.59	3.55
30–39	2.6	2.05	2.9	1.54	2.92	1.89	3.18	2.93	2.58	1.88	2.83	2.47	2.89	1.82	4.32	4.82
40–49	2.67	2.15	3.35	2.79	3.2	2.15	3.12	1.77	2.93	2.89	2.96	2.43	3.07	2.25	3.55	3.13
50–59	3.19	3.29	3.62	2.93	3.47	2.45	3.32	2.3	3.29	3.37	3.35	2.82	3.5	3.28	4.26	4.96
60–69	3.88	3.79	4.43	4.41	3.94	3.88	3.81	2.46	3.55	3.17	3.57	2.98	4.54	4.24	3.97	3.3
70–79	4.36	3.6	5.04	3.31	4.67	3.59	4.72	3.75	4.37	3.46	4.59	3.94	5.43	4.36	5.9	4.15
80+	4.93	3.98	6.55	9.22	4.88	2.99	5.73	3.68	5.1	3.47	4.91	3.45	5.27	3.5	4.14	2.48
Age	Function															
	Foot and ankle		General		Hand and wrist		Joint replacement		Spine		Sports		Trauma		MISC	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
18–19	43.32	10.66	47.82	11.47	48.91	9.09	44.97	9.05	44.47	8.99	47.47	10.76	44.12	12.91	49.26	11.81
20–29	41.86	10.16	48.27	10.41	49.18	10.1	46.43	11.1	42.08	7.98	44.9	9.76	39.12	11.88	47.26	12.43
30–39	40.11	10.33	44.14	10.01	46.91	10.05	42.7	9.56	39.75	7.5	43.19	9.28	38.54	10.21	45.48	12.61
40–49	39.61	9.51	41.45	8.79	45.47	10.33	43.98	9.75	39.13	8.27	42.72	8.73	37.11	9.64	45.25	12.45
50–59	39.6	9.01	42.94	9.07	44.91	9.64	43.19	9.26	38.42	7.57	42.63	8.28	37.76	9.35	45.94	11.08
60–69	39.25	9.1	42.33	7.75	45.18	9.12	43.37	8.51	38.8	8	42.17	8.15	37.76	10.32	45.48	9.45
70–79	38.79	9.28	42.4	8.58	43.5	8.87	39.58	8.51	36.47	7.48	40.75	7.86	36.15	10.29	45.51	10.46
80+	35.12	9.1	38.88	10.61	39.97	8.94	36.5	8.83	35.47	7	37.77	8.04	31.16	8.52	37.32	8.03
Age	Pain															
	Foot and ankle		General		Hand and wrist		Joint replacement		Spine		Sports		Trauma		MISC	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
18–19	54.72	8.15	54.75	9.28	54.47	10	58.61	9.32	59.13	6.87	54.92	8.22	55.15	11.05	59.01	7.56
20–29	57.96	8.77	55.59	7.5	55.91	8.67	57.44	7.76	61.15	7.21	56.77	7.92	59.56	10.32	59.19	8.47
30–39	60.54	8.62	57.06	8.16	58.54	8.63	61.3	7.4	63.68	6.75	59.68	8.19	60.05	8.79	59.83	7.62
40–49	59.58	8.97	61.14	9.3	59.6	8.93	59.5	8.84	64.09	7.82	60.02	7.49	62.36	8.2	59.46	10.01
50–59	58.79	8.92	58.2	8.73	58.96	8.56	60.51	7.2	63.67	7.39	59.71	7.5	61.54	8.56	57.93	9.69
60–69	58.37	8.86	57.55	8.61	57.13	8.26	57.06	8.82	62.27	7.59	59.09	7.38	59.77	8.92	57.14	9.34
70–79	57.03	9.32	58.7	6.44	56.41	8.12	58.18	8.71	62.89	7.39	58.59	7.64	58.52	10.14	54.9	10.38
80+	59.05	8.68	57.42	7.22	56.14	7.93	58.4	8.34	62.28	6.96	59.29	7.96	59.07	9.26	60.4	8.34
Age	Depression															
	Foot and ankle		General		Hand and wrist		Joint replacement		Spine		Sports		Trauma		MISC	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
18–19	47.25	8.72	48.31	8.92	44.43	10.28	45.02	10.66	48.61	11.07	43.66	9.23	47.33	10.49	45.21	10.16
20–29	48.71	10.79	45.1	9.03	46.78	10.8	46.2	10.56	51.07	10.85	46.07	9.73	51.84	10.34	44.88	8.69
30–39	49.76	10.85	46.43	11.17	48.45	10.78	49.33	11.13	53.17	10.26	49.47	10.48	50.39	11.19	47.5	9.69
40–49	49.55	10.63	51.87	11.4	49.41	11.47	50.21	10.93	53.68	10.67	48.39	10.2	53.42	10.47	47.01	10.78
50–59	49.01	9.6	51.93	10.45	49.51	10.47	50.07	9.74	53.99	10.31	48.87	10.08	52.01	9.87	47.31	9.35
60–69	49.14	9.77	49.42	8.68	48.46	9.76	48.67	9.31	51.17	9.93	47.84	9.17	50.45	9.67	46.62	8.6
70–79	47.71	9.16	44.46	7.97	47.99	9.05	48.19	9.23	50.89	9.49	48.26	8.89	49.23	10.76	44.94	8.63
80+	50.63	9.88	49.87	9.45	48.3	8.96	50.14	9.62	51.5	9.3	49.71	9.17	52.02	9.51	47.79	7.7

Notes: The table indicates assessment duration given in minutes (+/- stdev) for each age group. Assessment scores for each of the three domains are given as an average T-Score (+/- stdev).

80 years of age were less likely to be administered the assessments and were also less likely to complete them. It has been shown that the elderly population has higher rates of survey refusal due to overall poor health, cognitive impairments, and lower levels of education. (Jacomb, Jorm, Korten, Christensen, & Henderson, 2002; Mihelic & Crimmins, 1997; Norton, Breitner, Welsh, & Wyse, 1994) Other investigators point to increase skepticism among the elderly and the need to build rapport in this population to foster feelings of trust, safety, and confidentiality. (Murphy, Schwerin, Eyerma, & Kennet, 2008) Differences in quit and abandon rates based on gender while statistically significant are likely not practically meaningful. The statistically significant decrease in completion rate coupled with the significant increase in declined rate overtime was concerning and pointed to

survey fatigue. However, when we examined the decline rate based on how recently the patient had been administered the assessments, no correlation was found. Future work might focus on the total number of administrations or an average number over a well-defined period of time. Differences in PROMIS scores based on age and specialty, perhaps unsurprisingly, showed that in general older individuals take longer to complete the assessments and have lower physical function. The interesting finding was that while there were statistically significant differences based on age for pain and depression, these did not approach the MCID. Differences greater than the MCID were found between patients being treated by different orthopaedic specialties. This is likely due to the nature of the underlying pathologies and/or injuries treated by those specialties rather than differences in

care, or any attributable bias. It is therefore a reasonable finding that trauma has both the lowest physical function average as well as the highest pain average.

#### 4.2. Electronic capture of PRO data

The routine use of PROs as standard clinical practice in the United States is very limited. (Black, 2013; Chen et al., 2013) The utility of PROs as routine measures in clinical practice used for individual treatment plans has yet to be adequately demonstrated. (Black, 2013) Our goals of quick and efficient collection and the real-time insertion of the data into the EMR were achieved. The power of routine collection offers great potential to elucidate many questions as to the comparative effectiveness of particular treatments when examining the aggregate data. (Speerin et al., 2014; Wei, Hawker, Jevsevar, & Bozic, 2015) We have demonstrated with the right implementation strategy PROs can be seamlessly integrated into a high-volume clinical setting. This assertion is supported by our high collection rates and by the low number of administration, staff, and patient complaints. While there was no formal complaint process, we found that none of bi-weekly focus group meetings focused on themes related to the burden of data collection (Table 1). These meetings acted as feedback and a chance for us to encourage the staff to keep up the effort. The greatest burden was placed upon the registration staff, having to explain the purpose of the survey instruments and the importance of collection at each (and every) office visit, coupled with the logistics of handing out, retrieving, and cleaning tablet computers. The dedication of staff and/or volunteers to assist with waiting room technical/logistical issues is critical. A key area of concern prior to implementation was the disruption of patient flow within the clinic. We found that allowing patients to finish while in the waiting room directly following check-in is preferable; however, if need be the tablet can accompany the patient in the clinic environment and be completed during other available wait times. The continuing challenge will be how to best integrate and utilize the information now available in our healthcare system. Proposed future work may focus on the identification of patient characteristics for which threshold scores on each assessment may trigger actionable steps or suggest alternative care plan options.

Hung and colleagues have published on the use of PROMIS in the clinical setting of foot and ankle practice and found it was effective. (Hung et al., 2013, 2014) Hung et al. published similar findings demonstrating the use of the PROMIS physical function item bank in spinal disorders. (Hung et al., 2014) One manuscript discussed PROMIS measures as representing a trans-Department of Health and Human Services effort to develop a standard set of measures for informing decision making in clinical research, practice, and health policy. (Reeve et al., 2007) The PROMIS CAT has been found to be

highly sensitive with orthopaedic diagnoses with very little floor or ceiling effects note (Fries et al., 2014; Hung et al., 2011, 2013, 2014; Papuga et al., 2014) It has been demonstrated that the PROMIS CAT surveys represent a low patient burden outcome measure in the orthopaedic clinic. (Hung et al., 2011, 2014; Papuga et al., 2014) The PROMIS CAT was found to be high precision with fewer questions which makes it more feasible for use in clinical practice. (Bjorner et al., 2014; Wagner, Schink, Bass, et al., 2015) Here we have shown that the low patient burden previously described for the PROMIS CAT combined with an effective implementation strategy can be low burden for the clinic staff as well.

Our choice to incorporate PROMIS CAT into the EMR was based on the broad applicability of the instruments, the low patient burden, and the improvements that IRT offered over other traditional legacy instruments. In a recent review of the literature, Coons et al. found that there is mounting evidence that electronic capture of function and symptom experience outperform unsupervised paper-based self-reports. (Coons et al., 2015) Electronic capture can lead to more accurate and complete data (Gwaltney, Shields, & Shiffman, 2008; Zbrozek et al., 2013), the avoidance of secondary data entry errors (Coons et al., 2015; Zbrozek et al., 2013), less administrative burden (Dale & Hagen, 2007; Greenwood, Hakim, Carson, & Doyle, 2006), high respondent acceptance (Bushnell, Reilly, Galani, et al., 2006; Fritz, Balhorn, Riek, Breil, & Dugas, 2012; Greenwood et al., 2006; Velikova et al., 1999), and potential cost savings. (Benedik et al., 2014; Fritz et al., 2012; Touvier, Mejean, Kesse-Guyot, et al., 2010; Uhlig, Seitz, Eter, Promesberger, & Busse, 2014) While other electronic capture software exists, the creation of a new software product was crucial in our ability minimally affect workflow, to deploy assessments in addition to PROMIS CAT, and have the results available for the provider in the electronic medical record immediately. The integration with the scheduling software eliminates the potential for several sources of error while also decreasing the burden upon registration staff. The registration staff has to simply pull up the provider and select that specific appointment. The use of the QR code technology eliminates several steps and additional potential for error. While primarily developed for the use of PROMIS CAT tools, this platform does allow for the incorporation of any survey-based assessments. The use of disease specific assessments may be critical in the broader distribution of PROs across the healthcare system. (Black, 2013; Chen et al., 2013; Nelson et al., 2015) Our new software methodology both lowers the potential burden of collection while eliminating sources of error and allow real-time access to the results.

The functionality of our new software incorporates robust tools for analyzing the acquired data in aggregate to identify potential problems in assessment



administration. An appropriate directive for use of the data beyond that of individual patient care must be established in order to realize the potential for clinic-wide or systemwide impact on healthcare in general. (Black, 2013; Chen et al., 2013; Cutler & Ghosh, 2012; Institute, 2001; Nelson et al., 2015) There is potential to use this type of data to compare specific treatments, providers, divisions, etc.; however, the first question to be asked is, does having access to this type of information change the way providers and patients choose and then respond to different types of treatment? The effect of having access to PRO data and the utilization of these data is key to understanding its potential.

### 4.3. Lessons learned

The stage-wise implementation strategy used here shows a steady increase in the number of patient encounters during which PROMIS surveys were administered. This staged approach to implementation allowed for a growing familiarity with the new processes prior to becoming available department wide. Initially staff were asked to use the process for specific providers only. In this way, if problems were encountered, the effect was isolated and the problem resolved. As we revised our processes and became more confident, we were able to include additional providers and sub-specialties. Several large decreases in the week-to-week numbers can be attributed to specific causes. During week 6 we had a wireless infrastructure failure, week 8 included a holiday recess so the clinic volume was low, and week 14 had low volume due other irregularities that led to a light clinic schedule (Figure 2).

Proper training and availability of resources are key to the engagement of crucial stakeholders in the adoption of any new technology. (Krau, 2015; May, 2015; Sand-Holzer, Deutsch, Frese, & Winter, 2015) The education of clinicians in the underlying principles of PROMIS CAT scoring and IRT provided sufficient evidence for the use of such instruments. Three specific points were addressed during the education of clinical staff: (1) generalizability of PROMIS scores, (2) basic concepts of IRT, and (3) patient education in use and availability of PROMIS scores. One area of focus going forward will be to demonstrate a sustained increase in provider use. Beyond familiarity with the instruments, continued education of providers will include the identification of specific instances in which the use of PRO data may be used to justify changes in the management of an individual's healthcare. In order to fully realize this type of impact a provider will need to be able to provide additional context to the scores, beyond how they relate to the general population. This context will include information about how an individual's scores relate to others with a similar demographics, history, diagnosis, and general state of health. Dissemination of aggregate data to provide this context must be organized and presented so that

it is easily understood by both provider and patient. Providers must also be given additional resources such as referral services if PRO data become actionable.

There were several lessons learned during this process. The importance of clear channels of communication that have broad distribution among all levels of clinical personal is key to effective implementation of new processes. Making sure that a consistent message is conveyed from top to bottom leaves little room for ambiguity. Keeping everyone up to date on short-term and long-term goals and the progress made was critical to our success. We also learned that staff not directly involved with specific processes have the potential to harm the effort by being unaware of their existence. Patient feedback reflected a desire to understand the purpose behind the assessments, the intended use of the data, and an apprehension about the tablet computer if they had never used one before. Patient voices also informed the focus group themes listed in Table 1 and helped develop specific actions to help address those concerns. The FAQ document (Appendix 1) that has been developed as a handout for patients given prior to completing an assessment has helped us to better inform patients. In general, patients are very willing to complete the assessments, as evident by the high completion rates reported in Table 2, and are interested in the results based on provider feedback given below. This was especially true in the patient's interaction with providers; it was important that every provider had access to the data and were able to talk about PROMIS results if/when patients asked.

### **Feedback from providers:**

The patients appreciate the 'tracking' of their scores over time and our interaction is enhanced. For example, if the patient has Achilles tendonitis and we look at the physical function scores, pain interference scores at the first visit, highlighting that the patient is not as active and has more pain than the average US population. We discuss options for treatment and agree on trying physical therapy and antiinflammatory medication. We see each other again in 1 month and look at the PRO scores. The patient reports that he doesn't know if he is significantly improved. We look at his scores and reports slight improvement with better function and less pain. With this information, we decide to continue with physical therapy for another month. The patient cost to continue includes the copy for therapy however he thinks this is worth the cost if he can avoid surgery. He returns again after another month and states he is better and his scores also reflect this. If we did not have the PRO to view at the first follow-up visit, the direction of care may have changed due to the patient's report of not feeling better but the very discriminating and sensitive PRO demonstrated some improvement. The direction of care change would have advanced to obtaining an MRI in preparation for surgery. Those conversations occur every day and truly impact the care we give to our patients.

If a patient is seen for ankle arthritis and has very low Physical Function score, I know from examining all ankle arthritis patients that this patient would benefit from surgery. Her scores are very low and often brace-wear and medication are not effective. I share this experience with the patient so she can make a good choice if considering a custom brace that would cost over \$800.00 dollars that might not be effective. No one wants to pay money on something that hasn't been shown to be effective.

When a patient has a depression score that is increasing while physical function and pain scores are improving, I ask the patient about what is going on, and can I help you? I show him his scores and ask what is concerning him? He might say that with his injury he hasn't been able to work and the rent is coming due. It helps me enter into a conversation that I might not have had before PRO. In this case, I was able to talk about light duty with his boss and helped him to meet the rent payment goal.

Registration and technical staff scripting was key to effectively keeping our message clear and consistent for the patients, and in turn for the staff. Script revisions based on stakeholder feedback were instrumental in the ability to come up with a script that was brief and informative, while allaying patient privacy fears. We found that words such as “survey,” “depression,” and “research” had negative connotations with some of our patients. We replaced them with assessment, mood, and standard of care and saw improved engagement.

The wireless infrastructure was strained during this pilot, and the addition of 100 additional devices proved to be a challenge. We saw intermittent connectivity issues throughout the pilot. Although they only affected a small percentage of encounters, the problems created frustration, for patients and staff. Some areas of the clinic had weak wireless signal and were unable to support the process if the tablets were used in those areas. There was also a concern of how many simultaneous connections a single access point was able to handle successfully, and provide sufficient functionality. An enhanced wireless survey was performed to understand and address these problems. One weakness of the current system is that it was developed to work in a specific healthcare system and with specific scheduling and EMR software. Customization would be needed in order for this system to be used in other healthcare systems.

The collection and use of patient-centered data has begun at our institution with the intention of increase in the value of healthcare to both patients and providers. Traditional measurements of value in healthcare are tied to complication rates, mortality rates, and monetary cost. One goal that has been stated as we move forward is to make use of the PROMIS data to drive down clinical variability based on provider preference and encourage the development of standardized treatment protocols. The use of standardized clinical protocols within departments should allow for a more accurate estimate

of clinical costs by which to standardized procedural costs. These standardized costs can then be compared to the expected change in patient outcomes for a particular treatment or procedure. Having these cost and outcome data available for both patient and provider prior to making healthcare decisions can help reveal the relative value of the treatment to an individual patient. The ability to provide this type of transparency to patients based on objective measures of outcomes that patients really care about could provide a competitive market advantage in the healthcare marketplace. However, there is considerable skepticism on the part of clinicians that standardization of care pathways is valuable. The ability to better quantify patient outcomes as we standardize care, and hopefully demonstrate that outcomes are preserved or improved will be instrumental in driving departmental behavior towards standardization.

There were several limitations of this study, including the preliminary nature of our short implementation window, and none of the metrics collected conclusively illustrate the long-term viability of such a collection. Future work should focus on the clinical usage rate with a focus on qualitative analysis to address the clinical impact on patient care as viewed both by the providers and the patients.

## 5. Conclusion

Here we have demonstrated that with the right implementation strategy, PROs can be effectively integrated into a high volume clinical setting with minimal burden on patients, staff, and providers. We believe that the valuable experience we have gained in this initial rollout will serve to help us to improve patient care. The availability of this type of data across our system opens up many opportunities for care improvement and research. The next important steps will be to assess provider usage and the efficacy of use of the PROMIS data during the clinical visit and its impact on patient care.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## References

- Benedik, E., Korousic, S. B., Simcic, M., Rogelj, I., Bratanič, B., Ding, E. L., & Mis, N. F. (2014). Comparison of paper- and web-based dietary records: A pilot study. *Annals of Nutrition & Metabolism*, 64, 156–166.
- Bjorner, J. B., Rose, M., Gandek, B., Stone, A. A., Junghaenel, D. U., & Ware, J. E., Jr (2014). Method of administration of PROMIS scales did not significantly impact score level, reliability, or validity. *Journal of Clinical Epidemiology*, 67, 108–113.
- Black, N. (2013). Patient reported outcome measures could help transform healthcare. *BMJ*, 346, f167.
- Broderick, J. E., Schneider, S., Junghaenel, D. U., Schwartz, J. E., & Stone, A. A. (2013). Validity and reliability of patient-reported outcomes measurement information system instruments in osteoarthritis. *Arthritis Care Res (Hoboken)*, 65, 1625–1633.

- Bushnell, D. M., Reilly, M. C., Galani, C., et al. (2006). Validation of electronic data capture of the Irritable Bowel Syndrome-Quality of Life Measure, the Work Productivity and Activity Impairment Questionnaire for Irritable Bowel Syndrome and the EuroQol. *Value Health*, 9, 98–105.
- Cella, D., Riley, W., Stone, A., et al. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*, 63, 1179–1194.
- Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., ... Rose, M. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. *Medical Care*, 45, S3–S11.
- Chen, J., Ou, L., & Hollis, S. J. (2013). A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Services Research*, 13, 211.
- Coons, S. J., Eremenco, S., Lundy, J. J., O'Donohoe, P., O'Gorman, H., & Malizia, W. (2015). Capturing patient-reported outcome (PRO) data electronically: the past, present, and promise of ePRO measurement in clinical trials. *Patient*, 8, 301–309.
- Cutler, D. M., & Ghosh, K. (2012). The potential for cost savings through bundled episode payments. *New England Journal of Medicine*, 366, 1075–1077.
- Dale, O., & Hagen, K. B. (2007). Despite technical problems personal digital assistants outperform pen and paper when collecting patient diary data. *Journal of Clinical Epidemiology*, 60, 8–17.
- Fries, J. F., Bruce, B., & Cella, D. (2005). The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. *Clinical and Experimental Rheumatology*, 23, S53–S57.
- Fries, J. F., Witter, J., Rose, M., Cella, D., Khanna, D., & Morgan-Dewitt, E. (2014). Item response theory, computerized adaptive testing, and PROMIS: assessment of physical function. *Journal of Rheumatology*, 41, 153–158.
- Fritz, F., Balhorn, S., Riek, M., Breil, B., & Dugas, M. (2012). Qualitative and quantitative evaluation of EHR-integrated mobile patient questionnaires regarding usability and cost-efficiency. *International Journal of Medical Informatics*, 81, 303–313.
- Greenwood, M. C., Hakim, A. J., Carson, E., & Doyle, D. V. (2006). Touch-screen computer systems in the rheumatology clinic offer a reliable and user-friendly means of collecting quality-of-life and outcome data from patients with rheumatoid arthritis. *Rheumatology*, 45, 66–71.
- Gwaltney, C. J., Shields, A. L., & Shiffman, S. (2008). Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health*, 11, 322–333.
- Hung, M., Baumhauer, J. F., Latt, L. D., Saltzman, C. L., Soohoo, N. F., & Hunt, K. J. (2013). Validation of PROMIS (R) Physical Function computerized adaptive tests for orthopaedic foot and ankle outcome research. *Clinical Orthopaedics and Related Research*, 471, 3466–3474.
- Hung, M., Clegg, D. O., Greene, T., & Saltzman, C. L. (2011). Evaluation of the PROMIS physical function item bank in orthopaedic patients. *Journal of Orthopaedic Research*, 29, 947–953.
- Hung, M., Baumhauer, J. F., Brodsky, J. W., Cheng, C., Ellis, S. J., Franklin, J. D., ... Saltzman, C. L. (2014a). Psychometric comparison of the PROMIS physical function CAT with the FAAM and FFI for measuring patient-reported outcomes. *Foot and Ankle International*, 35, 592–599.
- Hung, M., Franklin, J. D., Hon, S. D., Cheng, C., Conrad, J., & Saltzman, C. L. (2014b). Time for a paradigm shift with computerized adaptive testing of general physical function outcomes measurements. *Foot and Ankle International*, 35, 1–7.
- Hung, M., Hon, S. D., Franklin, J. D., et al. (2014c). Psychometric properties of the PROMIS physical function item bank in patients with spinal disorders. *Spine*, 39, 158–163.
- Institute of Medicine. (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington DC: National Academy of Science.
- Jacomb, P. A., Jorm, A. F., Korten, A. E., Christensen, H., & Henderson, A. S. (2002). Predictors of refusal to participate: a longitudinal health survey of the elderly in Australia. *BMC Public Health*, 2, 4.
- Khanna, D., Maranian, P., Rothrock, N., et al. (2012). Feasibility and construct validity of PROMIS and “legacy” instruments in an academic scleroderma clinic. *Value Health*, 15, 128–134.
- Krau, S. D. (2015). Technology in nursing: the mandate for new implementation and adoption approaches. *Nursing Clinics*, 50, 11–12.
- May, C. R. (2015). Making sense of technology adoption in healthcare: Meso-level considerations. *BMC Medicine*, 13, 92.
- Mihelic, A. H., & Crimmins, E. M. (1997). Loss to follow-up in a sample of Americans 70 years of age and older: the LSOA 1984–1990. *Journals of Gerontology*, 52B, S37–S48.
- Murphy, J., Schwerin, M., Eyerman, J., & Kennet, J. (2008). Barriers to survey participation among older adults in the national survey on drug use and health: The importance of establishing trust. *Survey Practice*, 1(2), 1–6.
- Nelson, E. C., Eftimovska, E., Lind, C., Hager, A., Wasson, J., & Lindblad, S. (2015). Patient reported outcome measures in practice. *BMJ*, 350, g7818.
- Norman, G. R., Sloan, J. A., & Wywich, K. W. (2003). Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Medical Care*, 41, 582–592.
- Norton, M. C., Breitner, J. C., Welsh, K. A., & Wyse, B. W. (1994). Characteristics of nonresponders in a community survey of the elderly. *Journal of the American Geriatrics Society*, 42, 1252–1256.
- Overbeek, C. L., Nota, S. P., Jayakumar, P., Hageman, M. G., & Ring, D. (2015). The PROMIS physical function correlates with the QuickDASH in patients with upper extremity illness. *Clinical Orthopaedics and Related Research*, 473, 311–317.
- Papuga, M. O., Beck, C. A., Kates, S. L., Schwarz, E. M., & Maloney, M. D. (2014). Validation of GAITRite and PROMIS as high-throughput physical function outcome measures following ACL reconstruction. *Journal of Orthopaedic Research*, 32, 793–801.
- Reeve, B. B., Burke, L. B., Chiang, Y. P., Clauser, S. B., Colpe, L. J., Elias, J. W., ... Moy, C. S. (2007). Enhancing measurement in health outcomes research supported by Agencies within the US Department of Health and Human Services. *Quality of Life Research*, 16(1), 175–186.
- Richard Chin, B. Y. L. (2008). *Introduction to Clinical Trial Statistics*. Burlington, MA: Elsevier.
- Sand-Holzer, M., Deutsch, T., Frese, T., & Winter, A. (2015). Predictors of students' self-reported adoption of a smartphone application for medical education in general practice. *BMC Medical Education*, 15, 91.



- Schick-Makaroff, K., & Molzahn, A. (2015). Strategies to use tablet computers for collection of electronic patient-reported outcomes. *Health Qual Life Outcomes*, 13, 2.
- Segal, C., Holve E., & Sabharwal, R. (2013). *Collecting and Using Patient- Reported Outcomes (PRO) for Comparative Effectiveness Research (CER) and Patient-Centered Outcomes Research (PCOR): Challenges and Opportunities. Issue Briefs and Reports.*
- Selby, J. V., Beal, A. C., & Frank, L. (2012). The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda. *JAMA*, 307, 1583–1584.
- Sloan, J. A., Symond, T., Vargas-Chanes, D., & Friendly, B. (2003). Practical guidelines for assessing the clinical significance of health-related quality of life changes within clinical trials. *Therapeutic Innovation and Regulatory Science*, 37, 23–31.
- Speerin, R., Slater, H., Li, L., Moore, K., Chan, M., Dreinhöfer, K., ... Briggs, A. M. (2014). Moving from evidence to practice: models of care for the prevention and management of musculoskeletal conditions. *Best Practice & Research Clinical Rheumatology*, 28, 479–515.
- Touvier, M., Mejean, C., Kesse-Guyot, E., et al. (2010). Comparison between web- based and paper versions of a self-administered anthropometric questionnaire. *European Journal of Epidemiology*, 25, 287–296.
- Tyser, A. R., Beckmann, J., Franklin, J. D., et al. (2014). Evaluation of the PROMIS physical function computer adaptive test in the upper extremity. *The Journal of Hand Surgery*, 39, 2047–2051.
- Uhlig, C. E., Seitz, B., Eter, N., Promesberger, J., & Busse, H. (2014). Efficiencies of Internet-based digital and paper-based scientific surveys and the estimated costs and time for different-sized cohorts. *PLoS ONE*, 9, e108441.
- Velikova, G., Wright, E. P., Smith, A. B., Cull, A., Gould, A., ... Forman, D. (1999). Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. *Journal of Clinical Oncology*, 17, 998–1007.
- Wagner, L. I., Schink, J., Bass, M., et al. (2015). Bringing PROMIS to practice: brief and precise symptom screening in ambulatory cancer care. *Cancer*, 121, 927–934.
- Wei, D. H., Hawker, G. A., Jevsevar, D. S., & Bozic, K. J. (2015). Improving value in musculoskeletal care delivery: AOA critical issues. *Journal of Bone & Joint Surgery*, 97, 769–774.
- Wu, A. W., & Snyder, C. (2011). Getting ready for patient-reported outcomes measures (PROMs) in clinical practice. *Healthcare Papers*, 11, 48–53.
- Yoon, S., Wilcox, A. B., & Bakken, S. (2013). Comparisons among health behavior surveys: implications for the design of informatics infrastructures that support comparative effectiveness research. *EGEMS*, 1, 1021.
- Zbrozek, A., Hebert, J., Gogates, G., Thorell, R., Dell, C., Molsen, E., ... Hines, S. (2013). Validation of electronic systems to collect patient-reported outcome (PRO) data-recommendations for clinical trial teams: Report of the ISPOR ePRO systems validation good research practices task force. *Value Health*, 16, 480–489.

## Appendix 1: Patient-reported outcomes measurement information system

### Frequently asked questions

#### PROMIS

Q. What is the purpose of the PROMIS health assessment?

A. The health assessment gives your provider a clear understanding/snapshot of how you are feeling today and helps them work with you in developing a treatment plan.

Q. Who is requesting this information and why?

A. Your physician is requesting this information and it is Standard of Care (like taking blood pressure).

Q. Where do my answers go?

A. Your answers are electronically uploaded into your medical record and are ready for today's visit.

Q. Can I see my answers in "My Chart"?

A. No, it is not linked to My Chart; however, your provider can show you YOUR information. Only your provider and other medical professionals with access to your medical record can view your answers.

Q. Am I answering this about my injury/surgery or my health in general?

A. Your general health, which includes what brought you here today. Your answers should be related to how you feel today, a "snapshot" in time.

Q. Do I have to do this every time I come? Why?

Yes. The Health Assessment is Standard of Care and updates your provider on how you are doing today.

Q. Are the questions the same every time?

A. The questions are generated based on the answers that you give. If your prior answer is the same, the next question will be the same. If you answer differently, it triggers a new question.

Q. How long does the assessment take?

A. It only takes a few minutes to complete the assessment.

Q. Why do I have to answer questions that don't seem to apply to me?

A. It is important for your provider to know if you are able to do the daily activities you normally do or enjoy doing.

Q. Can I decline?

A. Yes, you can opt-out if you wish; however, this is your opportunity to update your provider on how you are feeling today.

This information can be used as part of your discussions with your doctor to develop your treatment plan and to monitor your progress.