


Check-in Phone Calls Six Weeks and Six Months After Musculoskeletal Specialty Care

Journal of Patient Experience
Volume 8: 1-6
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23743735211038777
journals.sagepub.com/home/jpx


Amelia E Mercado¹ , Amanda Gonzalez, MD¹,
David Ring, MD, PhD¹ , Karl Koenig, MD¹, and Tom J Crijns, MD¹

Abstract

Health organizations want to measure patient-reported outcome measures (PROMs) longitudinally. Two medical assistants called people selectively at 6 weeks and routinely at 6 months to obtain PROMs as part of routine musculoskeletal specialty care. Only 47% of patients (123 of 263) were reached by phone at 6 weeks and 41% (246 of 594) at 6 months. Spanish speakers were more likely to answer. Eighty-nine percent of people reached at 6 weeks and 76% reached at 6 months were willing to complete PROMs. Speaking Spanish, older age, and fewer symptoms of depression were associated with completing PROMs when reached by phone at 6 weeks. Women and Spanish-speaking patients were more likely to participate at 6 months. The observation that most people do not participate in longitudinal PROM measurement in a system that devotes notable resources in the setting of musculoskeletal specialty care—with some variation by psychosocial factors—suggests that longitudinal PROM measurement may be infeasible. Future work can explore alternative measures of the value of care that account for the large percentage of people that deprioritize care after a single specialty visit.

Keywords

patient-reported outcomes, attrition, loss to follow-up, depression, anxiety

Introduction

Patient-reported outcome measures (PROMs) are questionnaires that quantify symptom intensity and activity intolerance. Quantification of the subjective aspects of illness using PROMs can be used in research to monitor treatment effectiveness and help determine the value of health resources. They can also be used with patients as part of shared-decision making (1–3). PROMs are increasingly used in musculoskeletal specialty care, perhaps in part because a large proportion of orthopedic surgery is discretionary and addresses quality of life (4). Health systems want to track PROMs to demonstrate the value of their work (5). Alternative payment models attempt to measure the quality or value of care, which is supported by longitudinal completion of PROMs (2).

Prior research suggests that the majority of patients see a musculoskeletal specialist a single time and that few are still in care 6 months after their initial visit (6,7). This poses a challenge for clinicians and health systems that want to track PROMs over some minimum time after a specialty care episode to help establish the value of care. It is well

known that research questionnaires sent via mail, email, or text have poor response rates (8). To inform attempts to improve participation we sought patient characteristics associated with completing PROMs remote from initial care.

Using a practice-wide database of upper-extremity patients, in a setting where an alternative payment model included longitudinal measurement of PROMs and two full-time bilingual medical assistants were hired for this purpose, we asked: (1) among patients who are scheduled for a 6-week or 6-month follow-up phone call, what factors are associated with the ability to successfully contact patients? (2) For patients who are successfully reached, what factors are associated with willingness to complete repeat PROMs? (3) What

¹Department of Surgery and Perioperative Care, Dell Medical School, The University of Texas at Austin, Austin, TX, USA

Corresponding Author:

David Ring, Department of Surgery and Perioperative Care, Dell Medical School, The University of Texas at Austin, Health Discovery Building, MC Z0800, Austin, TX, USA.
Email: david.ring@austin.utexas.edu



factors are associated with completing repeat PROMs for all patients?

Methods

Study Design and Setting

In this institutional review board-approved study, we performed a retrospective database study of all adult (≥ 18 years) patients with upper-extremity conditions who were scheduled for a 6-week and/or 6-month follow-up phone call after their initial office visit between October 17, 2017, and January 29, 2019. A notable percentage of the patients in this unit are part of an alternative payment model with a reporting metric based on longitudinal measurement of PROMs designed to help measure the value and quality of care. During the time of this study, the organization hired two full-time, bilingual medical assistants to contact all patients routinely 6 months after an initial specialty care visit, with selective use of a 6-week call as an alternative to a return office visit or for patients the team wants to monitor more closely. Patients who were not called by the medical assistants at any time point and patients who spoke languages other than English or Spanish were excluded from the study.

Two hundred sixty-three patients were called 6 weeks after their initial visit; the mean age was 52 ± 14 years, 63% were women, and 65% were English speaking and 35% Spanish speaking (Table 1). At 6 weeks, 47% of patients (123 of 263) were successfully reached by phone. A total of 595 patients had a new patient visit during the study period and we attempted to reach all of them by phone 6 months after the visit to complete PROMs. The mean age of this cohort was 51 ± 14 years, 62% were women and 69% were English speaking and 31% were Spanish speaking. At 6 months, 41% of patients (246 of 594) were successfully reached by phone.

Measurements

Patients were asked to complete surveys at baseline, at 6 weeks, and 6 months after their initial visit. Arm-specific physical limitations were measured with the Quick Disabilities of the Arm, Shoulder, and Hand (qDASH), an abbreviated version of the DASH consisting of 11 items, with scores ranging from 0 (no disability) to 100 (most severe disability) (9). Coping strategies in response to nociception were measured using the two question version of the Pain Self-Efficacy Questionnaire (PSEQ) (10). Higher scores indicate better accommodation of pain. We measured symptoms of depression using the Patient Health Questionnaire-2 (PHQ-2), which inquires about the frequency of depressed mood and anhedonia in the past 2 weeks. Each item is scored in a 4-point Likert-scale, ranging from 0 (not at all) to 3 (nearly every day), with higher scores indicating more symptoms of depression (11). Patients with scores >3 on the PHQ-2 were asked to

Table 1. Patient Demographics.

Variables	Phone call	
	6-week	6-month
N	263	594
Age	52 ± 14	51 ± 14
Gender		
Women	167 (63%)	371 (62%)
Men	96 (37%)	223 (38%)
Language		
English	170 (65%)	407 (69%)
Spanish	93 (35%)	187 (31%)
City		
Austin	215 (81%)	464 (78%)
Other	48 (19%)	130 (22%)
GAD-7 (baseline)	0 (0–2)	0 (0–2)
PHQ-2/9 (baseline)	0 (0–2)	0 (0–2)
PSEQ-2 (baseline)	8 (5–10)	7 (5–10)
qDASH		
Baseline	50 (32–68)	55 (34–70)
6 weeks	38 (15–55)	43 (23–61)
6 months	23 (9–45)	27 (9–46)
Reached		
No	140 (53%)	348 (59%)
Yes	123 (47%)	246 (41%)
Completed survey		
Yes	88 (72%)	130 (53%)
No	35 (28%)	116 (47%)

Continuous variables as mean \pm standard deviation or median (interquartile range); discrete variables as number (percentage).

Abbreviations: GAD-7, Generalized Anxiety Disorder-7; PHQ-2/9, Patient Health Questionnaire-2/9; PSEQ-2, Pain Self-Efficacy Questionnaire-2; qDASH, quick disabilities of the arm, shoulder, and hand.

complete the 9-item version. The Generalized Anxiety Disorder-7 (GAD-7) was used to measure symptoms of anxiety in the last 2 weeks on a continuum, with a higher score indicating greater symptoms of anxiety (12). Additionally, patients completed a questionnaire of basic demographics (e.g., age, gender, and ZIP code). A research assistant marked whether they were able to contact the patient by phone, and, if reached, whether patients agreed to complete the survey (qDASH).

Statistical Analysis

Continuous variables were presented as mean \pm standard deviation or median (interquartile range). Discrete variables were presented as frequencies with percentages. To detect variables associated with reaching patients at each time point and their willingness to complete the survey, we conducted bivariate analyses, accounting for age, gender, language (English compared to Spanish), city (Austin compared to other), GAD-7 at baseline, PHQ-2/9 at baseline, PSEQ at baseline, and qDASH at 6 weeks or 6 months. We used the Student *t*-test or Mann-Whitney *U* test for continuous variables, where appropriate. For categorical variables, Fisher's exact were used. City had to be dichotomized into

“Austin” and “other” due to low numbers. All variables with $P < .10$ were moved to multivariable logistic regression. We reported Odds ratios, standard errors, and P values. All P values $< .05$ were considered statistically significant.

Results

In bivariate analysis, greater odds of reaching patients at 6 weeks was associated with women and the Spanish language (Appendix in the Supplemental Material). In multivariable logistic regression, greater odds of contacting patients at 6 weeks after an in-person visit was independently associated

with the Spanish language alone (Table 2). Greater odds of reaching patients at 6 months was associated with Spanish language and living in Austin in bivariate and multivariable analysis (Table 3).

When patients were successfully reached, greater odds of completing the survey at 6 weeks was associated with older age, women, speaking Spanish, fewer symptoms of anxiety, and fewer symptoms of depression in bivariate analysis. In multivariable analysis, greater odds of completing PROMs at 6 weeks were independently associated with older age, Spanish language, and lower symptoms of depression at baseline (Table 4). Greater odds of patients

Table 2. Multivariable Logistic Regression Analysis of Factors Associated with Completing the Survey at 6 Weeks.

Variables	Odds ratio (95% confidence interval)	Standard error	P-value	Pseudo R ²
Age	1.0 (0.99 to 1.0)	0.011	.051	
Gender				
Women	Reference value			
Men	0.61 (0.32 to 1.2)	0.20	.141	
Language				
Spanish	Reference value			
English	0.24 (0.13 to 0.44)	0.075	<.001	
City				
Other	Reference value			
Austin	1.5 (0.66 to 3.3)	0.60	.349	
Generalized Anxiety Disorder-7 (GAD-7; baseline) ^a	0.95 (0.88 to 1.0)	0.035	.177	

Bold values indicate statistical significance, $P < .05$. N = 229.

^aThirty-four baseline GAD-7 scores were missing.

Table 3. Multivariable Logistic Regression Analysis of Factors Associated with Completing the Survey at 6 Months.

Variables	Odds ratio (95% confidence interval)	Standard error	P-value	Pseudo R ²
				0.041
Gender				
Women	Reference value			
Men	0.63 (0.41 to 0.98)	0.14	.040	
Language				
Spanish	Reference value			
English	0.41 (0.27 to 0.61)	0.085	<.001	

Bold values indicate statistical significance, $P < .05$. N = 594.

Table 4. Multivariable Logistic Regression Analysis of Factors Associated with the Ability to Reach Patients by Phone at 6 Weeks.

Variables	Odds ratio (95% confidence interval)	Standard error	P-value	Pseudo R ²
				0.065
Gender				
Women	Reference value			
Men	0.63 (0.37 to 1.1)	0.17	.095	
Language				
Spanish	Reference value			
English	0.33 (0.19 to 0.56)	0.090	<.001	

Bold values indicate statistical significance, $P < .05$. N = 263.

completing the survey at 6 months were associated with women and speaking Spanish in both bivariate and multivariable analysis (Table 5).

Among all patients, greater odds of completing PROMs at 6 weeks was associated with older age, women, Spanish language, living in Austin, and fewer symptoms of anxiety. In multivariable analysis, greater odds of completing the survey were associated with speaking Spanish (Table 6). Greater odds of completing the survey at 6 months were associated with women and speaking Spanish in bivariate and multivariable analysis (Table 7).

Discussion

Health care organizations would like to measure PROMs longitudinally to be able to measure the value of their care, but it is not clear this is feasible. In research settings, there is some evidence that phone calls have higher response rates than other forms of communication (8), but in our experience, response rates are low no matter the method of communication. In the setting of an alternative payment model that includes longitudinal PROM measurement as a quality/value metric, substantial resources were devoted, and yet it was difficult to reach

Table 5. Multivariable Logistic Regression Analysis of Factors Associated with the Ability to Reach Patients by Phone at 6 Months.

Variables	Odds ratio (95% confidence interval)	Standard error	P-value	Pseudo R ²
				0.018
Language				
Spanish	<i>Reference value</i>			
English	0.59 (0.42 to 0.84)	0.11	.003	
City				
Other				
Austin	1.6 (1.0 to 2.4)	0.33	.035	

Bold values indicate statistical significance, $P < .05$. N = 594.

Table 6. Multivariable Logistic Regression Analysis of Factors Associated with Completing the Survey at 6 Weeks for Patients Who were Reached by Phone.

Variables	Odds ratio (95% confidence interval)	Standard error	P-value	Pseudo R ²
				0.17
Age	1.05 (1.0 to 1.1)	0.020	.012	
Gender				
Women	<i>Reference value</i>			
Men	0.49 (0.17 to 1.4)	0.26	.178	
Language				
Spanish	<i>Reference value</i>			
English	0.31 (0.11 to 0.85)	0.16	.023	
Patient Health Questionnaire-2/9 (PHQ-2/9; baseline) ^a	0.90 (0.82 to 0.98)	0.040	.015	

Bold values indicate statistical significance, $P < .05$.

^aThe Generalized Anxiety Disorder-7 was omitted from the model due to multicollinearity with the PHQ-2/9 ($\rho = 0.66$; $P < .001$). N = 107. Sixteen baseline PHQ-2/9 scores were missing.

Table 7. Multivariable Logistic Regression Analysis of Factors Associated with Completing the Survey at 6 Months for Patients Who were Reached by Phone.

Variables	Odds ratio (95% confidence interval)	Standard error	P-value	Pseudo R ²
				0.045
Gender				
Women	<i>Reference value</i>			
Men	0.57 (0.33 to 0.99)	0.16	.046	
Language				
Spanish	<i>Reference value</i>			
English	0.44 (0.26 to 0.76)	0.12	.003	

Bold values indicate statistical significance, $P < .05$. N = 246.

people, and participation was limited, with some variation by psychosocial factors.

This study has several limitations that should be kept in mind. First, readers should take care when interpreting the data regarding patients receiving 6-week phone calls (which was selective) and patients receiving 6-month phone calls (routine). That said, the similarities in the rate of response and factors associated with participation are notable. Second, the specific rates may apply best to our setting, although, based on parallels with research in several settings in the United States, they seem representative. And the statistical associations only depend on variation within the sample and are likely more easily reproducible. Third, there are socioeconomic factors that may influence whether patients can be reached by phone (e.g., not being able to pay the phone bill) and we could not account for these factors in our analysis. Finally, while our research assistants called patients into the evenings and occasionally on weekends, the fact that most of the calls were during business hours may bias response in favor of patients who are retired, who have flexible jobs, or who are unemployed.

The finding that Spanish speakers were more likely to answer a 6-week and 6-month phone call and complete PROMs, may indicate the influence of personal and cultural factors on patient interest in participating in research or differences in care utilization (13,14). Cultural differences between Spanish and English-speaking patients—such as greater deference or respect, greater appreciation for the care, or other factors—may account for the difference in response rate, but we did not study these factors. Within our practice, a noteworthy proportion of clinicians are proficient in Spanish, which may increase the response rates. There is evidence suggesting that Spanish speakers tend to be more satisfied by language-concordant relationships with their physician (15–17).

The observation among the subset of patients who were reached by phone that fewer symptoms of depression and speaking Spanish were associated with completing PROMs suggests that longitudinal collection of PROMs will be less successful among people experiencing mental health opportunities. It is known that greater symptoms of depression are associated with greater symptom intensity and magnitude of incapability (18–23), which may be related to willingness, or perceived capability, to participate. These findings are in line with the notion that there are important psychological differences between people who stay in care and those who do not. The overall low participation in the longitudinal collection of PROMs is consistent with the published experience of others. People may not participate unless they feel it benefits them directly or unless they feel gratitude for the benefit received in the form of a diagnostic or therapeutic intervention.

This study of patients with upper-extremity illness found that participation in longitudinal completion of PROMs is low even when patients receive a personal call in their preferred language. We also found that the most important factor influencing participation is psychosocial. We know from other studies that most patients see a specialist once

and then deprioritize care (24). The current study adds that people also tend not to answer a phone call from the office and that even if reached, may decline to repeat PROMs. It would be useful to know whether people that deprioritize care or participation in longitudinal PROM measurement had their needs met, felt the care was not useful and worthwhile to continue, or if the symptoms have resolved or been accommodated. Future research might also address strategies for improving participation in longitudinal completion of PROMs in musculoskeletal specialty care. And it is possible that longitudinal tracking of PROMs is not possible and other strategies such as text messages with a facile response addressing whether or not a person's needs have been met. These findings may also support efforts to transform the way we monitor care toward more objective and feasible measures of health (e.g., tracking of mobility and activity using smartphones).

Authors' Note

This study was approved by the institutional review board at the University of Texas at Austin (FWA #00002030). This article does not contain any studies with human or animal subjects. Informed consent for patient information to be published in this article was not obtained due to a waiver of informed consent because the data is anonymous.



Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: David Ring received royalties from Wright Medical (Memphis, TN, USA) for elbow plates of less than USD10,000 per year and Skeletal Dynamics for an internal joint stabilizer elbow between USD10,000 and 100,000 per year. David Ring is a Deputy Editor for *Hand and Wrist, Journal of Orthopaedic Trauma, and Clinical Orthopaedics and Related Research* and has received or may receive payments or benefits amounting to USD5000 per year. David Ring received honoraria from meetings of the AO North America (Wayne, PA, USA), AO International (Davos, Switzerland), and various hospitals and universities.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Amelia E Mercado  <https://orcid.org/0000-0002-5948-4068>
David Ring  <https://orcid.org/0000-0001-5947-5316>

Supplemental Material

Supplemental material for this article is available online.

References

1. Porter ME. What is value in health care? *N Engl J Med*. 2010;363(26):2477-81. doi:10.1056/NEJMp1011024
2. Squitieri L, Bozic KJ, Pusic AL. The role of patient-reported outcome measures in value-based payment reform. *Value Heal*. 2017;20(6):834-6. doi:10.1016/j.jval.2017.02.003

3. Basch E. Patient-Reported outcomes—harnessing Patients' voices to improve clinical care. *N Engl J Med* 2017;376. 2017;376(2):105-8. doi:10.1056/nejmp1611252
4. Horn ME, Reinke EK, Couce LJ, Reeve BB, Ledbetter L, George SZ. Reporting and utilization of patient-reported outcomes measurement information system® (PROMIS®) measures in orthopedic research and practice: a systematic review. *J Orthop Surg Res.* 2020;15(1):533. doi:10.1186/s13018-020-02068-9
5. Field J, Holmes MM, Newell D. PROMs data: can it be used to make decisions for individual patients? A narrative review. *Patient Relat Outcome Meas.* 2019;10:23-41. doi:10.2147/PROM.S.156291
6. Crijns TJ, Bernstein DN, Teunis T, Gonzalez RM, Wilbur D, Ring D. The association between symptoms of depression and office visits in patients with nontraumatic upper-extremity illness. *J Hand Surg Am.* 2020;45(2):159.d1-159.e8. doi:10.1016/j.jhsa.2019.03.019
7. Bernstein DN, Crijns TJ, Mahmood B, Ring D, Hammert WC, Mahmood B. Patient characteristics, treatment, and presenting PROMIS scores associated with number of office visits for traumatic hand and wrist conditions. *Clin Orthop Relat Res.* 2019;477(10):2345-55. doi:10.1097/CORR.0000000000000742
8. Nota SPFT, Strooker JA, Ring D. Differences in response rates between mail, e-mail, and telephone follow-up in hand surgery research. *Hand.* 2014;9(4):504-10. doi:10.1007/s11552-014-9618-x
9. Beaton DE, Wright JG, Katz JN, Upper Extremity Collaborative Group. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Jt Surg Ser A.* 2005;87(5):1038-46. doi:10.2106/JBJS.D.02060
10. Nicholas MK, McGuire BE, Asghari A. A 2-item short form of the pain self-efficacy questionnaire: development and psychometric evaluation of PSEQ-2. *J Pain.* 2015;16(2):153-63. doi:10.1016/j.jpain.2014.11.002
11. Kroenke K, Spitzer RL, Williams JBW. The patient health questionnaire-2: validity of a two-item depression screener. *Med Care.* 2010;363(26):2477-81. doi:10.1097/01.MLR.0000093487.78664.3C
12. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092-97. doi:10.1001/archinte.166.10.1092
13. Menendez ME, Ring D. Factors associated with non-attendance at a hand surgery appointment. *Hand.* 2015;10(2):221-6. doi:10.1007/s11552-014-9685-z
14. Berg PWLT, Ring D. Patients lost to follow-up after metacarpal fractures. *J Hand Surg Am.* 2012;37(1):42-6. doi:10.1016/j.jhsa.2011.08.003
15. Ferguson WJ, Candib LM. Culture, language, and the doctor-patient relationship. *Fam Med.* 2002;34(5):353-61.
16. Schouten BC, Meeuwesen L. Cultural differences in medical communication: a review of the literature. *Patient Educ Couns.* 2006;64(1-3):21-34. doi:10.1016/j.pec.2005.11.014
17. Seijo R, Gomez H, Freidenberg J. Language as a communication barrier in medical care for Hispanic patients. *Hisp J Behav Sci.* 1991;13(4):363-76. doi:10.1177/07399863910134001
18. Duivenvoorden T, Vissers MM, Verhaar JAN, Busschbach JJV, Gosens T, Bloem RM. Anxiety and depressive symptoms before and after total hip and knee arthroplasty: a prospective multicentre study. *Osteoarthritis Cartil.* 2013; 21(12):1834-40. doi:10.1016/j.joca.2013.08.022
19. La A, Nadarajah V, Jauregui JJ, Shield WP, Medina SH, Dubina AG. Clinical characteristics associated with depression or anxiety among patients presenting for knee surgery. *J Clin Orthop Trauma.* 2020;11(Suppl 1):S164-70. doi:10.1016/j.jcot.2019.08.009
20. Zautra AJ, Smith BW. Depression and reactivity to stress in older women with rheumatoid arthritis and osteoarthritis. *Psychosom Med.* 2001;63(4):687-96. doi:10.1097/00006842-200107000-00022
21. Vissers MM, Bussmann JB, Verhaar JAN, Busschbach JJV, Bierma-Zeinstra SMA, Rejiman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. *Semin Arthritis Rheum.* 2012;41(4):576-88. doi:10.1016/j.semarthrit.2011.07.003
22. Vincent HK, Hagen JE, Zdziarski-Horodyski LA, Patrick M, Sadasivan KK, Guenther R. Patient-reported outcomes measurement information system outcome measures and mental health in orthopaedic trauma patients during early recovery. *J Orthop Trauma.* 2018;32(9):467-73. doi:10.1097/BOT.0000000000001245
23. Werner BC, Wong AC, Chang B, Craig EV, Dines DM, Warren RF. Depression and patient-reported outcomes following total shoulder arthroplasty. *J Bone Joint Surg Am.* 2017;99(8):688-95. doi:10.2106/JBJS.16.00541
24. Crijns TJ, Ring D, Valencia V. Factors associated with the cost of care for the most common atraumatic painful upper extremity conditions. *J Hand Surg Am.* 2019;44(11):989.e1-e18. doi:10.1016/j.jhsa.2019.01.001