7-15-2016

Level of Patient-Physician Agreement in Assessment of Change Following Conservative Rehabilitation for Shoulder Pain

Stephanie D. Moore-Reed  
*California State University, Fresno*

W. Ben Kibler  
*Shoulder Center of Kentucky*

Heather M. Bush  
*University of Kentucky, heather.bush@uky.edu*

Timothy L. Uhl  
*University of Kentucky, tluhl2@uky.edu*

[Click here to let us know how access to this document benefits you.](https://uknowledge.uky.edu/biostatistics_facpub)

Follow this and additional works at: [https://uknowledge.uky.edu/biostatistics_facpub](https://uknowledge.uky.edu/biostatistics_facpub)

Part of the [Biostatistics Commons](https://uknowledge.uky.edu/biostatistics_facpub), and the [Rehabilitation and Therapy Commons](https://uknowledge.uky.edu/biostatistics_facpub)

Repository Citation

[https://uknowledge.uky.edu/biostatistics_facpub/22](https://uknowledge.uky.edu/biostatistics_facpub/22)

This Article is brought to you for free and open access by the Biostatistics at UKnowledge. It has been accepted for inclusion in Biostatistics Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
Patient-Physician agreement

1 Level of patient-physician agreement in assessment of change following conservative rehabilitation for shoulder pain

3 Keywords: shoulder pain; orthopaedics; self report questionnaires; conservative treatment; rehabilitation
Abstract

Background: Assessment of health-related status has been shown to vary between patients and physicians, but the degree of patient-physician discordance in assessment of change in status is unknown.

Methods: Ninety-nine patients with shoulder dysfunction underwent a standardized physician examination and completed several self-reported questionnaires. All patients were prescribed the same physical therapy intervention. Six weeks later the patients returned to the physician, when self-report questionnaires were re-assessed and the Global Rating of Change (GROC) was completed by the patient. The physician completed the GROC retrospectively. To determine agreement between patient and physician, Intra-Class Correlation Coefficient (ICC) and Pearson’s r using the 15-point GROC and weighted kappa using a consolidated 3-point GROC were calculated.

Results: Utilizing the 15-point GROC, complete agreement was observed in 37/99 patients (37%). ICC and Pearson’s r between patient and physician were 0.62 and 0.63 respectively. Utilizing a consolidated 3-point GROC, complete agreement was observed in 76/99 patients (77%). Weighted kappa was 0.62.

Discussion: Assessment of change reported by the patient demonstrates moderate to good agreement with physician assessment. These findings indicate that the GROC does reflect and represent similar assessment of change in health status by patients and physicians. This can aid discussion of both past treatment results and future treatment plans.
Patient-Physician agreement

Introduction

Health-related assessment ratings have been shown to vary between patients and clinicians, resulting in patient-clinician discordance. This discordance has been reported in assessments of disease severity\textsuperscript{1-6}, physical functioning\textsuperscript{5,7-12}, pain\textsuperscript{8,11,13} and quality of life\textsuperscript{13} in a variety of acute and chronic pathologies and select musculoskeletal disorders. In general, patients tend to rate themselves as being more severely impacted compared to physician ratings.\textsuperscript{1,2,6,8} However, there is some evidence that this may vary depending upon the pathology being examined.\textsuperscript{5,9,13} The magnitude of disagreement and whether clinicians overestimate or underestimate impairments and disease severity appears to vary based on the disease.\textsuperscript{5,9,13} This may reflect that clinicians tend to predetermine the effects a health condition will have on a patient based on the perceived generalized severity of the condition, rather than the individual patient’s characteristics.

Determining the most “true” assessment of a patient’s health or healing status can be challenging because physicians and patients are likely to factor different information into their judgment. Physicians are often thought to consider pain as a secondary result of a pathology or anatomic abnormality.\textsuperscript{14} Evidence suggests physicians use their clinical experience,\textsuperscript{1,2} the patient’s disease duration\textsuperscript{15}, and objective findings (e.g. clinical signs and symptoms and laboratory tests)\textsuperscript{1,3,4,14-16} to determine their assessment. Patients, on the other hand, may not understand abnormalities explained by laboratory tests or diagnostic imaging\textsuperscript{1}, and judge severity of their injuries on their individual experience\textsuperscript{17}. Patients also sense pain in a multifactorial manner that may be experienced even in the absence of pathology\textsuperscript{14} and factor pain into their assessment\textsuperscript{4,6}. A study of patients with lupus identified that patient-reported pain accounted for 20% of the variance in patient-reported disease activity, but was that it was not a significant
patient-physician agreement is a predictor of physician reported disease activity. Additionally, patients and physicians may have different expectations with regard to the progression or outcome of the intervention, or what constitutes a satisfactory progression in treatment or a good outcome.

The impact of patient-physician discordance in orthopaedics and shoulder surgery is relatively unknown. While discordance has been examined in one-time assessments of disease state or impairment, few studies have examined the effect these differences in perception of the results of treatment may have on the assessment of change over time or outcome following an intervention. Two studies have reported fair to good patient-clinician agreement in patients with low back pain and disorders of the neck-shoulder region. Agreement on assessment of outcome has been examined relative to pain and overall satisfaction, but only in a cohort of post-operative patients following total hip arthroplasty. The need to examine the agreement or discordance is important, since this is not an issue of whose assessment is right or wrong; each perspective is equally valid. The patient’s perspective should always be considered by the health care provider as the patient is actually experiencing the treatment and is affected by the results. However, clinicians are responsible for the content, timing, and direction of treatment and are therefore most influential in guiding the patients’ course of medical treatment. Determining whether a patient has improved is an important factor in making treatment decisions for both the physician and patient. If there is discordance in patient-reported and physician-reported assessment of change among patients seeking medical care for shoulder pain, strategies should be developed to improve agreement and communication to manage the discordance and perhaps develop other tools to assess change that will minimize discordance. Therefore, the purpose of this study was to examine the level of agreement between patient and physician assessment of change, using a Global Rating of Change (GROC) scale in a cohort of patients being treated for
We hypothesized that there would be moderate (66 to 75%) agreement between patient-reported and physician-reported assessment of change.

**Materials and Methods**

**Subjects**

Data from 99 subjects were used in this analysis (age = 41±12 years, height = 175±10 cm, weight = 84±19 kg, 66 males). These data come from a larger study in which patients were enrolled prospectively over two years. Of the 191 eligible subjects (220 patients approached, 211 enrolled, 20 withdrew), 99 had all data required for the present analysis. Patients reporting to the Lexington Clinic Orthopaedics and Sports Medicine Center with shoulder pain were identified as potential subjects. Patients were eligible for enrollment if they presented with clinical history consistent with dysfunction due to musculoskeletal shoulder injury, reported pain with overhead activity and were between 15 and 60 years of age. Patients were excluded if they demonstrated signs and symptoms consistent with cervical radiculopathy, adhesive capsulitis, glenohumeral arthritis or reported tingling/numbness in the upper extremity, surgery on the involved shoulder within the past year, or steroid injection within the last month.

Patients who met the criteria and consented to participate underwent a full standardized examination by the physician and completed a battery of self-reported questionnaires including a numeric pain rating scale (NPRS; 0=no pain, 10=worst pain) and the Quick Disabilities of the Arm, Shoulder and Hand (QuickDASH; 0 = no disability, 100 = severe disability). All patients read and signed an informed consent form prior to enrollment in the study that was approved by the institutional review boards of the University of Kentucky and Lexington Clinic. Subjects were prescribed physical therapy and provided with a standardized rehabilitation protocol to take to the therapist of their choosing.
Patient-Physician agreement

Procedures

Patient-oriented assessment of change was collected prospectively at the time of physician follow up (6±1 weeks). Subjects completed a global rating of change (GROC) to assess perceived improvement. The GROC is a 15 item scale ranging from “a very great deal worse” to “a very great deal better” (Figure 1). Subjects were instructed to select the statement that best represented their perceived change in functional status since the initial evaluation. Physician assessment of change was done retrospectively at the end of the enrollment period. The treating physician, an orthopaedic surgeon, was provided with clinical notes from both the initial evaluation and follow up visit for each patient and completed the same 15-point GROC scale. Intra-rater reliability was excellent (ICC = 0.929) and was established by having the physician rate the same 10 subjects at two separate times, with a minimum of one week between ratings.

Data Reduction

The 15-point GROC was further consolidated into a 3-point scale by collapsing response options into “better” (GROC score ≥ +3), no change (-2 to +2), and “worse” (GROC score ≤ -3) based on previously reported cutoffs used to identify clinically meaningful improvement. Providing patients (or clinicians) with too many options may be of concern as the individual may have difficulty attaching meaning to each separate response choice. By treating the 15-point scale as continuous, Intra-Class Correlation Coefficient (ICC) and Pearson’s r could be conducted, while the consolidated 3-point scale allowed for confirmation of the findings with weighted kappa using a more simplified scale of better/no change/worse.

Statistical Analysis
To assess patient-physician agreement, ICC, Pearson’s r correlation coefficient, and linear weighted kappa were calculated. ICC and Pearson’s r were calculated using the responses on the 15-point GROC. ICCs were interpreted according to the following: <0.40 Poor, 0.04-0.75 Fair to Good, >0.75 Excellent. Linear weighted kappa was calculated using the consolidated 3-point scale (better, no change, worse). The strength of agreement for kappa was interpreted according to the following: <0.00 Poor, 0.00-0.20 Slight, 0.21-0.40 Fair, 0.41-0.60 Moderate, 0.61-0.80 Substantial, 0.81-0.99 Almost Perfect. Maximum kappa was calculated according to Sim and Wright. The maximum kappa value provides a more meaningful reference value for interpretation because inadequate variation in the data can result in artificially low kappa values.

**Results**

Utilizing the 15-point GROC scale, complete agreement between patient-reported and physician-reported GROC score was observed in 37/99 patients (37%). ICC and Pearson’s r were 0.62 and 0.63 respectively. Utilizing the consolidated 3-point scale (better, no change, worse), complete agreement was observed in 76/99 patients (77%). Weighted kappa was 0.62 with a maximum weighted kappa was determined to be 0.95. Bivariate relationship between patient-reported and physician-reported GROC scores is depicted in a scatterplot (Figure 2).

**Discussion**

One of the keystones of the doctor-patient relationship is that they are both in agreement regarding the results of treatment and the direction of future care. This requires agreement on the status of these treatment efforts. One key element would be the change in functional status as a result of treatment. This study examined the patient-physician agreement or discordance related to assessment of change following rehabilitation in patients with shoulder pain. Our hypothesis of moderate agreement was supported, indicating that the Global Rating of Change scale appears...
to reflect and represent the same degree of change perceived by each group. Overall, we observed moderate to good agreement. Our findings indicate similar patient-physician agreement compared to previous research. Patient-physician agreement reported in the literature ranges from 58 to 77%.1, 4, 6, 7, 10, 16, 28 Our finding of 37% complete patient-physician agreement on the 15-point GROC scale was expected to be lower because complete agreement was necessary. Using the 3-point scale we were able to examine more global agreement, i.e. did the patient and physician agree that the patient was better, the same or worse, rather than matching exactly to a particular point on a 15-point scale. Complete patient-physician agreement using this 3-point scale was 77%, which is at the high end of, though consistent with, previous reports. Our assessment of weighted kappa (0.62) is also higher than previous reports to assess patient-clinician agreement (range 0.09 to 0.39).3, 10, 12

Our findings of higher agreement than previous literature may be because our patients did not report high pain severity or disability. Discordance between patients and physicians is known to be greater and more common in patients with more severe ratings of disease activity, impairment or pain.2, 4, 10, 18 The mean rating of current pain on the NPRS was 4±2 at initial evaluation and 3±2 at follow up. The mean QuickDASH at initial exam was 38±18 indicating our patients were approximately 40% disabled at initial evaluation. At follow up, patients improved by an average of 8±15 points on the QuickDASH. Our sample seems to represent the typical population of shoulder pain patients as our level of pain and disability are consistent with patients with shoulder pain seeking care from an orthopaedic surgeon.29-31

Limited research explores agreement or discordance in ratings of change over time in functional health status. Patients with rheumatoid arthritis rated their change in global function over 3 months, as did their treating physician.19 The authors observed a patient-physician
relationship (ICC = 0.64, r = 0.63) very similar to the current study’s findings. Patients with heart disease were asked to use a 7-point “transition index scale” that appears quite similar to the GROC to assess change in health-related quality of life in patients with heart disease. These authors identified poor agreement (k=0.09 to 0.23) between patients and physicians. The low agreement may be due to the type of data collected and compared. A single global assessment made by the physician was compared to multiple domains assessed by the patients. In the current study the same global assessment was performed by both the patient and physician, which seems to result in higher agreement in the present study and in previous research.

In the only previous study to assess patient-physician agreement in change over time involving an orthopaedic population, patient-reported assessment of pain and overall satisfaction following total hip arthroplasty was compared to physician assessment using a visual analog scale (VAS). Differences in patient and physician ratings of pain were statistically significantly different (1.7±2.6cm and 1.1±1.8cm respectively), though the difference was only 0.6cm on the VAS. Reports of overall satisfaction between patient and physician were not significantly different (8.6±2.1 and 8.8±1.7cm respectively). The authors did note that patient-physician agreement was notably worse among the patients with high pain or low satisfaction. While this was an orthopaedic population, the cohort was post-surgical and the authors did not provide an assessment of agreement (e.g., kappa, ICC), making it difficult to draw direct comparisons to the current study. Our data provide the first examination of assessment of change following conservative rehabilitation in an orthopaedic population.

We used a 15-point GROC to assess perceived change. The “global”, less specific nature of the GROC allows the patient to base their response on what is most important to them. This was ideal for addressing the purpose of the present study in that we wanted to identify if...
differences existed between perceptions of patients and clinicians. Test-retest reliability of the GROC within 24 hours was excellent in patients with musculoskeletal disorders (ICC range 0.90 to 0.99).\textsuperscript{32} One limitation of a global rating of change assessment is that it requires the patient to recall their previous condition with respect to their current status.\textsuperscript{23} It has been suggested that GROC scores may be influenced by current status as follow up time increases.\textsuperscript{32} The 3-point GROC showed a much higher percentage of complete agreement between patient and physician evaluations, probably due to limiting the available options. It may serve as a better basis for discussion between the patient and physician regarding the results of treatment, and therefore help to guide the discussion about future treatment plans.

Limitations
A few limitations of this study should be noted in order to interpret these results accurately. First, patients completed the GROC at the time of their visit, while the physician completed the GROC retrospectively at the end of the enrollment period. The physician had his own notes to refer to when completing the GROC but it may have been more timely to have the physician rate the patient using the GROC scale immediately following the visit. However, it was felt that a longer time interval could provide a more objective analysis of the amount of change, and doing the evaluations at one time would improve the consistency of the ratings. Additionally, inclusion of a single physician may limit the generalizability of the results and validation of the findings with additional physicians should be performed. Our assessments examined change over time from baseline to follow up. While all patients were prescribed a standardized physical therapy intervention, several variables could have factored into the results including expectation of treatment success, patient satisfaction with outcome or physician services\textsuperscript{18} and adherence to therapy. Future studies should account for
those variables to further explain the patient-clinician relationship with regard to agreement on health-related assessment.

Conclusion

Our results indicate that physician-reported assessment of change demonstrated moderate to good agreement with patient-reported assessment of change in patients with orthopaedic shoulder pain, which supported our hypothesis of moderate agreement between the two parties. This indicates that patient and physician are for the most part on the same page in how the patient is responding to a non-operative intervention which supports there is limited discordance in treatment of orthopaedic conditions. The results suggest the Global Rating of Change can be used to represent both the patient and physician assessment of the results of treatment. It can serve as an effective means to facilitate the patient-physician dialogue, linking both stakeholders’ perceptions of the treatment so that both can understand the perceptions of the treatment, the results of the treatments, and the need for and direction of future treatments.
REFERENCES


Patient-Physician agreement

Figure legends

Figure 1: Global Rating of Change Scale

Figure 2: Patient-Physician Agreement Plot

Points that fall within the green (“better”, n=37), yellow (“no change”, n=35) and red (“worse”, n=4) boxes represent that the patient and physician both rated the patient in the same category. Points that fall outside of the boxes represent disagreement between the patient and physician (n=23). The values represent the number of patients represented by that data point.