



4-1-2017

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

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Follow this and additional works at: https://uknowledge.uky.edu/biostatistics_facpub

 Part of the [Biostatistics Commons](#), and the [Rehabilitation and Therapy Commons](#)

Repository Citation

Moore-Reed, Stephanie D.; Kibler, W. Ben; Bush, Heather M.; and Uhl, Timothy L., "Level of Patient-Physician Agreement in Assessment of Change Following Conservative Rehabilitation for Shoulder Pain" (2017). *Biostatistics Faculty Publications*. 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.

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

Notes/Citation Information

Published in *Shoulder & Elbow*, v. 9, issue 2, p. 127-132.

Copyright © 2016 by The British Elbow & Shoulder Society

The copyright holder has granted the permission for posting the article here.

The document available for download is the authors' post-peer-review final draft of the article.

Digital Object Identifier (DOI)

<https://doi.org/10.1177/1758573216658799>

Patient-Physician agreement

- 1 **Level of patient-physician agreement in assessment of change following conservative**
- 2 **rehabilitation for shoulder pain**
- 3 **Keywords:** *shoulder pain; orthopaedics; self report questionnaires; conservative treatment;*
- 4 *rehabilitation*

5 **Abstract**

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

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

17 **Results:** Utilizing the 15-point GROC, complete agreement was observed in 37/99 patients
18 (37%). ICC and Pearson's r between patient and physician were 0.62 and 0.63 respectively.
19 Utilizing a consolidated 3-point GROC, complete agreement was observed in 76/99 patients
20 (77%). Weighted kappa was 0.62.

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

25 **Introduction**

26
27 Health-related assessment ratings have been shown to vary between patients and
28 clinicians, resulting in patient-clinician discordance. This discordance has been reported in
29 assessments of disease severity¹⁻⁶, physical functioning,^{5, 7-12} pain^{8, 11, 13} and quality of life¹³ in a
30 variety of acute and chronic pathologies and select musculoskeletal disorders. In general,
31 patients tend to rate themselves as being more severely impacted compared to physician ratings.^{1,}
32 ^{2, 6, 8} However, there is some evidence that this may vary depending upon the pathology being
33 examined.^{5, 9, 13} The magnitude of disagreement and whether clinicians overestimate or
34 underestimate impairments and disease severity appears to vary based on the disease.^{5, 9, 13} This
35 may reflect that clinicians tend to predetermine the effects a health condition will have on a
36 patient based on the perceived generalized severity of the condition, rather than the individual
37 patient's characteristics.

38 Determining the most "true" assessment of a patient's health or healing status can be
39 challenging because physicians and patients are likely to factor different information into their
40 judgment. Physicians are often thought to consider pain as a secondary result of a pathology or
41 anatomic abnormality.¹⁴ Evidence suggests physicians use their clinical experience,^{1, 2} the
42 patient's disease duration¹⁵, and objective findings (e.g. clinical signs and symptoms and
43 laboratory tests)^{1, 3, 4, 14-16} to determine their assessment. Patients, on the other hand, may not
44 understand abnormalities explained by laboratory tests or diagnostic imaging¹, and judge severity
45 of their injuries on their individual experience¹⁷. Patients also sense pain in a multifactorial
46 manner that may be experienced even in the absence of pathology¹⁴ and factor pain into their
47 assessment^{4, 6}. A study of patients with lupus identified that patient-reported pain accounted for
48 20% of the variance in patient-reported disease activity, but was that it was not a significant

Patient-Physician agreement

49 predictor of physician reported disease activity.⁴ Additionally, patients and physicians may have
50 different expectations with regard to the progression or outcome of the intervention, or what
51 constitutes a satisfactory progression in treatment or a good outcome.¹⁸

52 The impact of patient-physician discordance in orthopaedics and shoulder surgery is
53 relatively unknown. While discordance has been examined in one-time assessments of disease
54 state or impairment, few studies have examined the effect these differences in perception of the
55 results of treatment may have on the assessment of change over time or outcome following an
56 intervention.^{3, 18, 19} Two studies have reported fair to good patient-clinician agreement in patients
57 with low back pain¹¹ and disorders of the neck-shoulder region¹². Agreement on assessment of
58 outcome has been examined relative to pain and overall satisfaction, but only in a cohort of post-
59 operative patients following total hip arthroplasty.¹⁸ The need to examine the agreement or
60 discordance is important, since this is not an issue of whose assessment is right or wrong; each
61 perspective is equally valid. The patient's perspective should always be considered by the health
62 care provider as the patient is actually experiencing the treatment and is affected by the results.
63 However, clinicians are responsible for the content, timing, and direction of treatment and are
64 therefore most influential in guiding the patients' course of medical treatment. Determining
65 whether a patient has improved is an important factor in making treatment decisions for both the
66 physician and patient. If there is discordance in patient-reported and physician-reported
67 assessment of change among patients seeking medical care for shoulder pain, strategies should
68 be developed to improve agreement and communication to manage the discordance and perhaps
69 develop other tools to assess change that will minimize discordance. Therefore, the purpose of
70 this study was to examine the level of agreement between patient and physician assessment of
71 change, using a Global Rating of Change (GROC) scale in a cohort of patients being treated for

Patient-Physician agreement

72 shoulder problems. We hypothesized that there would be moderate (66 to 75%) agreement
73 between patient-reported and physician-reported assessment of change.

74 **Materials and Methods**

75 Subjects

76 Data from 99 subjects were used in this analysis (age = 41 ± 12 years, height = 175 ± 10
77 cm, weight = 84 ± 19 kg, 66 males). These data come from a larger study in which patients were
78 enrolled prospectively over two years. Of the 191 eligible subjects (220 patients approached,
79 211 enrolled, 20 withdrew), 99 had all data required for the present analysis. Patients reporting
80 to the [REDACTED] with shoulder pain were
81 identified as potential subjects. Patients were eligible for enrollment if they presented with
82 clinical history consistent with dysfunction due to musculoskeletal shoulder injury, reported pain
83 with overhead activity and were between 15 and 60 years of age. Patients were excluded if they
84 demonstrated signs and symptoms consistent with cervical radiculopathy²⁰, adhesive capsulitis²¹,
85 glenohumeral arthritis²² or reported tingling/numbness in the upper extremity, surgery on the
86 involved shoulder within the past year, or steroid injection within the last month.

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

Patient-Physician agreement

95 Procedures

96 Patient-oriented assessment of change was collected prospectively at the time of
97 physician follow up (6±1 weeks). Subjects completed a global rating of change (GROC) to
98 assess perceived improvement. The GROC is a 15 item scale ranging from “a very great deal
99 worse” to “a very great deal better” (Figure 1).²³ Subjects were instructed to select the statement
100 that best represented their perceived change in functional status since the initial evaluation.
101 Physician assessment of change was done retrospectively at the end of the enrollment period.
102 The treating physician, an orthopaedic surgeon, (██████) was provided with clinical notes from
103 both the initial evaluation and follow up visit for each patient and completed the same 15-point
104 GROC scale. Intra-rater reliability was excellent (ICC = 0.929) and was established by having
105 the physician rate the same 10 subjects at two separate times, with a minimum of one week
106 between ratings.

107 Data Reduction

108 The 15-point GROC was further consolidated into a 3-point scale by collapsing response
109 options into “better” (GROC score $\geq +3$), no change (-2 to +2), and “worse” (GROC score ≤ -3)
110 based on previously reported cutoffs used to identify clinically meaningful improvement.²⁴
111 Providing patients (or clinicians) with too many options may be of concern as the individual may
112 have difficulty attaching meaning to each separate response choice.²³ By treating the 15-point
113 scale as continuous, Intra-Class Correlation Coefficient (ICC) and Pearson’s r could be
114 conducted, while the consolidated 3-point scale allowed for confirmation of the findings with
115 weighted kappa using a more simplified scale of better/no change/worse.

116 Statistical Analysis

Patient-Physician agreement

117 To assess patient-physician agreement, ICC, Pearson's r correlation coefficient, and
118 linear weighted kappa were calculated. ICC and Pearson's r were calculated using the responses
119 on the 15-point GROC. ICCs were interpreted according to the following: <0.40 Poor, 0.04-0.75
120 Fair to Good, >0.75 Excellent.²⁵ Linear weighted kappa was calculated using the consolidated 3-
121 point scale (better, no change, worse). The strength of agreement for kappa was interpreted
122 according to the following: <0.00 Poor, 0.00-0.20 Slight, 0.21-0.40 Fair, 0.41-0.60 Moderate,
123 0.61-0.80 Substantial, 0.81-0.99 Almost Perfect.²⁶ Maximum kappa was calculated according to
124 Sim and Wright.²⁷ The maximum kappa value provides a more meaningful reference value for
125 interpretation because inadequate variation in the data can result in artificially low kappa
126 values.²⁷

127 **Results**

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

134 **Discussion**

135 One of the keystones of the doctor-patient relationship is that they are both in agreement
136 regarding the results of treatment and the direction of future care. This requires agreement on the
137 status of these treatment efforts. One key element would be the change in functional status as a
138 result of treatment. This study examined the patient-physician agreement or discordance related
139 to assessment of change following rehabilitation in patients with shoulder pain. Our hypothesis
140 of moderate agreement was supported, indicating that the Global Rating of Change scale appears

Patient-Physician agreement

141 to reflect and represent the same degree of change perceived by each group. Overall, we
142 observed moderate to good agreement. Our findings indicate similar patient-physician agreement
143 compared to previous research. Patient-physician agreement reported in the literature ranges
144 from 58 to 77%.^{1, 4, 6, 7, 10, 16, 28} Our finding of 37% complete patient-physician agreement on the
145 15-point GROC scale was expected to be lower because complete agreement was necessary.
146 Using the 3-point scale we were able to examine more global agreement, i.e. did the patient and
147 physician agree that the patient was better, the same or worse, rather than matching exactly to a
148 particular point on a 15-point scale. Complete patient-physician agreement using this 3-point
149 scale was 77%, which is at the high end of, though consistent with, previous reports. Our
150 assessment of weighted kappa (0.62) is also higher than previous reports to assess patient-
151 clinician agreement (range 0.09 to 0.39).^{3, 10, 12}

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

161 Limited research explores agreement or discordance in ratings of change over time in
162 functional health status. Patients with rheumatoid arthritis rated their change in global function
163 over 3 months, as did their treating physician.¹⁹ The authors observed a patient-physician

Patient-Physician agreement

164 relationship (ICC = 0.64, $r = 0.63$) very similar to the current study's findings.¹⁹ Patients with
165 heart disease were asked to use a 7-point "transition index scale" that appears quite similar to the
166 GROC to assess change in health-related quality of life in patients with heart disease.³ These
167 authors identified poor agreement ($k=0.09$ to 0.23) between patients and physicians. The low
168 agreement may be due to the type of data collected and compared. A single global assessment
169 made by the physician was compared to multiple domains assessed by the patients³. In the
170 current study the same global assessment was performed by both the patient and physician,
171 which seems to result in higher agreement in the present study and in previous research¹⁹.

172 In the only previous study to assess patient-physician agreement in change over time
173 involving an orthopaedic population, patient-reported assessment of pain and overall satisfaction
174 following total hip arthroplasty was compared to physician assessment using a visual analog
175 scale (VAS).¹⁸ Differences in patient and physician ratings of pain were statistically
176 significantly different ($1.7\pm 2.6\text{cm}$ and $1.1\pm 1.8\text{cm}$ respectively), though the difference was only
177 0.6cm on the VAS. Reports of overall satisfaction between patient and physician were not
178 significantly different (8.6 ± 2.1 and $8.8\pm 1.7\text{cm}$ respectively). The authors did note that patient-
179 physician agreement was notably worse among the patients with high pain or low satisfaction.
180 While this was an orthopaedic population, the cohort was post-surgical and the authors did not
181 provide an assessment of agreement (e.g., kappa, ICC), making it difficult to draw direct
182 comparisons to the current study. Our data provide the first examination of assessment of
183 change following conservative rehabilitation in an orthopaedic population.

184 We used a 15-point GROC to assess perceived change. The "global", less specific nature
185 of the GROC allows the patient to base their response on what is most important to them.²³ This
186 was ideal for addressing the purpose of the present study in that we wanted to identify if

Patient-Physician agreement

187 differences existed between perceptions of patients and clinicians. Test-retest reliability of the
188 GROC within 24 hours was excellent in patients with musculoskeletal disorders (ICC range 0.90
189 to 0.99).³² One limitation of a global rating of change assessment is that it requires the patient to
190 recall their previous condition with respect to their current status.²³ It has been suggested that
191 GROC scores may be influenced by current status as follow up time increases.³² The 3-point
192 GROC showed a much higher percentage of complete agreement between patient and physician
193 evaluations, probably due to limiting the available options. It may serve as a better basis for
194 discussion between the patient and physician regarding the results of treatment, and therefore
195 help to guide the discussion about future treatment plans.

196 Limitations

197 A few limitations of this study should be noted in order to interpret these results
198 accurately. First, patients completed the GROC at the time of their visit, while the physician
199 completed the GROC retrospectively at the end of the enrollment period. The physician had his
200 own notes to refer to when completing the GROC but it may have been more timely to have the
201 physician rate the patient using the GROC scale immediately following the visit. However, it
202 was felt that a longer time interval could provide a more objective analysis of the amount of
203 change, and doing the evaluations at one time would improve the consistency of the ratings.
204 Additionally, inclusion of a single physician may limit the generalizability of the results and
205 validation of the findings with additional physicians should be performed.

206 Our assessments examined change over time from baseline to follow up. While all
207 patients were prescribed a standardized physical therapy intervention, several variables could
208 have factored into the results including expectation of treatment success, patient satisfaction with
209 outcome or physician services¹⁸ and adherence to therapy. Future studies should account for

Patient-Physician agreement

210 those variables to further explain the patient-clinician relationship with regard to agreement on
211 health-related assessment.

212 Conclusion

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

223 **REFERENCES**

- 224
- 225 1. Alarcon GS, McGwin G, Jr., Brooks K, et al. Systemic lupus erythematosus in three
226 ethnic groups. XI. Sources of discrepancy in perception of disease activity: a comparison of
227 physician and patient visual analog scale scores. *Arthritis Rheum.* 2002; 47: 408-13.
- 228 2. Dandorfer SW, Rech J, Manger B, Schett G and Englbrecht M. Differences in the
229 patient's and the physician's perspective of disease in psoriatic arthritis. *Semin Arthritis Rheum.*
230 2012; 42: 32-41.
- 231 3. Wyrwich KW, Metz SM, Kroenke K, Tierney WM, Babu AN and Wolinsky FD.
232 Triangulating patient and clinician perspectives on clinically important differences in health-
233 related quality of life among patients with heart disease. *Health Serv Res.* 2007; 42: 2257-74;
234 discussion 94-323.
- 235 4. Yen JC, Abrahamowicz M, Dobkin PL, Clarke AE, Battista RN and Fortin PR.
236 Determinants of discordance between patients and physicians in their assessment of lupus
237 disease activity. *J Rheumatol.* 2003; 30: 1967-76.
- 238 5. Sewitch MJ, Abrahamowicz M, Bitton A, et al. Psychosocial correlates of patient-
239 physician discordance in inflammatory bowel disease. *Am J Gastroenterol.* 2002; 97: 2174-83.
- 240 6. Khan NA, Spencer HJ, Abda E, et al. Determinants of discordance in patients' and
241 physicians' rating of rheumatoid arthritis disease activity. *Arthritis Care Res (Hoboken).* 2012;
242 64: 206-14.
- 243 7. Berkanovic E, Hurwicz ML and Lachenbruch PA. Concordant and discrepant views of
244 patients' physical functioning. *Arthritis Care Res.* 1995; 8: 94-101.
- 245 8. Dobkin PL, De Civita M, Abrahamowicz M, et al. Patient-physician discordance in
246 fibromyalgia. *J Rheumatol.* 2003; 30: 1326-34.
- 247 9. Hidding A, van Santen M, De Klerk E, et al. Comparison between self-report measures
248 and clinical observations of functional disability in ankylosing spondylitis, rheumatoid arthritis
249 and fibromyalgia. *J Rheumatol.* 1994; 21: 818-23.
- 250 10. Kwok CK, O'Connor GT, Regan-Smith MG, et al. Concordance between clinician and
251 patient assessment of physical and mental health status. *J Rheumatol.* 1992; 19: 1031-7.
- 252 11. Perreault K and Dionne CE. Patient-physiotherapist agreement in low back pain. *J Pain.*
253 2005; 6: 817-28.
- 254 12. Perreault N, Brisson C, Dionne CE, Montreuil S and Punnett L. Agreement between a
255 self-administered questionnaire on musculoskeletal disorders of the neck-shoulder region and a
256 physical examination. *BMC musculoskeletal disorders.* 2008; 9: 34.
- 257 13. Chassany O, Le-Jeunne P, Duracinsky M, Schwalm MS and Mathieu M. Discrepancies
258 between patient-reported outcomes and clinician-reported outcomes in chronic venous disease,
259 irritable bowel syndrome, and peripheral arterial occlusive disease. *Value Health.* 2006; 9: 39-46.
- 260 14. Turk DC and Okifuji A. Assessment of patients' reporting of pain: an integrated
261 perspective. *Lancet.* 1999; 353: 1784-8.
- 262 15. Hudson M, Impens A, Baron M, et al. Discordance between patient and physician
263 assessments of disease severity in systemic sclerosis. *J Rheumatol.* 2010; 37: 2307-12.
- 264 16. Neville C, Clarke AE, Joseph L, Belisle P, Ferland D and Fortin PR. Learning from
265 discordance in patient and physician global assessments of systemic lupus erythematosus disease
266 activity. *J Rheumatol.* 2000; 27: 675-9.

- 267 17. Platt FW and Keating KN. Differences in physician and patient perceptions of
268 uncomplicated UTI symptom severity: understanding the communication gap. *Int J Clin Pract.*
269 2007; 61: 303-8.
- 270 18. Lieberman JR, Dorey F, Shekelle P, et al. Differences between patients' and physicians'
271 evaluations of outcome after total hip arthroplasty. *J Bone Joint Surg Am.* 1996; 78: 835-8.
- 272 19. Hanly JG, Mosher D, Sutton E, Weerasinghe S and Theriault D. Self-assessment of
273 disease activity by patients with rheumatoid arthritis. *J Rheumatol.* 1996; 23: 1531-8.
- 274 20. Wainner RS, Fritz JM, Irrgang JJ, Boninger ML, Delitto A and Allison S. Reliability and
275 diagnostic accuracy of the clinical examination and patient self-report measures for cervical
276 radiculopathy. *Spine.* 2003; 28: 52-62.
- 277 21. Griggs SM, Ahn A and Green A. Idiopathic adhesive capsulitis. A prospective functional
278 outcome study of nonoperative treatment. *J Bone Joint Surg Am.* 2000; 82-A: 1398-407.
- 279 22. Kelley MJ and Ramsey ML. Osteoarthritis and traumatic arthritis of the shoulder. *J Hand*
280 *Ther.* 2000; 13: 148-62.
- 281 23. Kamper SJ, Maher CG and Mackay G. Global rating of change scales: a review of
282 strengths and weaknesses and considerations for design. *J Man Manip Ther.* 2009; 17: 163-70.
- 283 24. Young IA, Cleland JA, Michener LA and Brown C. Reliability, construct validity, and
284 responsiveness of the neck disability index, patient-specific functional scale, and numeric pain
285 rating scale in patients with cervical radiculopathy. *American Journal Of Physical Medicine &*
286 *Rehabilitation / Association Of Academic Physiatrists.* 2010; 89: 831-9.
- 287 25. Fleiss JL. *The Design and Analysis of Clinical Experiments.* New York, NY: Wiley,
288 1986.
- 289 26. Landis JR and Koch GG. The measurement of observer agreement for categorical data.
290 *Biometrics.* 1977; 33: 159-74.
- 291 27. Sim J and Wright CC. The kappa statistic in reliability studies: use, interpretation, and
292 sample size requirements. *Phys Ther.* 2005; 85: 257-68.
- 293 28. Scheuer E, Steurer J and Buddeberg C. Predictors of differences in symptom perception
294 of older patients and their doctors. *Fam Pract.* 2002; 19: 357-61.
- 295 29. Polson K, Reid D, McNair PJ and Larmer P. Responsiveness, minimal importance
296 difference and minimal detectable change scores of the shortened disability arm shoulder hand
297 (QuickDASH) questionnaire. *Man Ther.* 2010; 15: 404-7.
- 298 30. Mintken PE, Glynn P and Cleland JA. Psychometric properties of the shortened
299 disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH) and Numeric Pain
300 Rating Scale in patients with shoulder pain. *J Shoulder Elbow Surg.* 2009; 18: 920-6.
- 301 31. Gummesson C, Ward MM and Atroshi I. The shortened disabilities of the arm, shoulder
302 and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-
303 length DASH. *BMC musculoskeletal disorders.* 2006; 7: 44.
- 304 32. Kamper SJ, Ostelo RW, Knol DL, Maher CG, de Vet HC and Hancock MJ. Global
305 Perceived Effect scales provided reliable assessments of health transition in people with
306 musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol.*
307 2010; 63: 760-6 e1.

308
309

Patient-Physician agreement

310 **Figure legends**

311 Figure 1: Global Rating of Change Scale

312

313 Figure 2: Patient-Physician Agreement Plot

314 Points that fall within the green (“better”, n=37), yellow (“no change”, n=35) and red (“worse”,
315 n=4) boxes represent that the patient and physician both rated the patient in the same category.

316 Points that fall outside of the boxes represent disagreement between the patient and physician
317 (n=23). The values represent the number of patients represented by that data point.

318

319