

# Clinimetrics corner: the Global Rating of Change Score (GRoC) poorly correlates with functional measures and is not temporally stable

**Craig Garrison<sup>1</sup>, Chad Cook<sup>2</sup>**

<sup>1</sup>Research Physical Therapist, Texas Health Ben Hogan Sports Medicine, Fort Worth, TX, USA, <sup>2</sup>Division of Physical Therapy, Walsh University, North Canton, OH, USA

The Global Rating of Change Score (GRoC) is a frequently used outcome measure that is used independently to measure improvements in a patient's condition or as an anchor for other outcomes measures. The tool has been criticized for recall bias, biases in administration, and for poor reliability over time. Our findings, captured from a sample of patients with shoulder impingement, suggest that all these concerns are of merit. Our results show poor correlation of the GRoC with functional measures after 2 and 3 weeks and decay of the associative stability of the GRoC from week to week.

**Keywords:** Global Rating of Change, Outcomes measures, Shoulder impingement

## Background

The Global Rating of Change Score (GRoC) is a frequently used outcome measure that has been used to independently score self-perceived improvement in a patient and has been used as an anchor method to determine minimal clinically important change scores. The GRoC is a single-item, recall-based questionnaire of well-being that is based on progress (or lack of progress) since an initial treatment encounter.<sup>1</sup> Patients are routinely asked to make global ratings on changes in regards to their level of shoulder well-being since the previous week's treatment on a 15-point self-report scale (from -7 to 7), although other scale values have been used. The GRoC has been the outcome of choice for patients with shoulder pain in previous studies.<sup>2,3</sup> The tool is easy to administer, easy to use, and is considered a bastion for outcomes measures.

The GRoC is a retrospective judgment tool and there are weaknesses that have been identified with these forms of outcomes measures.<sup>4,5</sup> First, recall bias may limit the reliability of the tool over time.<sup>5</sup> Second, retrospective tools are prone to bias, specifically those associated with social responsibility or mode of administrative influences that are consciously or subconsciously supplied by the clinician.<sup>6-8</sup> Last, the

tools allow the patient to score their perceived recovery using any construct they choose,<sup>9</sup> which is a strength and a weakness. Thus, elements such as pain or factors that are not considered important by the treating clinician may be the factors selected by the patient when determining recovery.

Recently, within a clinical context, we have noticed the lack of temporal stability of individual GRoC measures for patients who were enrolled in our shoulder study. Patient GRoC findings fluctuated widely from week to week, and in many cases, the actual values presented during the weeks before the final week (which was taken as the final GRoC score for the analysis) differed dramatically from the later values. In addition, clinically, there did not seem to be a longitudinal correlation between GRoC scores and independent measures of function, measures that have historically served to determine who truly has benefited from a specific intervention. This finding prompted the investigation of the temporal stability of the GRoC and whether the weekly GRoC values were related to weekly American Shoulder and Elbow Surgeon's Scale (ASES) change scores and whether or not individuals indicated the ability to achieve the three Patient Specific Functional Scale measures they identified as challenges before the initiation of rehabilitation. We hypothesized that the GRoC individual weekly measures did not correlate with the functional measures and that a weekly finding of a

Correspondence to: C Cook, 2020 East Maple Street, North Canton, OH 44720. Email: ccook@walsh.edu

GRoC does not have consistent associative stability with functional measures over time.

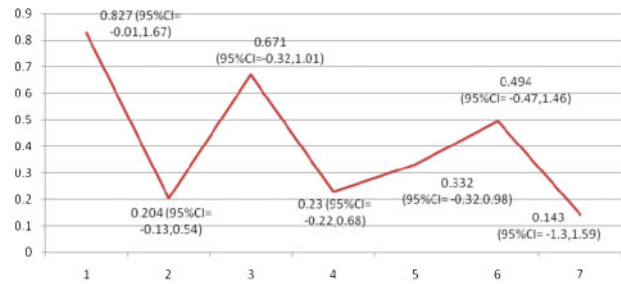
**Methods**

Data for this secondary analysis were obtained from a prospective, longitudinal trial involving a standardized treatment program. Details of the study are outlined elsewhere.<sup>10</sup> The original study was designed to identify associative variables for prognosis of patients who received physical therapy for shoulder impingement syndrome. All patients in the study exhibited some or all of the following criteria: (1) manifestation by gradual onset and characterization of a dull ache at the anterolateral aspect of the shoulder often with radiation to the deltoid insertion; (2) patient complaints with overhead activity; and (3) the condition is worsened with resisted abduction, resisted external rotation, overhead positioning of the arm, or direct pressure against the shoulder such as lying on it. All subjects were diagnosed with impingement syndrome by an attending physician and the diagnosis was confirmed by the treating physical therapist.

Outcomes data germane to this study in addition to the GRoC included the ASES shoulder questionnaire<sup>11</sup> and the Patient Specific Functional Scale (PFPS).<sup>12</sup> The ASES was developed by the Research Committee of the American Shoulder and Elbow Surgeons and was adopted by the membership as a standardized form for measurement of musculoskeletal function.<sup>11</sup> The questionnaire, which includes two theoretical subscales, pain and function/disability,<sup>11</sup> has been shown to possess reliability,<sup>13,14</sup> construct validity with the SF-36,<sup>14,15</sup> and responsiveness when moderate to large effect sizes are present during treatment.<sup>14</sup> In addition, a minimally clinically important difference has been demonstrated as a change of 6.4 ASES points.<sup>14,15</sup> The PFPS<sup>12</sup> uses three questions that are designed to assess difficulties in functional activities that are specific to that person. The original PFPS allows users to rate the level of difficulty of each activity from 0 (unable to perform) to 10 (able to perform at pre-injury level). The same three activities are summed to provide a total score of 0–30. We modified the PFPS within our study by asking individuals to pick an activity they could not do and then at each week asking whether or not they were currently able to do the task. All outcomes including the GRoC were collected weekly by a third party, non-clinician.

*Statistical analysis*

We created regression coefficients for each time period by analyzing the association of biweekly measures for the GRoC, with the assumption that the two are related to one another and that progression (improvement) is a linear phenomenon. Biweekly regression



**Figure 1** Linear association of weekly GRoC measures (stability) over time.

coefficients were plotted over time and the coefficient and 95% confidence interval measures were reported.

Median values and low and high range values for each GRoC weekly score were calculated as well as mean and confidence interval measures. In addition, we plotted the correlations between weekly GRoC measures and independent weekly ASES and PFPS measures (all weekly), over time.

**Results**

Figure 1 outlines the temporal stability of the GRoC after assessment of the influence of the prior week’s GRoC scores with the subsequent week. The seven timeframes (8 different weeks) show an inconsistent pattern of stability and that the prior week’s score do not always drive the follow-up week. In addition, the very wide confidence intervals suggest a high degree of variability in the findings as well. The weekly GRoC median values ranged from a low of 0 to a high of 3. Mean values were also dispersed significantly from week to week with very wide confidence intervals (Table 1).

Correlational findings with the functional measures of ASES and PFPS yielded significant correlations in the first 2 weeks and only two instances of association after that point (Table 2). The correlations were statistically significant but relatively low.

**Discussion**

Our findings support our hypothesis that the GRoC is not temporally stable in that a particular finding 1 week ago is not associative to functional results the following week. This suggests that patient’s report of global change in their condition may fluctuate widely from one week to another and that GRoC outcomes are not linear and progressive, but instead fluctuate

**Table 1** Median and mean weekly GRoC measures

	Median (range)	Mean (95% confidence interval)
Week 1	2 (–4–7)	2.4 (–1.83–2.96)
Week 2	2 (–4–7)	2.3 (–1.57–3.02)
Week 3	3 (–4–7)	2.4 (–1.63–3.16)
Week 4	3 (–4–7)	1.9 (–0.90–2.89)
Week 5	0 (0–6)	1.5 (–0.40–2.59)
Week 6	2 (–3–6)	2.3 (–0.57–4.02)
Week 7	2 (–1–6)	2.6 (–0.93–4.2)
Week 8	0 (–2–0)	0.7 (–0.61–2.01)

**Table 2** GRoC correlations with weekly ASES and weekly PFPS measures

	Comparison of weekly GRoC to weekly ASES change score	Comparison of weekly GRoC to obtained of defined function 1	Comparison of weekly GRoC to obtained of defined function 2	Comparison of weekly GRoC to obtained of defined function 3
Week 1	0.31**	0.26*	0.35**	0.34**
Week 2	0.29*	0.38**	-0.43**	-0.24*
Week 3	0.27*	0.15	-0.00	-0.20
Week 4	0.23	0.15	-0.19	-0.25
Week 5	0.13	0.36	-0.17	-0.34
Week 6	0.31	0.40	-0.79**	-0.26
Week 7	-0.05	-0.20	-0.34	-0.64
Week 8	0.9	-0.50	0.50	1.0**

Note: \*Significance at 0.05.

\*\*Significance at 0.01.

widely over multi-week period. This wide degree of variability could affect the overall outcome at the end point of a trial and could influence a minimally clinical important change score if used as an anchor point in a study. This deserves further investigation to whether this phenomenon is unique to our study or if it is consistent, yet unexplored in others.

Our most notable finding is that the GRoC, as a whole, does not correlate with the ASES or PFPS measures after 2 weeks. This is in contrast to others who have found a notable correlation, albeit at 4 weeks only.<sup>16</sup> We found a small correlation at 1 and 2 weeks for all three measures and for week 3 with the ASES, in addition to two rather specious findings of correlation after that point. This suggests two possible considerations. First, the GRoC may measure a construct outside of domain of function. Since it is a multidimensional questionnaire, one could argue that the construct is specific to each patient and that this finding has limited impact. However, what may be more compelling possibility is that the correlation with function declines significantly after week 2 for the PFPS and after week 3 for the ASES. This suggests that concerns associated with recall decline or reliability of the GRoC may be warranted. As a reminder, the ASES and PFPS are both measures that are designed to capture findings at the present point of time, whereas the GRoC requests that individuals remember back before their therapy was initiated.

We witnessed a decline in median GroC values over time, specifically after the third and fourth weeks of care. This may mean that the chances for identifying a self-perceived improvement reduces over time and that studies that examine GRoC-related outcomes are less inclined to see positive outcomes with long term versus short term. Since a number of studies have used the GRoC for short-term analyses only, it is unlikely that this phenomenon has been observed during care in these studies. If our study limited data collection to less than 4 weeks, we too would have not identified this finding.

## Conclusions and Key Points

Further work is needed to determine the true value of the GRoC as an outcome measure and in turn as an anchor measure. Several key points were identified in this study:

1. There is fluctuant temporal stability of the GRoC from week to week.
2. There is poor correlation between the GRoC and functional measures.
3. The GRoC is only correlated to functional measures up to 3 weeks.

## References

- 1 Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Controlled Clin Trials*. 1989;10:407-15.
- 2 Garrison JC, Shanley E, Thigpen C, Hegedus E, Cook C. Between-session changes predict overall perception of improvement but not functional improvement in patients with shoulder impingement syndrome seen for physical therapy: an observational study. *Physiother Theory Pract*. 2011;27(2):137-45.
- 3 Mintken PE, Glynn P, Cleland JA. Psychometric properties of the shortened disabilities of the arm, shoulder, and hand questionnaire (QuickDASH) and numeric pain rating scale in patients with shoulder pain. *J Shoulder Elbow Surg*. 2009;18:920-6.
- 4 Ross M. Relation of implicit theories to the construction of personal histories. *Psychol Rev*. 1989;96:341-57.
- 5 Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol*. 1997;50:869-79.
- 6 Schmitt JC, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol*. 2004;57:1008-18.
- 7 Schmitt JC, Di Fabio RP. The validity of prospective and retrospective global change criterion measures. *Arch Phys Med Rehabil*. 2005;86:2270-6.
- 8 Cook C. Mode of administration bias. *J Man Manip Ther*. 2010;18(2):61-3.
- 9 Kamper S, Maher C, Mackay G. Global Rating of Change Scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther*. 2009;17(3):163-70.
- 10 Garrison JC, Shanley E, Thigpen C, Hegedus E, Cook C. Between-session changes predict overall perception of improvement but not functional improvement in patients with shoulder impingement syndrome seen for physical therapy: an observational study. *Physiother Theory Pract*. 2011;27:137-45.
- 11 Cook C, Hegedus E, Goode A, Mina C, Pietrobon R, Higgins LD. Relative validity of the modified American Shoulder and Elbow Surgeons (M-ASES) questionnaire using item response theory. *Rheumatol Int*. 2008;28:217-23.
- 12 Westaway MD, Stratford PW, Binkley JM. The patient-specific functional scale: validation of its use in persons with neck dysfunction. *J Orthop Sports Phys Ther*. 1998;27:331-8.
- 13 McClure P, Michener L. Measures of adult shoulder function. *Arthritis Rheum*. 2003;49:50-8.

- 14 Kocher MS, Horan MP, Briggs KK, Richardson TR, O'Holleran J, Hawkins RJ. Reliability, validity, and responsiveness of the American Shoulder and Elbow Surgeons subjective shoulder scale in patients with shoulder instability, rotator cuff disease and glenohumeral arthritis. *J Bone Joint Surg.* 2005;87:2006–11.
- 15 Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section: reliability, validity, and responsiveness. *J Shoulder Elbow Surg.* 2002;11:587–94.
- 16 Fritz JM, Hebert J, Koppenhaver S, Parent E. Beyond minimally important change. *Spine (Phila Pa 1976).* 2009;34:2803–9.