Assessing results after distal radius fracture treatment – a comparison of objective and subjective tools

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Geriatric Orthopaedic Surgery &amp; Rehabilitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>GOS-11-0013.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Original Manuscript</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Fragility Fractures, Geriatric Trauma, Physical Medicine and Rehabilitation, Trauma surgery, Upper extremity surgery</td>
</tr>
<tr>
<td>Abstract:</td>
<td>Functional outcomes following distal radius fractures are directly influenced by the choice of outcome assessment instruments used. Our objective was to compare scoring systems in measuring patient functional outcomes, and determine which scoring system compared most favourably with the widely-used Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire.</td>
</tr>
</tbody>
</table>
Assessing results after distal radius fracture treatment – a comparison of objective and subjective tools

The authors declared no potential conflicts of interest with respect to the authorship and/or publication of this article.
Abstract

Objectives: Functional outcomes following distal radius fractures are directly influenced by the choice of outcome assessment instruments used. Our objective was to compare scoring systems in measuring patient functional outcomes, and determine which scoring system compared most favourably with the widely-used Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire.

Methods: 108 patients between May 2004 and November 2006 were treated operatively following distal radius fractures. Follow-up was at three months, six months, one year and two years post-surgery, during which anatomical and functional assessments were performed. Patient outcomes were recorded using DASH, the Green and O’Brien system, Gartland and Werley system and Sarmiento radiological scoring system.

Results: There was a stronger correlation between the Green and O’Brien scoring system and DASH ($r = -0.54$) than Gartland and Werley and DASH ($r = 0.44$). The Green and O’Brien scoring system was more demanding so patients rated ‘excellent’ or ‘good’ had better functional outcome than those bearing the same grade in the Gartland and Werley system. Nonetheless the Green and O’Brien score and Gartland and Werley score showed good correlation with each other ($r = 0.66$). The Sarmiento radiological score had no significant correlation with any of the other scoring systems. Significant predictors of the DASH score were function ($r = 0.42$), power grip ($r = 0.41$), pain ($r = 0.37$) and range of motion ($r = 0.28$).

Conclusion: The Green and O’Brien scoring system correlated most strongly with the DASH score. Radiological scoring (reflecting anatomical deformity) was not significantly correlated with functional outcome. Whilst subjective parameters ‘pain’ and ‘function’ are influenced by psychosocial factors and thus highly variable, it is paramount to include subjective tools in outcome assessment in future studies on wrist fractures.
Keywords
Distal radius fractures; outcome assessment; patient self-assessment; elderly
Introduction
Distal radius fractures are the commonest fractures in the elderly. The measurement of results of treatment is dependent on the type of outcome assessment used. It is important to know which assessment tool is the most useful to the surgeon and the patient, and which outcome measures are the most reliable in reflecting disability and function after trauma.

The correlation between objective and subjective measures have shown to be variable, although many of these reports have been based on relatively small sample sizes [1-5]. Traditional measures of objective variables like grip strength and range of motion do not always accurately correlate to patient-reported pain and function [6]. Although scoring systems like the Gartland and Werley score and the Green and O'Brien score have come into favour, substantial differences still exist between these scoring systems [4]. In recent years the most-widely used instrument in evaluating upper extremity outcome is the DASH patient-rated health questionnaire [5, 7].

According to kinesiological theory, the upper extremity operates as a single functional unit. The DASH questionnaire, although neither side- nor joint-specific, is highly responsive to change in assessment of function following distal radius fractures [8].

This was a retrospective study in a tertiary teaching hospital looking at how closely objective and subjective measurements reflected patient functional outcome following open reduction and internal fixation of distal radius fractures between May 2004 and November 2006. Our objective was to compare different assessment tools in measuring patient outcomes, and to see which scoring system compared most favourably to the DASH score.

Materials and Methods

Between May 2004 and November 2006, 108 patients with distal radial fractures were treated using the 2.4-mm and 3.5mm locking plate fixation (Synthes, Switzerland). The indication for surgery was a displaced fracture of the distal radius following unsuccessful closed reduction or fracture with intra-articular disruption. Patients with an operative procedure performed one month after the initial injury and patients with polytrauma with an injury severity score of >16 were excluded [9].

Patient demographics and characteristics

Twenty seven of the 135 patients were lost on follow-up or had incomplete
assessments performed and were excluded from the study. Of the 108 patients, there were 42 (39%) men and 66 (61%) women, with a mean age of 55.5 years (range 13-90 years). A 2.4mm-plate was used in 62 patients (57%) and 3.5mm-plate on 46 (43%) patients.

Follow-up Protocol

All patients were seen in an outpatients follow-up clinic two weeks after the operation for wound check and suture removal. Subsequent follow-up was at three months, six months, one year and two years post-surgery. Anatomical and functional assessments were performed at an average of 20.6 months (range 3 to 26 months).

Anatomical Assessment

Postero-anterior and lateral radiographs were taken at each follow-up visit and measurements were recorded using the method developed by Kreder et al [10]. On the postero-anterior films, radial length, radial angle, articular step-off and gap were measured. On the lateral film, palmar tilt and articular step-off and gap were measured. A radiological score was derived from the above measurements – in this study Sarmiento et al’s modification of the Lidström and Frykman radiological classification was used [11].

Functional Assessment

Three evaluation tools were used in this study. The Cooney’s modification of the Green and O’Brien score [12] is an examiner-rated assessment of pain, functional status, range of motion and grip strength. Each of the four parameters are given a weighting of 25 points, giving a total score out of 100. With excellent being 90-100, good 80-89, fair 65-79 and poor <65.

The demerit system of Gartland and Werley [13] is a mixed subjective and objective assessment that includes residual deformity (3 points), subjective evaluation (6 points), objective evaluation based on range of movement (5 points) and complications including pain (5 points). With excellent being 0-2, good 3-8, fair 9-20, poor ≥ 21.

The DASH questionnaire is a patient-rated tool and is the most validated measure of
upper extremity functional status [14]. Questions are based on daily activities, symptoms including pain, and an optional work and sports/performing arts module. A final score is calculated, ranging from 0 (no disability) to 100 (the most severe disability). Thus a higher score indicates greater disability. A validated Chinese version of the DASH questionnaire was used in this study [15].

Statistical analyses

Continuous variables were described using means, standard deviations (SD) and ranges.

The Pearson product-moment correlation coefficient was used to evaluate the association in scores between the outcome instruments.

The Spearman rho correlation was used to evaluate association in the final grading (excellent, good, fair, poor) between different outcome instruments.

Multiple linear regression analysis was performed to identify which predictor variables were significantly associated with outcome score (DASH).

P values of <0.01 were regarded as significant. All analyses were carried out using the SPSS software package (version 16; SPSS Inc., Chicago, IL.)

Results

Of the 108 patients there were 46 under 55 year-olds (43%) and 62 over 55 year-olds (57%). Their mean DASH scores were 12.5 (SD = 15.3) and 12.1 (SD = 15.1) respectively. As the under 55 year-old and over 55 year-old age groups showed no significant difference between their mean DASH scores (Table 1), no further breakdown of their respective correlations to the grading instruments were carried out; Patients of all ages were analysed together (Table 2).

A. Correlation in scores between outcome instruments
A moderate correlation was seen between scores in Green and O’Brien and DASH (r = -0.54, p = 0.01) as well as between Gartland and Werley and DASH (r = 0.44, p = 0.01).
Between the two clinician-based scoring systems (Green and O’Brien and Gartland and Werley), there was good correlation (r = -0.75, p = 0.01).
B. Correlation in final grading between outcome instruments

As seen in Table 3, there was moderate rank correlation between final Green and O’Brien grade and DASH (r = -0.42, p = 0.01) but a weaker correlation between Gartland and Werley and DASH (r = 0.31, p = 0.01). No correlation was shown between the Sarmiento score and DASH (r = -0.045, p = 0.652). Correlation between Green and O’Brien and Garland and Werley scoring systems was strong, however (r = 0.66, p = 0.01).

C. Predictors of clinical outcome

Multiple linear regression analysis identified function (0.42, p < 0.01), power grip (r = 0.41, p < 0.01), pain (r = 0.37, p < 0.01) and range of motion (r = 0.28, p = 0.02) as significant predictors to the DASH score (Table 4).

Discussion and Conclusions

Distal radius fractures are amongst the most common fractures encountered by orthopaedic surgeons as they are the commonest fractures in the elderly. Although the initial description by Colles stated that there is little functional compromise in its aftermath, for years this has been a subject of debate and there is no clear consensus regarding its appropriate treatment. Scoring systems help clinicians evaluate the necessity for operative treatment by looking at patients’ function outcome but correlation of scoring systems to the DASH score – the most widely used health questionnaire for upper extremity outcome – are variable.

We have shown in this study that the Green and O’Brien scoring system has a stronger correlation to DASH in both its raw score and final grading than that of Gartland and Werley and DASH. Table 2 also shows that patients had a higher average DASH score (i.e. less disability) when graded with the Green and O’Brien system compared to using the Gartland and Werley system. Therefore patients rated ‘excellent’ or ‘good’ have better functional outcome than those bearing the same grade in the latter scoring system. Green and O’Brien is more demanding a scoring system, requiring 100% normal grip and function for the maximum mark to be awarded in those parameters.

Reliance on certain variables can impact on their correlation with the DASH score. We have shown that significant predictors of the DASH score were found to include power grip, pain, function and range of motion – the exact 4 parameters used in the
Green and O’Brien score. The Gartland and Werley score, however, takes into account other parameters such as residual deformity and complications (nerve complications) but not grip strength. It is unsure as to how significant the contribution of these factors is towards the DASH score.

The Sarmiento radiological score has proven to have no correlation with the DASH score. This suggests that radiographic outcome and thus anatomical deformity has little effect towards functional outcome. These findings are consistent with results from previous studies [3, 16]. This may explain the comparatively weak correlation of Gartland and Werley score with DASH. It should be noted that all the subjects in this study were treated operatively, thus few patients were left with substantial residual deformity.

In the Gartland and Werley score as well as the DASH score, the ‘pain’ and ‘function’ parameters are subjective rather than physician-rated. Perception of pain and function by patients are known to be strongly influenced by psychosocial factors [17]. Although such parameters are important it can produce highly variable results. There is less of such variability in the Green and O’Brien score. Patient self-assessment questionnaires rely highly on compliance; Often lower completion rates can be expected of longer questionnaires and in older and frailer populations, particularly patients with poor hand and wrist function. This can pose a potential selection bias in this wrist fracture population.

Recently there been increasing in popularity in the use of the PRWE (patient-rated wrist evaluation) score [6, 18]. It comprises of 2 subscales – pain and function – and was developed specifically for patients with wrist fractures. Following validity, reliability and responsiveness testing, it has proven to be a robust scoring system correlating well to patient functional outcome. However the PRWE score was not included in our study as it was only recently introduced and had not been widely adopted during the period of data collection.

Given the weight of patient-rated factors in influencing final outcome, a suggestion drawn from the results of this study is that a subjective tool should always be included as part of outcome assessment in all future studies in distal radius fractures, even if produces variability in results. This is applicable for both young and osteoporotic adults.

The weakness of this study is that the follow-up times were variable hence patients
might have been in different stages of rehabilitation. Also, we have only carried out linear regression analysis on the DASH score. It would be interesting to carry out a similar analysis for the Green and O’Brien and Gartland and Werley scores to find out the impact of each individual factor on each scoring system. In terms of statistical analyses, in this study only linear regression was performed – we made the assumption that the parameters (age, function, pain, range of motion, etc.) formed a linear relationship with the DASH score. However, we have found that significant collinearity exists between variables, for example, between age and range of motion (-0.37), pain and function (0.41), plate type and range of motion (-0.64). High correlations between variables may pose the problem of multicollinearity in regression analysis. Fractures of the distal radius are common in older people, particularly in post-menopausal women. Often they are the result of low energy trauma. Since rehabilitation potential is highly variable with age, it would be useful to see whether significant differences exist by using the same outcome assessment instruments in younger and older populations.
A.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Green &amp; O’Brien</th>
<th>Gartland and Werley</th>
<th>Sarmiento</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>8.9 (11.8)</td>
<td>9.9 (12.8)</td>
<td>13.4 (15.8)</td>
</tr>
<tr>
<td>Good</td>
<td>19.0 (19.0)</td>
<td>31.5 (22.0)</td>
<td>10.5 (14.6)</td>
</tr>
<tr>
<td>Fair</td>
<td>29.5 (23.7)</td>
<td>25.0 (-)</td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>34.1 (-)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Green &amp; O’Brien</th>
<th>Gartland and Werley</th>
<th>Sarmiento</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>7.5 (10.7)</td>
<td>9.1 (11.7)</td>
<td>12.8 (16.4)</td>
</tr>
<tr>
<td>Good</td>
<td>15.6 (14.3)</td>
<td>16.7 (17.9)</td>
<td>9.2 (9.3)</td>
</tr>
<tr>
<td>Fair</td>
<td>16.7 (15.3)</td>
<td>30.3 (28.8)</td>
<td>17.1 (24.1)</td>
</tr>
<tr>
<td>Poor</td>
<td>47.7 (19.8)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Mean DASH scores (standard deviation) for each final grade when using different scoring systems. A) in under 55-year olds, and B) in over 55-year olds. (-): no standard deviation available as n=1.
<table>
<thead>
<tr>
<th>Grade</th>
<th>Green &amp; O’Brien</th>
<th>Gartland and Werley</th>
<th>Sarmiento</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>8.2 (11.1)</td>
<td>9.5 (12.2)</td>
<td>15.6 (17.8)</td>
</tr>
<tr>
<td>Good</td>
<td>16.9 (11.7)</td>
<td>20.2 (17.4)</td>
<td>6.5 (7.2)</td>
</tr>
<tr>
<td>Fair</td>
<td>20.7 (18.3)</td>
<td>29.0 (23.7)</td>
<td>17.1 (24.1)</td>
</tr>
<tr>
<td>Poor</td>
<td>44.3 (17.6)</td>
<td>24.0 (19.8)</td>
<td>23.6 (22.1)</td>
</tr>
</tbody>
</table>

Table 2: Mean DASH scores (standard deviation) for each final grade when using different scoring systems – includes patients of all ages.
<table>
<thead>
<tr>
<th></th>
<th>Sarmiento</th>
<th>DASH</th>
<th>Green &amp; O'Brien Grade</th>
<th>Gartland &amp; Werley Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Correlation Coefficient</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sarmiento</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.045</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DASH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.642</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Green &amp; O'Brien Grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.455</td>
<td></td>
<td></td>
<td>.662**</td>
</tr>
<tr>
<td><strong>Gartland &amp; Werley Grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.791</td>
<td>.001</td>
<td>.000</td>
<td>.000</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).**

Table 3: Spearman's rank correlation between final grading in Green and O'Brien, Gartland and Werley, Sarmiento and DASH.
### DASH score

<table>
<thead>
<tr>
<th>Correlation</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.95</td>
</tr>
<tr>
<td>Gender</td>
<td>0.056</td>
</tr>
<tr>
<td>Fracture type</td>
<td>-0.071</td>
</tr>
<tr>
<td>Plate</td>
<td>0.059</td>
</tr>
<tr>
<td>Power Grip %</td>
<td>-0.406</td>
</tr>
<tr>
<td>Pain</td>
<td>-0.365</td>
</tr>
<tr>
<td>Function</td>
<td>-0.415</td>
</tr>
<tr>
<td>ROM</td>
<td>-0.278</td>
</tr>
<tr>
<td>Length of follow-up</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Table 4: Regression analysis of predictor variables in Green and O’Brien, Gartland and Werley, Sarmiento and DASH scores.
References


Appendix 1:
Green and O’Brien Score (Cooney’s modification) [12]

I. Pain (25 points)
- 25  None
- 20  Mild, occasional
- 15  Moderate, tolerable
- 0   Severe or intolerable

II. Range of motion (25 points): flexion + extension, % of normal
- 25  100
- 15  75-99
- 10  50-74
-  5  25-49
-  0  0-24

III. Grip strength (25 points), % of normal
- 25  100
- 15  75-99
- 10  50-74
-  5  25-49
-  0  0-24

IV. Activities (25 points)
- 25  Returned to regular employment
- 20  Restricted employment
- 15  Able to work but unemployed
-  0  Unable to work because of pain

Final result
- 90-100  Excellent
- 80-89   Good
- 65-79   Fair
- <65     Poor
Appendix 2:

Gartland and Werley Score (demerit system) [13]

I. Subjective evaluation – 6
   - Excellent: no pain, disability or limitation of motion (0)
   - Good: occasional pain, slight limitation of motion and no disability (2)
   - Fair: occasional pain, some limitation of motion, feeling of weakness in wrist, no particular disability if careful and activities slightly restricted (4)
   - Poor: pain, limitation of motions, disability and activities more or less markedly restricted (6)

II. Objective evaluation – 5
   - Loss of dorsiflexion (5)
   - Ulnar deviation (3)
   - Supination (2)
   - Palmar flexion (1)
   - Radial deviation (1)
   - Circumduction (1)
   - Distal radioulnar joint (1)

III. Residual deformity – 3
   - Prominent ulnar styloid (1)
   - Residual dorsal tilt (2)
   - Radial deviation of hand (2-3)

IV. Complications – 5
   - Arthritic changes
     - Minimal (1)
     - Minimal with pain (3)
     - Moderate (2)
     - Moderate with pain (4)
     - Severe (3)
     - Severe with pain (5)
   - Nerve complications (median) (1-3)
   - Poor finger function due to cast (1-2)

Final result

0-2 Excellent
3-8 Good
9-20 Fair
≥21 Poor
Appendix 3:
Sarmiento Radiological Score (modified from Lidström and Frykman) [11]

Excellent  No or insignificant deformity
Dorsal angulation $\geq 0$ degrees
Shortening of $< 3$mm
Loss of radial deviation $< 4$ degrees

Good  Slight deformity
Dorsal angulation of 1-10 degrees
Shortening of 3-6mm
Loss of radial deviation 5-9 degrees

Fair  Moderate deformity
Dorsal angulation of 11-14 degrees
Shortening of 7-11mm
Loss of radial deviation 10-14 degrees

Poor  Severe deformity
Dorsal angulation of $\geq 15$ degrees
Shortening of $\geq 12$mm
Loss of radial deviation $> 15$ degrees

(Average radial deviation of 23 degrees)