Original Article

Chertsey Outcome Score for Trauma: Development and validation of a new unifying patient reported outcome measure for orthopaedic trauma

Efthymios Iliopoulos*, Sujit Agarwal, Arshad Khaleel

The Rowley Bristow Unit, Ashford and St Peter's Hospitals NHS Foundation Trust, Chertsey, United Kingdom

Purpose: Lately there is an increasing tendency of using Patient Reported Outcome Measures (PROMs) as a final indicator of the outcome of many surgical treatments in orthopaedics and in other medical specialties. Currently there are many outcome scores in orthopaedics and most of them are site specific. In the contrary there is a lack of trauma specific outcome scores.

Methods: We have designed a new PROM especially for orthopaedic trauma patients, in order to measure in what extent the patients manage to return to their pre-injury state. This score uses as baseline the pre-injury status of the patient and has the aim to determine the percentage of rehabilitation after treatment for any injury.

Results: A total of 60 Chertsey Outcome Score for Trauma (COST) questionnaires were gathered in our outpatients department. The participants were 57% male (aged 46.81 years ± 18.5 years) and the questionnaires collected at mean 10 months post-injury. A Cronbach's Alpha value of 0.89 was identified for the whole construct. The three dimensions of the scale had good internal consistency as well (Cronbach's Alpha test values 0.74, 0.84 and 0.81 for symptoms, function and mental status respectively). Strong/moderate correlation (Spearman's Rho test 0.43 to 0.65) was observed between the respective physical/mental dimensions of the COST and SF-12v2 questionnaires.

Conclusion: There is a need among the orthopaedic trauma society for a specific PROM of trauma. COST is a useful and easy to use tool for every trauma surgeon.

© 2017 Daping Hospital and the Research Institute of Surgery of the Third Military Medical University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Lately there is an increasing tendency of using Patient Reported Outcome Measures (PROMs) as a final indicator of the outcome of many surgical treatments in orthopaedics and in other medical specialties. Hip and knee joint replacements, varicose vein and groin hernia surgery are at the moment the procedures which are included to the Department of Health's PROMs programme.1 Currently there are many outcome scores in orthopaedics and most of them are site specific. In the contrary there is a lack of trauma specific outcome scores. The most commonly used PROMs are the generic health status and quality of life EQ5-D and SF-36,2,3 but these scores focus in the patient's general health and quality of life and they are not specific for trauma. The Glasgow Outcome Scale (GOS) is another commonly used score for measuring the rehabilitation status after head injuries.4 This fact leads the trauma surgeons to use a score which was initially designed for measuring the outcome of elective operations such as the Oxford scores, the Manchester/Oxford foot score, the DASH and the Constant shoulder score.5–8

We have designed a new PROM especially for orthopaedic trauma patients, in order to measure in what extent the patients manage to return to their pre-injury state. This score uses as baseline the pre-injury status of the patient and has the aim to determine the percentage of rehabilitation after treatment for any injury. This PROM is not site specific and can be used for every trauma patient. It uses simple wording and design and can be easily and quickly filled by the patients either in the outpatient clinics or by phone conversation. The aim of the present study was to provide a valid, easy to use, outcome questionnaire to the orthopaedic trauma society.

* Corresponding author.
E-mail address: iliopoulosE@gmail.com (E. Iliopoulos).

Peer review under responsibility of Daping Hospital and the Research Institute of Surgery of the Third Military Medical University.

https://doi.org/10.1016/j.cjtee.2017.08.006

1008-1275/© 2017 Daping Hospital and the Research Institute of Surgery of the Third Military Medical University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Materials and methods

Questionnaire design

The design of the questionnaire was agreed by a discussion group where doctors, nurses and patients were involved. A pilot questionnaire was distributed to ten patients and feedback was collected from them. The questionnaire was changed accordingly to its final version and distributed amongst the trauma surgeons of the department for the validation process.

The new outcome score consists of eleven simple questions. The first ten questions are using the 5-point Likert scale and the final question a scale from zero to ten. The questionnaire can be divided into three dimensions (symptoms, function and mental status); each one of them consists of three questions. The patient is asked to reply in what extent he agrees or disagrees with the statements about his present status comparing to his/her pre-injury status (Table 1).

Every question scores from 0 points to maximum 4 points and the final question scores from 0 to 10 points. All the scores are summed (a score 0–50) and doubled in order to form the final score. The final COST score can range from 0 to 100, where 0 is the worst perceived outcome from the patient and the 100 the best (equal to pre-injury level). This scoring system makes it easier for the examiner to visualise the percentage (%) of the recovery of the patient in each individual follow-up appointment. The three dimensions of the questionnaire have their own score as well.

Data collection

The data collection was performed from May 2016 until December 2016. Questionnaires were collected in the outpatient clinics of the hospital by the orthopaedic trauma surgeons of the department. Inclusion criteria were patients who had upper or lower limb trauma and attended the clinics for their final clinical evaluation after treatment (either operative or conservative). Patients who were under ongoing treatment were excluded from the study.

Instruments

The patients were asked to complete the COST and the SF-12v2 questionnaire. SF-12v2 questionnaire was used to evaluate the validity of the new construct. SF-12v2 consists of twelve questions and has two outcome scores, the Physical Component Score (PCS) and the Mental Component Score (MCS). In both dimensions higher scores correspond to better outcomes.5 Middle aged US British and Australian residence have normal scores of 50 in both scales with standard deviation of nine.3,10,11 SF-12 was used to assess the ‘good’ result of the COST questionnaire.

Internal consistency

Internal consistency was evaluated by using the Cronbach’s alpha test for each question and each dimension of the COST questionnaire.12 Cronbach’s alpha coefficient scores between 0.7 and 0.9 indicates high internal consistency, but values higher than 0.9 may represent redundancy of the items.13–15 The total COST score consistency with each COST dimension was also evaluated.

Construct validity

Non-parametric correlation (Spearman’s Rho) was used to evaluate the validity of the new construct. SF-12v2 questionnaire was used for the validation and a strong/moderate correlation was expected for the similar dimensions of the COST questionnaire.16

Statistical analysis

Statistical analysis was performed by using the statistical program SPSS ver.17.0 (SPSS Inc, Chicago, IL, USA). Statistical significance considered to be 0.05.

Results

A total of 60 patients aged 46.81 ± 18.5 years (57% male) took part to the study by completing the COST and the SF-12 questionnaire. The response rate was 95.2% (60/63 questionnaires). The participants were treated in our hospital for various trauma conditions and the data collection was conducted at the end of their treatment (mean 10 months after the injury). Table 2 summarises the descriptive data of the cohort and COST scores for each injury. The mean COST score was 58 ± 18.5 and the mean SF-12 PCS and MCS score was 43.5 ± 9.2 and 53.3 ± 9.6 respectively. Table 3 summarises the answers of each individual item of the COST questionnaire.

Internal consistency

Strong internal consistency was observed for the COST questionnaire as an evaluation tool. Cronbach’s alpha coefficient score

| Table 1 | Questionnaire design: Please complete the following questions about your condition the last two weeks. Fill the answers by ticking the box to the answer which is more close to your current condition. |
|---|---|
| **Symptoms** | Strongly disagree | Disagree | Neither agree nor disagree | Agree | Strongly agree |
| 1 I am pain-free |  |  |  |  |  |
| 2 I do not have stiffness |  |  |  |  |  |
| 3 Swelling settled down completely |  |  |  |  |  |
| **Function** |  |  |  |  |  |
| 4 I have returned to work at the same level as before injury |  |  |  |  |  |
| 5 I run/walk/climb stairs as before |  |  |  |  |  |
| 6 I have returned to sport/leisure activities at the same level as before |  |  |  |  |  |
| **Mental Status** |  |  |  |  |  |
| 7 I can do things as I want to |  |  |  |  |  |
| 8 I feel handicapped |  |  |  |  |  |
| 9 My emotional problems are interfering with my social life |  |  |  |  |  |
| **Overall** |  |  |  |  |  |
| 10 Generally I have managed to return to normal |  |  |  |  |  |
| 11 Please state in what extend you are as before the injury (10 is as you were before the injury) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
for the COST scale was 0.89. The internal consistency was evaluated for each of the three dimensions of the questionnaire, symptoms, function and mental status and was found to be strong as well, with values of 0.74, 0.84 and 0.81 respectively. Each item-total correlation was ranged from 0.87 to 0.89 (Table 3).

Construct validity

The construct validity was tested by using the Spearman’s Rho correlation coefficients. The COST questionnaire has moderate/strong correlation with the SF-12v2 scores. Spearman’s correlations are 0.62 and 0.43 with the physical (PCS) and mental (MCS) subscales of the SF-12v2 respectively. Correlations between physical activity SF-12v2 scale (PCS) and COST’s symptoms and function dimensions are 0.56 and 0.65 respectively. Mental dimension of the COST correlation has a value of 0.45 with the mental component of the SF-12v2 (MCS).

Scoring

The method of scoring the questionnaire is mentioned above. In order to determine the cut off for the ‘good’ result of the scale, the data from the SF-12v2 were used. Patients who had SF-12v2 score over 50 in both physical and mental scales were selected as this is the normal score for the middle aged population.11 Only 16 patients met these criteria. The total COST mean score of these patients was 77.1 with range from 52 to 98. The standard deviation was 12. The ‘good’ result was calculated to be 60 and above (mean score – 1.5 SD).

Discussion

This study demonstrated the development and validation of a new PROM for measuring the outcomes of orthopaedic trauma patients. The main novelty of the questionnaire is to treat trauma as pathology and by using the pre-injury status as a baseline can measure the extent of rehabilitation and return to pre-injury status of the patients. All the questions are simple and non-site specific, which makes it useful to all trauma cases. By using it as an outcome score for different injuries makes it possible to evaluate and compare the results.

Until now the measurement of the outcome after an injury was performed by using surrogate scales, depending on the site of the injury. More specifically the rehabilitation/outcome of injuries around the knee were evaluated with the International Knee Documentation Committee (IKDC) Subjective Knee Form,17 the Oxford knee score18 or the Knee Injury and Osteoarthritis Outcome Score (KOOS)19 questionnaire. For injuries around the foot and ankle, FADI20 or the Foot Functional Index21 scores were used. For hip injuries the Harris Hip score22 and Oxford Hip Score (OHS)23 were widely used. Shoulder and other upper limb pathologies were measured by the quick DASH24 and constant25 scores. Non site specific scores such as SF-36 and SF-1226,27 and EQ-5d scores28 have been used for measuring the impact of the injury in the daily activities. These questionnaires have been developed to measure the general health and activity of the patients and are borrowed by the trauma physicians because of lack of a trauma specific scale. The COST questionnaire is having an answer to that need and can be used for all trauma pathologies because it is not site specific.

As mentioned above COST has a high internal consistency (0.89) and each separate dimension of the questionnaire (symptoms, function and mental status) had high internal consistency as well. Similar values have been reported at the development of other widely used questionnaires. It has been reported for SF-12v2 internal consistency >0.8 for both components.16 DASH questionnaire showed internal consistency of 0.9122 and the KOOS subscales ranged from 0.74 to 0.96.30 International Knee Documentation Committee (IKDC) Subjective Knee Form had 0.87 and 0.88 for symptom and knee articulation and activity level dimensions respectively.31

SF-12v2 was used for the validation of the new scale. SF-12v2 was chosen as it is not site specific and measures the general physical activity and mental health of an individual. It is not developed to measure the outcome of trauma patients but at this situation, it was the most relevant questionnaire for the purpose of validating the COST questionnaire. SF-12v2 has been used to determine the ‘good’ outcome of the COST questionnaire, by correlating the two scales.

There are some potential limitations of the present study. In this study the COST questionnaire was collected only from patients who had simple orthopaedic injuries at their final follow-up appointment. The use of the score in different trauma conditions could illuminate different aspects of the present questionnaire. A new study on poly-trauma patients with potential nonorthopaedic injuries would be interesting. Another limitation of this study was that the scores were collected once for each patient at the final follow-up appointment. This did not allow investigating any possible floor – ceiling effect of the scale.

### Table 2
COST scores of the cohort.

| Gender   | No | %    |
|----------|----|------|
| Male     | 34 | 56.7 |
| Female   | 26 | 43.3 |

| Injury – Fracture | Frequencies (Operative/Conservative management) | COST score (mean ±SD) |
|-------------------|-----------------------------------------------|------------------------|
| Shoulder/Humerus/Elbow | 6 (2/4) | 67.3 ± 23.2 |
| Forearm/Wrist/Hand | 10 (6/4) | 66.6 ± 15.9 |
| Hip/Femur/knee | 18 (15/3) | 54.6 ± 18.3 |
| Tibia/Ankle | 16 (10/6) | 56.5 ± 16.9 |
| Foot | 10 (8/2) | 50 ± 18.5 |
| Total | 60 (41/19) | 57.6 ± 18.5 |

### Table 3
Descriptive statistics and Cronbach’s Alpha for each item of the COST scale.

| Question No | Min | Max | Mean | SD | Alpha |
|-------------|-----|-----|------|----|-------|
| 1           | 0   | 4   | 1.82 | 1.1 | 0.89  |
| 2           | 0   | 4   | 1.75 | 1.13| 0.88  |
| 3           | 0   | 4   | 2.18 | 1.3 | 0.89  |
| 4           | 0   | 4   | 2.02 | 1.19| 0.87  |
| 5           | 0   | 4   | 1.87 | 1.38| 0.88  |
| 6           | 0   | 4   | 1.40 | 1.18| 0.87  |
| 7           | 0   | 4   | 2.42 | 1.09| 0.88  |
| 8           | 1   | 4   | 2.78 | 1.04| 0.88  |
| 9           | 0   | 4   | 3.15 | 1.06| 0.88  |
| 10          | 0   | 4   | 2.55 | 1   | 0.87  |
| 11          | 1   | 10  | 6.92 | 1.73| 0.88  |

* Cronbach’s Alpha if this item is removed.
For further development and use of the COST questionnaire, it is important to be used in larger scale studies. This was not the purpose of this study, as the main aim was to measure the internal consistency and validate the scale. It could be used to monitor the rehabilitation process of the trauma patients by using it on each follow-up appointment with the responsible doctor or the physiotherapist. The fact that it is not site specific can be valuable to compare the different impact that different injuries inflict to patients’ symptoms, function and mental status.

Conclusion

There is a need among the trauma society for a specific PROM of Trauma. COST is a useful and easy to use tool for every orthopaedic trauma surgeon. It is valid with high internal consistency.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.cjtee.2017.08.006.

References

1. DoH GB. Equity and Excellence: Liberating the NHS. White Paper; 2015. Available at: https://www.gov.uk/government/publications/liberating-the-nhs-white-paper.
2. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20:1727–1736.
3. Ware JE, Snow KK, Kosinski M, et al. Development of the MOS Short-Form Health Survey: the MOS 36-Item Short-Form Health Survey (SF-36). Med Care. 1990;28:402–420.
4. Jennett B, Bond M. Assessment of outcome after severe brain damage. Lancet. 1975;1:480–484.
5. Murray DW, Fitzpatrick R, Rogers K, et al. The use of the Oxford hip and knee scores. J Bone Joint Surg Br. 2007;89:1010–1014.
6. Dawson J, Boller I, Doll H, et al. The MOSFQ patient-reported questionnaire: assessment of data quality, reliability and validity in relation to foot and ankle surgery. Foot (Edinb). 2011;21:92–102.
7. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). Am J Ind Med. 1996;26:602–608.
8. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. Clin Orthop Relat Res. 1987 Jan;214:160–164.
9. Ware Jr J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996;34:220–233.
10. Lyons RA, Fielder H, Littlepage BN. Measuring health status with the SF-36: the need for regional norms. J Public Health Med. 1995;17:46–50.
11. Mishra G, Schofield MJ. Norms for the physical and mental health component summary scores of the SF-36 for young, middle-aged and older Australian women. Qual Life Res. 1998;7:215–220.
12. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16:297–334.
13. Boyle GJ. Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? Pers Individ Dif. 1991;12:291–294.
14. Nunnally JC, Bernstein I. Psychometric Theory. Vol. 1. Rdecpicksfarg; 1994:701.
15. Portney LG, Watkins MP. Foundations of Clinical Research: Application to Practice. CONN, Norwalk: Applet Langer; 1993.
16. Cheak-Zamora NC, Wyrwich KW, McBride TD. Reliability and validity of the SF-12v2 in the medical expenditure panel survey. Qual Life Res. 2008;16:727–735.
17. Irsgard J, Anderson AF, Boland AL, et al. Development and validation of the international knee documentation committee subjective knee form. Am J Sports Med. 2001;29:600–613.
18. Dawson J, Fitzpatrick R, Murray D, et al. Questionnaire on the perceptions of patients about total knee replacement. J Bone Joint Surg Br. 1998;80:63–69.
19. Roos EM, Roos HP, Lohmander LS, et al. Knee injury and osteoarthritis outcome score (KOOS)—development of a self-administered outcome measure. J Orthop Sports Phys Ther. 1998;28:88–96.
20. Martin RI, Burdett RG, Irrgang JJ. Development of the foot and ankle disability index (FADI). J Orthop Sport Phys Ther. 1999;29:A32–A33.
21. Budiman-Mak E, Conrad KJ, Roach KE. The Foot Function Index: a measure of foot pain and disability. J Clin Epidemiol. 1991;44:561–570.
22. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. J Bone Joint Surg Am. 1969;51:737–755.
23. Wyke V, Learmonth ID, Cavendish VJ. The Oxford hip score: the patient’s perspective. Health Qual Life Outcomes. 2005;3:66.
24. Gummesson C, Ward MM, Atrosi i. The shortened disabilities of the arm, shoulder and hand questionnaire (Quick DASH): validity and reliability based on responses within the full-length DASH. BMC Musculoskelet Disord. 2006;7:44.
25. Constant CR, Gerber C, Emery RH, et al. A review of the Constant score: modifications and guidelines for its use. J Shoulder Elbow Surg. 2008;17:355–361.
26. Kiely JM, Brazel KJ, Guse CE, et al. Correlation of SF-12 and SF-36 in a trauma population. J Surg Res. 2006;122:214–218.
27. Jenkinson C, Layte R. Development and testing of the UK SF-12 (short form health survey). J Heal Serv Res Policy. 1997;2:14–18.
28. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care. 2005;43:203–220.
29. Kitis A, Celik E, Aslan UB, et al. DASH questionnaire for the analysis of musculoskeletal symptoms in industry workers: a validity and reliability study. Osteoarthr Cartil. 2009;17:1434–1439.
30. Bekkers JE, de Winde TS, Rajmakers NJ, et al. Validation of the knee injury and osteoarthritis outcome score (KOOS) for the treatment of focal cartilage lesions. Osteoarthr Cartil. 2009;17:1392–1398.
31. Higgins LD, Taylor MK, Park D, et al. Reliability and validity of the international knee documentation committee (IKDC) subjective knee form. J Bone Joint Spine. 2007;74:594–599.