75759_Auto_Edited.docx

Prospective Study

Patient reported outcome measures in anterior cruciate ligament rupture and reconstruction: the significance of outcome score prediction

Oday Al-Dadah, Lee Shepstone, Simon T Donell

Abstract

BACKGROUND

Numerous anterior cruciate ligament (ACL) clinical outcome measures exist. However, the result of one score does not equate to the findings of another even when evaluating the same patient group.

AIM

This study aimed to investigate if statistically derived formulae can be used to predict the outcome of one knee scoring system when the result of another is known in patients with ACL rupture before and after reconstruction.

METHODS

Fifty patients with ACL rupture were evaluated using nine clinical

outcome measures. These included Tegner Activity Score, Lysholm Knee Score, Cincinnati Knee Score, International Knee Documentation Committee (IKDC) Objective Knee Score, Tapper and Hoover Meniscal Grading Score, IKDC Subjective Knee Score, Knee Outcome Survey - Activities of Daily Living Scale (KOS-ADLS), Short Form-12 Item Health Survey and Knee Injury and Osteoarthritis Outcome Score. Thirty-four patients underwent an ACL reconstruction and were reassessed post-operatively.

RESULTS

The mean total of each of the nine outcome scores appreciably differed from each other. Significant correlations and regressions were found between most of the outcome scores and were stronger post-operatively. The strongest correlation was found between Cincinnati and KOS-ADLS ($r = 0.91$, $p<0.001$). The strongest regression formula was also found between Cincinnati and KOS-ADLS ($R^2=0.84$, $p<0.001$).

CONCLUSION

The formulae produced from this study can be used to predict the outcome of one knee score when the results of the other are known. These formulae could facilitate the conduct of systematic reviews and meta-analysis in studies relating to ACL injuries by allowing the pooling of substantially more data.

INTRODUCTION

Patient reported outcome measures (PROMs) can be used to determine injury severity and evaluate the effectiveness of treatment. PROMs can quantify the end results of interventions and focus on the patients' experiences, preferences and values. Clinical outcome scores have an important academic and clinical role in all fields of medicine as they are patient centred. PROMs can assess impairment and disability. Impairment is the physiological or anatomical loss or abnormality of structure or function at the organ level (i.e. reduced range of joint movement or increased joint translation). Disability is the functional limitation consequent to impairment which restricts the ability to perform certain activities (i.e. walking, running, participating in sport). Handicap is the physical disadvantage incurred in the context of the individual as a result of impairment and disability[1]. An anterior cruciate ligament (ACL) rupture can give rise to excessive knee joint laxity (impairment) which can result in difficulty with fast cutting actions (disability) and so can be a handicap for a professional athlete but not necessarily for a sedentary office worker. Outcome measures (also known as

instruments) often take the form of questionnaires which include a standardised set of questions and response choices which yield data that are amenable to further statistical analysis. Each questionnaire is comprised of a series of items. Each item represents a single question or statement along with its standardised set of responses. The final scores in many instruments are usually calculated by summing the answers to each of the individual question items. The total scores in some outcome measures can be graded and expressed as excellent, good, fair or poor.

PROMs can be broadly categorised into generic, disease-specific, clinician-completed and patient-completed instruments. The use of these instruments in clinical research allows the patients' perspective to be taken into consideration when investigating a disease process or evaluating the results of an intervention. Although traditionally end-points such as plain radiographs, measured ligament laxity and clinical findings have been used as the primary outcome measures, an increasing emphasis on the use of health-related quality of life instruments is emerging in the conduct of clinical trials. This is reflected by the dramatic increase in the number of validated clinical outcome measures reported in the literature today.

Between 1984 and 1997 over 200 articles were published relating to ACL injuries according to a review article[2]. There were 54 distinctly different outcome measures identified that were specifically designed for assessing ACL injuries. This indicates that there is no single agreed 'gold standard' PROM relating to ACL outcome research. O'Donoghue[3] described the first outcome score used to assess the results of ACL surgery in 1955. This was a clinician-completed rating scale which included an objective examination and a 100-point questionnaire completed by the interviewer. In order to evaluate patients with ACL injuries, many more individual clinical outcome measures have been created.

As a result, numerous ACL knee scoring systems exist in the literature. However, the result of one outcome measure does not equate to the findings of another even when evaluating the same patient group. In a prospective study, Bollen et al[4] assessed a group of patients with ACL injuries and found that the subjects scored consistently

higher on the Lysholm knee score than on the Cincinnati score. Other authors[5,6] have found similar discrepancies among comparisons of various other validated knee outcome scores.

We conducted a prospective longitudinal study analysing PROM data in patients with ACL rupture before and after reconstructive knee surgery. The primary aim of this study was to assess the statistical correlation between all the clinical outcome scoring systems. The secondary aim of this study was to investigate if statistically derived formulae from regression analysis can be used to predict the outcome of one knee scoring system when the result of another is known.

MATERIALS AND METHODS

Full approval was received for the study from the Research Ethics Committee and the Research Governance Committee. All subjects signed informed consent forms to participate. This therapeutic study is a prospective longitudinal cohort study and formed part of the first author's Doctorate thesis. Some data points in this study also served as data in the therapeutic arm of another case-control study submitted for publication.

A total of 50 subjects were recruited to the study. Their demographics are detailed in Table I. The mean time from injury to clinic review was 63 wk (SD=59). An ACL rupture was diagnosed by clinical history and examination and MRI scan of the injured knee for all patients. The diagnosis was confirmed at the time of knee arthroscopy. Clinical history and examination confirmed a normal contra-lateral knee. The flow of patients through the study is illustrated in Figure 1. Four patients with delayed surgical intervention postponed their operation for personal reasons (i.e. work or university commitments). Of the 34 patients who underwent ACL reconstruction, 25 had an ipsilateral middle third bone-patella tendon-bone autograft and nine had an ipsilateral quadrupled hamstring autograft. At the time of surgery 11 patients were found to have a concomitant medial meniscal tear, eight patients had a lateral meniscal

tear and 11 patients had both a medial and a lateral meniscal tear. The mean time to follow-up was 14 wk (SD=4) following surgery.

Inclusion criteria were subjects 16 to 45 years of age. Exclusion criteria included patients with a concomitant posterior cruciate ligament (PCL), medial collateral ligament (MCL) or lateral collateral ligament (LCL) tear of the knee, significant history of ankle or hip pathology, lumbar spine symptoms (including radiculopathy in either limb), neurological or vestibular disease, diabetes or regular use of opiate analgesics.

A total of nine clinical outcome measures were used in this study. Five were clinician- completed instruments and four were patient-completed instruments. These instruments were chosen because they are the most commonly used in the literature with the exception of the Tapper and Hoover Grading Score which was included as it is the only outcome measure specifically developed to assess meniscal injuries. All of the above clinical outcome measures have been validated for use in assessing patients with knee injuries. The clinician-completed knee scores were undertaken at the time of the subjects' attendance at the research clinic. The patient-completed knee scores were mailed to the subjects approximately 7 days prior to their attendance at the research clinic. Therefore, the participants completed these outcome measures in their own time and provided a completely uninfluenced evaluation and perception of their functional knee impairment. All subjects were assessed with these outcome measures at baseline (pre-operatively) and reassessed post-operatively (for the subjects who were followed-up after surgery).

The patient reported outcome measures investigated in this study included:

Tegner Activity Score[7]

Lysholm Knee Score[7]

Cincinnati Knee Score[8-10]

International Knee Documentation Committee (IKDC) Examination Score[11,12]

Tapper and Hoover Meniscal Grading Score[13] (T&H)

International Knee Documentation Committee (IKDC) Subjective Knee Score[14,15]

Knee Outcome Survey - Activities of Daily Living Scale[16] (KOS-ADLS)

Short Form - 12 Item Health Survey[17] (SF-12)

Knee Injury and Osteoarthritis Outcome Score[18,19] (KOOS)

Statistical Analysis

A *post*-hoc power calculation for this study was derived from the results of the longitudinal within-group data of the Lysholm score as detailed in Table 2. The sample size of 34 subjects based on a conventional type I error of 5% with a within-group mean difference of 13.6 and a within-group standard deviation of 12.8 yielded a statistical power calculation of 99.1% for this study. All continuous data variables displayed a normal distribution as verified by both plotted histograms and the Shapiro-Wilks test. The results were evaluated using the Pearson product moment correlation test and the linear and multiple linear regression tests to analyse the continuous variables. The results of both the IKDC Examination score and the T&H score were categorical ordinal variables and the appropriate non-parametric statistical test (Spearman rank-order correlation test) was used for their analysis. The level of statistical significance was set at p< 0.05. Statistical analysis was performed using SPSS for Windows version 25.0 (IBM Corp., Armonk, New York). The power calculation was performed using Minitab statistical software version 19 (Minitab LLC, State College, PA).

RESULTS

The mean and mode averages for each of the clinical outcome measures (continuous and categorical variables respectively) are displayed in Table II.

Table III presents the results of the correlation analysis between each of the knee outcome scores (continuous variables) *pre*-operatively. In general, a significant correlation was found between most of the knee outcome scores with the strongest correlation being between the Lysholm and the Cincinnati scores. The SF-12 MCS score was found to be the weakest correlate variable overall.

Table IV presents the results of the correlation analysis between each of the knee outcome scores (categorical with continuous variables) *pre*-operatively. The Tapper and

Hoover Meniscal Grading score was found to have a significant correlation with all of the knee outcome scores except for the SF-12 MCS score. The IKDC examination score had a poorer correlation with all the knee outcome scores compared to the Tapper and Hoover Meniscal Grading score. There was also no correlation found between the IKDC examination score and the Tapper and Hoover Meniscal Grading score.

Table V presents the results of the correlation analysis between each of the knee outcome scores (continuous variables) *post*-operatively. Overall a significant correlation was found between most of the knee outcome scores with the strongest correlation being between the Cincinnati and the KOS-ADLS scores. It is also evident that in general, the *post*-operative correlations are stronger in comparison to the *pre*-operative results.

Table VI presents the results of the correlation analysis between each of the knee outcome scores (categorical with continuous variables) *post*-operatively. The Tapper and Hoover Meniscal Grading score was found to have a significant correlation with all of the knee outcome scores and had a stronger correlation with each knee score compared to the IKDC examination score. There was also no correlation found between the IKDC examination score and the Tapper and Hoover Meniscal Grading score.

Figure 2 displays the results of the linear regression analysis between the knee outcome measures (continuous variables) *pre*-operatively which produce one overall outcome result. The stated formulae can be used to predict the outcome of a knee score when the result of the other is known. The Lysholm vs. Cincinnati knee score comparison yielded the strongest regression coefficient ($R^2=0.68$). The Tegner score was found to be the weakest regression variable overall ($R^2<0.3$).

Table VII shows the results of the multiple linear regression analysis between the knee outcome measures (continuous variables) *pre*-operatively which produce two or more outcome results (i.e. SF-12 and KOOS scores). The stated formulae can be used to predict the outcome of a knee score when the results of the other variables are known. The KOS-ADLS vs. KOOS knee score comparison yielded the strongest regression

coefficient (R^2=0.74). The Tegner score was found to be the weakest comparator overall (R^2<0.3).

Figure 3 displays the results of the linear regression analysis between the knee outcome measures (continuous variables) *post*-operatively which produce one overall outcome result. The outcome of one knee score can be predicted by the formulae when the result of the other is known. The Cincinnati vs. KOS-ADLS knee score comparison yielded the strongest regression coefficient (R^2=0.84). The Tegner score was found to be the weakest comparator overall (R^2<0.3). It is also evident that in general, the *post*-operative regression analyses are stronger in comparison to the *pre*-operative results.

Table VIII shows the results of the multiple linear regression analysis between the knee outcome measures (continuous variables) *post*-operatively which produce two or more outcome results (i.e. SF-12 and KOOS scores). The outcome of one knee score can be predicted by the formulae when the results of the other variables are known. The KOS-ADLS vs. KOOS knee score comparison yielded the strongest regression coefficient (R^2=0.87). The Tegner score was again found to be the weakest comparator overall. It is apparent that the *post*-operative regression analyses are stronger in comparison to the *pre*-operative results.

DISCUSSION

Significant correlations were found between most of the clinical outcome scores before and after surgery. The strength of the correlations was higher post-operatively. Further statistical analysis produced formulae which allowed the outcome of one knee score to be calculated based on the results of the other outcome measures used in this study in patients with ACL ruptures.

A longitudinal approach was undertaken to test the hypotheses in this study. This allowed the correlation results pre-operatively to be compared directly to that of the post-operative findings of the same individuals. The strength of the correlations was found to be greater following ACL reconstruction. This may be due to a more uniform comparison from time of surgery to clinic assessment post-operatively as compared to

the greater diversity with regards to time of injury to clinic review pre-operatively of the ACL patients. Most of the patients with an ACL rupture had chronic injuries however some subjects had relatively acute ruptures which may have had a bearing on the results of the outcome measures prior to surgery. This could explain the slightly lower correlation between the pre-operative knee scores as compared to the post-operative results. The strongest correlation was found between the Lysholm and the Cincinnati knee scores ($r = 0.83$) pre-operatively. The weakest overall comparator before surgery was the SF-12, in particular the MCS sub-score. This may be explained by the fact that the SF-12 is a generic outcome measure while all the other eight instruments are disease-specific to knee pathology. Post-operatively the strongest correlation was found between the Cincinnati and the KOS-ADLS scores ($r = 0.91$). The IKDC objective examination score is a more elaborate and detailed outcome measure than the Tapper and Hoover Meniscal grading system which was originally designed to assess meniscal tears. However, the latter outcome measure was found to have a stronger correlation with all the other knee scores than the IKDC objective score before and after surgery.

Linear and multiple linear regression analyses were used to generate predictive formulae which allowed the outcome of one knee score to be calculated based on the result of another instrument. These formulae could facilitate the conduct of systematic reviews and meta-analysis in studies relating to ACL injuries by allowing the pooling of data of the results of different knee scoring systems. Similar to the correlation analyses, the results of the regression analyses were stronger post-operatively as compared to the pre-operative findings. The main weakness of this component of the study was the regression analysis results pertaining to the Tegner activity score which was consistently found to be the weakest variable pre-operatively ($R^2 < 0.3$) and post-operatively. A small regression coefficient (i.e. value near to 0) implies that the explanatory variable X (i.e. Tegner activity score) can only account for and explain a small proportion of the total variation of the response variable Y (i.e. Lysholm score, $R^2 = 0.11$) when the results are fitted into the regression equation ($Y = a + bX$ were a = intercept and b = slope).

There are many clinical outcome measures available which can be used in association with ACL injuries. Bollen et al[4] compared the results obtained from the Lysholm and Cincinnati knee scores in patients with ACL deficient knees and found that the latter scale consistently produced lower scores for each patient as compared to that of the Lysholm knee score. This was also noted in the present study both in the pre-operative and post-operative results. They used regression techniques to produce a "rate of exchange" which yielded the formula Lysholm = 30 + (0.72 x Cincinnati). Risberg et al[20] evaluated the Lysholm and Cincinnati scores in patients with ACL ruptures and also found a similar result from their regression analysis yielding the formula Lysholm = 27.8 + (0.73 x Cincinnati). The pre-operative results of the present study are in keeping with these findings as the regression analysis produced the formula Lysholm = 25.4 + (0.72 x Cincinnati) with $R^2 = 0.68$. However, the post-operative regression analysis yielded the formula Lysholm = 8.1 + (1.02 x Cincinnati) with a stronger regression coefficient ($R^2 = 0.80$). Both the Lysholm and the Cincinnati knee scores produce results that are continuous variables. They can be converted into overall categorical ratings (i.e. excellent, good, fair or poor). However, in the present study the results were kept in their original raw continuous data format in order to facilitate the linear regression analysis. Sgaglione et al[21] evaluated knee scoring instruments in patients who underwent ACL reconstruction. These included the Lysholm, Tegner and Cincinnati scores. They also found the scores obtained from the latter outcome measure were lower than the results obtained from the Lysholm scores for each individual patient. Furthermore, they found the results of the knee scores were inflated when the raw scores were converted to categorical ratings. They found that in general, the use of clinical outcome measures can lead to higher scores in patients with low activity levels as compared to subjects who are more active and place higher demands on their knee and so consequently experience greater symptoms. However, this can be accounted for by the inclusion of the Tegner activity score which takes into consideration the activity level of the subject.

In general, there are a number of factors which can influence the end result for each PROM score as reported by the individual patient themselves. These include the patient's age, gender, level of athletic commitment, type of sport as the intricacies of many sports are different, chronicity of condition, type of surgery, patients that had opted out of surgery, ease of return-to-sport and level of return-to-sport all may affect the questionnaire scores.

As yet there is no single outcome measure that is universally considered as the solitary gold standard and therefore many studies use a combination of instruments when evaluating the results of their intervention. As different knee scoring systems yield different results, it is consequently difficult to analyse and compare the relative success of different interventions. This limitation is magnified when attempting to conduct a meta-analysis on a particular topic relating to ACL injury or surgery as the use of different outcome measures in each study limit the capacity to which the outcome data can be pooled together and statistically analysed. In the absence of a single uniform method of evaluation, the formulae produced from the present study relating to the more commonly used knee scoring systems can therefore allow for a more direct interpretation of different clinical outcome measures. This can therefore facilitate the conduct of meta-analyses and the comparison of efficacy of different interventions.

CONCLUSION

Significant correlations were found between most of the clinical outcome measures used in this study with the strength of the correlations being greater post-operatively. Statistically derived formulae produced from this study can be used to predict the outcome of one knee score when the results of the other are known. These formulae could facilitate the conduct of systematic reviews and meta-analysis in studies relating to ACL injuries by allowing the pooling of substantially more data of the most commonly used knee outcome scores.

ARTICLE HIGHLIGHTS

Research background

Many different types of clinical outcome scores exist regarding the anterior cruciate ligament (ACL).

Research motivation

To evaluate how the commonly used patient reported outcome scores (PROMs) differ from each other in the context of ACL injuries.

Research objectives

To develop mathematical formulae which will allow the results of one score to be calculated from the results of the other.

Research methods

PROM data was collected from patients before and after ACL reconstruction surgery and statistically analyzed using correlation and regression tests.

Research results

Statistically significant results for both the correlation and regression analyses were found between most of the outcome scores and were generally stronger following surgery.

Research conclusions

The mathematical formulae produced from this study can be used to predict the outcome of one knee score when the results of the other are known.

Research perspectives

These mathematical formulae can facilitate the conduct of systematic reviews and meta-analysis in studies relating to ACL surgery by allowing the pooling of substantially more data.

75759_Auto_Edited.docx

ORIGINALITY REPORT

7%

SIMILARITY INDEX

PRIMARY SOURCES

1   Oday Al-Dadah, Lee Shepstone, Simon T. Donell. "Clinical outcome measures in anterior cruciate ligament reconstruction: Clinician vs patient completed knee scores", The Surgeon, 2021
    Crossref
                                                                   151 words — 4%

2   Neng‑Ping Gao, Oday Al‑Dadah. "Comparison of day‑case versus inpatient uni‑compartmental knee replacement", Musculoskeletal Care, 2022
    Crossref
                                                                   104 words — 3%

3   Martine T. Dolan, Natalie A. Lowenstein, Jamie E. Collins, Elizabeth G. Matzkin. "Majority of patients find sleep patterns return to normal 6 months following rotator cuff repair", Journal of Shoulder and Elbow Surgery, 2022
    Crossref
                                                                   13 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 12 WORDS