The influence of age and unreported symptoms on the Oxford Shoulder Score

John Dabis, Henry B. Colaço, Helen Hingston,
Magnus Arnander, Duncan Tennent and Eyiemi O. Pearse

Abstract

Background: One potential limitation of interpreting the Oxford Shoulder Score (OSS) in longitudinal studies is that the observed score may be influenced by age and other variables, which may change over time. The purpose of the present study was to investigate the influence of increasing age and unreported non-shoulder upper limb and neck symptoms on the OSS.

Methods: We collected OSS data from a sample of our ethnically diverse local population. All subjects indicated whether they suffered from any neck, shoulder, elbow or wrist symptoms for which they had not sought a medical opinion. Those reporting no symptoms formed the asymptomatic group.

Results: We found a significant decline in OSS with increasing age in the whole study population, as well as in both the asymptomatic and symptomatic groups with previously unreported symptoms: Spearman correlation coefficient = -0.27 , -0.28 and -0.33 respectively ($p < 0.001$). The median OSS in the asymptomatic group was 48 [interquartile range (IQR) 48 to 48]. This was significantly higher than the symptomatic group, with a median OSS of 46 (IQR 40 to 47) ($p < 0.001$).

Conclusions: We found the OSS to be affected by non-shoulder upper-limb and neck pathology as well as age. Within the limitations of the OSS, the differences we found do not exceed the minimal important change.

Keywords

age, influence, noise, non-shoulder upper limb, Oxford Shoulder Score

Date received: 17th March 2017; revised: 23rd September 2017; accepted: 26th September 2017

Introduction

The Oxford Shoulder Score (OSS) is a widely used validated patient related outcome measure (PROM). The score is calculated from responses to 12 questions using a five-point Likert-type scale. The original maximum score was 12/60, although this was later revised to be more intuitive, with a maximum score of 48/48.^{1,2}

The OSS was developed to detect change and to assess outcomes in the setting of a randomized controlled trial.¹ The process of randomization in a comparative study ensures an equal distribution of confounding variables in each of the study groups and, as such, negates the effect of any potential confounding variables on outcome. This neutralizing effect does not occur in case series and cohort studies in which confounding factors can influence measured outcomes.³ Despite this, the OSS is used increasingly in cohort studies, audits and joint registries for

longitudinal assessment. Since April 2012, in addition to monitoring survivorship and revision rates, the United Kingdom National Joint Registry has used the OSS to record PROMs in patients undergoing shoulder arthroplasty. One potential drawback regarding the use of the OSS in longitudinal studies is that variables changing with time may influence the observed score. To our knowledge, this is an issue that has not been addressed previously.

One variable which, by definition, changes with time is age. The originators did suggest that all outcome scores tend to worsen with advancing age and a

St Georges University Hospital, Blackshaw Road, Tooting, UK

Corresponding author:

John Dabis, St George's University Hospital, Blackshaw Road, Tooting, London SW17 0QT, UK.
Email: johndabis1@gmail.com

'normal score' may be less than 48. The influence of age on the OSS is controversial and previous studies suggest that age may or may not have an influence on the OSS.^{4,5} To neutralize the influence of age, other scoring systems such as the Constant score can be adjusted and normalized for age.^{6,7} There is no such age-adjustment with the OSS and a concern is that any observed decline in the OSS over time may be as much attributable to aging as it is to a decline in the effectiveness of a particular intervention.

Another factor that may influence the OSS and change with time is the presence or absence of non-shoulder upper limb or neck symptoms. Although the Disabilities of the Arm, Shoulder and Hand (DASH)⁸ score is valid for patients with any upper extremity disease or disorder, the presence of concurrent non-shoulder upper limb pathology or 'noise' is not specified for both the DASH and OSS.

In view of the widespread adoption of the OSS for reporting of follow-up data, it is important to determine whether such variables influence the OSS and to quantify any such influence. The present study investigates the influence of age and non-shoulder upper limb and neck symptoms on the OSS in the local population. We hypothesize that increasing age and the presence of nonshoulder upper limb or neck symptoms will have a negative influence on the measured OSS.

Materials and methods

To assess the influence of age and non-shoulder upper limb pathology on the OSS, we aimed to examine whether the OSS in our ethnically diverse local population varied with age or with the presence or absence of nonshoulder upper limb pathology.

We sampled our local population by approaching individuals attending our hospital for conditions other than musculoskeletal problems, relatives and friends accompanying them and non-orthopaedic members of staff. Individuals were over the age of 16 years and were able to understand the written OSS questionnaire in English. With respect to the shoulder in particular, we considered individuals to have 'normal' shoulders if they were not receiving any form of treatment for their shoulder and if they had not sought medical attention because of a shoulder problem. Because there is no accepted definition of 'normal', we used these parameters to define a normal shoulder. All individuals who were currently awaiting assessment or receiving treatment for an upper limb or neck condition were excluded from the study.

We developed a questionnaire that we piloted on 50 individuals. The questionnaire consisted of basic demographic data including sex, age and hand

dominance of each respondent. The second part of the questionnaire recorded the OSS for each shoulder and established whether participants had any unreported upper limb symptoms for which they were not seeking medical attention or receiving treatment. Following an analysis of the responses, we modified the questionnaire to clarify whether individuals reporting upper limb problems had neck, shoulder, elbow, wrist or hand symptoms.

Statistical analysis was performed using SPSS, version 22 (IBM Corp., Armonk, NY, USA). The data were not normally distributed and averages were expressed as medians with interquartile ranges. The Spearman correlation coefficient was used to analyze the relationship between continuous variables. The Mann–Whitney *U*-test was used to analyze unpaired data and the Wilcoxon signed-rank test was used to analyze paired data. The Kruskal–Wallis Test with post-hoc Bonferroni correction was used to compare data when patients were grouped into decade age groups and the effect size was expressed as an eta square value. Categorical variables were analyzed using Fisher's exact test. $p < 0.05$ was considered statistically significant for observed differences.

Results

In total, 422 individuals completed the questionnaire over a 6-week period. Sixty were excluded because they were currently awaiting assessment of or undergoing treatment for an upper limb or neck condition. Of the eligible 362 respondents included in the study population, there were 174 males and 188 females, with a median age of 51 years (range 16 years to 88 years).

In total, 255 of 362 individuals reported no neck or upper limb symptoms in either limb, and the separate scores for both limbs ($n = 510$) were therefore included in the analysis (Group 1: asymptomatic) (Figure 1). The remaining 107 of 362 individuals reported either upper limb or neck symptoms for which they were not seeking medical attention or receiving treatment. The influence of upper limb symptoms on the OSS in the contralateral upper limb has not been determined previously. A Wilcoxon signed-rank test revealed a significant difference between the OSS in symptomatic and the OSS in contralateral limbs in individuals with unilateral symptoms ($p < 0.001$). In addition, neck symptoms could potentially affect one or both upper limbs. Consequently, the limb with the lowest OSS was used in our analysis and formed the symptomatic group ($n = 107$). Of the 107 symptomatic individuals (Group 2: symptomatic), 71 reported shoulder symptoms, 15 reported neck symptoms, 10 reported elbow symptoms and 11 reported hand and wrist symptoms. In total, 617 ($510 + 107$) responses were included in the analysis.

Figure 1. Flow diagram representing patient groups.

Upper limb pathology and the OSS

The median OSS in the asymptomatic group ($n = 510$) was 48 [interquartile range (IQR) 48 to 48]; this was significantly higher than the median OSS of 46 (IQR 40 to 47) in the symptomatic group ($n = 107$, $p < 0.001$; Mann–Whitney U -test). The median OSS in individuals reporting shoulder symptoms ($n = 71$) was 46 (IQR 39 to 47). For other upper limb non-shoulder or neck symptoms ($n = 36$), the OSS was 46 (IQR 41 to 48).

Age and the OSS

Analysis of all included responses ($n = 617$) demonstrated a decline in the OSS with age (Figure 2a); the Spearman correlation coefficient was -0.270 ($p < 0.001$). This negative relationship between age and the OSS was also seen when the asymptomatic and symptomatic groups were analyzed separately (Spearman correlation coefficients of -0.280 and -0.333 , respectively ($p < 0.001$) (Figure 2b, c).

We found significant differences between some decade age groups, although this was not consistent across the entire age range (Figure 3a). Analysis of

the whole study group ($n = 617$) with the Kruskal–Wallis test using the Bonferroni correction demonstrated significant differences only between individuals in the second to fifdecades and individuals in the seventh and eighth decade, with eta square of 0.10 ($p < 0.001$). The findings were similar when the asymptomatic group was analyzed separately and, in the symptomatic group, a significant difference was only found between fourth and eighth decades.

However, when the study group was divided into individuals under 50 years or over 50 years old, there was a clear statistically significant difference in OSS between age groups (Mann–Whitney U -test $p < 0.001$) (Figure 3b). When considering the whole study population of 362 individuals, 83.9% under the age of 50 ($n = 174$) reported the maximum OSS of 48, whereas only 56.4% of individuals over 50 years ($n = 188$) reported the maximum OSS in both limbs; this difference was statistically significant ($p < 0.001$, Fisher’s exact test). When the asymptomatic group of individuals ($n = 255$) was analyzed separately, 100% under the age of 50 ($n = 133$) reported the maximum OSS of 48, whereas only 76.2% of individuals over 50 years ($n = 122$) reported the maximum OSS in both limbs;

Figure 2. (a) Correlation between OSS and Age in the study population (n=617). (b) Correlation between OSS and Age in asymptomatic Group 1 (n=510). (c) Correlation between OSS and Age in symptomatic Group 2 (n=107).

this difference was also statistically significant ($p < 0.001$, Fisher's exact test).

Discussion

The OSS was developed for the assessment of outcomes on shoulder surgery in the context of a randomized trial. It has been reported previously that the OSS, in what is considered to be a local population, is not necessarily the maximum possible score. Using the original version of the OSS (12 to 60 maximum), Younis et al.⁴ found a wide range in their sample of normal individuals with a mean score of 15.5 (range 12 to 47). Similarly, Clement et al.⁵ found a median OSS of 16 (range 12 to 42) in their sample of normal shoulders. Using the revised OSS (0 to 48 maximum), we found a median OSS of 48 (maximum) but with a similar range (range 11 to 48).

Absolute values of the OSS are therefore of questionable reliability and should not be used either to assess

outcome or to determine the threshold for intervention. Randomization minimizes bias and the influence of variables that may impact the OSS. However, the use of the OSS has been extended to cohort and longitudinal studies. We acknowledge that the OSS was developed to detect change and that, in the present study, we have used it for a purpose for which it was not developed (i.e. absolute values). However, we have done so to assess the appropriateness of its use in cohort studies and registries. Although the effect of confounding variables on the OSS in a given individual is unlikely to change when the time interval between the initial assessment and the follow-up assessment is short, the occurrence of such variables, such as the presence of other upper limb pathology, may change over an extended period of time, and this could influence the OSS.

Although the OSS was designed to be joint-specific so that it should be as sensitive to the outcome of shoulder surgery as possible and be influenced as little as

Figure 3. (a). Boxplot of OSS v decade age group (Median, IQR, outliers). (b) Boxplot of OSS v Age category: under 50 & over 50 years (median, IQR, outliers).

possible by other co-morbidities, we accept that it is impossible to totally eradicate the effect of co-morbidities (also termed noise).² We were particularly interested to determine whether nonshoulder upper limb or neck symptoms could influence the OSS

because this has not been established in the literature. Other PROMs such as the DASH score cannot be used in the presence of other upper limb pathology because they are not considered sufficiently specific.⁸ There is no such requirement for exclusion of individuals with

nonshoulder upper limb pathology when using the OSS to assess shoulder function. The prevalence of upper limb symptoms has been reported to be high in the general population; Roquelaure et al.⁹ reported high prevalence rates of upper extremity symptoms in their study population of 2,685 workers. Of these, 31% worked primarily in manufacturing industries, whereas the remaining 69% did predominantly nonphysical jobs. Approximately 58% of women and 53% of men had experienced upper-extremity symptoms during the preceding 12 months, and 35% of women and 27% of men experienced these symptoms during the week before the survey. Symptoms often overlapped between two anatomic regions within the same individual, particularly the neck and shoulder regions. There was a significant change in the number of individuals experiencing upper limb symptoms within 1 year, suggesting that noise from concurrent non-shoulder pathology could affect observed scores, even with studies of a relatively short follow-up interval.

As expected, we found a lower median OSS in individuals reporting shoulder symptoms than we did in the 255 completely asymptomatic individuals: median OSS 48 (IQR 48 to 48) compared with 46 (IQR 39 to 47), respectively. However, we also found a lower median OSS in the 36 individuals who reported nonshoulder upper limb or neck symptoms than we did in the 255 completely asymptomatic individuals: median OSS 48 (IQR 48 to 48) compared to 46 (IQR 41 to 48) respectively. This suggests that other upper limb and neck symptoms may influence the OSS, although we were unable to apply statistical analysis to assess this as a result of disparate group sizes and thus could not determine whether this was a statistically significant finding.

The influence of age on the OSS is uncertain; Clement et al.⁵ found a significant relationship between age and OSS with a worse score with advancing age (Spearman $r=0.62$). They also found widening confidence intervals with advancing age indicating a greater variability of scores with age. Another study by Younis et al.⁴ failed to establish a significant relationship between age and the OSS, although they employed a different methodological approach.

We demonstrated a statistically significant negative correlation between age and OSS, in both the total study population ($n=362$) and the truly asymptomatic individuals (Group 1), with a Spearman's correlation coefficient of -0.27 and -0.28 , respectively. We attempted to quantify the decline; however, there was no statistically significant difference in OSS when adjacent decades were compared with the Kruskal–Wallis test using the Bonferroni correction. This suggests that, although there is a statistically significant correlation between age and the OSS, it is unlikely that the OSS

will change sufficiently between decades to be of clinical significance.

However, we found a clear significant difference between individuals aged under 50 years and those aged over 50 years: 100% of individuals aged under 50 years reported the maximum OSS of 48; however, only 76.2% of individuals over 50 years reported the maximum OSS. There was a statistically significant difference in median scores between the two groups. This finding suggests that the OSS may be more reliable in individuals below the age of 50 years than those over the age of 50 years. Individuals over the age of 50 years are more likely to have other co-morbidities, which may affect the answers to the less joint-specific questions in the OSS, such as 'Could you carry a tray containing a plate of food across a room?' and 'Could you do the household shopping on your own?'

It is worthy of note that Van Kampen et al.¹⁰ have determined that the smallest detectable change not resulting from measurement error in the OSS is 6.0. Within the limitations of the OSS and the interference of age and other 'noise', the differences that we found do not exceed the minimal important change (MIC). This is not widely applied to analysis in most studies. Our findings are in keeping with a significant level of 'noise' when determining the OSS. We suggest that an age-related decline in score should be accounted for by an OSS correction for age and possibly also for non-shoulder upper limb pathology. The OSS can continue to be used for long-term cohort studies and registry data, if the score is interpreted with caution. Otherwise, we need to accept up to a 12.5% error in our measurement of the OSS.

In conclusion, we found the OSS to be affected by nonshoulder upper-limb and neck pathology, as well as age. The effect of both on the OSS is statistically significant and, although unlikely to be of clinical significance in short-term to mid-term longitudinal studies, it may become more significant in longer-term studies, particularly as patient age increases over 50 years. We therefore contend that we either accept a high error rate for the OSS (up to 12.5%), which could render many studies meaningless, or we should adapt the OSS to minimize noise and improve its reliability as an outcome measure.

Acknowledgements

All the authors have been actively involved in the planning and enactment of the present study, and have also assisted with the preparation of the submitted article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethical review and patient consent

No ethical approval was required.

References

1. Dawson J, Fitzpatrick R and Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br* 1996; 78: 593–600.
2. Dawson J, Rogers K, Fitzpatrick R, et al. The Oxford shoulder score revisited. *Arch Orthop Trauma Surg* 2009; 129: 119–123.
3. Rolfson O, Bohm E, Franklin P, et al. Patient-reported outcome measures in arthroplasty registries. *Acta Orthop* 2016; 87(Suppl 1): 9–23.
4. Younis F, Sultan J, Dix S, et al. The range of the Oxford Shoulder Score in the asymptomatic population: a marker for postoperative improvement. *Ann R Coll Surg Engl* 2011; 93: 629–633.
5. Clement ND and Court-Brown CM. Oxford shoulder score in a normal population. *Int J Shoulder Surg* 2014; 8: 10–14.
6. Katolik LI, Romeo AA, Cole BJ, et al. Normalization of the Constant score. *J Shoulder Elbow Surg* 2005; 14: 279–285.
7. Yian EH, Ramappa AJ, Arneberg O, et al. The Constant score in normal shoulders. *J Shoulder Elbow Surg* 2005; 14: 128–133.
8. Hudak PL, Amadio PC and Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand). The Upper Extremity Collaborative Group (UECG). *Am J Ind Med* 1996; 29: 602–608.
9. Roquelaure Y, Ha C, Leclerc A, et al. Epidemiologic surveillance of upper-extremity musculoskeletal disorders in the working population. *Arthritis Rheum* 2006; 55: 765–78.
10. Van Kampen DA, Willems WJ, Van Beers LWAH, et al. Determination and comparison of the smallest detectable change (SDC) and the minimal important change (MIC) of four-shoulder patient-reported outcome measures (PROMs). *J Orthop Surg Res* 2013; 8: 40.