The reproducibility of measurement of shoulder movement

Paul D Triffitt, Clare Wildin and Daniel Hajioff

To study the reproducibility of the measurement of shoulder movement, we have examined a series of 64 patients with and without shoulder problems, measuring active elevation, abduction, and external rotation in adduction using an inclinometer. The difference within which readings by different observers were expected to lie for 95% of the pairs of observations ranged from 24° to 33° for different movements in asymptomatic shoulders and from 24° to 41° in those with unilateral shoulder symptoms awaiting surgery.

University Department of Orthopaedic Surgery, The Glenfield Hospital, Leicester, LE3 9QP, U.K. Tel +44 116–2871471. Fax –2321702
Submitted 98-12-04. Accepted 99-05-11

The accurate measurement of movement of joints is frequently difficult, for example as a result of problems in establishing the axes of the limb bones and those of joint movement (Lea and Gerhardt 1995). Few studies have addressed reproducibility of the measurement of shoulder movement. Passive glenohumeral abduction and external rotation measured by 2 observers in 10 women were found to give correlation coefficients of 0.72 and 0.45, respectively (Clarke et al. 1974). Measurement of active external rotation in abduction by 4 observers in 12 asymptomatic subjects was found to give a small variation, with a standard deviation of under 5° (Boone et al. 1978). Correlation coefficients between 0.26 and 0.89 were found between pairs of examiners for 7 passive ranges of movement in 50 subjects most of whom had shoulder symptoms (Riddle et al. 1987). We analysed reproducibility in terms of the mean within-subject standard deviation, which gives a more useful quantitative measure than a correlation coefficient (Bland and Altman 1996).

Patients and methods

We studied 64 patients (34 men) with a mean age of 54 (17–88) years. 19 patients were scheduled for shoulder surgery, with diagnoses of a cuff tear in 5, glenohumeral osteoarthrosis in 5, capsulitis in 3, instability in 2, and impingement, acromioclavicular osteolysis, glenohumeral rheumatoid arthritis, and synovial chondromatosis in 1 case each. In 6 cases, the contralateral side was symptomatic. 45 patients were scheduled for orthopedic procedures in other areas. Of these patients, 3 gave a history of shoulder symptoms on direct questioning.

Patients were examined consecutively in an orthopedic preadmission clinic. Their dominance, diagnosis, and history of shoulder symptoms were recorded, together with the side affected. Three active ranges of movement were measured in each shoulder: elevation, i.e., the maximum angle between the humerus and the vertical axis of the trunk that could be attained, which is usually in a plane between true flexion and true abduction (Hawkins and Bokor 1990); abduction in a plane approximately 30° in front of the coronal plane; and external rotation with the shoulder adducted and the elbow flexed to 90°, measured according to the angle subtended by the ulna. For the first two movements, the patient was sitting to reduce spinal motion, and for the third, the patient was supine. Measurements were done with an inclinometer and recorded to the nearest 2°. One measurement of each movement was made by each of at least 2 and at most 4 of a panel of 5 examiners, according to availability, each of whom was unaware of the readings of the other examiners. The

panel comprised three junior physicians, one middle-grade surgeon, and a consultant orthopedic surgeon, thus representing a range of experience. In 84% of cases, readings included those by both the middle-grade and senior surgeons.

The data were analysed by calculation of the mean within-subject standard deviation (Bland and Altman 1996). Reproducibility limits were then calculated to give the value less than or equal to which the difference between two test results may be expected to lie with a probability of 95%. This is given by the within-subject standard deviation multiplied by $1.96 \times \sqrt{2}$ (= 2.77).

Results

The mean elevation for all subjects and examiners was 151 (39–179)°, the mean abduction was 146 (25–79)° and the mean external rotation was 45 (–14 to 73)°.

In the group as a whole and in asymptomatic subjects, the difference within which two readings were expected to lie for 95% of the observations varied between 24° and 33° (Table 1). In those scheduled for surgery and without contralateral symptoms, the figures varied between 24° and 41° (Table 2). The variation in symptomatic limbs tended to be greater than that in the asymptomatic limbs.

As the within-subject standard deviation for each movement approximated 10°, the 95% confidence interval for any individual observation was approximately ± 1.96 × 10°, i.e., approximately ± 20°.

Discussion

We found considerable interobserver variation in the measurement of active movement of the shoulder, in both symptomatic and asymptomatic subjects. The reproducibility limits give a more realistic estimate of variation than a correlation coefficient, which is more difficult to interpret and which can remain high despite considerable variation. Such misleading results in this application have been noted previously (Bland and Altman 1986).

Table 1. Reproducibility limits (degrees), or the difference within which two readings are expected to lie for 95% of the pairs of observations

	All patients (n 64)	NSS (n 42)	RDom (n 38)
Elevation R	27	24	24
Elevation L	29	31	31
Abduction R	32	33	34
Abduction L	32	32	33
External rotation R	31	26	26
External rotation L	25	27	26

NSS neither shoulder symptomatic, RDom right dominant patients with neither shoulder symptomatic, R right, L left.

Table 2. Reproducibility limits (degrees) for shoulder surgical patients without contralateral symptoms (n 13)

	Limit
Elevation A	36
Elevation N	27
Abduction A	34
Abduction N	29
External rotation A	41
External rotation N	24

A affected side, N unaffected side.

There was a trend for more variation in shoulders listed for surgery. Limitation of movement by pain may be expected to vary, and results of studies of asymptomatic subjects should be interpreted cautiously in relation to clinical practice.

The variation we found probably had several sources. Use of an inclinometer rather than a goniometer largely solved the problem of establishing the axis of rotation of the joint, but problems of establishing the longitudinal axis of the humerus and of stopping spinal movement remained. The amount of effort undertaken by each patient probably also varied. Interobserver variation could have been reduced by each observer taking more than one measurement and then averaging the results, as has been done in some previous studies. In most clinical circumstances, however, this is not practicable, and we consider that the current method is appropriate for clinical use. We suggest that the movements investigated in this study can be recorded to the nearest 10° without loss of accuracy.

Bland J M, Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; i: 307-10.

Bland J M, Altman D G. Measurement error. Br Med J 1996; 312: 1654.

Boone D C, Azen S P, Lin C-M, Spence C, Baron C, Lee L. Reliability of goniometric measurements. Phys Ther 1978; 58: 1355-60.

Clarke G R, Willis L A, Fish W W, Nichols P J R. Assessment of movement at the gleno-humeral joint. Orthopaedics (Oxford) 1974; 7: 55-71.

Hawkins R J, Bokor D J. Clinical evaluation of shoulder problems. In: The Shoulder (Eds. Rockwood C R Jr and Matsen F A III). W.B. Saunders Co., Philadelphia 1990; 4: 160.

Lea R D, Gerhardt J J. Range-of-motion measurements. J Bone Joint Surg (Am) 1995; 77: 784-98.

Riddle D L, Rothstein J M, Lamb R L. Goniometric reliability in a clinical setting. Shoulder measurements. Phys Ther 1987; 67: 668-73.