

Outcome evaluation in shoulder surgery using 3D kinematics sensors

Brian Coley^{a,*}, Brigitte M. Jolles^b, Alain Farron^b, Aline Bourgeois^b,
François Nussbaumer^b, Claude Pichonnaz^b, Kamiar Aminian^a

^aLaboratory of Movement Analysis and Measurement, Ecole Polytechnique Fédérale de Lausanne (EPFL),
STI-Bat. ELG, CH-1015 Lausanne, Switzerland

^bHôpital Orthopédique de la Suisse Romande, CHUV, University of Lausanne, Switzerland

Received 13 October 2005; received in revised form 1 June 2006; accepted 12 June 2006

Abstract

A new method of scoring systems for the functional assessment of the shoulder is presented. 3D accelerometers and gyroscopes attached on the humerus were used to differentiate a healthy from a painful shoulder. The method was first tested on 10 healthy volunteer subjects with no shoulder pathology. The system was then tested on 10 patients with unilateral shoulder pathology (rotator cuff disease, osteoarthritis) before and after surgery (3, 6 months). In order to evaluate the system, nine tests based on the Simple Shoulder Test (SST) were performed on each shoulder for each patient. Three scores were defined: the P score was based on the angular velocities and accelerations of the humerus; the RAV score was based only on the angular velocities of the humerus; the M score was based on the sum of all moments of the humerus. Our kinematic scores indicated significant differences between baseline and follow-up ($p < 0.05$) and differentiated between patients with varying severity of the same condition. We demonstrated a reliable technique of evaluating shoulder pathology and the results of surgery. © 2006 Elsevier B.V. All rights reserved.

Keywords: Outcome evaluation; Accelerometers and gyroscopes; Clinical scores; Ambulatory system

1. Introduction

The importance of monitoring the results of medical procedures has long been recognized in surgery and particularly in orthopaedic surgery. Outcome assessment has been given new impetus during the past decade as the emphasis has shifted from the era of expansion and technical development to one of assessment and accountability. There are over 20 different assessment methods for the functional outcome of shoulder procedures [1]. Some of these (such as the Disabilities of the Arm and Shoulder score (DASH) [2] and the Simple Shoulder Test score (SST) [3]) are widely used, although none has been universally accepted. These instruments assign a score to the patient using questionnaires based on separate domains: pain, function and overall satisfaction. Albeit validated, these instruments give only

subjective scores and therefore give an incomplete answer on patient's shoulder evaluation.

Objective assessments like radiographs [4,5] provide a static image of the shoulder but do not measure its dynamic function. Although laboratory measurements (video-based motion analysis) [6–8] provide complete 3D kinematics of the shoulder, they require a dedicated laboratory and assume that data measured for a short period are representative of usual performance. This constraint, together with the cost of this technology and the time needed for the analysis has restricted its use in clinical practice. In this study we propose a different approach: measuring 3D kinematics from body fixed sensors using an ambulatory recording device. The goal of this study was two-fold: to find objective parameters (scores) for the assessment of shoulder function based on body fixed kinematics sensors and to evaluate the effectiveness of these parameters in quantifying the kinematic differences between healthy and affected shoulders. By validating such approach, we would provide the clinician with a system to assess shoulder function using an objective score.

* Corresponding author. Tel.: +41 21 693 47 73; fax: +41 21 693 69 15.
E-mail address: brian.coley@epfl.ch (B. Coley).

Table 1
Summary of the nine tests carried out for painful and healthy shoulders

Tests	Description
1	Rest position
2	Hand to the back
3	Hand behind the head
4	Object ahead
5	4 kg in abduction
6	8 kg along the body
7	Hand to the opposite shoulder
8	Change a bulb
9	Object on side (Elbow in 90°, ext./int. rotation)

The subject is in standing position.

2. Methods

2.1. Subjects and materials

Ten healthy subjects (25.1 years old \pm 4.1) and 10 patients with unilateral shoulder pathology (7 rotator cuff disease (7 rotator cuff repair)/3 osteoarthritis (3 prosthetic shoulder arthroplasty); 4 women, 6 men; 62.4 years old \pm 10.4) were studied. Nine tests representing movements during daily activity based on the Simple Shoulder Test were carried out for both shoulders (Table 1) before surgery, 3 and 6 months after surgery. These tests were also carried out twice with 1 year interval on the same healthy subjects. Each test lasted 20 s and was video filmed for further validation of the movements and assessment of any falsely recorded movements.

In this study, one module with three miniature capacitive gyroscopes (*Analog device, ADXRS 250, \pm 400°/s*) and three miniature accelerometers (*Analog device, ADXL 210, \pm 5 g*) were fixed by a patch on the humerus (Fig. 1). This way, the sensors measured the anterior elevation–extension, abduction–adduction and internal–external rotation of the

shoulder. The signal from the sensors was amplified and low-pass filtered (cut-off frequency: 17 Hz) to remove noise [12,13]. The sensors and their conditioning electronics were packaged in a very small box (25 mm \times 25 mm \times 13 mm). All signals were digitized at 200Hz sampling rate and recorded by the data logger (Physilog[®], BioAGM, CH) carried on the subject's waist.

The Simple Shoulder Test and the Disabilities of the Arm and Shoulder score were filled out by each subject [1] to estimate the validity of our method. The SST consists of 12 questions with “yes or no” answer. DASH is a 30-item questionnaire designed to evaluate upper extremity-related symptoms and to measure functional status at the level of disability. The SST and DASH scores are both validated scores and patient-reported outcomes measures. There are over nine different validation studies for the DASH score [17,18] and three validation studies for the SST [19].

2.2. Angles estimation

Internal and external rotational movements (roll), extension and anterior elevation movements (pitch) and abduction and adduction movements (yaw) were estimated from 3D accelerometers and 3D gyroscopes. The accelerometers measure the gravity component, and using this feature, it is possible to measure the segment orientation when it is motionless [9]. Drift and DC components of the angular velocities were removed using wavelet transformation and considering the initial and final orientation of the segment based on the acceleration signals. The 3D angles were obtained after integration of the three angular velocities. Fig. 2 shows the flow chart of the 3D angles estimation.

As reference system, a Zebris CMS-HS ultrasound-based motion measurement system was used [14]. This system

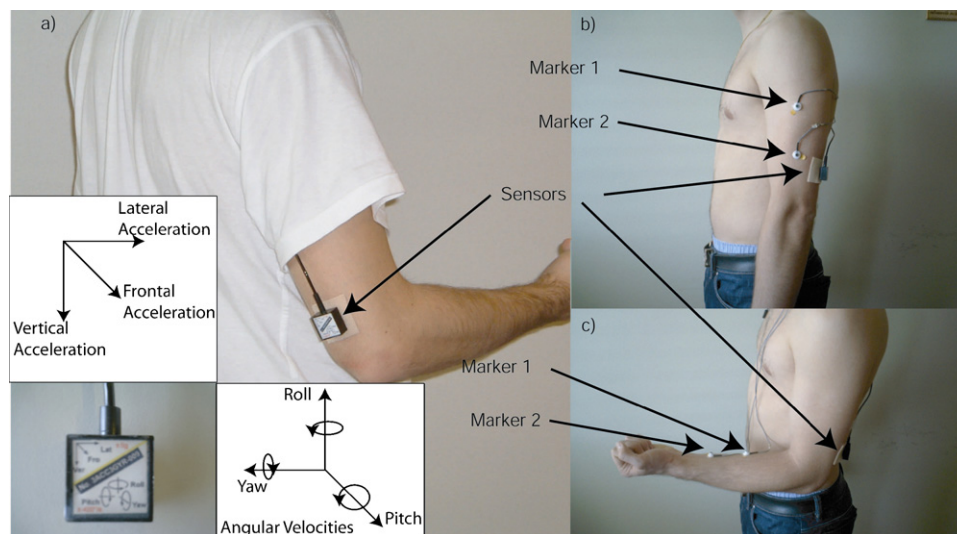


Fig. 1. (a) Position of the kinematic sensors module including 3D gyroscope and 3D accelerometer. (b) Position of the reference markers for abduction/adduction (yaw), flexion/elevation (pitch) rotation. (c) Position of the reference markers for internal and external rotation (roll). The reference markers from the reference system were used for assessing our kinematic system.

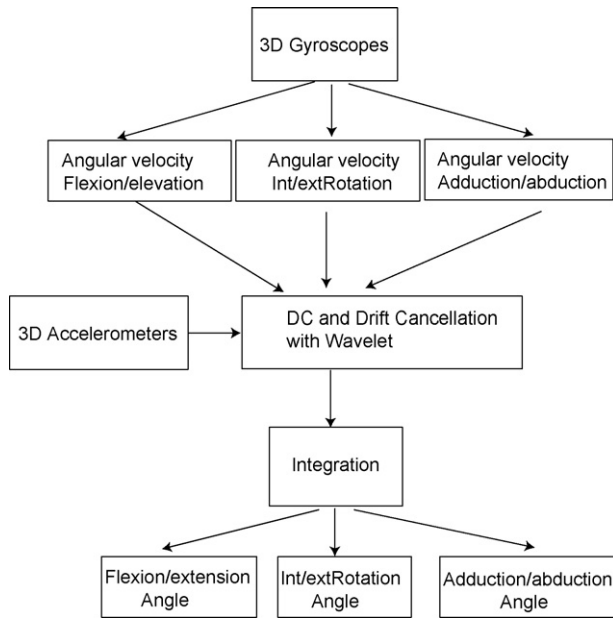


Fig. 2. Flow chart for the angle estimation. Angles were estimated from the integral of angular velocity and by considering initial and final orientation from the accelerometers.

consists of three fixed sonic emitters which send out a burst of ultrasound, and receivers placed on body segments. The time taken for the burst to reach each receiver is recorded. Using this delay, the distances between the receiver and each emitter can be calculated from the sound velocity. Knowing the distance from three emitters, the coordinates of the receiver placed on body segment can be computed by triangulation with an absolute accuracy better than 1.0 mm [15,16] with a sampling rate of 100 Hz. In this study, two ultrasound receivers were attached over the same segment (humerus) (marker 1, marker 2). Spatial marker positions (x , y , z) were recorded and used for calculation of humerus orientation angles. Synchronization between the reference and the Physilog systems was performed by electrical trigger. The angle data obtained by the body-fixed sensors were down sampled to 100 Hz for comparison purpose. The flexion/extension and abduction/adduction angles of the humerus were estimated using the spatial coordinates of the microphone markers on the humerus (Fig. 1b). The internal/external rotation angles of the humerus were estimated using the spatial coordinates of the microphone markers on the radius (Fig. 1c). Basic movements, such as anterior flexion–extension, abduction, adduction and internal/external rotation, were performed with our system and the reference system on 10 healthy subjects to assess the accuracy of our angle estimation method.

2.3. RAV score algorithm

Our second method was to estimate the difference of kinematics between the healthy and painful shoulders. It was based only on the angular velocities of the humerus. The 3D range of angular velocity (RAV) was calculated by the

difference between the maximum and the minimum of angular velocity ($^{\circ}/s$) measured by 3D gyroscopes during each test in internal and external rotational (roll), flexion/extension (pitch) and abduction/adduction (yaw) directions for each subject. The RAVr parameter was estimated as the average of the sum of the RAV in the three axis of rotation.

$$RAVr = \frac{\sum_{\text{roll,pitch,yaw}} \text{range}(\text{angular velocity})}{3} \quad (1)$$

The difference between a healthy and a painful shoulder ($\Delta RAVr$) was expressed as the percentage of RAV of the healthy shoulder ($\Delta RAVr$).

$$\Delta RAVr = \frac{RAV_{\text{healthy}} - RAV_{\text{painful}}}{RAV_{\text{healthy}}} \quad (2)$$

The RAV score is defined as the average of the $\Delta RAVr$ over all nine tests.

$$RAV \text{ score} = 1 - \text{mean} \left[\sum_{\text{Test}=1}^9 \Delta RAVr \right] \times 100 [\%] \quad (3)$$

2.4. P score algorithm

The main idea was to observe the relationship between humerus acceleration and angular velocities. Fig. 3 shows the difference between the healthy and the painful side for one axis and a patient. In order to estimate the difference between both sides, we calculated for each test the surface inside the curve for both sides. The simplest estimation of this surface was to calculate the area of the rectangle, which circumscribes the curve corresponding to the product of the acceleration range by the angular velocity range (Fig. 3).

$$Pr = \sum_{\text{roll,pitch,yaw}} \text{range}(\text{acceleration}) \cdot \text{range}(\text{angular velocity}) \quad (4)$$

We calculated this surface for each axis for both sides and added these to obtain a parameter called Pr for a healthy and a painful side. By considering that the product of angular velocity and acceleration is related to power of movement, we can therefore assume that P is a power dependent quantity. This parameter can also be considered as the control of the humerus velocity by its acceleration.

The difference between the Pr parameter of a healthy and a painful side relative by the healthy side was considered as ΔPr parameter.

$$\Delta Pr = \frac{P_{\text{healthy}} - P_{\text{painful}}}{P_{\text{healthy}}} \quad (5)$$

The first score is defined as the average of the ΔPr over all nine tests.

$$P \text{ score} = 1 - \text{mean} \left[\sum_{\text{Test}=1}^9 \Delta Pr \right] \times 100 [\%] \quad (6)$$

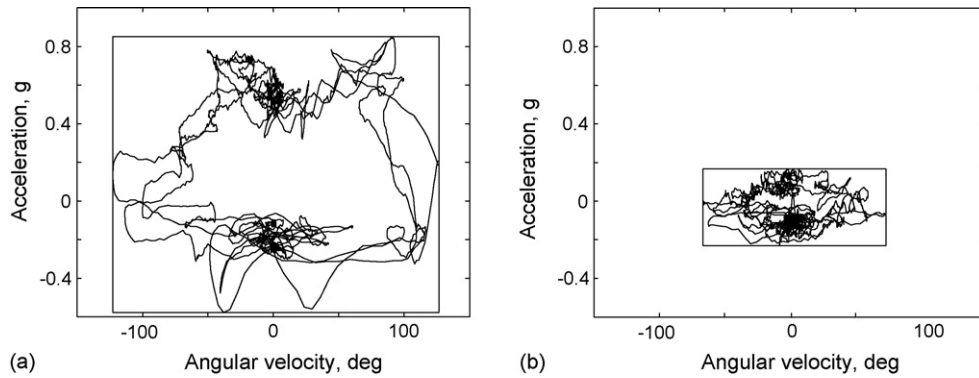


Fig. 3. Humerus acceleration as a function of its angular velocity for a patient. (a) The trace represents the humerus acceleration vs. angular velocity for the healthy side. (b) The trace represents the humerus acceleration vs. angular velocity for the painful side. The rectangle, which circumscribes the curve, corresponds to the product of the acceleration range by the angular velocity range (Pr).

Comparing to RAV where only angular velocities were used, P score used both the angular velocities and the accelerations of the humerus.

2.5. M score algorithm

Our last method was to estimate the difference of moments \vec{M} between the healthy and the painful shoulder; it was based on the angular velocities $\vec{\omega}$ of the humerus and the anthropometric data of the patient. The equation of the sum of all moments on a body segment can be expressed by [10]. \vec{M} was defined as the moment of the humerus (Eq. (7)), I as the inertia matrix (Eq. (8)).

$$\vec{M} = I \cdot \vec{\omega} + \vec{\omega} \times (I \cdot \vec{\omega}) \quad (7)$$

$$I = \begin{bmatrix} I_{\text{pitch}} & 0 & 0 \\ 0 & I_{\text{roll}} & 0 \\ 0 & 0 & I_{\text{yaw}} \end{bmatrix} \quad (8)$$

Using the mathematical definition of moment of inertia from Vaughan et al. [11] and the anthropometric data of the patient (length of the humerus: L_h , circumference of the biceps: C_h , mass of the humerus: m), the relationship of the moment of inertia about flexion/extension (I_{pitch}), the moment of inertia about abduction/adduction (I_{yaw}) and the moment of inertia about internal/external rotation (I_{roll}) can be derived (Eq. (9)).

$$I_{\text{pitch}} = \frac{m \cdot (0.076C_h^2 + L_h^2)}{12} \quad (9)$$

$$I_{\text{roll}} = \frac{m \cdot C_h^2}{8\pi^2}$$

$$I_{\text{yaw}} = I_{\text{pitch}}$$

We used this method to evaluate the difference between the healthy and the painful shoulder, calculating the maximum of the norm of the moment (noted by $|||$) during each test for each shoulder.

$$\Delta M = \max |||M_{\text{healthy}}||| - \max |||M_{\text{painful}}||| \quad (10)$$

The difference between the healthy and the painful shoulder was expressed as the percentage of the moment of the healthy shoulder.

$$\Delta Mr = \frac{\Delta M}{\max |||M_{\text{healthy}}|||} \quad (11)$$

The M score is defined as the average of the ΔMr over all nine tests.

$$M \text{ score} = 1 - \text{mean} \left[\sum_{\text{Test}=1}^9 \Delta Mr \right] \times 100 [\%] \quad (12)$$

A subject with a total mobility of his/her shoulder will have a M score, a RAV score and P score of 100% and a patient without any mobility of his/her shoulder will have a M score, a RAV score and a P score of 0%.

2.6. Statistical analysis

The Wilcoxon matched pairs signed rank sum test was used as a non-parametric hypothesis test to show if there were significant differences (at a significance level 5%) between baseline versus 3 months, and baseline versus 6 months for 10 patients.

The Wilcoxon rank sum test was used as a non-parametric hypothesis test to show if there were significant differences between baseline versus 10 control subjects, 3 months versus 10 control subjects and 6 months versus 10 control subjects.

3. Results

3.1. Angle estimation

Fig. 4 shows the angles of the basic movements for the reference system Zebris and the kinematic sensors. The proposed method offered accurate estimation of shoulder angles. The results of all the tests (Table 2) were very close to those of the reference system presenting a small average

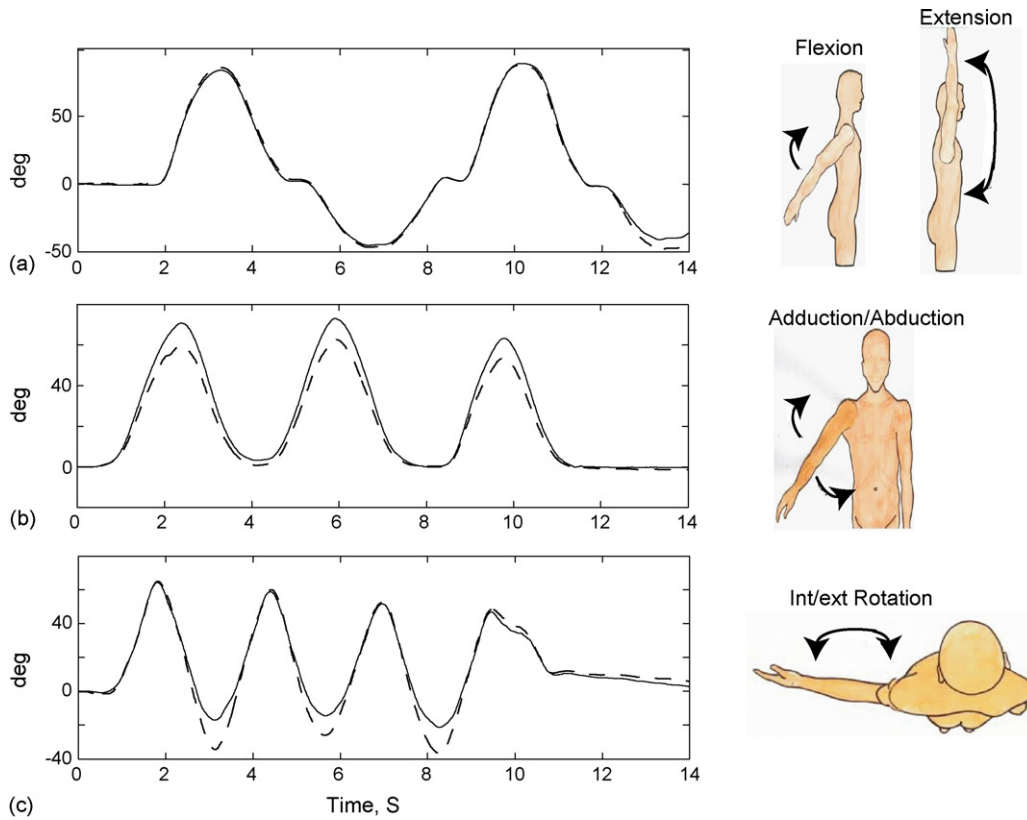


Fig. 4. Angles estimation compared to the reference system Zebris. (a) Flexion/extension. (b) Abduction/adduction. (c) Internal/external rotation. Dashed line: reference system. Solid line: kinematic sensors.

error in RMS (5.81°), mean (1.80°) and standard deviation (4.82°) of the difference signal, reflecting accurate and precise estimation, respectively; and excellent correlation coefficient (0.99) values reflected highly linear response.

3.2. P score

Fig. 5(a1 and b1) show the comparison of P parameters between a patient and a control subject for the nine tests

realized. It can be observed that for the patient (Fig. 5a1) the P parameter is higher for the healthy side than the painful side for all tests. But for the healthy subject (Fig. 5b1) the Pr parameter is approximately equal between the right and the left shoulder for each test. Table 3 shows the P score for a healthy subject. The P score for the healthy subjects ranged from 85% to 97% (mean: 92%), which is twice compared to patients before surgery (Tables 3 and 4).

Table 2
Comparison between humerus angles obtained from kinematic sensors and reference system for 10 subjects

	Flexion/elevation error ($^\circ$)			<i>r</i>	Abduction/adduction error ($^\circ$)			<i>r</i>	Rotation int./ext. error ($^\circ$)			<i>r</i>
	RMS	Mean	S.D.		RMS	Mean	S.D.		RMS	Mean	S.D.	
Subject 1	2.50	-0.45	2.47	0.9986	2.95	-2.20	1.97	0.9968	3.19	0.58	3.13	0.9983
Subject 2	5.64	-3.08	4.72	0.9936	3.83	3.34	1.88	0.9940	2.38	-0.95	2.19	0.9972
Subject 3	4.86	6.25	3.36	0.9888	5.53	-4.08	3.63	0.9994	5.72	-1.90	5.39	0.9865
Subject 4	7.49	6.48	7.29	0.9970	9.61	8.59	6.37	0.9653	8.04	-3.97	6.69	0.9491
Subject 5	7.25	6.02	6.90	0.9945	5.21	2.54	3.63	0.9880	7.99	1.32	7.88	0.9829
Subject 6	7.17	4.40	5.16	0.9953	8.97	6.55	8.52	0.9863	6.25	-5.92	4.61	0.9657
Subject 7	6.59	4.42	5.01	0.9962	1.41	0.48	1.33	0.9993	3.71	-4.49	3.57	0.9739
Subject 8	8.66	2.95	7.16	0.9984	3.62	0.31	3.58	0.9976	5.82	2.25	3.37	0.9950
Subject 9	6.56	5.16	6.44	0.9975	7.80	7.98	5.55	0.9849	6.50	2.68	6.10	0.9971
Subject 10	10.03	4.26	9.09	0.9989	1.12	0.09	1.10	0.9991	7.81	4.32	6.51	0.9933
Mean	6.68	3.64	5.76	0.9959	5.01	2.36	3.76	0.9911	5.74	-0.61	4.94	0.9839

The error represents the root mean square (RMS), mean and S.D. of the difference between reference and our measuring device. 'r' represents the correlation coefficient between the two measuring systems.

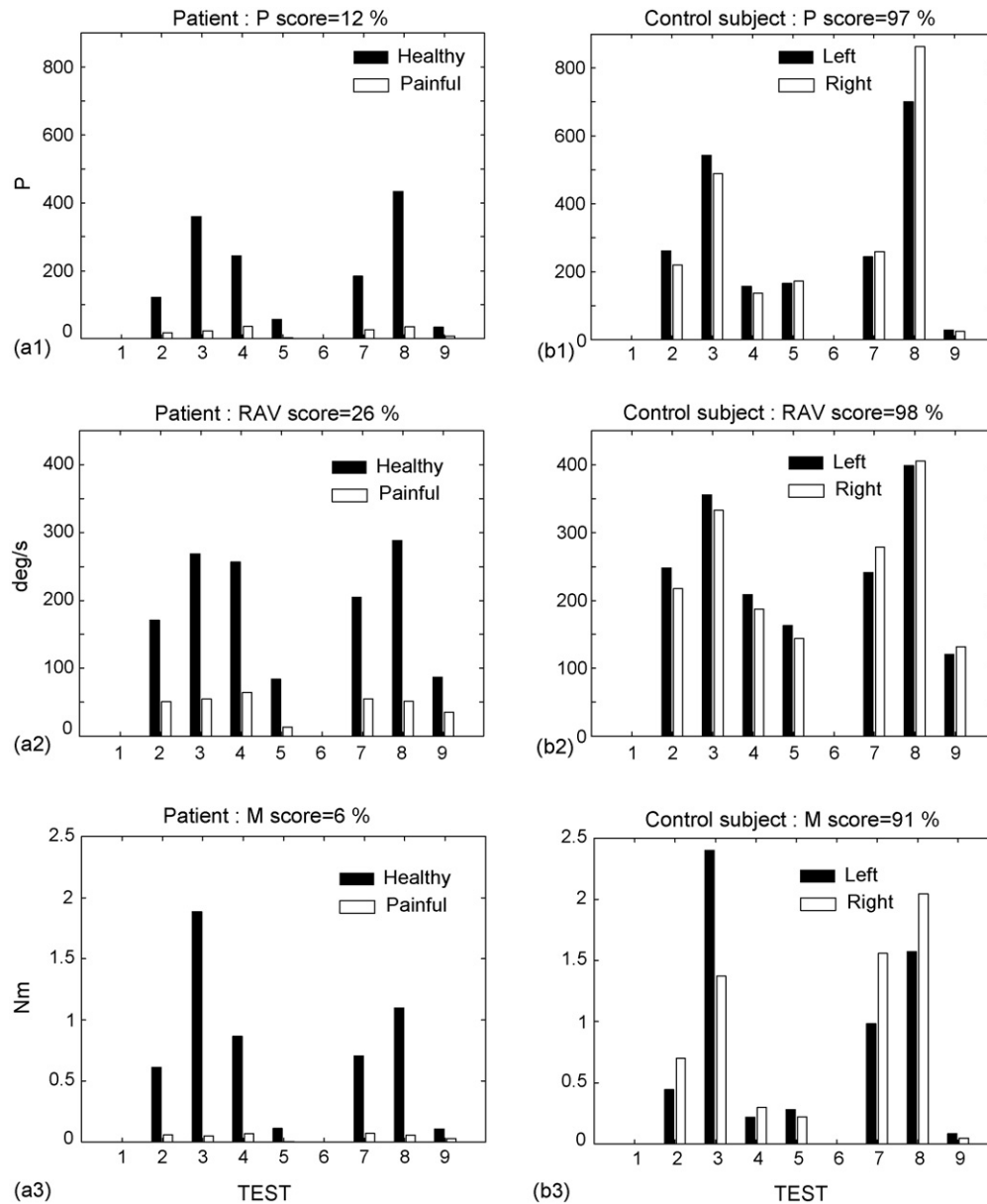


Fig. 5. Pr parameter for a patient (a1) and a control subject (b1). RAVr parameter for a patient (a2) and a control subject (b2). Mr parameter for a patient (a3) and a control subject (b3).

Table 3 shows all the results in comparison with the baseline (before surgery). The Wilcoxon matched pairs signed rank sum test indicates that significant differences were found between the P score at baseline versus the P score at 3 months and the P score at baseline versus the P score at 6 months ($p < 0.05$).

The P score average was 46%, 67% and 72%, respectively, at baseline, 3 month and 6 month after surgery. Fig. 6(a) shows the improvement of the P score after surgery in comparison to the baseline values and the control subjects.

We observed significant differences between the P score at the baseline versus the P score of the healthy subjects and the P score at 3 month versus the P score of the healthy

subjects, but no significant differences were found between the P score at 6 month versus the P score of the healthy subjects ($p = 0.074$).

3.3. RAV score

Fig. 5(a2 and b2) shows the comparison of RAV parameters between a patient and a control subject for the nine (Editor's question: for the nine ... does not refer to anything—please clarify). The RAV parameter is higher for the healthy side than the painful side for all tests (Fig. 5a2). But for a healthy subject (Fig. 5b2) the Δ RAV parameter is approximately similar between the right and the left shoulder for each test. The RAV score for healthy subject

Table 3
DASH, SST, P score, RAV score and M score for patients before surgery (baseline) and at 3, 6 months after surgery

Patients	1	2	3	4	5	6	7	8	9	10	Wilcoxon test
RAV score baseline	42	80	69	70	66	5	50	64	84	59	
RAV score 3 months	87	94	79	98	76	81	62	60	94	76	$p = 0.0039$
RAV score 6 months	87	93	93	94	70	95	54	66	97	76	$p = 0.0020$
P score baseline	28	75	57	62	48	3	36	38	67	48	
P score 3 months	70	74	82	91	67	61	42	39	88	59	$p = 0.0059$
P score 6 months	76	67	98	93	58	97	33	39	87	69	$p = 0.0195$
M score baseline	22	51	48	42	36	22	15	25	55	25	
M score 3 months	64	90	59	37	65	63	31	44	69	64	$p = 0.0041$
M score 6 months	66	83	97	44	52	70	23	42	86	60	$p = 0.0020$
Dash baseline	137	91	47	74	93	75	93	128	79	47	
Dash 3 months	137	101	34	49	80	74	115	78	50	65	NS
Dash 6 months	94	93	34	32	81	54	110	72	54	38	$p = 0.0273$
SST baseline	0	7	9	5	1	5	1	1	4	6	
SST 3 months	0	3	11	11	6	6	1	3	5	2	NS
SST 6 months	5	4	11	10	6	9	1	3	7	10	$p = 0.0234$

NS indicates that no significant differences were found at 5%. The DASH (30 is “very good mobility” and 150 is “very poor mobility”), SST (0 is “very poor mobility” and 12 is “very good mobility”).

ranged from 87% to 99% (mean: 94%). While this score was in average 59% for patients preoperatively (Tables 3 and 4).

Significant differences were found between the RAV score at baseline and at 3 months, as well as between the RAV score at baseline and at 6 months ($p < 0.05$).

The average RAV score was, respectively, 81% and 83% at 3 months and 6 months after surgery (Table 3). Fig. 6(b) shows the improvement of RAV score after surgery in comparison to the baseline values and the control subjects.

The RAV score of the healthy subjects was significantly higher than the RAV score at baseline as well as the RAV score at 3 month, but significant differences were also found between the RAV score at 6 months and the RAV score of the healthy subjects ($p = 0.037$).

Table 4
DASH, SST, P score, RAV score and M score for healthy subjects

Subjects	P score (%)	RAV score (%)	M score (%)
1	91(7)	94(5)	91(2)
2	96(-12)	99(-14)	87(3)
3	93(-4)	98(-4)	88(3)
4	94(3)	98(-1)	82(2)
5	96(-3)	91(5)	97(-9)
6	93(-11)	95(5)	86(12)
7	97(-13)	95(-8)	95(-15)
8	90(10)	96(1)	93(-3)
9	93(5)	93(6)	72(17)
10	98(-9)	96(-9)	89(5)
Mean $\Delta(1 - 2)$	-2.7	-1.4	0.7
S.D. $\Delta(1 - 2)$	8.5	7.1	9.4

For all the healthy subjects: the SST was 12 and the DASH was 30. In brackets: difference between the baseline and the 1 year measurements ($\Delta(1 - 2)$).

3.4. M score

Fig. 5(a3 and b3) show the comparison of moment in Newton-meter (Nm) between a patient and a control subject for the nine. The moments are higher for the healthy side than the painful side for all tests (Fig. 5a3); while the moments are similar between the right and the left shoulder for healthy subjects (Fig. 5b3). The M score for healthy subjects ranged from 82% to 97% (mean: 88%), which is more than twice the average for the patients preoperatively (Tables 3 and 4).

The M score at baseline was significantly lower than the M score at 3 months as well as at 6 months ($p < 0.05$).

Table 3 shows all the results in comparison with the baseline. The M score average was, respectively, 59% and 62% at 3 months and 6 months after surgery. Fig. 6(c) shows the improvement of the M score after surgery in comparison to the baseline values and the control subjects.

We observed that there were significant differences between the M score at the baseline versus the M score of the healthy subjects and the M score at 3 month versus the M score of the healthy subjects, but significant differences were also found between the M score at 6 month versus the M score of the healthy subjects ($p = 0.009$).

4. Discussion

Previous studies on shoulder outcome evaluation used questionnaires and imposed movements. Kirkley et al. [1] presented the differences between scoring systems for the functional assessment of the shoulder. They observed that many of the items may seem irrelevant to patients with specific conditions and none has been accepted as the universal standard. In some cases, the patients could not

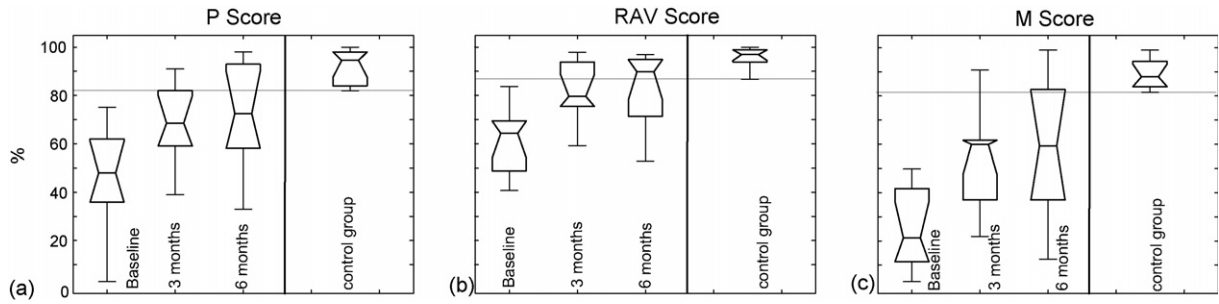


Fig. 6. Box plot for the P score (a), the RAV score (b) and the M score (c). Boxes contain 50% of the results and lines represent the range. Dashed line shows the limit for healthy subjects.

understand the questions and were unable to answer or answered erroneously. The DASH instrument is a questionnaire which may carry some of the above problems. Furthermore, it may be influenced by the psychological condition of the patient. Due to the dichotomous response option (yes or no), the SST instrument is likely to have poor sensitivity to differentiate between patients with varying severity of the same condition [1].

Our outcome evaluation of shoulder surgery was based on objective scores derived from accurate 3D measurement (Table 2) of shoulder kinematics on healthy and affected individuals performing specific tasks. These scores are based on acceleration and angular velocity rather than angles measurements. Though angles can be estimated accurately with our system, they have not shown pertinent changes

between a healthy and a painful shoulder. Fig. 7 shows the 3D angles for a patient for the test no. 2, where the subject moved his hand to the back. The angular ranges are rather larger for the painful side in comparison to the healthy side for the abduction/adduction (yaw) and flexion/extension (pitch) axis. This observation shows that the patient has a strategy to minimize the pain by accomplishing a longer path than normal for the painful shoulder to perform the same movement. However, this is not the case for all patients, since every patient has a different movement strategy to reduce the shoulder pain. Therefore, it was not possible to use the angle magnitude as an objective parameter to quantify the difference between healthy and painful shoulders.

This paper proposed three different scores: the P score based on a combination of accelerations and angular

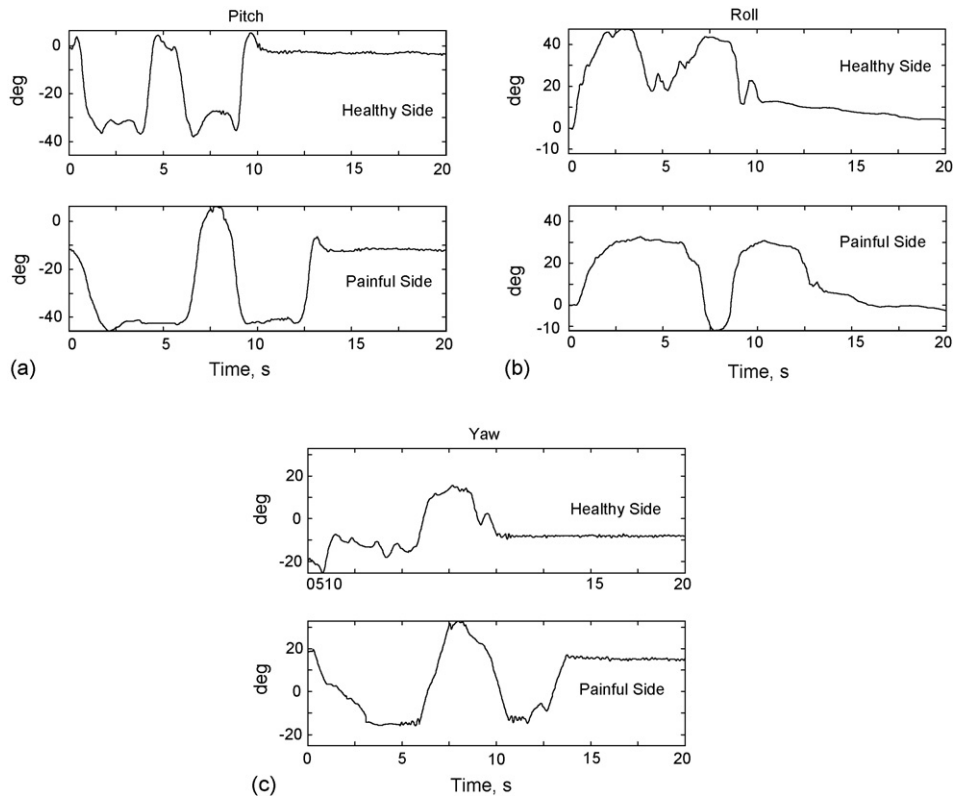


Fig. 7. Humerus angles for test 2, consisting to move the hand to the back. Healthy humerus angles and painful humerus angles in flexion/elevation (pitch) (a), in internal/external rotation (roll) (b) and in abduction/adduction (yaw) (c).

velocities, the RAV score based on the differences of angular velocities range and the M score based on the sum of all moments of the humerus. These scores represent a way to assess shoulder function based on quantification of the kinematic differences between healthy and painful shoulders. Fig. 6 shows the comparison between baseline, 3, 6 months after surgery for the three scores. For all the patients, shoulder mobility increased significantly after surgery (Table 3). In addition, the scores are clearly distinct between healthy subjects and patients with a painful shoulder at baseline without any overlapping of the confidence intervals (Fig. 6).

Table 3 also shows the results of the Wilcoxon matched pairs signed rank sum test for the clinical scores (DASH, SST). While kinematic scores showed significant differences between baseline and follow-up time ($p < 0.02$), the clinical scores (DASH, SST) showed no significant differences between baseline and 3 months but the differences became significant at 6 months ($p < 0.03$). These results suggest that our kinematic scores may be more sensitive to the functional changes than the clinical scores, even at 3 months after surgery.

Table 3 shows that patient 7 had poor clinical scores after surgery. He had an inflammatory capsulitis after 6 months. The kinematic scores also detected this post-operative complication with changes consistent with the patient suffering with pain while performing some movements. Another complication involved patient 8 who suffered chronic dislocation. His clinical scores were improved but the kinematic scores were equal to the baseline, expressing the poor mobility of this patient.

By producing objective scores based on 3D kinematics of the shoulder our system assessed the shoulder function. However, it cannot yet be used for the diagnosis of complex pathology or to differentiate between pathologies. Our score is not related directly to pain but to its effect on mobility. For example, if a patient experiences shoulder pain and restricted range of motion; our system will detect this lack of function. However, in the absence of recovery of shoulder functional even if the pain is removed after surgery our scores will remain low.

It is noteworthy that those three scores compare the patient's affected and non-affected shoulder only if the pathology is unilateral. Further study with more measurements on patients is needed to be able the use of these scores in the assessment of shoulder function independently of the pathology.

Patients were selected with unilateral symptomatic shoulders. However, rotator cuff pathology on the unaffected side could not be excluded. For this reason, the first comparison of the scores was made on an intra-patient basis along time. However, if the unaffected shoulder is asymptomatic, it represents the same concept of reference for all the patients: the goal of function recovery after surgery, taking into account their shoulder joint evolution with age. Based on this concept, we performed comparisons across patients.

Concerning the sensor attachment some precautions should be taken. Firstly, in order to reduce the effects of skin artefact a sticking elastic band was used to fix the sensors. In addition, the module was placed on the distal and posterior part of the humerus where there is less skin movement and where the sensor can fully detect rotation of the humerus. In fact, if the sensor is positioned at the top of the humerus (near the humeral head), the internal/external rotation cannot be measured.

In order to estimate the repeatability of the system, measurements were repeated on the 10 control subjects after 1 year. The comparison between the two measurements showed low difference (less than 3% in average with S.D. less than 10%) (Table 4).

The proposed scores are clinically meaningful. The RAV score represents the velocity of the humerus. The P score shows how the patient controls the velocity of his humerus using a combination of accelerations and angular velocities. The M score represents the sum of all moments on the shoulder. Based on this study and the limited sample size, it is difficult to decide which score is more adapted (Editor's question: clarify the use of the word adapted. Do you mean practical for clinical use/adaptable?). To answer this question we would require to study more subjects, to perform a clinical validation by considering the type of pathology as well as to assess the results of these scores during long-term monitoring of daily activity. Our proposed scores could be used in the long-term monitoring of shoulder kinematics in daily activity. By recognizing physical activity using additional sensors [9,20] it would be possible to provide a better evaluation of shoulder mobility and therefore offer a score which would be more reliable since it would be based on natural and voluntary activity of the patients. Moreover, using one sensor module on each humerus and one of the three scores, it should be possible to compare painful and healthy shoulders during daily activity. In this respect, our proposed system appears particularly effective: the sensors have low power consumption (4 mA) and the standard batteries, the system allows to record up to 12 h with a memory of 80 MB. Monitoring the subjects in their usual environment with minimal interference would therefore be possible, in contrast with other systems that require the use of a laboratory.

5. Conclusion

Based on kinematics of the patients, we were able to define three objective scores for the assessment of shoulder function and for the quantification of the differences of kinematics between healthy and painful shoulders.

The proposed system has the potential to be used during daily activities as well as before and after shoulder surgery and to provide a useful outcome.

Acknowledgements

The authors are grateful to J. Gramiger and P. Morel for the design of the ambulatory system. This work was supported by Swiss National Foundation (PNR 53) Grant no. 405340-104752/1.

References

- [1] Kirkley A, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. *Arthro Rel Surg* 2003;19(10):1109–20.
- [2] Jester A, Harth A, Wind G, Germann G, Sauerbier M. Disabilities of the arm, shoulder and hand (DASH) questionnaire: determining functional activity profiles in patients with upper extremity disorders. *J Hand Surg (British and European Volume 2005)* 2004;30B(1):23–8.
- [3] Lippitt SB, Harryman DT, Matsen FA. A practical tool for evaluating function: the Simple Shoulder Test. In: Matsen FA, Fu FH, Hawkins RJ, editors. *The shoulder: a balance of mobility and stability*. Rosemont, IL: Am Acad Ortho Surg; 1992. p. 501–18.
- [4] Nakagawa Y, Hyakuna K, Otani S, Hashitani M, Nakamura T. Epidemiologic study of glenohumeral osteoarthritis with plain radiography. *J Shoulder Elbow Surg* 1999;8(6):580–4.
- [5] Fraenkel L, Lavalley M, Felson D. The use of radiographs to evaluate shoulder pain in the ED. *Am J Emerg Med* 1998;16(6):560–3.
- [6] Gagnon M, Larrivé A, Desjardins P. Strategies of loads tilts and shoulders positioning in asymmetrical lifting. A concomitant evaluation of the reference systems axes. *J Clin Biomech* 2000;15:478–88.
- [7] Gagnon M, Delisle A, Desjardin P. Biomechanical differences between best and worst performances in repeated free asymmetrical lifts. *Int J Ind Ergo* 2002;29:73–83.
- [8] Hienz H, Grobel K, Offner G. Real-time hand-arm motion analysis using a single video camera. In: *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*; 1996. p. 323–7.
- [9] Najafi B, Aminian K, Paraschiv-Ionescu A, Loew F, Büla C, Robert P. Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. *IEEE Trans Bio Eng* 2003;50:711–23.
- [10] Van den Bogert A, Read L, Nigg B. A method for inverse dynamic analysis using accelerometry. *J Biomech* 1996;29:949–54.
- [11] Vaughan C, Davis B, O'Connor J. *Dynamics of human gait* Human Kinetics Publishers; 1992.
- [12] Aminian K. Human movement capture and their clinical applications. In: Begg RK, Palaniswami M, editors. *Computational intelligence for movement sciences: neural networks, support vector machines and other emerging techniques*. USA: Idea Group Inc.; 2006 [Chapter 3].
- [13] Mathie MJ, Coster A, Lovell NH, Celler BG. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiol Meas* 2004;25:R1–20.
- [14] Kiss RM, Kocsis L, Knoll Z. Joint kinematics and spatial temporal parameters of gait measured by an ultrasound-based system. *Med Eng Phys* 2004;26:611–20.
- [15] Z.M. GmbH, Measuring system for 3D-motion analysis CMS-HS, vol. 8/99. Tech. Data and Operating Instructions Text Release 1999.
- [16] Overhoff HM, Lazovic DM, Liebing M, Macher C. Total knee arthroplasty: coordinate system definition and planning based on 3-D ultrasound image volumes, vol. 1230 International Congress Series; 2001. pp. 292–9.
- [17] SooHoo NF, McDonald AP, Seiler 3rd JG, McGillivray GR. Evaluation of the construct validity of the DASH questionnaire by correlation to the SF-36. *J Hand Surg [Am]* 2002;27(3):537–41.
- [18] Gummesson C, Atroshi I, Ekdahl C. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *BMC Musculoskelet Disord* 2003;4:11.
- [19] Beaton D, Richards RR. Assessing the reliability and responsiveness of 5 shoulder questionnaires. *J Shoulder Elbow Surg* 1998;7(6):565–72.
- [20] Paraschiv-Ionescu A, Buchser EE, Rutschmann B, Najafi B, Aminian K. Ambulatory system for the quantitative and qualitative analysis of gait and posture in chronic pain patients treated with spinal cord stimulation. *Gait Posture* 2004;20(2):113–25.