Acta Orthop. Belg., 2012, **78**, 751-756

ORIGINAL STUDY

The accuracy of retrospectively measured range of motion in knee arthroplasty

Sina Babazadeh, Michelle M. Dowsey, James D. Stone, Peter F.M. Choong

From St. Vincent's Hospital, University of Melbourne, Australia

In retrospective studies the range of motion of the knee is gathered from existing clinical notes or databases. This study aims to assess the validity of this retrospective data. The range of motion of 48 patients was assessed using a goniometer and compared to that entered in the patient notes by the examiner during a routine clinical examination, without the examiner being aware. The range of motion of a further 20 patients was subsequently assessed and compared to the findings of the same clinical examiners but this time with the examiner being aware. When the examiner was unaware of the study, the accuracy of the measured range of motion was clinically unacceptable. When the examiner was aware of this study, the accuracy of the measure was much improved. The patient's body mass index can affect the accuracy of visual estimation. Visual estimation of range of motion can be accurate and correlates well to that measured by a goniometer if the examiner attempts to be accurate. However range of motion that is routinely measured in a clinical setting and entered into the patient notes is not accurate and should not be relied on in future research.

Keywords : knee ; arthritis ; retrospective ; prospective ; flexion.

INTRODUCTION

The range of motion of the knee is an important criterion for predicting the outcome of a joint replacement, and subsequently judging its success (*6,9,12*). Maximum knee flexion also plays a vital role in research when comparing patients and prostheses (*3,11,13*).

Many methods have been employed to measure a patient's range of motion. The most commonly used method in orthopaedics is clinical examination using visual appraisal. When conducting formal research, a goniometer is more commonly used (*7*). Currently the gold standard is a lateral radiograph of the knee in full flexion. This is highly impractical, with a heavy time and financial burden and unnecessary radiation exposure (*1,5*). The goniometer has been compared to radiographic measurements of flexion and has been found to be reliable and valid, with excellent inter-observer reliability (correlation coefficient of 0.99) (*7*). When standardized, measuring flexion using visual estimation and

■ Sina Babazadeh, MB, BS, Research fellow.
■ Michelle M. Dowsey, RN, PhD, Head Arthritis Research.
■ James D. Stoney, MB, BS, Consultant Orthopaedic Surgeon.
■ Peter F. M. Choong, MB, BS, MD, Sir Hugh Devine Professor of Surgery and Head, Department of Surgery.
Department of Orthopaedics, St. Vincent's Hospital (Melbourne), University of Melbourne, Australia.
Correspondence : Prof Peter F.M. Choong : Sir Hugh Devine Professor of Surgery and Head, Department of Surgery, The University of Melbourne and Director of Orthopaedics St. Vincent's Hospital (Melbourne) PO Box 2900, Fitzroy, Victoria 3065, Australia. E-mail : sarcoma@bigpond.net.au

No benefits or funds were received in support of this study. The authors report no conflict of interests.

goniometer compare well with flexion measured from radiographs with a correlation of 0.86 and 0.96 respectively (*5*). Intra-observer correlation between visual estimation and measurement using a goniometer is also surprisingly good (0.92) (*5*).

The accuracy of un-standardized retrospectively collected data is not clear. Retrospective studies often use patient notes or a patient database to determine range of motion (*3,4,8,11,13,14*). These values, especially in the pre-operative setting, are almost invariably gauged from a visual estimation, usually by an individual who is unaware ; the findings will be used in future research. Retrospectively collected pre-operative values cannot be verified, however these values are often considered factual and reliable when utilized for research and audit.

The aims of this study are to scientifically asses the accuracy of retrospectively collected range of motion data in the setting of arthritis and knee arthroplasty, and to assess if increased accuracy can be achieved by prospectively measuring data. Our hypothesis is that retrospectively collected data is inaccurate and should not be used in arthroplasty research assessing range of motion.

MATERIALS AND METHODS

This prospective trial consists of two phases. During the initial phase, the accuracy of range of motion data that is routinely collected during patient examination was assessed using a goniometer. In the second phase, the accuracy of visual estimation of range of motion was compared to a goniometer.

Phase one

Forty-eight patients were recruited at a large tertiary centre specializing in arthroplasty surgery. All patients were recruited whilst attending a routine visit. All patients were either consented to undergo, or had recently undergone a total knee joint replacement.

As part of a routine review, the surgeon and physiotherapist gauged the patient's range of motion using visual estimation and documented the results within the patient's file. Once entered into the patient notes, this data is subsequently placed on a database, which can be used to audit patient's clinical outcomes.

After their visit was completed, the patients subsequently had their range of movement reassessed using a goniometer. This was performed in another examination room in a separate part of the building by a researcher, with the surgeon and physiotherapist unaware. All patients consented.

The same standard 2-arm goniometer (12-inch arms with one degree markings) was used on all patients. In all cases the patient was adequately exposed from hip to feet and asked to lie supine on an examination bed, with the knee being measured closest to the examiner. In all cases the anatomical landmarks used were the greater trochanter proximally, the lateral condyle of the femur as the midpoint and the lateral malleolus distally. The patient was first asked to maximally extend the knee, and then maximally flex the knee to record the active range of movement. The patient was asked three times whether flexion was maximal, so that he/she had an opportunity to increase active flexion as much as possible. All patients were measured three times.

The surgeon, physiotherapist and nursing staff were kept blinded. The researcher was blinded to the range of motion values recorded by the surgeon and physiotherapist, and recorded his findings separately.

Phase two

At the conclusion of the initial phase, a further 20 patients were recruited. During this phase of the trial, the surgeon was made aware that their measurements would be compared to measurements taken using a goniometer and was asked to be as accurate as possible when undertaking their clinical examination. Range of motion of these twenty patients was recorded by the surgeons using visual estimation.

These patients were then re-examined by the researcher using a goniometer in the same standard manner as described above. The researcher was kept blinded to the values measured by the surgeons.

Statistical analysis

Statistical analysis was carried out using SigmaStat version 3.5 (SPSS Inc., Chicago, Illinois). A Pearson Product Moment Correlation test was used to determine intra-observer correlation for the researcher and inter-observer correlation between observers. Agreement was assessed using the Altman-Bland graphical method (*2*). Continuous parametric data was analysed using one-way analysis of variance (ANOVA), while non-parametric data were tested using Mann-Whitney non-parametric rank-sum test. A p-value < 0.05 was considered significant.

RESULTS

Sixty-eight patients were consented for this trial in total. No patient was excluded. The mean age of the participants was 69.9 years, with a range of 46 to 84. Forty-one of the patients were female, and all suffered from osteoarthritis. Sixty-two were pre-operative ; the remainder were post-operative, ranging from 6 weeks to 3 years. Patients were analysed in two groups, as described above.

Intra-observer correlation of the three measurements made by the researcher using a goniometer was noted to be excellent with a correlation coefficient of 0.90 for flexion and 0.95 for extension ($p < 0.001$).

Phase one

Inter-observer correlation

Results for range of motion as measured by the surgeon, physiotherapist using visual estimation, and by the researcher using a goniometer can be found in table I.

Inter-observer correlation was sought between the three observers using a Pearson Product Moment Correlation test. These results can be found in table II.

Agreement between goniometer and the surgeon can be observed using figure 1 for flexion. Graphical representation of agreement for extension and goniometer versus physiotherapist were similar.

The initial cohort was then split into two groups, separated by whether flexion as measured by the goniometer was recorded as 90 or more degrees (group one : flexible knees), or under 90 degrees (group 2 : stiff knees). These two groups were then compared in regards to whether the visual observers (being surgeon and physiotherapist), under or over-estimated the flexion range of these patients. Results can be found in table III.

A similar analysis was carried out using measurements taken in extension. The cohort was split into those recording extension values of 0 degrees or more (that is able to hyperextend, group 3), and those that had a flexion contracture of any degree

Table I. — Initial blinded group, comparing the mean, range and SD between the three observers in flexion and extension

	Researcher (goniometer)	Surgeon (visual)	Physiotherapist (visual)
Flexion Mean	102	102	103
Range	65-135	40-120	85-130
SD	16.5	12.4	15.6
Extension Mean	-7.83	-5.3	-2.2
Range	-30-5	-40-10	-20-10
SD	7.1	8.2	5.5

Table II. — Inter-observer correlation, comparing flexion (bottom) and extension (top)

Correlation Coefficients	Researcher (goniometer)	Surgeon (visual)	Physiotherapist (visual)
Researcher (goniometer)		0.704 (Extension) P < 0.001	0.620 (Extension) P < 0.001
Surgeon (visual)	0.602 (Flexion) P < 0.001		0.438 (Extension) P = 0.005
Physiotherapist (visual)	0.563 (Flexion) P < 0.001	0.494 (Flexion) P = 0.001	

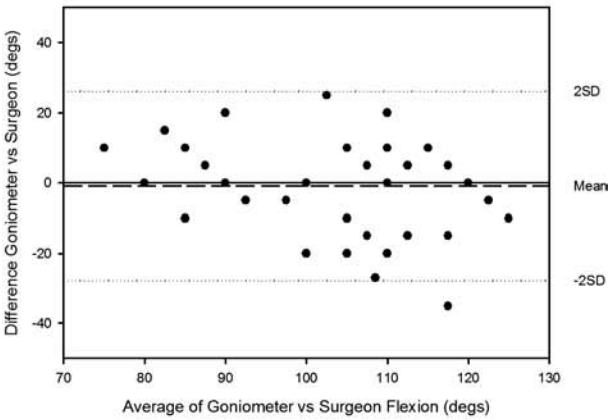

Fig. 1. — Graph demonstrating agreement between flexion as measured by goniometer and as measured by the surgeon in phase one. The x-axis represents the mean of the two measures, while the y-axis represents the difference between the two measures as per the Altman-Bland method. Lines represent two standard deviations from the mean.

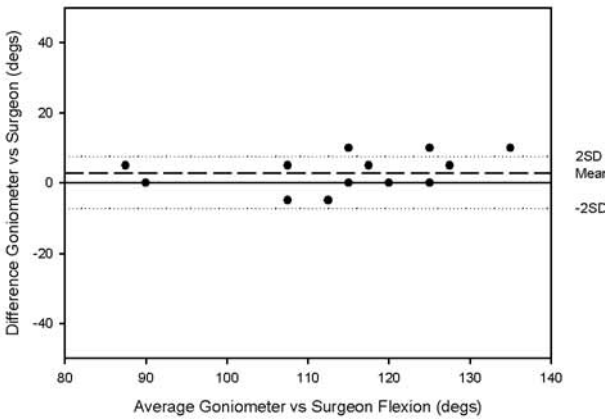

Fig. 2. — Graph demonstrating agreement between flexion as measured by goniometer and as measured by the surgeon in phase two. The x-axis represents the mean of the two measures, while the y-axis represents the difference between the two measures as per the Altman-Bland method. Lines represent two standard deviations from the mean.

(that is an inability to fully straighten the limb, group 4). Results can be found in table III.

Accuracy between surgeons

When comparing the accuracy of measurements in extension or flexion, between surgeons, no surgeon was significantly more or less accurate than this colleague (Flexion p = 0.749, extension p = 0.485).

Accuracy of pre- versus post-operative flexion

When comparing pre and post-operative knees, neither flexion nor extension results were more or less accurate, either measured by the surgeon or physiotherapist (Flexion p = 0.397, extension p = 0.302).

Accuracy versus BMI

An association between inaccuracy of range of motion measurements and obesity was also sought. Initially a linear regression revealed a trend existing between the measured body mass index (BMI) and the inaccuracy of flexion measured (p = 0.033). This was further supported by a one way analysis on ranks when the cohort was split into those with a BMI 30 or above and those under 30 (Table IV).

Phase two

The results of the second group of 20 patients can be found in table V. Inter-observer correlation was sought between the surgeon and researcher using a Pearson Product Moment Correlation test. These results can be found in table VI and visualized using figures 1 & 2.

DISCUSSION

This is the first study assessing the validity of retrospectively collected clinical data in ortho-

Table III. — Accuracy versus degree of flexion and extension. (negative values indicate underestimation)

Measured by goniometer	Difference as measured by Surgeon	Difference as measured by Physiotherapist	P-value
Flexion 90 or more	-4°	-2.7°	P = 0.001
Flexion < 90 degrees	+10.4°	+13°	P = 0.004
Extension 0 or more	-0.5°	+0.5°	P = 0.054
Extension < 0	+3.5°	+8°	P < 0.001

Table IV. — One way analysis on ranks comparing those with BMI under 30 to those with BMI equal to or over 30

	BMI < 30	BMI ≥ 30	P-value
Surgeon-			
Flexion	10° under-estimated	3° over-estimated	0.006*
Extension	0°	5° over-estimated	0.852
Physio-			
Flexion	8° under-estimated	6° over-estimated	0.008*
Extension	5° over-estimated	5° over-estimated	0.779

* = significant.

Table V. — Comparing flexion and range for both flexion and extension as measured by the surgeon using visual estimation and the researcher using a goniometer

	Researcher (goniometer)	Surgeon (visual)
Flexion Mean	111°	116°
Range	85°-130°	90°-140°
SD	13.2°	13.8°
Extension Mean	-11°	-7°
Range	-30°-5°	-25°-0°

Table VI. — Correlation between goniometer and visual estimation made by surgeons when aware of study

	Goniometer versus Visual	P-value
Flexion	0.934	< 0.001
Extension	0.974	< 0.001

paedics. The results of this study indicate that the retrospective measurement of range-of-motion is not accurate, regardless of whether it is recorded by the surgeon or physiotherapist. When individuals are recording range-of-motion in a routine fashion, the observers tends to gravitate their results towards 90° of flexion and 0° of extension. There is a trend to over-estimate the flexion of knees that are below 90 degrees. These trends are apparent regardless of the surgeon, or the timing of the review, either pre- or post-operative. Visual estimation of range-of-motion was found to be comparable to that measured by goniometer when the observer attempts to be accurate.

Reviewing the graphical representation of the data in figures 1 and 2, it can be seen that the 2-standard deviation lines in phase one are clinically unacceptable, approaching 25 to 30° for flexion (Fig. 1) and 10 to 15° for extension. Accuracy improved significantly in phase two with the 2 SD lines being at more clinically acceptable levels (Fig. 2).

Edwards *et al* (*5*) compared flexion as measured either by visual estimation, goniometer or lateral radiograph of the knee in 27 knees. They found that the inter- and intra-observer correlation was high between goniometer and visual estimation, when compared to the lateral radiograph (0.92 and 0.79 respectively). In Edwards' study, it was noted that the goniometer tended to under-estimate the flexion angle, especially with increasing flexion. A noteworthy difference between Edwards' study and this study is that the observers in this study were unaware that their measures were being analysed. This was an attempt to better reflect the recorded data that is commonly used in retrospective studies.

There are a number of possible reasons explaining the inaccuracy of the retrospectively collected data.

In a busy clinical setting, examination of patients is often far from ideal. Frequently the patient is examined without proper exposure and even occasionally without being asked to move to an examination bed. On a number of occasions it has been witnessed that range-of-motion has been measured whilst the patient is seated.

If a knee is flexed to less than 90 degrees then in general it is considered stiff. This may be a reason why the visual estimates of flexion gravitate towards the 90 degree mark. Also, when the knee is examined with the patient seated, flexion much greater than 90 degrees is usually limited by the chair and floor.

The increased accuracy once the observers were made aware of the study could be explained by the "Hawthorne effect" (*10*). This term refers to an observed modification of behaviour by subjects simply in response to the knowledge that they are being studied. In this case, a significant improvement in accuracy of measured range of motion was noted when the observer was made aware that his measurements were being scrutinized. This highlights the ability of visual estimation of range of motion to be accurate should the observer attempt to be so.

The BMI of the patient affects the accuracy of visual estimation of flexion. A lower BMI tended to result in an under-estimation of the flexion range, and conversely a higher BMI tended to result in an over-estimation. Similar results were reported by Austin *et al* (*1*), but not found by Edward *et al* (*5*).

This study did have some limitations. Only one observer measured the range of motion using a goniometer. This was done to standardize measurements across all patients and to minimize the footprint of this study. Also other observers for this study were not standardized. Data from four separate surgeons and multiple physiotherapists were used. This is in part an attempt to replicate how data used in retrospective studies are normally obtained.

To conclude, retrospectively collected data incorporating range of motion cannot be considered accurate unless the examiner makes a conscious attempt to ensure accuracy. Accuracy considerably improves when the observer knows his results will be analysed. The observer should also be aware that the patient's BMI will affect accuracy of the measurement.

REFERENCES

1. **Austin MS, Ghanem E, Joshi A *et al.*** The assessment of intraoperative prosthetic knee range of motion using two methods. *J Arthroplasty* 2008 ; 23 : 515-521.
2. **Bland JM, Altman DG.** Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986 ; 1 (8476) : 307-310.
3. **Cho SD, Youm YS, Park KB.** Three- to six-year follow-up results after high-flexion total knee arthroplasty : can we allow passive deep knee bending ? *Knee Surg Sports Traumatol Arthrosc* 2011 ; 19 : 899-903.
4. **Deschamps G, Khiami F, Catonne Y *et al.*** Total knee arthroplasty for osteoarthritis secondary to extra-articular malunions. *Orthop Traumatol Surg Res* 2010 ; 96 : 849-855.
5. **Edwards JZ, Greene KA, Davis RS *et al*.** Measuring flexion in knee arthroplasty patients. *J Arthroplasty* 2004 ; 19 : 369-372.
6. **Giori NJ, Giori KL, Woolson ST *et al.*** Measurement of perioperative flexion-extension mechanics of the knee joint. *J Arthroplasty* 2001 ; 16 : 877-881.
7. **Gogia PP, Braatz JH, Rose SJ, Norton BJ.** Reliability and validity of goniometric measurements at the knee. *Phys Ther* 1987 ; 67 : 192-195.
8. **Kamath AF, Horneff JG, Gaffney V, Israelite CL, Nelson CL.** Ethnic and gender differences in the functional disparities after primary total knee arthroplasty. *Clin Orthop Relat Res* 2010 ; 468 : 3355-3361.
9. **Kawamura H, Bourne RB.** Factors affecting range of flexion after total knee arthroplasty. *J Orthop Sci* 2001 ; 6 : 248-252.
10. **Landsberger HA.** *Hawthorne revisited. Management and the worker : its critics, and developments in human relations in industry*. Ithaca, N.Y., : Cornell University ; 1958.
11. **Malik A, Salas A, Ben Ari J, Ma Y, Gonzalez Della Valle A.** Range of motion and function are similar in patients undergoing TKA with posterior stabilised and high-flexion inserts. *Int Orthop* 2010 ; 34 : 965-972.
12. **Maloney WJ, Schurman DJ.** The effects of implant design on range of motion after total knee arthroplasty. Total condylar versus posterior stabilized total condylar designs. *Clin Orthop Relat Res* 1992 ; 278 : 147-152.
13. **Massin P, Dupuy FR, Khlifi H *et al.*** Does hyperflex total knee design improve postoperative active flexion ? *Orthop Traumatol Surg Res* 2011 ; 96 : 376-380.
14. **Seon JK, Park SJ, Yoon TR *et al.*** The effect of anteroposterior laxity on the range of movement and knee function following a cruciate-retaining total knee replacement. *J Bone Joint Surg* 2010 ; 92-B : 1090-1095.