X-ray Technologists' Reproducibility from Automated Measurements of the Medial Tibiofemoral Joint Space Width in Knee Osteoarthritis for a Multicenter, Multinational Clinical Trial

J. CHRIS BUCKLAND-WRIGHT, CHARLES F. BIRD, CATHY A. RITTER-HRNCIRIK, GARY A. CLINE, CAROL TONKIN, THOMAS N. HANGARTNER, RUPERT J. WARD, JOAN M. MEYER, and MICHAEL P. MEREDITH

ABSTRACT. ***Objective.*** To determine the reproducibility of x-ray technologists, 26 in North America (NA), 24 in Europe (EU), in reliably repositioning patient's osteoarthritic (OA) knees, from computerized measurements of minimum joint space width (JSW) and reproducibility in joint repositioning, during their training for the clinical trial.

Methods. Technologists from 12 NA and 12 EU clinical radiology units received identical training, at one site on each continent, in performing the fluoroscopically assisted semiflexed knee examination and in quality control criteria (QCC) for film acceptance. Subjects recruited were 129 in NA and 70 in EU, with both knees radiographed for some subjects. Each technologist radiographed 5 OA knees and repeated the process on the same knees 2 to 7 days later. Minimum medial JSW was measured at a single center on digitized images with computer software that corrected for radiographic magnification. Technologists' reproducibility in joint repositioning and JSW measurement was determined from the difference between test and retest.

Results. In all, only 3/50 technologists failed qualification criteria with a repeat-film JSW difference > 0.50 mm. The mean, standard deviation (SD) of the difference in JSW between test/retest for 146 NA film-pairs of –0.020 (0.16) mm was not statistically different from that in 120 EU film-pairs: –0.001 (0.18) mm. In NA and EU 45% of examinations achieved high quality, i.e., JSW difference between repeat films < 0.1 mm, and 92% achieved excellent to good quality with a difference between repeat films < 0.3 mm. NA and EU technologists' reproducibility was unaffected by subject's sex, age, and degree of JSW loss. Reproducibility in joint reposition for all technologists was excellent.

Conclusion. Between-continent precision of JSW measurements from all accepted pairs of semiflexed views was excellent to very good and similar to the high technical quality achieved in the authors' original report. The value of training incorporating both test/retest radiographs and film QCC is essential for the high technical quality required for multinational clinical trials. (J Rheumatol 2003;30:329–38)

Key Indexing Terms:
OSTEOARTHRITIS KNEE RADIOLOGY CLINICAL TRIAL

From the Department of Applied Clinical Anatomy, King's College London, School of Biomedical Sciences, London, UK; Departments of Arthritic New Drug Development and Biometrics and Statistical Sciences, Health Care Research Center, Procter & Gamble Pharmaceuticals, Cincinnati, Ohio; and BioMedical Imaging Laboratory, Wright State University, Dayton, Ohio, USA.

J.C. Buckland-Wright, PhD, DSc, Professor of Radiological Anatomy, University of London, King's College London; C.F. Bird, BSc, MB, BS, Research Associate; R.J. Ward, BSc, MPh, Research Associate, Department of Applied Clinical Anatomy, King's College London; J.M. Meyer, PhD; C.A. Ritter-Hrncirik, BS, Arthritic New Drug Development Department, Procter & Gamble Pharmaceuticals; M.P. Meredith, PhD, Senior Research Fellow; G.A. Cline, PhD, Senior Biostatistician, Biometrics and Statistical Sciences Department, Procter & Gamble Pharmaceuticals; C. Tonkin, DCR, Senior Radiographer, Department of Applied Clinical Anatomy, King's College London; T.N. Hangartner, PhD, Brage Golding Distinguished Professor of Research, Wright State University.

Address reprint requests to Prof. J.C. Buckland-Wright, Department of Applied Clinical Anatomy, King's College London, Hodgkin Building, School of Biomedical Sciences, London Bridge, London SE1 1UL, UK. E-mail: chris.buckland-wright@kcl.ac.uk

Submitted February 5, 2002; revision accepted August 1, 2002.

This article is one of a series reporting the results of a 2 year multicenter, placebo controlled, double blind, randomized dose-response study conducted in Europe and North America of risedronate treatment in patients with mild to moderate medial knee osteoarthritis (OA). The primary efficacy objective is to assess whether risedronate can slow the progression of joint space narrowing (JSN) in the signal knee relative to placebo treated patients. We report the reproducibility of computerized measurements of minimum joint space width (JSW) in the medial tibiofemoral compartment of OA knees radiographed in the fluoroscopically assisted semiflexed view, obtained by x-ray technologists in Europe and North America during their training for this multicenter clinical trial.

Guidelines for clinical trials in patients with knee OA recommend the use of radiography[1,2] for determining the response to treatment, in which the minimum medial tibiofemoral compartment JSW or interbone distance is a

surrogate measure for the thickness of articular cartilage. Measurement of this feature assesses the combination of both the material thickness of articular cartilage and its compressibility when the joint is radiographed in the standing position[3]. In clinical trials, radiographic protocols ensure consistency of radioanatomic positioning of the joint both within and between patients at repeat examination over time. Protocols that ensure that the knee is anatomically repositioned correctly at each examination use fluoroscopy, prior to obtaining the x-ray image, to visualize the knee during the procedure in order to align the central ray of the x-ray beam so that it is parallel with the surface of the medial tibial plateau and that the tibial spines are centrally placed with respect to the femoral intercondylar notch. Of the different radiographic protocols that employ fluoroscopy for positioning the knee, i.e., semiflexed[4,5], extended[6,7], and tunnel or schuss[8] views, it is only in the standing semiflexed view that the site of the medial compartment JSW measurement is in the center of the compartment and coincides with the region of normal functional load in the joint[9]. At this site minimum JSW reliably assesses cartilage thickness and status[3,9]. This method, in combination with the computer software for obtaining minimum JSW measurement in digitized radiographic images[10], affords notable precision of this measurement in the tibiofemoral compartment of OA knees[5,11-14].

The possibility of disseminating this protocol so as to be performed with similar reproducibility within clinical radiology units has been confirmed[13]. OA and normal knees were imaged twice within one week by each of 2 teams of x-ray technologists who had been trained in the semiflexed view[4,5]. Although the authors did not achieve the same estimates of measurement precision as that reported by Buckland-Wright, *et al*[5], they commented that this was probably due to the uniformly high level of technical quality of radiographs in the original report[5]. For a large clinical trial of the size required in the present multicenter study conducted in Europe and North America, in which the primary outcome measure is to employ the same technology (measurements of the medial compartment minimum JSW taken from fluoroscopically positioned standing semiflexed knee radiographs), considerable attention has to be paid to the standardization and quality of the training received by each of the x-ray technologists employed in the trial.

We report the measurement precision obtained by 26 North American and 24 European x-ray technologists who followed an identical training program at separate radiology facilities on the 2 continents at Miami Valley Hospital, Dayton, Ohio, and at Guy's Hospital, London. Each technologist had to meet the criteria for qualification determined by the Qualification Protocol for the study, whose aim was "to ensure that all participating technologists could position and repeat precisely the standing semiflexed knee technique"[4,5]. They were provided with an illustrated training manual in the language of the country in which they worked, as well as lectures and practicals on the methodology and the clinical trial. They each received training in the procedures necessary to maintain high quality control of their performance. Further, each technologist had to radiograph the OA knee of 5 patients during training and to repeat the process under supervision within a 2 to 7 day period.

MATERIALS AND METHODS

Approval for the study was obtained at each institution at which training was undertaken. This was provided by the Ethical Committee of Guy's and St. Thomas's Hospital Trust, London, and by the Institutional Review Boards of Miami Valley Hospital and Wright State University, Dayton, Ohio.

Patients. All technologists were trained on patients who had knee OA. One hundred twenty-nine patients were recruited at the Miami Valley Hospital (68 women, 61 men) with a median age of 60 (range 40–79) years, and 70 at Guy's and St. Thomas's Hospital (43 women, 27 men) with a median age of 63 (range 40–78) years. All had OA of the medial tibiofemoral compartment, as evidenced by a clinical history of knee OA and radiography, with knee pain on most days during at least one month out of 3 prior to the initial visit. Patients were excluded if they had lateral compartment disease only and if they had no JSW remaining in the medial compartment, i.e., bone on bone, when radiographed in the semiflexed view, since the total JSW loss is not always detected in the extended knee view[9].

North American and European clinical x-ray technologists. Two clinical x-ray technologists were provided by each of the 10 and 3 from 2 regional radiographic facilities participating in the clinical trial in North America (NA). The 26 technologists (22 women, 4 men) had a mean age of 39 (SD 12, range 21–55) years and a mean of 15 years net experience (SD 11.3, range 1–32). Similarly, 2 technologists were provided by the 12 European (EU) regional radiographic facilities from 2 sites in the United Kingdom, Germany, and Italy and one site in Ireland, France, The Netherlands, Slovenia, Czech Republic, and Sweden. The 24 technologists (20 women, 4 men) had a mean age of 42 (SD 10.4, range 26–58) years and a mean of 18 years net experience (SD 9.8, range 4–36). All technologists traveled to their respective training centers in NA and Europe.

Standardization of the NA and European training. To standardize the training between NA and Europe, 2 technologists from Dayton who were to undertake the instruction in NA visited Guy's Hospital for an initial introduction and training by the senior technologist (CT) responsible for this method. At the commencement of the NA training program, CT visited Dayton to formally instruct the 2 technologists so that they met the qualification criteria for participation in the clinical trial. Following their training, their performance as instructors was supervised by CT for one week.

Radiographic procedure. This was undertaken according to the protocol by Buckland-Wright[4,5]. All radiographs were obtained with equipment comprising an overhead fluoroscopic image intensifier and an x-ray tube located under the examining table, which was tilted vertically, permitting an upright weight bearing examination, with patient support provided by the handles fitted to the edge of the table (Figure 1). The fine-focus setting was selected on the x-ray control panel. Each knee was imaged separately on a 9 × 10 inch film placed in a film holder located in front of the image intensifier. This ensured that, as soon as the knee had been screened into the correct position, there was no delay in obtaining the image. Radiography of the knee in the semiflexed view can be obtained with either an overhead or underbed fluoroscopic tube, since the perspective of the knee radiographed in either the anteroposterior (AP) or posteroanterior (PA) views remains similar, provided the x-ray tube is operated at the fine-focus setting and the image is corrected for radiographic magnification[5].

Radiography of the knee in the standing semiflexed view[5]. Prior to radiography a metal ball (5 mm diameter) encased in methyl methacrylate was

Figure 1. The radiographic examination table tilted vertically, with the fluoroscopic image intensifier and film carrier to the right of the patient standing on a platform. The x-ray tube is located under the examination table.

taped to the side of the knee for calculating radiographic magnification. Each knee was flexed until the tibial plateau was horizontal relative to the floor, parallel to the central x-ray beam, and perpendicular to the x-ray film. The center of the joint, defined by the joint space, was aligned with the center of the x-ray beam with the aid of the tube's positioning light. The precise position of the knee was obtained visually with the aid of fluoroscopy. With the heel fixed, the foot was internally or externally rotated until the tibial spines appeared centrally placed relative to the femoral notch; then the knee was flexed to achieve superimposition (± 1 mm) of the anterior and posterior margins of the medial tibial plateau. The average duration of fluoroscopic examination at the first visit was ≤ 2 minutes. The length of this fluoroscopic examination was necessary for teaching the technologists. Immediately after x-ray exposure, the outline of the foot was drawn on a large sheet of paper taped to the platform, to facilitate joint repositioning and to reduce time taken during fluoroscopy at subsequent visits to < 30 seconds. The position of the sheet of paper on the platform, at repeat examination, was achieved by aligning marks placed on the sheet at the initial examination with permanent marks visible on the platform.

Training procedure. Eight was the maximum number of technologists attending a training session lasting about one week. The training commenced with a slide presentation describing the scientific basis to the radiographic method and quality control criteria (QCC) (Table 1). The technologists were provided with their own R and L markers bearing personal identification numbers, magnification markers as described above, and a detailed manual. The manual described the radiographic process and quality control issues and contained illustrations of the optimum knee position and the technologist's exclusion criteria of OA knees, i.e., images with tibiofemoral bone on bone and lateral compartment disease. After seeing a short video on the technique, the technologists commenced hands-on training. Each one radiographed the OA knee of 5 patients under instruction from the technologist responsible for training. After a period of 2 but not more than 7 days, the technologist radiographed the same knees of the 5 patients in the presence of the instructing technologist. During the latter procedure the instructing technologist did not interfere, guide, or comment upon the trainee's performance, but simply monitored their work. At each patient visit, a maximum of 3 attempts was permitted to achieve an image meeting the QCC. Throughout the training, 20% of the patients had one and 5% had 2 repeat exposures. At the repeat examination, where more than one radiograph was acquired, the image closest to the previous radiograph was accepted. Failure to meet the QCC criterion and/or that for JSW measurement described below for any one of the 5 patient examinations resulted in the technologist failing the training session. Any technologist failing 3 sessions was deemed ineligible for qualification. Throughout the training, technologists not under instruction were encouraged to observe their colleagues, thereby increasing their experience. During the week, the technologists also received further instruction on the QCC (Table 1). They undertook, on 2 separate days, self-assessment of 26 knee films comprising the entire range of errors listed in the QCC. The technologist also received short lectures on OA and good clinical practice.

Quality control. All films produced by the technologists were sent to the Regional Quality Control Centre, where they were inspected by an experienced radiologist who assessed their acceptability according to the QCC (Table 1). All accepted films were sent to the Central Analysis Facility at King's College London for JSW measurement.

Table 1. Quality control criteria for radiograph acceptance: technologist's checklist.

Once the radiograph has been developed, check the following points. The radiograph must meet all the following criteria to ensure that it conforms to the quality control standards necessary for this study.

1. Anterior and posterior margins of the tibial plateau must be superimposed as closely as possible (i.e., 1 mm of each other at the mid-point of the medial compartment).
2. The anterior rim of the tibial plateau must be aligned with the posterior rim of the tibial plateau, rather than any posterior osteophytes.
3. The tibial spines must be central relative to the femoral notch.
4. There should be no evidence of patient movement.
5. The knee joint should be in the center of the film. Only one knee should appear on each film.
6. The long bones should be parallel to the long axis of the film (i.e., the patient ID should appear at the top of the film).
7. The metal sphere should appear level with the head of the fibula.
8. The metal sphere should not be too close to the edge of the x-ray image (the metal ball should be at least 5 mm from the edge).
9. The beam coning should not cut off the metal sphere.
10. The metal sphere should not overlie any bone or the joint space.
11. The metal ball should have distinct edges and be clearly circular to the naked eye.
12. The patient ID label should be legible and correctly positioned, without overlapping the edges of the film. Only ONE label per film.
13. The patient ID label should contain the required information.
14. The appropriate Right or Left marker should be present, correctly positioned adjacent to the patient ID label, and bear the correct radiographer ID.
15. No superfluous markings (e.g., writing or punch marks) should be present.
16. The radiograph should be neither over nor under-exposed. The outer margins of the femur and tibia should be clearly visible.

Joint space width measurement. A fully automated computerized method of JSW measurement was used in this study that did not depend upon manual intervention other than to feed the films into a Lumiscan 200 laser film digitizer (Lumisys Inc., Sunnyvale, CA, USA). Before digitization all films were bar coded to ensure that on digitization the computer database linked patient/visit data to the JSW measurement obtained from each radiograph. Each film was scanned twice; on the first the software programs located and measured the diameter of the metal ball. On the second scan, the film was digitized to a scale corrected for the effect of radiographic magnification. Minimum JSW was measured in the medial compartment from each of the radiographs using the automated computerized method of measurement[10,11]. Occasionally manual intervention was required in those instances where the radiographic quality of the film prevented operation of the automatic JSW measurement software. The coefficient of variation for the reproducibility of the software to measure medial compartment JSW had been determined previously as 1% for test-retest repeat radiographs of the knee in the semiflexed position[5].

Qualification of the technologist. To qualify for the clinical trial each technologist had to produce pairs of radiographic images of the same 5 patient knees that met the QCC and in which all film pairs had to have a difference in minimum medial compartment JSW of ≤ 0.5 mm. This threshold of acceptability was based upon published data[13] in which the underlying standard deviation of technologists' intra-subject differences in JSW was assumed to be between 0.25 and 0.33 mm for conventional radiographic equipment[13]. Under these assumptions, a sample size of 5 subjects per radiographic technologist will provide high probability that a technologist with a variance of 2 to 3 times that acceptable for qualification into the clinical trial will not be able to reliably repeat their measurement within the cutoff of 0.5 mm for the intra-subject JSW difference for all 5 subjects; whereas a well trained technologist had a high probability of successfully completing 5 subjects.

Joint flexion in the anterior-posterior plane. Accurate radioanatomical position of the joint relative to the central ray of the x-ray beam is obtained when the beam is parallel to the tibial plateau and the anterior and posterior margins of the medial tibial plateau appear superimposed in the radiographic image[4,5]. The inaccuracy of the medial tibial plateau leveling was measured by the distance between the anterior and posterior rims of the tibial plateau and the floor of the tibial plateau. Because of the image quality in standard radiographs, this method was devised to overcome the difficulty in reliably measuring the distance between the tibial rims, particularly where these are close together. Measurements were taken using a PC-driven Kontron digitization program (KS-100; Kontron Electronik GmbH, Eching, Germany) linked to a back-illuminated digitization tablet and a cross-wire cursor. The coefficient of variation (CV) for this method for repeat linear measurements was 3.7%[14]. For the study, measurements were taken from each of the anterior and posterior rims of the plateau to a point midway along the superior margin of the bright radio-dense line corresponding to the floor of the tibial plateau (Figure 2). Since it was not always possible to distinguish between the rims (due to their superimposition in the radiograph), the absolute difference of these distances was used as the "tibial leveling measurement." The median and 95% CI of the tibial leveling measurements obtained from the first and second radiographic visits were calculated for each knee, and their analysis is described below.

Data analysis. Pairs of films for which the radiographer did not meet the entrance criteria of a repeat-film JSW difference > 0.50 were not included in the analyses with the exception of the comparison between NA and EU, where the analyses were done with and without the failed pairs of films. The method described by Bland and Altman[15] was used to estimate the root mean square error (RMSE) from an ANOVA model with the knee as the explanatory variable and the JSW as the dependent variable. The CV is calculated by dividing the RMSE from the model by mean JSW. The significance level was 0.05 for all tests and all tests were 2 sided.

Mean and SD of the difference in minimum medial compartment JSW measurement between the test and retest were calculated for the film pairs at the NA and EU sites. The ANOVA model was used to test for the difference in the mean of the difference in JSW values between NA and EU. Bartlett's test was used to assess the difference in the mean and SD of the difference in JSW values between NA and EU. The average JSW between test/retest was calculated to determine the range of JSW in patients on the 2 continents; the ANOVA model was used to determine any statistical difference between the means.

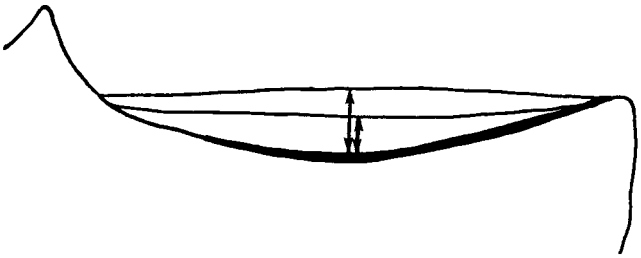

Figure 2. The radiographic appearance of the medial tibial plateau showing the measurements taken, in each radiograph, from the floor of the articular fossa to the anterior and posterior rims, respectively. Correct radioanatomic position was obtained when both rims appeared superimposed in the radiographic image.

The effect of patient characteristics in NA and EU upon the reproducibility of JSW measurement was evaluated for sex and age. The mean and 95% confidence interval (CI) minimum medial compartment JSW and the mean and SD of the difference in JSW between test and retest was calculated for the knees in each category, the ANOVA model was used to determine any difference in the means between the different sex and age groups. Bartlett's test was used to assess the difference in the mean and SD of the difference in JSW values between the sex and age groups.

Between-radiographic training site reproducibility of automated measurements of minimum JSW in the medial tibiofemoral compartment was determined among the paired radiographs for all technologists. A further analysis was performed among pairs of radiographs in which the radiographic quality was graded according to the test/retest difference in JSW measurements into those corresponding to technologists who had an excellent and a good performance, as represented by a difference in JSW measurement between paired films of < 0.1 and < 0.3 mm, respectively. The ANOVA model was used to test for the differences in the mean of the difference in JSW values between the categories. Bartlett's test was used to evaluate the difference in the mean and SD of the difference in JSW values between the categories.

For medial compartment tibial plateau leveling with respect to the alignment of the x-ray beam, the absolute difference between anterior and posterior tibial rim heights was calculated. For each pair of repeat visit radiographs, the mean and SD of these 2 measures were calculated and used to represent the degree and the reproducibility of tibial plateau leveling, respectively. For both these measures the lower boundary of zero corresponded to cases where the 2 rims were perfectly superimposed. Because these means and SD were not normally distributed, medians and 95% CI were used to summarize their distribution. The Kruskal-Wallis test was used to assess the statistical significance of the difference in tibial plateau leveling between NA and EU.

RESULTS

All technologists met the qualification criteria apart from 3 in NA whose difference in the minimum medial compartment JSW measurement between repeat radiographs of the same OA knee, in one out of 5 patient examinations, was > 0.5 mm. The failed technologists did not participate in the trial. The mean (SD) of the difference in JSW measurement

between the test and retest for the 146 film pairs obtained at the NA site was –0.020 (0.159) mm, and for the 120 film pairs obtained at the EU site was –0.001 (0.180) mm. There was no statistically significant difference between NA and EU in either the mean JSW difference (p = 0.37) or the standard deviations (p = 0.15) between the test and retest films. Additionally, when the 3 failed radiographs were included in the analysis, there was no statistically significant difference between NA [n = 149; –0.0003 (0.212) mm] and EU in either the mean JSW difference or the standard deviations between the test and retest films. This similarity in NA and EU technologists in their performance during training, as determined by the difference in JSW between test and retest radiographs, is illustrated in Figure 3.

JSW measurement in patients at NA and EU training centers. The average JSW between the test/retest was used to determine the patient's JSW. The range in JSW in patients examined on the 2 continents showed remarkable similarity (Figure 4), with no statistically significant difference between the values from the 2 sites. The larger values for minimum JSW (> 4.5 mm) represent the measurements obtained from the healthy knees of patients in whom both knees had been imaged and of which only one joint was diseased.

Factors affecting JSW measurement reproducibility. The effect of patient characteristics upon the reproducibility of JSW measurement was evaluated. The results presented in Table 2 show that there was no statistically significant difference in the JSW measurement reproducibility either between sexes (p = 0.63 for the mean, p = 0.42 for the SD) or between the different age groups (p = 0.19 for the mean, p = 0.78 for SD). Further, no statistically significant difference in the JSW measurement reproducibility was detected between any of the categories of JSN (p = 0.66 for the mean, p = 0.68 for SD), indicating that the technologists could reliably reposition the knee at repeat visits and that this was independent of the degree of joint space narrowing and hence the severity of knee OA (Table 3).

Combined site reproducibility. The values for the reproducibility of JSW measurements among the paired radiographs for all technologists at the training sites in NA and EU were extremely high in the entire sample and in the radiographs, reflecting the different levels of technologists' radiographic quality (Table 4). Following the technologists' training, 46% of the examinations achieved a radiographic quality that was excellent in that the difference in JSW measurements between repeat radiographs was < 0.1 mm, and 92% of the examinations achieved a radiographic quality that ranged from excellent to good, i.e., difference in JSW measurement between repeat films < 0.3 mm. The difference in performance between the 3 different levels of radiographic quality was statistically significantly different for both the mean (p = 0.036) and SD (p = 0.0001) of the degree of JSW difference between the test and retest films. Among all films the RMSE was 0.119. Additionally, when the 3 failed radiographs were included in the analysis, the RMSE was 0.140.

Reproducibility in radioanatomic joint positioning. The

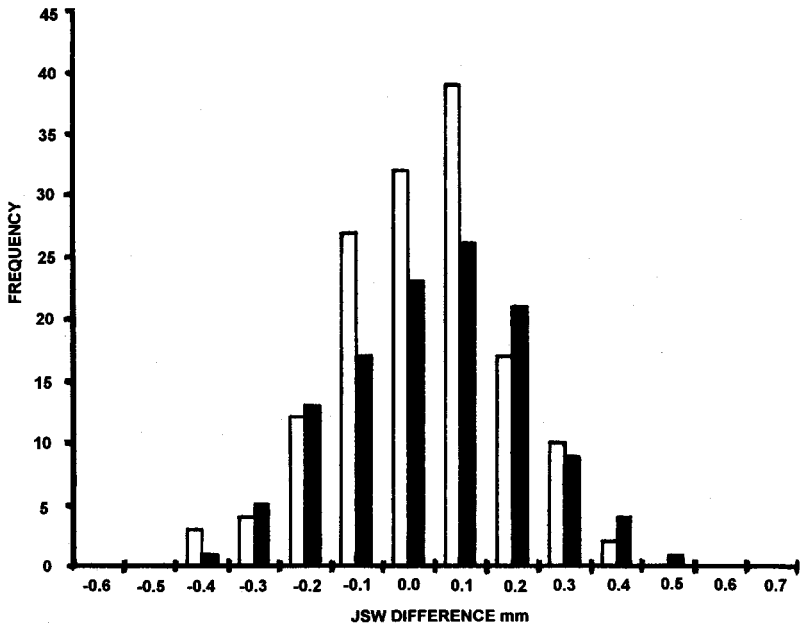

Figure 3. Similarity in the North American (white columns) and European (black columns) technologists in their performance during training, as determined by the difference in the minimum medial compartment JSW measurement between repeat radiographs of the same OA knee.

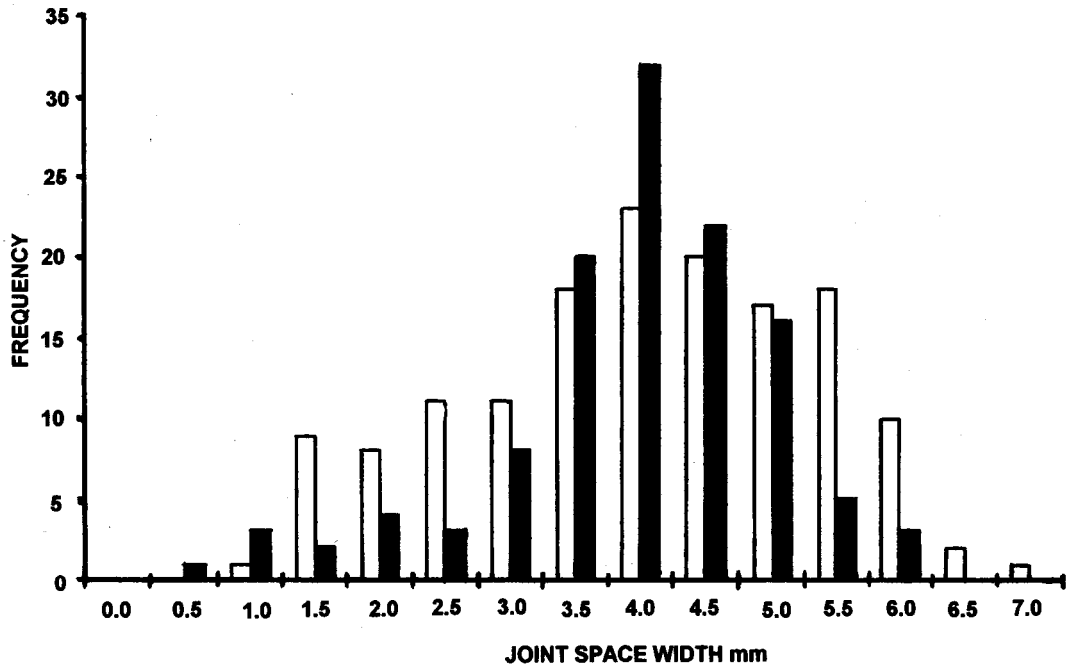

Figure 4. The range in minimum medial compartment JSW measurements in patients radiographed at the North American (white columns) and European (black columns) radiographic facilities were remarkably similar.

Table 2. The effect of patient characteristics in NA and EU upon the reproducibility of JSW measurement was evaluated for sex and age. The number, mean and 95% confidence interval (CI) minimum medial compartment JSW is given for the knees in each category, together with the mean and SD of the difference in JSW between test and retest. There was no statistically significant difference in the JSW measurement reproducibility between either sex or the different age groups.

Sex/Age Group	No. of Knees	Mean JSW (95% CI), mm	JSW Difference for Test/Retest Mean, mm	SD, mm
Female	148	3.6 (3.4, 3.8)	–0.016	0.174
Male	118	3.9 (3.7, 4.2)	–0.006	0.162
< 50 yrs	44	4.2 (3.9, 4.6)	–0.039	0.171
≥ 50 < 60 yrs	79	3.9 (3.6, 4.2)	–0.034	0.173
≥ 60 < 70 yrs	75	3.6 (3.3, 3.8)	0.0010	0.172
≥ 70 yrs	68	3.4 (3.2, 3.7)	0.008	0.155

technologists' reproducibility in the radioanatomic repositioning of the knees gave a median (95% CI) of the absolute value of the difference in tibial rim alignment between test and retest for the NA site of 0.03 (0.03, 0.04) mm and for the EU site of 0.25 (0.05, 0.29) mm. Although there was a statistically significant difference in this value between NA and EU ($p = 0.0062$) due to a greater variability in tibial plateau leveling in the latter, nevertheless, all technologists achieved good to excellent reproducibility in tibial rim alignment in obtaining superimposition of ≤ 1 mm of the anterior and posterior margins of the medial tibial plateau (item 1 of QCC, Table 1). The mean and SD of the JSW measurement between paired films on test/retest was grouped according to the different levels of radiographic quality corresponding to differences in tibial rim alignment between paired films of < 0.1, ≥ 0.1 and < 0.3, ≥ 0.3 mm, respectively (Table 5). This showed that the narrow range of variability in tibial plateau leveling had no statistically significant effect upon the precision of the JSW measurement, as shown by the similarity of the value of JSW SD in the different categories.

Table 3. For all patients, the effect of the degree of joint space narrowing (JSN) upon the reproducibility of JSW measurement was assessed. The number of knees is given for each category, together with the mean and SD of the difference in JSW between test and retest. No statistically significant difference in the JSW measurement reproducibility was detected between any of the categories of JSN.

Degree of JSN	No. of Knees	JSW Difference for Test/Retest (mm) Mean	SD
< 1.0	6	0.010	0.114
≥ 1 < 2	21	–0.037	0.134
≥ 2 < 3	32	–0.043	0.166
≥ 3 < 4	90	0.000	0.179
≥ 4 < 5	77	0.006	0.174
≥ 5 < 6	37	–0.040	0.159
≥ 6 < 7	3	0.023	0.156

Table 4. Combined radiographic training site reproducibility of automated measurements of minimum JSW in the medial tibiofemoral compartment among all paired radiographs and among pairs of radiographs in which the radiographic quality was graded, according to the test/retest difference in JSW measurements, into those representing excellent and good, corresponding to differences in JSW measurement between paired films of < 0.1 and < 0.3 mm, respectively.

Paired Radiographs	JSW Absolute Value of Difference Between Paired Radiographs	No. of Knees	JSW Difference for Test/Retest Mean mm	SD mm	CV, %	Root MSE
All pairs		266	–0.012	0.169	3.18	0.119
Excellent quality	< 0.1	121	0.005	0.057	1.08	0.040
Good quality	< 0.3	244	–0.004	0.139	2.62	0.098

MSE: mean square error.

Table 5. Combined radiographic training site reproducibility of automated measurements of minimum JSW in the medial compartment among all paired radiographs and among pairs of radiographs in which the radiographic quality was graded according to the absolute difference in tibial rim alignment into those representing excellent, good, and moderate, corresponding to absolute difference in tibial rim alignment between paired films of < 0.1, ≥ 0.1 and < 0.3 and > 0.3 mm, respectively.

Paired Radiographs	Absolute Value of Tibial Rim Alignment Difference Between Paired Films	No. of Knees	JSW Difference for Test/Retest Mean mm	SD mm	CV, %	Root MSE
All pairs		266	–0.012	0.17	3.18	0.119
Excellent quality	< 0.1	145	–0.012	0.18	3.36	0.124
Good quality	≥ 0.1 < 0.3	32	–0.026	0.13	2.46	0.094
Moderate quality	≥ 0.3	89	–0.005	0.17	3.14	0.119

MSE: mean square error.

DISCUSSION

In clinical trials of knee OA, where radiography is used to evaluate changes in articular cartilage thickness in the diseased medial tibiofemoral compartment and in which the primary outcome measure is derived from the difference in the rate of joint space narrowing (JSN) between treatment groups, it is essential to use validated standardized radiographic and measurement protocols[1]. Of the published radiographic methods that ensure reproducible radioanatomic positioning employing fluoroscopy to achieve precise reposition both within and between patients, only one has been evaluated in a field test undertaken in NA to compare its transferability from the laboratory of origin in the UK. Mazzuca, *et al* showed that the technology of fluoroscopic positioning of the knee in the semiflexed view is transferable between laboratories[13]. However, in practice, they found that they were unable to achieve quite the same levels of JSW measurement reproducibility as that originally reported[5]. They attributed this to 2 factors, the variable quality of the radiographs as a possible reflection of the adequacy of their training and their technologists' inability to effectively perform quality control assessment of their films, leading the authors to recommend that quality control should be assigned to an independent observer such as a musculoskeletal radiologist[13]. With respect to the technologists, they were not convinced that "more extensive training and/or higher standards of demonstrated proficiency of the technologists would have increased the overall technical quality of the radiographs..."[13], a conclusion that this study does not support since improved technologists' training resulted in a quality of technical performance that was good to excellent.

The training of a large number of technologists (50) for this multinational clinical trial, recruited from a geographically wide distribution in both NA and EU, required careful attention to the standard of training to ensure consistency. The use of an experienced technologist familiar with the methodology facilitated this objective. The results show each qualifying technologist achieved a high level of overall technical quality. The values for the precision they attained (SD 0.16 mm for NA and 0.18 mm for EU) are almost identical to that reported in the original study (SD 0.19 mm) describing this fluoro-assisted semiflexed view of the knee[5]. Thus we confirm that more extensive and higher standards of training of technologists do result in overall high technical quality. It was only those technologists who had received this training that were the sole operators for the ongoing multinational, multicenter clinical trial. During the course of this trial, experienced senior technologists who routinely visit their respective sites in NA and EU are monitoring the radiographic standards of the technologists.

The setting of a threshold value for designating the acceptance/rejection of a technologist's technical quality permitted those that were likely to underperform in the

clinic to be either excluded or retrained. In particular, it was important to give the technologists the responsibility for the quality of the radiographic image and to ensure that it met the criteria listed in the QCC (Table 1). Training technologists in this aspect of the procedure enhanced their awareness of the different elements that determine the quality of an image they produce. As a consequence of introducing QCC into the training, the Regional Quality Control Centre passed all training films as acceptable. It was at the Central Analysis Facility at King's College London that measurement of JSW in the test/retest films identified the 3 technologists who failed the qualification criterion.

Since the operational characteristics of the x-ray units can vary between the different radiographic facilities, technologists were made aware that the knee could be radiographed in either the AP or PA views. In either, the image of the knee was similar, and any differences due to radiographic magnification were corrected by means of the metal ball marker. Further, it was emphasized that the film carrier on the front of the fluoroscopic tube must be used to obtain the radiographic image, so as to minimize the time between fluoroscopic positioning of the knee and acquiring the final image, and thus reduce any risk of patient movement. Alternative methods were contraindicated, such as those units using 2 x-ray tubes, one to screen the knee into position that is moved out and replaced by a tube used to obtain the image. This process is slower and also increases the risk of patient movement.

Our results confirmed those of the previous study[13] that reproducibility of JSW measurement in the test/retest radiographs was unrelated to the patient's age, sex (Table 2), and radiographic severity of OA (Table 3). In this instance, we used an improved method for categorizing the degree of disease severity compared to that of Kellgren and Lawrence[16], by employing degrees of JSN measured in millimeters (Table 3).

Reproducibility in radioanatomic joint positioning. Under fluoroscopic examination, reproducibly repositioning the joint radioanatomically relative to the x-ray beam affects the reproducibility of JSW measurements required in clinical trials[5]. Two components define the radioanatomic position of the knee: the degree of internal/external rotation and the alignment of the x-ray beam with the joint space so that the tibial plateau is level and parallel to the central ray of the beam. Here, joint rotation is minimized by centering the tibial spines relative to the femoral notch, and use of the foot map ensures reproducible repositioning. However, tibial plateau alignment is dependent upon the technologist's skill in aligning, under fluoroscopic examination, the medial tibial plateau with the central ray of the beam such that the anterior and posterior margins of the plateau appear superimposed. Comparison between radiography visits showed that all technologists were able to reproducibly reposition patients' knees within an absolute value of the difference in tibial rim alignment of ≤ 1 mm, as required by the QCC (Table 1). Although the technologists at the EU sites had a significantly greater variability in tibial plateau leveling than those in NA, overall the reproducibility in tibial plateau leveling achieved by all technologists was excellent. Further, the relationship between the degree of tibial plateau leveling and the reproducibility in JSW measurement between paired films (Table 5) showed that the latter was unaffected by variation in tibial rim alignment. Consequently, the variations in JSW measurement reproducibility reported in Table 4 were not due to positioning defects such as joint rotation, since this was controlled by the use of the foot map, and we have shown previously that rotation has to be very large (> 15°) to have any effect upon JSW measurement precision[10], or to small variations in tibial rim alignment, as shown in Table 5, but to other factors affecting the radiographic appearance of the medial compartment. These factors may include, *inter alia*, joint laxity from articular cartilage loss or altered biomechanical status of the articular cartilage. Determining which factors may produce differences in radiographic quality would require further studies.

Implications for multicenter clinical trials in knee OA. The standards of radiographic performance achieved during the technologists' training at the NA (SD of repeated measurements = 0.16 mm, CV 3.0%) and EU (SD of repeated measurements = 0.18 mm, CV 3.5%) sites are similar to or better than measurements reported by Buckland-Wright, *et al*[5] in the original description of this method for positioning the knee in the semiflexed view (SD of repeated measurements = 0.19 mm, CV 5.5%). These results contrast with those described in the initial field test of the reproducibility of this method[13], where the analogous measurements were larger (standard error of the mean for JSW measures repeated within-unit = 0.32 mm, CV 8.7%). The results from the present study confirm the suggestion that good reproducibility of JSW measurements in test/retest radiographs is due to the uniformly high level of technical quality of the technologists[13]. We consider that the success we have had in training the technologists to a high level of technical quality has been the requirement of each technologist to undertake test/retest radiographs of 5 patients' knees in order to qualify for admission as a study technologist and in receiving practical QCC training. To date we are unaware of any other clinical trial that has followed this procedure.

As a consequence of the consistency in the quality of technical performance in reproducibly repositioning knees in the semiflexed position, there will be a large effect on the power and/or efficiency of a multicenter structure-modifying OA drug (SMOAD) trial to detect JSN in knee OA. The following illustration estimates the number of knees/group required for a 2 year trial with a statistical power of 80% (2 tailed $\alpha = 0.05$) to detect a 30% reduction in the rate of JSN in the SMOAD treatment group in

comparison with the placebo group[13], assuming a mean rate of JSN in the placebo group of 0.20 mm/yr suggested by the literature[17,18] and a conservative estimate for SD for the rate of JSN of 0.45 mm, which is based on an increase in the value previously used in such calculations of SD = 0.25 for the rate of JSN[5,17]. Under these circumstances, 222 knees/group (Figure 5) would be required to detect a therapeutic effect.

By adjusting the component of the variance term in the sample size formula that is due to measurement error, it is possible to estimate the sample size over a range of values for technologists' reproducibility (Figure 5). If the SD of the mean for repeat measures SD_m = 0.057 mm (CV 1.08%, RMSE 0.040), as was obtained from radiographs that were consistently of "excellent" quality (Table 4), the number of knees required to finish the study would be 195 per group (Figure 5C). Where the technical quality was "good" (Table 4), corresponding to SD_m = 0.139 mm (CV 2.62%, RMSE 0.098) attained in 92% of the pairs of radiographs in this study, the number of knees required would increase to 213 per group (Figure 5B). These sample sizes are for studies in which the qualification threshold for technologists' performance is increasingly stringent.

It is on the basis of these estimates that the numbers of patients for SMOAD clinical trials should be determined, since it takes into account variability likely to accrue throughout the course of a study, rather than possibly optimizing figures achieved during the training program. Importantly, among the extraneous variables that would contribute to JSW measurement precision is the likelihood that the high reliability achieved by the technologists in a training session may not be retained once they return to their own radiographic facility. Such factors that affect the performance of technologists in the field are the subject of a separate publication. This method of determining patient numbers based on a conservative scenario has the advantage that the primary radiographic outcome measure of the difference in the rate of JSN between the active and placebo arm is more likely to achieve the desired statistical power than one calculated simply on the values obtained at the time of training the technologists. In addition, Figure 5 shows that even if JSW measurement precision could be reduced to zero, a significant number of knees are required in each group due to population variability.

In conclusion, this study has shown that a well devised training program for technologists incorporating both the radiographic procedure and quality control assessment of the radiographs they produce can result in uniform production of films of excellent to very good technical quality. We consider that training technologists so that they undertake quality control of their own radiographs enhances their standard of radiography and encourages them to become involved in and contribute toward the clinical trial. Further, this study has shown that the same degree of technical

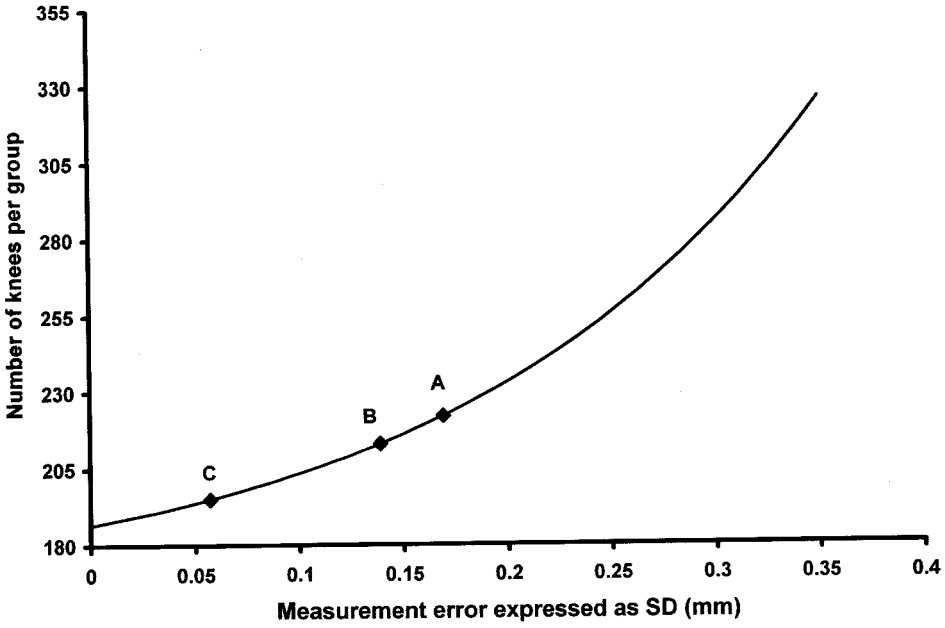

Figure 5. Number of knees/group required in a hypothetical structure-modifying OA drug trial to detect a 30% therapeutic effect within a 2 year study period from minimum JSW measured in the standing semiflexed view for a range of values for JSW measurement error (the SD of the mean for repeat measures). The sample sizes for studies in which qualification threshold for technologists' performance is increasingly stringent are identified as that required with all technologists (A) and those that are "good" (B) and "excellent" (C), corresponding to a SD of the mean for repeat measures of 0.169 mm, 0.139 mm, and 0.057 mm, respectively.

quality can be attained at 2 separate sites on either side of the Atlantic. The international nature of this radiographic procedure is evidenced by the successful training of technologists from the different countries within the EU. During the technologists' training period, the role of an independent assessor at the Regional Quality Control Centre to undertake quality control of the radiographs was found not to be necessary, although their role will be essential to monitor radiograph quality during the study period. Finally, the data from this study reaffirm the validity that the semiflexed view and computerized measurement will permit the design and conduct of SMOAD trials with marked sensitivity to JSN over a 2 year period. It is also necessary to consider, in any sample size calculation, the effect of extraneous variables that may arise over the course of the study period. Calculations that include these effects are more likely, by study end, to achieve statistical power that might otherwise not have been attained.

REFERENCES

1. Altman R, Brandt K, Hochberg M, Moskowitz R. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from the Task Force of the Osteoarthritis Research Society. Osteoarthritis Cart 1996;4:217-43.
2. Lequesne M, Brandt K, Bellamy R, Moskowitz R, Menkes CJ, Pelletier J-P. Guidelines for testing slow acting drugs in OA. Proceedings Vth Joint WHO and ILAR Task Force Meeting. J Rheumatol 1994; 21 Suppl 41:65-73.
3. Buckland-Wright JC, Macfarlane DG, Lynch JA, Jasani MK, Bradshaw CR. Joint space width measures cartilage thickness in osteoarthritis of the knee: high resolution plain film and double contrast macroradiographic investigation. Ann Rheum Dis 1995;54:263-8.
4. Buckland-Wright JC. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. Osteoarthritis Cart 1995;3 Suppl A:71-80.
5. Buckland-Wright JC, Macfarlane DG, Williams SA, Ward RJ. Accuracy and precision of joint space width measurements in standard and macro-radiographs of osteoarthritic knees. Ann Rheum Dis 1995;54:872-80.
6. Ravaud P, Auleley G, Chastang C, et al. Knee joint space width measurement: an experimental study of the influence of radiographic procedure and joint positioning. Br J Rheumatol 1996;35:761-6.
7. Ravaud P, Graudeau B, Auleley G, et al. Radiographic assessment of knee osteoarthritis: reproducibility and sensitivity to change. J Rheumatol 1996;23:1756-64.
8. Piperno M, Hellio M-P, Conrozier T, Bochu M, Mathieu P, Vignon E. Quantitative evaluation of joint space width in femorotibial osteoarthritis: comparison of three radiographic views. Osteoarthritis Cart 1998;6:252-9.
9. Buckland-Wright C. Radiographic assessment of osteoarthritis: comparison between existing methodologies. Osteoarthritis Cart 1999;7:430-3.
10. Lynch JA, Buckland-Wright JC, Macfarlane DG. Precision of joint space width measurement in knee osteoarthritis from digital image analysis of high definition macroradiographs. Osteoarthritis Cart 1993;1:209-18.
11. Buckland-Wright JC, Macfarlane DG, Jasani MK, Lynch JA. Quantitative microfocal radiographic assessment of osteoarthritis of the knee from weight bearing tunnel and semi-flexed standing views. J Rheumatol 1994;21:1734-41.
12. Buckland-Wright JC, Macfarlane DG, Lynch JA, Jasani MK. Quantitative microfocal radiography detects changes in OA knee joint space width in patients in placebo-controlled trial of NSAID therapy. J Rheumatol 1995;22:937-43.
13. Mazzuca SA, Brandt KD, Buckland-Wright JC, et al. Field test of the reproducibility of automated measurements of medial tibiofemoral joint space width derived from standardized knee radiographs. J Rheumatol 1999;26:1359-65.
14. Buckland-Wright JC, Wolfe F, Ward RJ, Flowers N, Hayne C. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP) and schuss views. J Rheumatol 1999;26:2664-74.
15. Bland JM, Altman DG. Measurement error. BMJ 1996;313:744.
16. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthrosis. Ann Rheum Dis 1957;16:494-501.
17. Buckland-Wright JC. Quantitation of radiographic changes. In: Brandt K, Lohmander S, Doherty M, editors. Textbook of osteoarthritis. Oxford: Oxford University Press: 1998;459-72, 578-80.
18. Mazzuca SA, Brandt KD, Katz BP. Is conventional radiography suitable for evaluation of a disease-modifying drug in patients with knee osteoarthritis? Osteoarthritis Cart 1997;5:217-26.