Université d'Ottawa · University of Ottawa

DERIVATION OF A RULE FOR PREDICTING UNSTABLE MENISCAL TEARS IN PATIENTS WITH OSTEOARTHRITIS OF THE KNEE

by

© Geoffrey F. Dervin, MD, FRCSC

Thesis submitted to the School of Graduate Studies and Research in partial fulfillment of the requirements for the M.Sc. degree in Epidemiology

University of Ottawa

January 17, 1999

0-612-38743-7

ABSTRACT

Arthroscopic surgery is the most commonly performed procedure for osteoarthritis of the knee , although outcomes are not consistently favourable. The resection of unstable meniscal tears appears to be associated with better outcomes following arthroscopy. Diagnosis of these tears is not precise and often relies on expensive technology such as magnetic resonance imaging. There are no available guidelines for the diagnosis of meniscal tears in patients with osteoarthritis of the knee.

The goal of the present study was to prospectively study a cohort of such patients referred for arthroscopy and develop a clinical prediction rule to select patients who would benefit from arthroscopy.

Patients with osteoarthritis of the knee refractory to conservative medical treatment were selected for arthroscopic evaluation. A standardized assessment protocol was administered to each patient and repeated by a second observer when feasible to permit calculation of the kappa coefficient(κ) for interobserver agreement. Arthroscopic determination of unstable meniscal tears was recorded by one observer who reviewed a video recording and was blinded to preoperative data based on objective criteria. Those variables which had the highest interobserver agreement and the strongest association with meniscal tear by univariable methods were then entered into a logistic regression model.

One hundred and fifty two patients were enrolled. There were 92 meniscal tears (77 medial and 15 lateral). Interobserver agreement between clinical fellows (1st observer) and the treating surgeon (2nd observer) was poor to fair ($\kappa < 0.4$) for all clinical variables except radiographic measures which were very good. Fellows and surgeons predicted unstable

meniscal tear preoperatively with equivalent accuracy of 60%. Surgeons were chosen as the source of predictor variables for the greater clinical relevance based on generalizability to current practicing surgeons. Those variables with univariate association of $p < 0.2$ and interobserver agreement of $\kappa > 0.15$ were entered into the logistic regression for prediction of a medial meniscal tear. The model yielded the following variables with their odds ratios (95% C.I.): history of swelling, 0.42 (0.21, 0.97); ballottable effusion, 0.38 (0.17, 0.91); and positive MacMurray test, 2.21(0.92, 4.5). The model was 69% accurate for all patients and 76% for those patients with advanced medial compartment osteoarthritis defined by joint space height of 2 mm or less.

This study underscored the difficulty in using clinical variables to predict unstable medial meniscal tears which were addressed at arthroscopy. The lack of interobserver agreement suggests that the model should not be adopted in clinical practice at this time for the findings may not be generalizable to other physician observers. A history of swelling and ballottable effusion were negative predictors of tears in this population while a positive MacMurray test was the only positive predictor which remained useful after adjusting for confounding in the regression model.

ACKNOWLEDGEMENTS

My wife, Joan Culliton, has been a strong and reliable supporter of this thesis from the completion of this course and throughout the production of this manuscript. She has sacrificed for my selfish pursuit and I will be indebted for her patience throughout this project.

I remain grateful to my thesis supervisors, Dr. Ian Stiell and Dr. George Wells, who persevered through the duration of the study with their comments and suggestions. I was fortunate to have such highly qualified insight into the construction of a decision rule.

My colleagues in the Orthopaedic division at the Ottawa General Hospital gave generously of their time and effort in recruiting patients and providing advice along the way. Thank you to the orthopaedic fellows who manned the preadmission clinics and performed the standardized assessments beyond their responsibilities. Special thanks to our chairman, Dr. Alan Giachino for his encouragement to forge ahead with this task and professional guidance.

Patient recruitment and data collection were diligently handled by Kelly Rody and Jenny Grabowski who ensured the study moved at a consistent pace and also provided subtle pressure to all participants for compliance with the protocol. Kim Morin ably helped with preparation of the manuscript usually with very short notice.

The Physicians Services Institute (Grant # 95 -33) and the Canadian Orthopaedic Foundation (Hip Hip Hooray) provided grant support for this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Statement of the Problem

Osteoarthritis is the most prevalent of the arthritides and is believed to afflict up to 5 million Canadians. The knee is commonly the most symptomatic joint and fortunately several medical and physical therapies exist for managing this condition.[1] The use of acetaminophen[2], nonsteroidal anti-inflammatory medications[3](NSAIDS), quadriceps strengthening[4], and a supervised walking program are all proven effective strategies for treatment. Transient benefits can be obtained with intra-articular steroid injection .[5,6] Viscosupplementation with intra-articular sodium hyaluronate injection appears to provide both symptomatic relief and may modify the rate of disease progression.[7,8] The clinical benefits appear to be of greater duration than those with intra-articular steroid use.[9]

Orthopaedic referral is reserved for patients who remain symptomatic after exhausting these measures. Surgical alternatives include arthroscopic lavage and/or debridement, osteotomy, and arthroplasty. Osteotomy of the proximal tibia or distal femur is indicated for patients with unicompartmental osteoarthritis and significant malalignment.[10] In appropriately selected patients, the clinical benefits may extend beyond 10 years.[11-13] The procedure is associated however with potential complications including under or overcorrection[14], neurovascular injury[15,16], and non-union.[17] Total joint arthroplasty is generally reserved for end stage tricompartmental joint destruction and generally affords excellent pain relief but does require activity modification.[18] Significant complications can also occur including venous thromboembolism, sepsis, and component loosening requiring revision. There is thus a large cohort of patients for whom these procedures are either contraindicated or less attractive and for whom arthroscopic debridement has been suggested as an alternative for failure of medical management.

Arthroscopic debridement is relatively more attractive than osteotomy or arthroplasty because it is usually an outpatient procedure with minimal morbidity, requiring less rehabilitative effort for the patient postoperatively. A recent review suggested less satisfactory results for knees with more advanced articular damage, severe tibiofemoral malalignment, and chronic symptoms irrespective of the intervention.[19] The variety of arthroscopic procedures and retrospective nature of much of the data is likely confusing to the practising arthritis surgeon. There is little consensus among orthopaedic surgeons as to who should be offered arthroscopy for osteoarthritis. In fact, a recent survey of Ontario's orthopaedic surgeons found considerable disagreement as to the utility of arthroscopy for osteoarthritis of the knee in three hypothetical case scenarios.[20] A definitive study of the role of arthroscopic debridement in osteoarthritis of the knee would be helpful to all clinicians managing these patients.

There were 16,118 knee arthroscopies performed in the province of Ontario between April 1, 1996 and March 30, 1997.[21] The percentage of those which were done for debridement of osteoarthritis is unknown. Inappropriate utilization of the procedure would seem to amplify the morbidity of this condition. Underutilization would leave some patients more disabled than is necessary. Overutilization would expose patients to needless anaesthetic and operative risks with little chance for improvement and constitute inefficient use of hospital resources. Further, the morbidity of these patients is extended if they were more appropriate candidates for osteotomy or arthroplasty. Several specific arthroscopic interventions have been described for the treatment of osteoarthritic knees.

1.2 Background: Arthroscopy for osteoarthritis of the knee

1.2.1 Saline lavage

Ike et al. [22] conducted a single blind randomized controlled trial of intra-articular tidal irrigation with 1000 cc of saline versus medical management with nonsteroidal and analgesic medications. They found superior clinical benefit for a variety of subjective and objective measures in the irrigation group up to 12 weeks follow-up. Livesly et al.[23] reported greater relief of pain extending to one year in patients undergoing saline lavage and physiotherapy than a control group of physiotherapy alone. Improvement in inflammatory signs persisted for 3 months. The study's validity has been challenged on the basis that it lacked randomization, omitted validated assessment scales for osteoarthritis of the knee, and had an uneven sex distribution in the treatment groups. Chang et al.[24] compared saline lavage with standard arthroscopic debridement (removal of loose articular cartilage, torn menisci, and proliferative synovium) and could not find a clinical difference at 3 and 12 months. The number of patients was small however and the lavage group did not have arthroscopic evaluation to ascertain the intraarticular pathology. Conversely, Hubbard [25] reported a single surgeon, randomized controlled trial of lavage alone versus lavage and debridement of chondral flaps. Abrasion and drilling of lesions was not performed. Only patients with isolated medial femoral condyle lesions Outerbridge grade III or IV were enrolled (meniscal pathology excluded) and the results clearly favoured debridement over simple washout as assessed by a modified Lysholm score.

1.2.2 Abrasion arthroplasty

Abrasion arthroplasty of full thickness chondral defects was espoused by L. Johnson [26] as having a role as a natural extension to open debridement first described by Pridie. [27]Johnson recommended a superficial abrasion of sclerotic articular lesions emphasising the retention of the cortical shell. He cautioned that aggressive abrasion to the depths of the cancellous bone or "spongialization" was not advisable and would yield unstable fibrous tissue. He recommended 8 weeks of postoperative non weight bearing and suggested contraindications were malalignment, instability, obesity, and inflammatory arthritis. Under ideal circumstances, he was able to show reparative fibrocartilage on histology and reappearance of joint space in some arthritic knees. He reported 74 of 95 patients improved in a retrospective review. Friedman et al.[28] compared patients undergoing debridement with abrasion arthroplasty to a matched historical control group in their own practice and reported the addition of abrasion improved more patients (53% vs. 32%) at one year. Others have not substantiated these results. Bert and Maschka [29] showed 66% good to excellent results at five years with debridement versus 51% for debridement and abrasion. The study was retrospective and patients did select themselves for treatment with abrasion reserved only for those agreeing to non weight bearing postoperatively. In a similarly designed study, Rand [30] also found better results at 3.8 years with simple debridement than abrasion arthroplasty (67% vs. 39%). The last two studies, while retrospective, offer concurrent controls to more reliably evaluate the value of abrasion arthroplasty and both authors recommended against performing this procedure.

1.2.3 Partial menisectomy

The menisci of the knee are wedge shaped structures of fibrocartilage which translate compression forces across the tibio-femoral joint into circumferential stresses which they absorb to protect the articular surfaces.[31] The adverse consequences of meniscal injury and subsequent removal were first described by Fairbank who reported radiographic changes of degenerative arthritis in patients who had undergone subtotal menisectomy.[32] Some meniscal tears are suitable for repair and several techniques have been described for tears in younger patients in the peripheral vascular zone of the meniscus.[33,34] The majority of symptomatic meniscal tears cannot be successfully repaired and are managed by partial menisectomy which is believed to be less deleterious than subtotal menisectomy.[35,36]

Arthroscopic partial menisectomy is a well-tolerated and effective procedure in patients aged 40 or more without significant degenerative change[37-41]. The role for resection of degenerative tears with coexisting articular wear is more contentious. Jones et al[42] retrospectively reviewed partial menisectomies in patients over 40 years of age and found considerable worse outcome in patients with degenerative tears (absence of trauma and fissured , horizontal cleavage tears) than those with traumatic tears (history of trauma, bucket handle or parrot beak tears). Though the study was small and retrospective, the authors recommended resection of only traumatic tears causing mechanical symptoms. Lotke et al.[43] reviewed their long term results (mean 10.8 yrs) of open medial menisectomy grouped according to preoperative radiographs. The outcome was satisfactory in 90 % of patients with normal radiographs but only 21 % for those with moderate or marked degenerative change. Jackson and Rouse [44] reported satisfactory short term results at a mean 2.5 yrs in 80 % of patients undergoing arthroscopic partial menisectomy in the presence of degenerative chondral change versus 95 % in those without degenerative change. The balance of evidence would suggest that partial menisectomy is an effective intervention in mild to moderate forms of osteoarthritis and that clinical criteria to predict the presence of unstable meniscal tears would be desirable.

1.3 Prediction Rules

1.3.1 General Considerations

Clinical prediction rules are derived from systematic clinical observations and designed to help clinicians use clinical findings to determine the accuracy of diagnostic tests, make estimates about prognosis or evaluate the effect of medical treatment. Wasson et al. conducted a review of articles relating to prediction or decision rules published in the years 1981 - 1984 in 4 leading medical journals.[45] They used the review as a backdrop to describe methodological standards for precise rules which could improve the efficiency and accuracy of physicians' judgements. Laupacis et al. performed a similar review of prediction rule publications for the years 1991 -1994 in the same medical journals.[46] They sought to determine the compliance with Wasson's methodological standards and suggested modifications to enhance the ultimate success of rules, that is the widespread adoption of these in different clinical settings and the change of medical practice. The Ottawa Ankle rules exemplify the principles of derivation[47], validation[48], and successful implementation[49]. Of even greater relevance was the ability to introduce these rules in other clinical settings[50] and ultimately confirm that the study had changed medical practice.[51] Rules which are concise and clinically relevant clearly have the greatest potential for widespread acceptance and more efficient utilization of medical resources. The following is an outline of the standards espoused by these two reviews to successfully develop prediction rules.

1.3.2 Outcome Selection

The outcome of a prediction rule should be both clearly defined and clinically important so that the results of the rule have the potential to effect a change in clinical practice and justify the effort in developing the rule. Assessment of the outcome should be blind to the predictor variables particularly where the outcome requires a subjective interpretation. Blind assessment also helps to minimize any preconceived bias by the investigator as to the role of the selected predictor variables as they relate to the outcome.

1.3.3 Predictor Variables

As with the outcome, the chosen predictor variables should be clearly defined to enhance their reproducibility in the derivation study. Variables defined in this way must at least be reliable in the controlled clinical environment where the rule is developed to help ensure successful implementation of the rule in other settings. Variables should also be clinically sensible to readers who are more likely to embrace a rule which includes variables with which they are familiar and can endorse. Similarly, those variables which do not enter the rule should be described so that readers who have a preconceived bias toward variables not ultimately included in the rule can understand why this was so. There are three common reasons for exclusion. First, the variable cannot be reproducibly assessed by different observers preventing consistent interpretation. Second, the variable may have no predictive association on its own. Finally, the predictive association of the variable may not enhance that of other variables when it is combined in the final rule. As with the outcome, the assessment should be performed blind to the outcome, which is usually accomplished with prospective data collection.

1.3.4 Patient Selection and Study Site

Inclusion and exclusion criteria for patient selection must be clearly defined so that readers can determine whether the rule can be applied to the patients they treat. Those characteristics which must be defined are age and sex and any other trait that would be clinically pertinent to the present study. This could include disease onset, severity, and previous treatment defined in established terms understood by clinicians in the field. The study site and setting should also be described to completely define the population in the derivation set to permit readers to evaluate how generalizable the rule is likely to be and address potential "referral filter bias" described by Sackett.[52].

1.3.5 Mathematical Techniques

The computational strategies used to develop the most relevant decision rule must be described. Univariate methods such as the two-by -two crosstabulation technique are very common but usually insufficient alone to account for the relationships among multiple clinically important variables. Multivariate methods such as logistic regression, discriminant analysis, and recursive partitioning are then required. The choice of technique depends on the nature of the predictor variables and the objective of the rule and should be described.

1.3.6 Sensibility

Feinstein defined sensibility of a rule as a composite of clarity of purpose, content validity, and ease of use.[53] Clarity of purpose refers to the impetus for and importance of the rule for clinicians and its intended applicability. Content validity ensures that the rule overtly appears to make sense to clinicians who wish to consider its use: the individual items appear appropriate based on existing knowledge and the sum of these are presented in a rule that would be deemed comprehensive. Ease of use is critical for clinicians overwhelmed with clinical information. For example, computational formats derived from regression equations with the use of continuous variables and the need for a calculator are cumbersome. The use of categorical cutpoints for continuous variables is one mechanism of simplifying the rule to allow for rapid application at the bedside.

1.3.7 Accuracy

The proportion of all patients correctly classified (accuracy) by the rule should be determined and reported. This should be described in more detail by providing the true positive rate (sensitivity) and the true negative rate (specificity) which are used differently by clinicians. High sensitivity is important for situations where the consequences of a false negative diagnosis are disastrous such as for a treatable form of cancer. Increased specificity is important where a false positive diagnosis would lead to unnecessary and potentially harmful clinical tests or treatment. It is apparent then, that each rule will have a different emphasis on sensitivity and specificity and the rationale for this should be described. Ultimately, clinicians are concerned with the positive and negative predictive values of rules (given as the probabilities of having or not having the outcome respectively). Predictive values are governed by sensitivity, specificity and prevalence of the outcome in the population being evaluated, re-emphasising the need for a clear description of the study population in the derivation set alluded to earlier (see 1.3.4).

1.3.8 Prospective Validation

Statistical techniques have been described which permit a cross-validation of the rule in the same population sample from which the rule was derived.[54,55] These are often referred to as "internal validation" and can provide an indication of how the rule will perform. "External validation" refers to the testing of the rule in a separate population from the initial derivation set. A new clinical setting provides the opportunity for definitive validation and confirms that adherence to all of the preceding principles has resulted in a rule which is stable and reliable in the new setting.[50]

1.3.9 Effect on patient care

The ultimate objective of a well-designed prediction rule would be widespread dissemination, successful implementation, and enduring effect on patient care. Such confirmation requires diligent study in different clinical settings where the rule was introduced after the prospective validation has been ensured. Verbeek et al.[51] confirmed that clinical practice had been changed after implementation of the Ottawa Ankle Rules in the assessment of patients with ankle injuries in the emergency department of a general hospital. Specifically, the ordering of unnecessary radiographs was prevented, waiting time in hospital decreased, and a cost saving was realized. Subsequent random audits are required to determine if the rule has permanently changed practice.

1.4 Prediction of meniscal tears: Previous studies

A literature search of Medline for the years 1966 – 1997 was performed using keywords meniscus, diagnosis, arthroscopy and articles which described prospective data collection, and meniscal diagnosis with clinical findings were reviewed. Daniel et al [56] showed that experienced clinicians were 79% accurate in diagnosing medial meniscal tears when relying on the clinical examination. They did not specify which features of the examination were most important and rather allowed clinicians to express a probability level of meniscal pathology. They did not feel a "point system" was valuable and believed the clinicians' judgement was more useful. The study fails to satisfy many of the criteria for prediction rules. It does not define predictor variables explicitly and was retrospective in nature. The outcome was determined partly by arthroscopy, the gold standard, in only 40% of cases, whereas the rest had arthrograms, which they found had an accuracy of 89% for medial and 85% for lateral tears. Finally, patients with radiographic osteoarthritis were excluded and restricted to patients aged 15 - 45.

Terry et al [57] demonstrated a similar accuracy rate of clinical diagnosis of 81% for a cohort of 206 consecutive patients in a Sports Medicine Clinic who were suspected to have internal derangement and were scheduled for arthroscopy. Interestingly, medial meniscus tears were diagnosed with 85% accuracy while lateral tears with only 58% accuracy. They excluded patients with significant ligamentous instability or severe gonarthrosis. Again, several methodological limitations are noted. First, the population is heterogeneous, with ages 12 -80 and clearly some cases of osteoarthritis were likely included, given the statement that only severe cases were excluded. The predictor variables were briefly described but only the senior author performed the examinations and no effort was made to establish the reliability of these variables. The outcomes were all types of internal derangement including meniscal tears but these were not explicitly defined nor was the outcome assessment blinded to the preoperative assessment.

Abdon et al [58] performed a prospective study of patients undergoing arthroscopy for suspected meniscal pathology and used discriminant analysis of a variety of clinical parameters. They found jointline tenderness and history of locking positive predictors while pain at rest, sick leave at time of surgery, and tenderness of medial patellar facet were negative predictors. The patients were described only by age (16 - 66) and gender (76% men), and no mention was made of their underlying arthritis or previous treatment. The predictor variables were described and patients were examined at two instances but no measure of observer agreement was given. Knee swelling was not evaluated in these patients which is a significant omission. It is not clear if the arthroscopic assessment was blinded. Though multivariate statistical modelling was used, there were 68 variables entered into the model for 88 tears, suggesting the data set was underpowered for the number of variables considered which could have led to overfitting of the model.

Fowler [59] prospectively studied 161 patients with knee pain and suspected meniscal pathology and found a combination of signs most valuable. Joint line pain was sensitive (85%) but not specific (29%). Pain on forced flexion had a 50% sensitivity and 68% specificity while the MacMurray test and block to extension had low sensitivity yet high specificity. Patients were aptly described for age, gender, duration of symptoms, and previous treatment but no mention made of underlying osteoarthritis, which is of relevance to the current study. Only 5 clinical signs were included in the assessment, and despite a mention of two separate assessments, no attempt to describe the interrator reliability was done. No mention was made of blinded outcome assessment and presumably this was not done. Meniscal tear was not defined explicitly. Only univariate statistical measures were used in this study.

In summary, the above studies have helped to outline which predictor variables may be of greatest relevance for prediction of meniscal tears, but the absence of reliability assessments and explicit definitions does not allow for generalizability of their conclusions. The variability of significant physical signs is likely attributable to the study population heterogeneity in the different studies. None of the studies were replicated in a new group of patients such that the stability of the models could not be confirmed. No study confined the prediction of meniscal tears to patients with established osteoarthritis nor did these studies perform subgroup analysis of osteoarthritic patients. It seems plausible that the clinical diagnosis of symptomatic unstable meniscal tears is complicated by the presence of concurrent symptomatic osteoarthritis of the knee and needs to be refined in this cohort of patients.

1.5 Goals and Objectives

1.5.1 Derivation of prediction rules for presence of unstable meniscal tears

The goals of this thesis are to develop prediction rules for the presence of unstable meniscal tears in osteoarthritis of the knee. We feel that there is valuable information to be gained from an effort to derive 2 separate rules. The first rule, to predict *any* unstable tear be it medial or lateral, is useful as it can potentially direct surgeons to a course of action, arthroscopy, for any patient with persisting symptoms from osteoarthritis. This assertion must be substantiated in the outcome phase of the study described in Section 1.5.2. The second rule, which is to predict unstable *medial* meniscal tears, recognizes the fact that most tears are medial and that clinicians are especially challenged diagnosing unstable medial meniscal tears particularly in the presence of advanced medial compartment osteoarthritic changes. A rule enhancing diagnostic accuracy would permit more directed therapy to the cause of symptoms - arthroscopy for unstable meniscal tears or a host of other options for degenerative articular cartilage damage alluded to earlier. The objectives for the present study can be summarized as follows:

a) Develop a standardized clinical assessment protocol of commonly used physical signs, symptoms and radiographs for osteoarthritis of the knee.

b) Determine the interrater reliability of these clinical variables.

c) Compare clinicians' (attending staff versus post residency fellows) accuracy in predicting the presence of unstable tears in these patients following a standardized assessment in order to evaluate the role of clinical experience.

d) Utilize multivariate statistical techniques to construct prediction rules identifying which symptomatic patients with osteoarthritis have an associated unstable meniscal tear requiring resection for:

i) any unstable meniscal tear (medial or lateral)

ii) an unstable medial meniscal tear in advanced medial compartment osteoarthritis

1.5.2 Effect of arthroscopic debridement for patients with osteoarthritis of the knee

This thesis is part of a larger study of considerable relevance to patients and clinicians, which is to define the effect of arthroscopic debridement as measured by disease specific and generic quality of life indices (described in 4.5.1). The results of most orthopaedic procedures are by convention reported after a minimum two-year follow-up which will be completed for all subjects enrolled in the present study by November, 1999. Ultimately, a prediction rule for appropriately selecting patients for this procedure will be most beneficial. The timeline for this study would not permit completion of the thesis and hence the decision made to predict unstable meniscal tears which has shown the strongest association with positive outcomes in previous studies of arthroscopy for osteoarthritis of the knee.

2. METHODS

2.1 Patient selection

All patients aged 40 to 75 referred to the Ottawa General Hospital Orthopaedic outpatient clinic between March, 1995 and November, 1997 with osteoarthritis of the knee (as defined by the American Rheumatological Association [60]) were considered for eligibility for the present study (Appendix 1). Patients with inflammatory or traumatic forms of osteoarthritis were excluded. The age criteria were chosen in accordance with most therapeutic studies of osteoarthritis. The Ottawa General Hospital is a tertiary care teaching institution, although the majority of patients referred to the Orthopaedic outpatient clinic come from general practitioners and is likely similar to the patient pool in a community hospital setting. Nearly 90% of all patients assessed for osteoarthritis of the knee were referred by local community family physicians. Patients who remained symptomatic despite supervised physical therapy and comprehensive medical management were considered for arthroscopy. The failed conservative management included oral and topical analgesics, nonsteroidal anti-inflammatory medications, and intra-articular cortisone injection in some people and was generally supervised by the referring physician. Selection of patients for surgery was left to the discretion of surgeons who had a range of experience between 3 and 25 years. Most were guided by their impression of mechanical symptoms and swelling as best indications for surgery and possible benefit. Those patients who accepted surgery were told that they would be approached by the research nurse to discuss their possible inclusion into the study. The study was explained to all prospective patients and their informed consent was required for participation in the study which was reviewed and approved by the Research Ethics Board of the Ottawa General Hospital (Appendix 2). The study was identified by the acronym SKOAP, for the

Study of Knee OsteoArthritis Prospective evaluation. The acronym was easily recalled as arthroscopy is generally referred to as a "scope".

2.2 Selection and definition of predictor variables

A review of pertinent literature involving prediction of meniscal tears in various cohorts of patients was performed and potentially relevant clinical variables selected. This was combined with a similar review of reports of arthroscopic menisectomy in osteoarthritis to obtain additional variables. A consensus meeting was held with the participating surgeons to agree on the selection of variables and their precise definition. A study manual which described the pertinent clinical variables was assembled and distributed to all study participants. The section describing the predictor variables is reproduced below.

2.2.1 Demographic Variables

a) Age, Gender

b) Work-related injury (Y/N)

c) NSAID/Analgesic regular daily use (Y/N)

d) BMI (Body Mass Index) rated as:

i) Overweight	BMI > 27
ii) Healthy weight	20 < BMI < 27
iii) Underweight	BMI < 20

2.2.2 Variables from history

a) Discrete injury (Y/N) The patient must recall an incident which provoked the current symptomatology above and beyond the chronic state, and has persisted beyond 2 weeks.

b) Pain:

i) Focal (patient must point with one finger to one specific area) (Y/N).

ii) Generalized (patient points to two or more areas or states "all over"). (Y/N)

iii) At rest; arising from chair; up/down stairs. (Y/N)

c) Swelling: Subjective sensation of activity related swelling. (Y/N)

d) Giving Way: The knee gives out or buckles suddenly with pivoting or turning. (Y/N)

e) Locking: Occurrence of being suddenly unable to extend the knee. (Y/N)

2.2.3 Variables from physical examination

a) Effusion:

i) Ballottable: With the knee maximally extended, one hand compresses the suprapatellar pouch as the other presses on the middle of the patella feeling for excursion greater than 5 mm. (Y/N)

ii) Sweep test: Milking the knee from inferomedial to superolateral and then sweeping the hand down lateral to the patella looking for a fluid wave medially. (Y/N)

iii) Visible by loss of medial patellar groove. (Y/N)

b) Joint Line Tenderness: Palpation:

i) Medial (Y/N) or ii) Lateral. (Y/N)

c) Joint Line Pain:

i) Forced, passive full flexion. (Y/N)

If yes, location: anteromedial, anterolateral, posteromedial, or posterolateral ?

ii) Forced, passive full extension. (Y/N)

If yes, location: anteromedial, anterolateral, posteromedial, or posterolateral ?

d) Lack of Full Passive Extension versus Contralateral side: Tested Supine: (Y/N)

e) Ligamentous Laxity: Rated as positive if greater than 5 mm joint space opening compared to contralateral side.

i) Varus, Valgus and Lachman at 30° flexion. (Y/N)

ii) Posterior Drawer at 90° flexion (Y/N)

f) Special Tests for Meniscal Irritability:

i) MacMurray Test: The knee is fully flexed and then an external rotation and extension is carried out to feel for a click or elicit pain indicating a tear of the posterior horn of the medial meniscus. Similarly, it is repeated with the leg in internal rotation to feel for a lateral meniscal tear. (Y/N)

ii) Steinman's Test: Flexing the knee to 90° and externally and internally rotating the tibia, causing increased pain at the join line. (Y/N)

iii) Circumduction in full flexion incorporating varus/valgus and internal/external rotation. (Y/N)

2.2.4 Variables from X-rays

Plain film radiographs were also obtained as variables because of our desire to include basic investigations as part of the determinants for the prediction rule , which we felt would have more clinical and cost effective relevance. Radiographs were obtained 1 week preoperatively in both the 3-foot standing anteroposterior (AP) and a 45° flexion weight bearing projection.[61] The 45° flexion posteroanterior radiograph was made with the femur angled 25° and the tibia 20° from the cassette, while the x-ray beam is centered at the superior pole of the patella and directed 10° caudad. The latter has been suggested as a more sensitive technique for joint space narrowing. A foot map was used in order to normalise rotation for both views and a foam wedge to control flexion for the 45° X-ray projection. This standardized technique was felt to be essential in order to enhance the reproducibility of the measurement. All radiographs were assessed by two observers who were blinded to the arthroscopic and clinical manifestations of the subjects. The features which were recorded included joint space in the lateral and medial tibiofemoral compartments, and anatomical tibiofemoral axis (in degrees).[62] Joint space in both compartments was measured at the narrowest point of contact of the tibiofemoral compartment.

2.3 Interobserver agreement and physicians' prediction

Subjects were assessed in the preadmission clinic of the hospital 7 - 10 days prior to surgery by a post graduate orthopaedic fellow using the standardized assessment protocol described. Where feasible, the operating surgeon repeated the standardized assessment the day of surgery in order to determine the reliability of the chosen clinical parameters for the study. All physicians were asked to predict to the nearest decile the percent probability of finding an unstable meniscal tear based on their clinical impression and a review of the radiographs.

The interobserver agreement compared the interpretation of predictor variables of 6 clinical fellows versus 7 orthopaedic staff using overall percent agreement and the kappa coefficient.[63] Agreement for radiographic indices was only tested between the principal investigator (GFD) and the research nurse (KR) for all patients for reasons of feasibility. The kappa coefficient is calculated as the percentage agreement expected beyond that of chance and is hence a better index of the reliability of the variable being studied.[63] The kappa coefficient is a more valid measure of concordance because with the exception of extremes of prevalence, the kappa coefficient is not as sensitive to the underlying prevalence of the variable being measured as would be the percent agreement. The latter measure would be deceptively high due to chance alone for variables with high prevalence. Adopting the standard adopted by Landis [64], the interpretation of kappa coefficients is commonly described as follows: (0 - .2 = slight, .2 - .4 = fair, .4 - .6 = moderate, .6 - .8 = substantial, and .8 - 1.0 = almost perfect). Reliability testing of predictor variables was restricted to interobserver which is usually less reproducible than intraobserver reliability and more clinically relevant in the formulation of a generalizable prediction rule.

2.4 Intervention

All subjects underwent arthroscopy of the knee under general or spinal anaesthesia. All surgery in this study was performed via standard anterolateral and anteromedial skin portals. The use of tourniquet was optional. A diagnostic arthroscopy was first carried out and videotaped prior to any intervention. The surgeon then performed the procedure which included resection of loose chondral flaps, unstable meniscal tears and synovectomy only where required for visualisation. Standard and motorised instruments were employed. A meniscal tear was considered unstable if it was full thickness longitudinal and displaceable, radial greater than 3 mm, parrot beak or complex. Abrasion arthroplasty or drilling was not performed. A final record of the arthroscopic procedure was made immediately prior to its conclusion.

2.5 Outcomes

2.5.1 Unstable meniscal tear

The presence of an unstable meniscal tear was selected as the most surgically pertinent outcome in this thesis. Resection of an unstable tear was deemed the most acceptable intervention in this population and one that allowed agreement amongst the participating surgeons. Much of the available data which is largely retrospective, suggests functional outcome post menisectomy is most correlated which articular cartilage changes [38,39,42-44] All arthroscopic procedures were recorded on videotape as a permanent record and evaluated by the principal investigator for the presence of an unstable tear. Ambiguous cases were resolved by a consensus of the principal investigator with one other participating surgeon, but not the operating surgeon.

2.5.2 Quality of life scores

The effect of the intervention described in Section 3.4 will be evaluated prospectively by two health related quality of life instruments preoperatively and at 3, 6, 12, and 24 months postoperative as part of the larger study and will not be reported here. Health related quality of life measures are categorized into two broad forms; generic and disease specific.

Generic scales are useful for their comprehensive evaluation and allow comparisons of interventions for unrelated conditions. The generic scale used in the present study was the SF-36, a self administered questionnaire containing 36 items representing eight dimensions.[65] The objective is to evaluate functional status, general well being, and overall health assessment. It has been used extensively, is easily administered and has been validated against existing health profile measures.[66]

Disease specific scales are generally more sensitive to change and are usually more relevant to the conditions they measure. The WOMAC (Western Ontario and McMaster Universities) osteoarthritis index is a disease specific quality of life instrument developed for assessment of patients with osteoarthritis of the knee and hip.[67] It is comprised of three sections; pain (5 questions), stiffness (2 questions), and physical function (17 questions). Its clinimetric properties have been tested in patients with osteoarthritis of the knee treated with non steroidal medication and total knee arthroplasty.[68,69] The pain dimension has proven particularly responsive , detecting significant improvement following a six week trial of nonsteroidal anti-inflammatory medication($p<0.001$).[68] The visual analog scale version was chosen for this study. French and English versions of both scales were obtained for the study and the research nurse was available for interpretative questions for the first administration of the test preoperatively.

2.5.3 Arthroscopic classification of articular cartilage damage

Several investigators have proposed classification schemes for articular cartilage pathology in the knee. Outerbridge was among the first when he described the macroscopic changes seen in patellae at the time of open meniscal surgery and described four grades.[70] Others have also proposed classification systems though none are altogether comprehensive, some omit lesions without surface changes[71,72], or do not incorporate surface area in the classification.[73,74] Noyes [75] proposed a classification which also incorporated specific descriptions to the area and location of articular lesions. While this was a welcomed effort toward classification, the ultimate scoring of lesions was subjectively determined based on the authors' impression of importance.

Dougados et al [76] devised a classification scheme for each tibiofemoral and the patellofemoral compartment which included depth of lesion, surface area involvement, and exact location. Their description of the depth of lesions was as follows:

Grade I; softening and swelling of cartilage

Grade II; superficial fissuring of the cartilage surface, velvet-like appearance

Grade III; deep fissuring reaching subchondral bone; including partially detached chondral flaps or crab- meat like appearance

Grade IV; erosion to exposed bone

Surface area involvement was estimated as a percentage of compartment involvement and location of the lesion recorded on an articular diagram (Appendix 3). The scale was validated by correlating severity of damage with the arthroscopist's global impression and with joint space narrowing on weight bearing X-rays. Multivariate regression was then used to devise a numerical score (continuous variable) and a tree structured regression to determine a grade (categorical variable). One shortcoming of the study was that the sample of patients was skewed toward lesser degrees of chondropathy. The classification was subsequently validated in a population of patients fulfilling the American College of Rheumatology clinical and radiographic criteria for osteoarthritis of the knee[60,77] similar to the patients enrolled in the present study. We therefore selected the French Society of Arthroscopy (SFA) scoring scheme for articular scoring and chose to use "grades" rather than "scores" as Ayral et al.[77] showed there was complete agreement in nine of ten cases between the same or different observers when using grade as a score.

2.6 Data Analysis

Data was recorded onto a customised form that was used for data entry by the study nurse (Appendix 4). Statistical analysis was performed using SPSS for Windows v. 6.1.3 (© SPSS Inc., 1995). Interobserver agreement of all clinical variables was measured with a kappa coefficient (κ) [64].

2.6.1 Univariate analysis

Univariate analysis was performed to determine the strength of association between each variable and the presence of any unstable meniscal tear, and a medial meniscal tear. Separate testing was used for the clinical fellows and the operating surgeons. Variables were dichotomised into clinically sensible cutpoints to simplify clinical use and ultimately rule acceptance. Test procedures were selected in accordance with the type of variable evaluated; by chi-square test for discrete data and unpaired two-tailed Student *t* test for continuous data (age).

2.6.2 Multivariate analysis

Those variables which were found to be most reliable ($\kappa > 0.15$) and associated with an unstable meniscal tear by univariate methods ($p < 0.2$) were entered into a logistic regression to develop the best model for predicting any unstable meniscal tear and a medial meniscal tear. The variables as tabulated by the surgeons were used to construct the model as it was felt these would have greater clinical relevance for practising orthopaedic surgeons of similar experience. Logistic regression was chosen over discriminant analysis because the latter assumes that predictor variables have a multivariate normal distribution which would require that they be mostly continuous, in contrast to the present study evaluating primarily nominal or dichotomous variables. The model was built using a forward stepwise technique until variables did not meet significance level either for entry or removal (0.10) using a likelihood-ratio test. This significance level was chosen so as to be liberal in the entry criteria and ensure that potentially valuable predictor variables would not be excluded. The Hosmer and Lemeshow goodness of fit test of the model was applied to test the accuracy of the model relative to observed events. Finally, a receiver operating characteristic (ROC) [78] curve was constructed using different probability thresholds for the model (which is by default (0.5)). The ROC curve provides for a range of sensitivity and specificity values and depicts the trade-off between the true positive and the false positive rates.

Recursive partitioning is a multivariable technique particularly suited to achieving 100% sensitivity (no false negatives), though often at the expense of specificity. The clinical relevance of unstable meniscal tears in osteoarthritis is not completely appreciated and therefore the clinical "cost" of false positives versus false negatives has not been thoroughly established. Thus we felt that a rule seeking the greatest overall accuracy was a more reasonable outcome than one seeking 100% specificity or sensitivity .

3. RESULTS

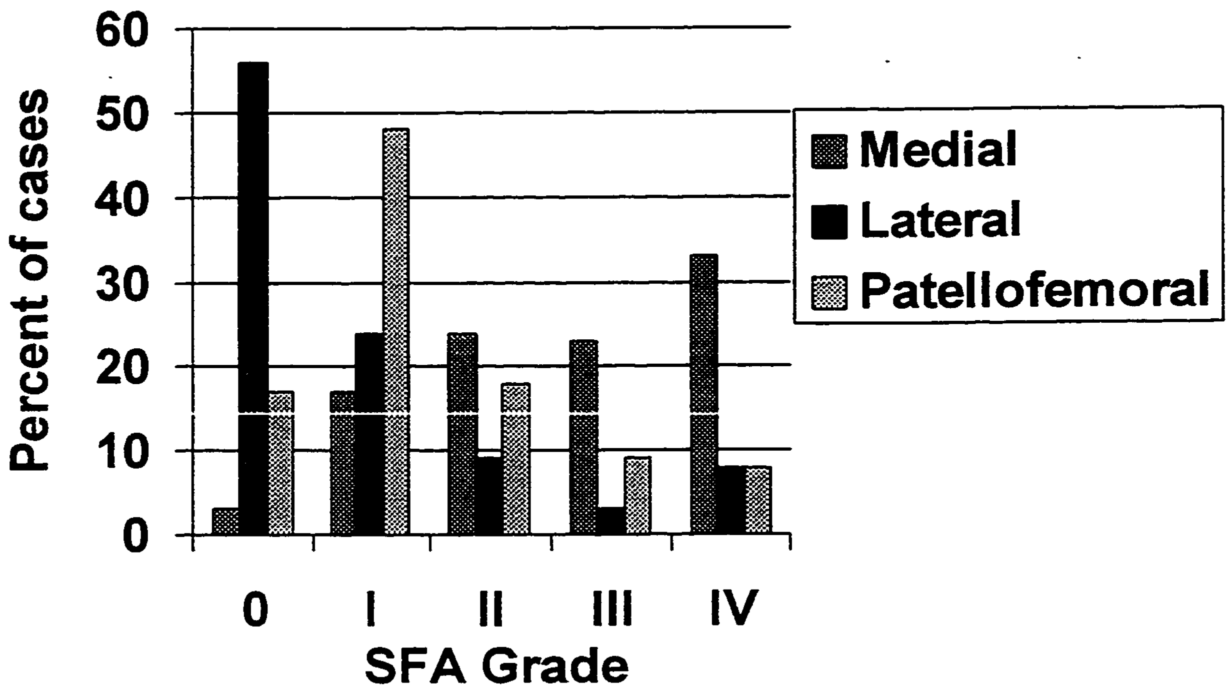
3.1 Descriptive statistics

Two hundred and seven patients were referred for admission into the study once the participating surgeon had scheduled arthroscopic surgery. Thirty five did not meet the criteria for established osteoarthritis as defined, and five others had previous surgery ; anterior cruciate ligament reconstruction (2), medial menisectomy (2) and osteochondritis dissecans (1). Of the remaining 167 patients who met eligibility criteria, 15 were ultimately excluded because of unsatisfactory video recording which could not be reliably interpreted post operatively by the assessor blinded to the patient's preoperative criteria. The characteristics of 152 patients who comprise the study cohort are described in Table 1. The mean age of subjects was 60.5 years and gender and body weight (as defined by categorical BMI) were equally represented. Pain was the most common presenting complaint, particularly with stair climbing and arising from a chair. The cohort demonstrated a wide range of disability as measured by the baseline WOMAC pain score of 24.2 ± 10.5 (mean $\pm$ S.D.). The range was 3 to 49 on a visual analog scale that spans 0 – 50. Ninety two patients had an unstable meniscal tear as determined independently by the principal investigator following videotape review. The vast majority were degenerative tears. None were felt to be repairable.

Table 1. Characteristics of 152 patients entered into study.	
Characteristic	
Age in yrs (mean ± SD)	60.5 ± 8.5
Range	43 - 75
Women (%)	51
Obesity (Body Mass Index > 27) (%)	51
Workers Compensation(%)	5
Current NSAID/analgesic use (%)	88
Unstable meniscal tear (%)	
Medial	49
Lateral	10
No tear	41

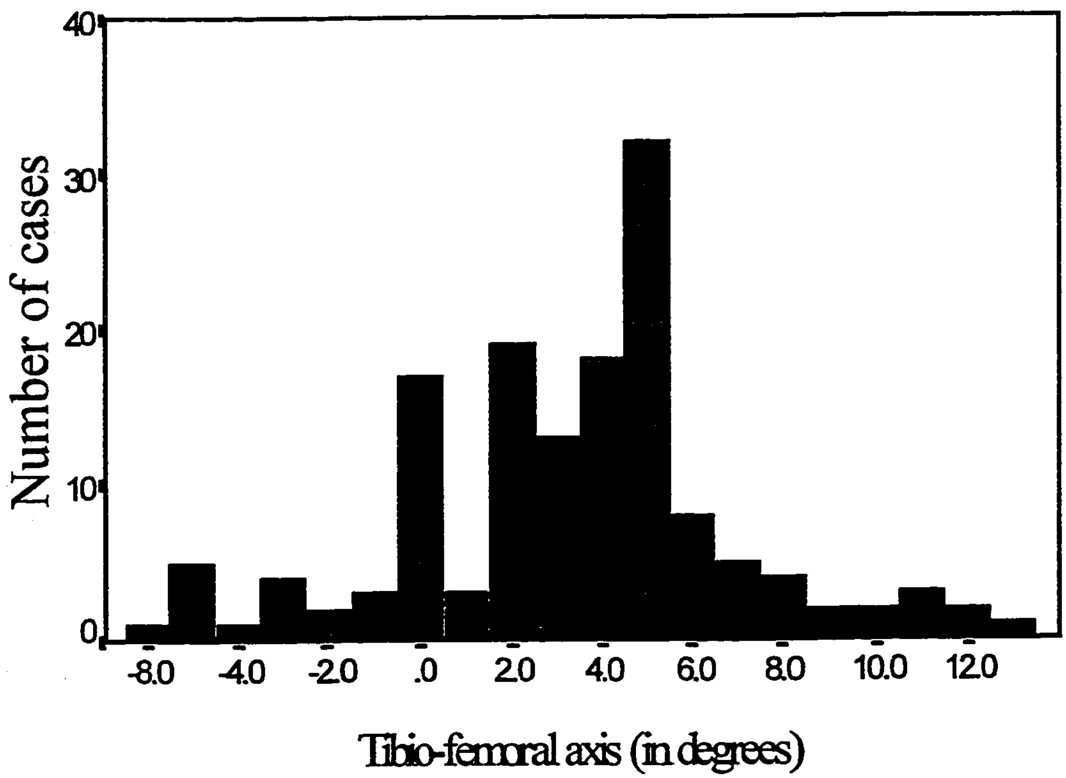
Figure 1 localizes the distribution of chondral damage severity graded arthroscopically for all cases. The medial compartment was considerably more damaged with 57 % showing Grade III or IV involvement. In contrast, both the lateral and the patellofemoral were less severely involved with only 13% and 17% Grade III and IV changes respectively. Accordingly, the spectrum of articular wear severity was more evenly distributed for the medial compartment than for the other 2 compartments. A record was not kept for those patients presenting to the Orthopaedic clinic for whom arthroscopy was not offered or offered and refused. Nonetheless, the distribution of pain symptoms and arthroscopic severity suggests that there was a complete range of disability and disease severity.

Figure 1. Distribution of French Society of Arthroscopy (SFA) articular cartilage grading by arthroscopy of knee compartment of enrolled patients.



The coronal tibio-femoral[62] measures for the three-foot standing film are depicted in Figure 2. The 25th, 50th and 75th percentile values were 2, 4 and 5 degrees respectively, illustrating a preponderance of knees with varus alignment. Varus alignment was defined as 3 degrees or less, neutral as between 4 and 8 degrees, and valgus as 9 degrees or more based on values established by Moreland with healthy adult males[62].

Figure 2. Histogram of standing radiograph coronal tibio-femoral alignment (in degrees) for 152 patients with osteoarthritis of the knee.



3.2 Physicians' prediction

Physicians in this study felt that the majority of patients presenting with persisting symptoms from osteoarthritis suffered from unstable meniscal tears. The fellows assessed all 152 patients while the staff assessed 130. There were no differences between the two samples with respect to age, sex, obesity, or incidence of unstable meniscal tear. Based on a probability of tear of 50% or greater, fellows indicated that 84% of patients had such tears whereas the staff felt that 85 % of patients had tears. Both groups overestimated the frequency of tears which was 59%. Figures 3 and 4 show the fellows' and staff physicians' predicted probabilities as a function of presence of unstable meniscal tear. There was no clinically significant difference among the clinicians with respect to diagnostic accuracy. It should be noted that for staff physicians, who selected patients for arthroscopy, 16 patients were felt to have only 30% or less probability of an unstable meniscal tear. This suggests that these patients were selected for surgery for reasons other than meniscal tear. It is possible that these patients were being considered for a more involved procedure such as osteotomy or knee arthroplasty and a trial of arthroscopy, a more benign procedure was being offered. It may also be that the attending physician felt the patient could benefit from arthroscopy for reasons other than meniscal tear. We did not survey the attitudes of physicians about their reasons for patient selection for surgery. Tables 2 and 3 show the classification performance of predicted tears by the fellows and the staff respectively showing remarkably similar performance characteristics for these two groups of physicians. Clinical experience did not appear to be an advantage for predicting tears. Using a threshold of 0.5 as positive prediction of a meniscal tear, the accuracy of prediction was 60% for both groups. The receiver operating characteristic (ROC) curve graphically depicts the accuracy of prediction and the trade off between true positive and false positive rates and is presented for both the clinical fellows and the attending staff (Figure 5). The shape of the curves show only slightly better determination than chance alone, which is the diagonal, which represents an area of 0.5 under the curve. The areas under the curve for both fellows and staff predictions were determined with their 95% confidence intervals; for fellows, 0.56 (0.47, 0.65) and for staff 0.62 (0.52, 0.72). No statistically significant difference was found between the group predictions ($p = 0.24$).

Figure 3. Clinical fellows' predicted probability of any unstable meniscal tear as a function of actual tear at arthroscopy.

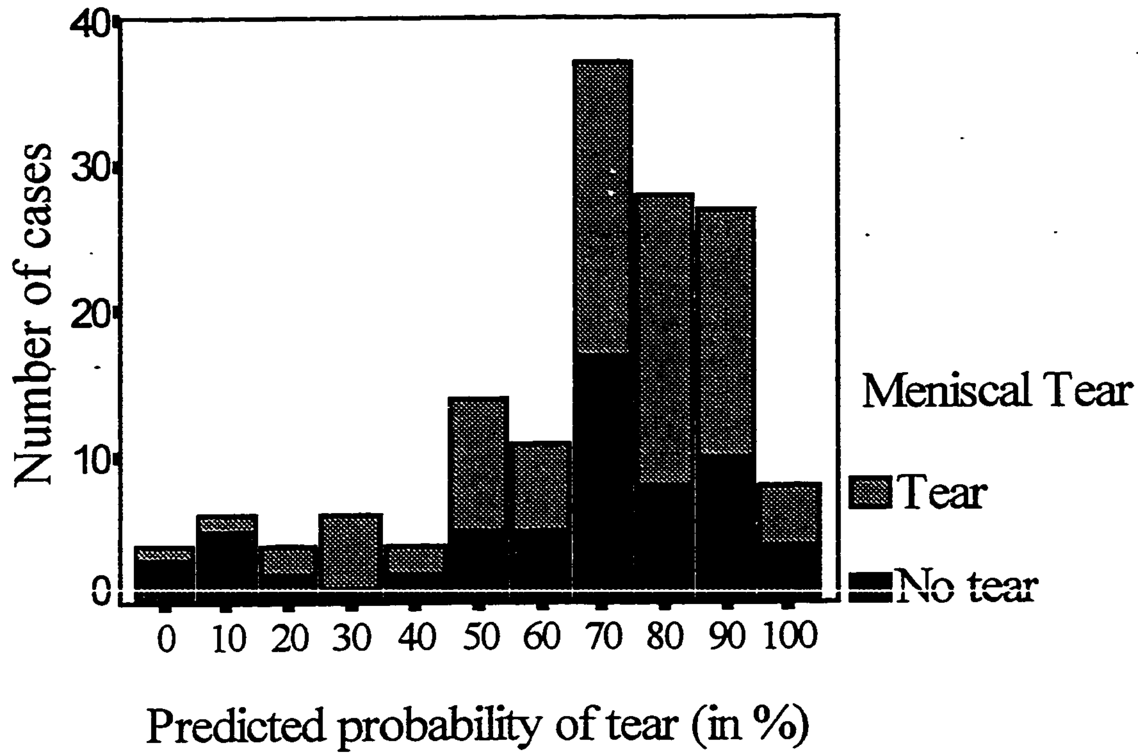


Table 2. Classification performance of fellows' prediction of unstable meniscal tear using 50% as probability threshold

		Observed Tear	
		Yes	No
Predicted Tear	Yes	76	49
	No	11	13
Total		149	
Sensitivity		**87%**	
Specificity		**21%**	
Accuracy		**60%**	
Posit. Pred. Value		**61%**	
Negat. Pred. Value		**54%**	

Figure 4. Staff predicted probability of any unstable meniscal tear as a function of actual tear at arthroscopy.

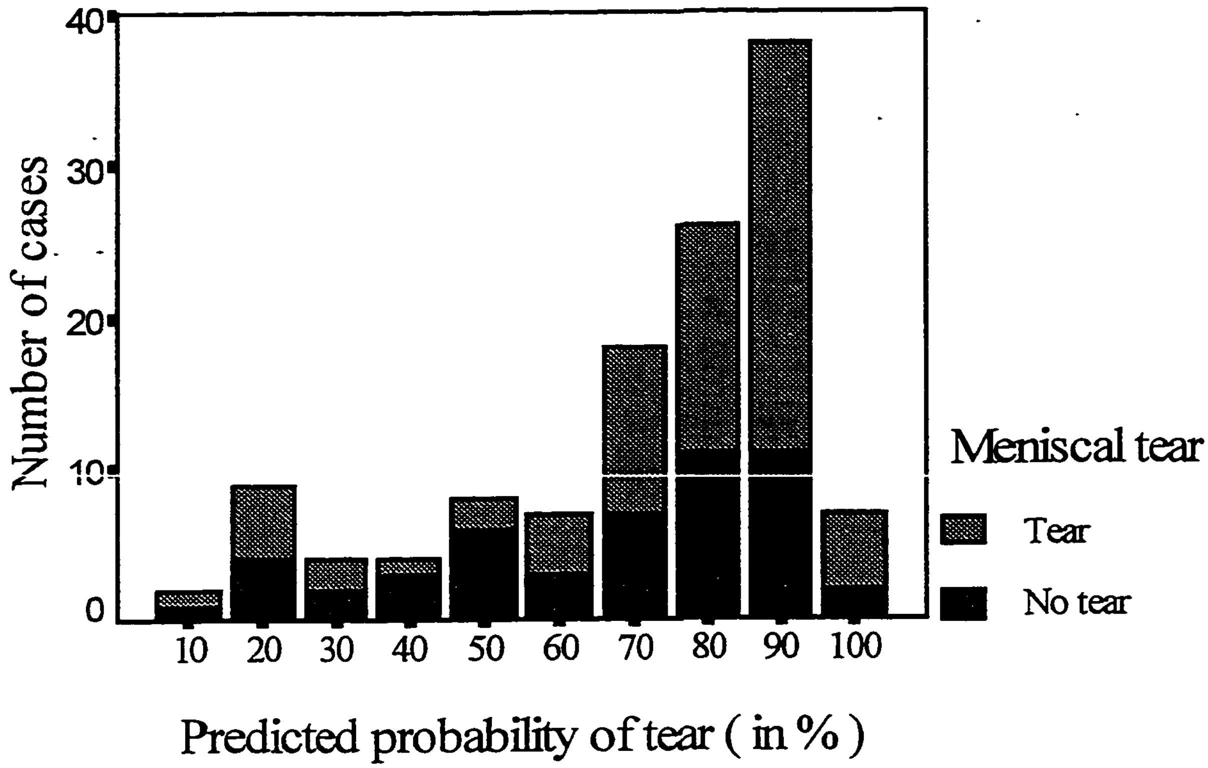
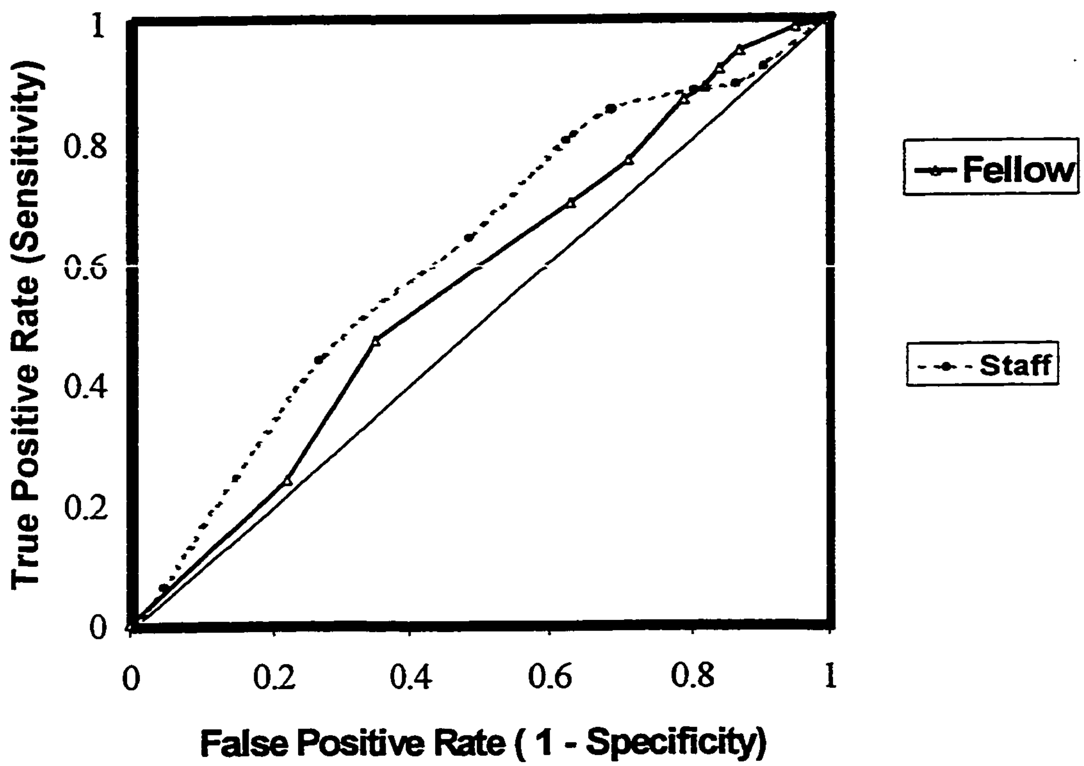


Table 3. Classification performance of staff prediction of unstable meniscal tear using 50% as probability threshold

		Observed Tear	
		Yes	No
Predicted Tear	Yes	64	40
	No	9	10
Total		123	
Sensitivity		**88%**	
Specificity		**20%**	
Accuracy		**60%**	
Posit. Pred. Value		**62%**	
Negat. Pred. Value		**53%**	

Figure 5. Receiver Operating Characteristic (ROC) curve of clinical fellow and staff prediction of unstable meniscal tears.



3.3 Interobserver reliability

The highest level of agreement for predictor variables by history and physical examination was 0.42 despite the production of a study manual which described each clinical variable in detail (Table 4). We found low kappa values of interobserver agreement for commonly cited tests such as the MacMurray test (0.15), circumduction/rotation at full flexion (0.22), and joint line tenderness (0.24). The extent of disagreement was disappointing as these variables are among the most often used clinical signs for evaluating internal derangements of the knee, a common presenting orthopaedic complaint. Agreement was much better for radiographic indices which were dichotomised for greater clinical relevance.

The poor Kappa values can be partly explained by the prevalence of the individual variables.[79] Thompson and Walter have shown that the Kappa statistic is a function of sensitivity, specificity, and prevalence of the measured variable. Prevalence assumes a greater relative importance when it nears 0 or 100%, and the value of kappa approaches 0 in either of these conditions. This was especially evident for the signs of ligament instability which were seen in 6 - 17% of cases ($\kappa < 0.02$)and at the other extreme, pain in flexion was recorded in 87% of cases ($\kappa = 0.07$). We sought to optimise this agreement a priori by providing a study manual with explicit description of all the variables, although a formal training session was not provided.

Table 4. Total interobserver and kappa agreements between participating staff and fellows for predictor variables in 135 patients with symptomatic osteoarthritis of the knee.

Predictor variable	Percent Agreement	Kappa (95% C.I.)
HISTORY		
Acute injury	61	0.21 (0.03, 0.39)
Swelling	69	0.33 (0.17, 0.49)
Giving way	60	0.12 (-0.04, 0.28)
Locking	80	0.44 (0.26, 0.62)
Pain:		
Generalised location	60	-0.03 (-0.15, 0.21)
Focal location	74	0.11 (-0.08, 0.30)
At rest	56	0.16 (0.0, 0.32)
Arising from chair	78	0.25 (0.05, 0.45)
With stairs	90	0.21 (-0.06, 0.48)
PHYSICAL FINDINGS		
Tenderness:		
Medial joint line	79	0.21 (0.01, 0.41)
Lateral joint line	70	0.25 (0.07, 0.43)
Pain with passive motion:		
Full Flexion	67	0.09 (-0.07, 0.25)
Anteromedial	59	0.14 (-0.05, 0.33)
Posteromedial	52	0.10 (-0.09, 0.29)
Anterolateral	65	0.18 (0.0, 0.37)
Posterolateral	72	0.05 (-0.17, 0.27)
Full Extension	57	0.10 (-0.08, 0.28)
Anteromedial	57	0.06 (-0.14, 0.26)
Posteromedial	48	-0.02 (-0.22, 0.18)
Anterolateral	71	0.02 (0.0, 0.40)
Posterolateral	76	0.01 (-0.17, 0.19)
Special Tests:		
MacMurray	59	0.16 (-0.01, 0.33)
Steinman	52	0.05 (-0.11, 0.21)
Circumduction	63	0.21 (0.05, 0.37)

Table 4. (continued)

Predictor variable	Percent Agreement	Kappa (95% C.I.)
Instability:		
Varus	93	0 (-0.18, 0.18)
Valgus	92	0.05 (-0.13, 0.23)
Lachman	96	-0.08 (-0.12, -0.04)
Abnormal Extension	50	0.07 (-0.05, 0.19)
Effusion:		
Visible	72	0.28 (0.10, 0.46)
Ballotment	62	0.19 (0.02, 0.36)
Sweep	58	0.14 (-0.04, 0.32)
RADIOGRAPHIC SIGNS		
3-Foot standing:		
Tibiofemoral varus (≤4°)	75	0.29 (0.13, 0.45)
Medial joint space ≤ 2 m	80	0.63 (0.49, 0.77)
Lateral joint space ≤ 3 mm	86	0.72 (0.49, 0.95)
45° Posteroanterior weight bearing:		
Tibiofemoral varus (≤4°)	88	0.63 (0.51, 0.75)
Medial joint space ≤ 2 mm	83	0.58 (0.44, 0.72)
Lateral joint space ≤ 2 mm	91	0.80 (0.62, 0.98)

3.4 Derivation of rule for all unstable meniscal tears.

3.4.1 Univariate analysis

Univariate analysis using chi square for nominal and independent Student's *t*-test for age were considered for the prediction of any meniscal tear (medial or lateral). Given the lack of strong agreement ($\kappa < 0.6$) between staff and fellows for most of the predictor variables, we elected to use the staff's assessment of predictor variables for entry into the model. We hypothesized that there was a level of experience which differentiated the staff from fellows and that the staff findings would be more generalizable to practising orthopaedic surgeons. The major finding which supported this premise was the association of MacMurray test which has been consistently regarded as the most provocative test and was more positively associated with tears in the staff (p=0.08) than the fellows (p=0.51). The negative consequence of this decision was a sacrifice of power in the study as we originally intended to derive the rule using the fellows' assessments which were intended to capture all of the patients in the study whereas staff assessments were encouraged whenever possible for reliability testing. Fortunately, there were enough tears to satisfy the requirements of a minimum 5 events per variable entered into a logistic regression.

The excluded patients did not differ significantly for age, sex or outcome. (Table 5) The values for univariate association with any unstable meniscal tear are given in Table 6.

Table 5. Characteristics of included and excluded patients from univariate and multivariate analysis.

Characteristic	Included	Excluded	P value
Number	130	22	
Mean age (S.D.)	60.2 (8.4)	57.2 (8.4)	NS*
Sex (M/F)	67/63	10/12	NS
Meniscal tear(%)			NS
No tear	42	42	
Medial tear	50	46	
Lateral tear	8	12	

* Not significant at alpha level of 0.05

Table 6. Univariate association of staff predictor variables versus ***all unstable meniscal tears*** at arthroscopy in 130 patients with osteoarthritis of the knee.

Clinical Finding	Resectable Tear		Chi Square	P Value
	Yes n=76 (%)	No n=54 (%)		
1. DEMOGRAPHICS				
Mean age (S.D.)	60 (10)*	59 (8)*	-	0.51**
Male	55	42	2.65	**0.10**
Obesity (BMI>27)	55	50	2.14	0.34
2. HISTORY				
Injury	37	35	0.07	0.79
Swelling	49	72	6.40	**0.01**
Giving Way	49	32	3.54	**0.06**
Locking	24	19	0.41	0.52
Pain:				
Generalised	65	72	0.78	0.38
Focal	81	94	4.36	**0.04**
Rest	40	58	4.11	**0.04**
Rising from chair	82	90	1.89	**0.17**
With stairs	89	94	1.34	0.25
3. PHYSICAL SIGNS				
Tenderness:				
Medial Joint Line	90	91	0.04	0.83
Lateral Joint Line	22	36	3.13	**0.08**
Pain:				
Passive Flexion	87	87	0.002	0.99
Anteromedial	51	64	2.07	**0.15**
Posteromedial	52	31	5.67	**0.02**
Anterolateral	18	25	0.87	0.35
Posterolateral	13	10	0.34	0.56

Table 6. (continued)

Clinical Finding	Resectable Tear Yes	Resectable Tear No	Chi Square	P Value
Passive Extension	67	67	0.002	0.96
Anteromedial	51	52	0.02	0.89
Posteromedial	31	20	1.71	**0.19**
Anterolateral	19	9	0.06	0.81
Posterolateral	3	11	2.66	**0.10**
Provocative Tests:				
MacMurray Test	67	52	3.08	**0.08**
Steinman	49	43	0.35	0.55
Circumduction	74	67	0.61	0.43
Instability:				
Varus	18	17	0.05	0.82
Valgus	17	18	0.06	0.81
Lachman	6	6	0.00	1.00
Abnormal Extension	60	45	2.64	**0.10**
Effusion:				
Sweep	40	50	1.27	0.26
Ballotment	24	39	3.30	**0.07**
Visible	30	31	0.04	0.84
4. RADIOGRAPHIC SIGNS				
3-Foot standing:				
Tibiofemoral varus (≤4°)	67	46	6.30	**0.01**
Medial joint space ≤ 2 mm	40	34	0.62	0.43
Lateral joint space ≤ 3 mm	1	11	6.90	**0.008**
45° Posteroanterior weight bearing:				
Tibiofemoral varus (≤4°)	67	50	4.50	**0.03**
Medial joint space ≤ 2 mm	49	35	2.80	**0.09**
Lateral joint space ≤ 2 mm	3	20	11.11	**<0.001**

* Numbers given are mean ± standard deviation
** Student's *t* -test

3.4.2 Multivariate analysis

Those variables with best univariate association ($p < 0.2$) and Kappa agreement of 0.15 or greater were made available to the model. (Table 7)

Table 7. Variables considered in the logistic regression analysis for prediction of all unstable meniscal tears

Predictor variable	chi square p value	kappa
HISTORY		
Swelling	0.01	0.33
Pain:		
At rest	0.04	0.16
Arising from chair	0.17	0.26
PHYSICAL FINDINGS		
Tenderness:		
Lateral joint line	0.08	0.25
Special Tests:		
MacMurray	0.08	0.16
Effusion:		
Ballotable	0.07	0.19
RADIOGRAPHIC SIGNS		
3-Foot standing:		
Tibiofemoral varus (≤4°)	0.01	0.29
Lateral joint space ≤ 3 mm	0.008	0.72
45° Posteroanterior weight bearing:		
Tibiofemoral varus (≤4°)	0.03	0.63
Medial joint space ≤ 2 mm	0.09	0.58
Lateral joint space ≤ 2 mm	<0.001	0.80

A forward stepwise model was used for 113 patients (after listwise deletion) with entry and removal criteria of 0.10 significance and the coefficients were established using maximum likelihood method. The analysis yielded 3 significant variables. The odds ratios for each variable are listed along with the 95% confidence intervals (Table 8). The model chi square is 20.82 (df = 3, p = 0.0001). The Hosmer - Lemeshow goodness of fit statistic was not significant (chi square = .165, df = 4, p = 0.99) indicating an adequate model based on the derivation set.

Table 8. Odds ratios and 95 % confidence intervals for predictor variables in the logistic regression model predicting any unstable meniscal tear.

Variable	Odds Ratio	95 % Confidence Interval
Positive MacMurray test	2.69	(1.11, 6.47)
History of swelling	0.24	(0.10, 0.61)
Lateral joint space >3 mm (45 °PA X-ray)	5.68	(1.29, 24.96)

The model yielded 2 variables, the MacMurray test (positive predictor) and a history of swelling (negative predictor) which are clinically sensible in the context of predicting an unstable meniscal tear. The radiographic measure is more difficult to understand given that tears may be either medial or lateral. While it had a sensitivity of 97% for any unstable tear, the specificity was only 20%. We felt that clinicians would not accept such a low specificity for a variable and we sought to further explore how the radiographs might aid in the accuracy of prediction. We also felt that the rule should provide the location of meniscal tear (medial or lateral) so that the diagnosis could be more specific.

3.5 Derivation for medial meniscal tears.

3.5.1 Rationale

The initial goal was to predict any unstable tear but the distribution of tears was significantly skewed in favour of medial tears (77 vs. 15). Given that many of the predictor variables were localized to one area of the joint (i.e. flexion pain, extension pain, joint line tenderness, it was assumed that there would be greater clinical relevance and sensibility to rules specific for a medial meniscal tear, which is more consistent with clinical practice where a specific diagnosis is sought. The small number of lateral meniscal tears would not allow sufficient statistical power to permit a stable logistic model and therefore no attempt was made to develop a separate decision rule for these tears.[80]

3.5.2 Statistical analysis

The methodology described in section 3.4 for deriving the rule for all tears was applied to derive a rule predicting only unstable medial meniscal tears. Those variables with best univariate association ($p< 0.2$,) and kappa agreement (≥ 0.15) were made available to the logistic regression analysis (Table 9,10).

A forward stepwise model was used with entry and removal criteria of 0.10 significance and the coefficients were estimated using maximum likelihood. One hundred and thirteen cases remained for consideration after listwise deletion for missing values. The analysis yielded 3 significant variables. A history of swelling and the presence of a ballottable effusion were predictors of no medial meniscal tear, while a positive MacMurray sign was a positive predictor of a medial meniscal tear. The odds ratios for each variable are listed along with the 95% confidence interval (Table11). The model chi square is 13.62 (df = 3, p = 0.0035). The Hosmer - Lemeshow goodness of fit statistic was not significant (p =0.58) indicating an adequate model based on the derivation set (Table 12).

Table 9. Univariate association of staff predictor variables versus ***unstable medial meniscal*** tear at arthroscopy in 130 patients with osteoarthritis of the knee.*

Clinical Finding	Resectable Tear Yes n=66 (%)	Resectable Tear No n=64 (%)	Chi Square 1 DF	P Value
1. DEMOGRAPHICS				
Age (mean ± S.D.)	59 (9)*	58 (9)*	-	0.51**
Male	57	43	3.10	**0.08**
Obesity (BMI>27)	54	50	0.64	0.73
2. HISTORY				
Injury	34	38	0.26	0.61
Swelling	48	69	6.16	**0.01**
Giving Way	42	42	0.006	0.94
Locking	22	22	0.002	0.96
Pain:				
Generalised	64	71	0.68	0.41
Focal	81	92	4.47	**0.06**
Rest	39	55	3.14	**0.07**
Rising from chair	81	89	1.55	0.21
With stairs	89	92	0.37	0.54
3. PHYSICAL SIGNS				
Tenderness:				
Medial Joint Line	89	91	0.10	0.75
Lateral Joint Line	24	31	0.69	0.41
Pain with passive motion:				
Passive Flexion	89	85	0.56	0.46
Anteromedial	58	68	1.28	0.26
Posteromedial	58	40	3.54	**0.06**
Anterolateral	24	24	0.001	1.00
Posterolateral	15	11	0.47	0.49

Table 9. Cont'd.

	Resectable Tear Yes	Resectable Tear No	Chi Square 1 DF	P Value
Passive Extension	69	66	0.14	0.71
Anteromedial	58	68	1.28	0.26
Posteromedial	43	30	1.77	**0.18**
Anterolateral	21	34	2.20	**0.14**
Posterolateral	4	11	1.76	**0.18**
Provocative Tests:				
MacMurray Test	66	53	1.77	**0.18**
Steinman	50	43	0.62	0.43
Circumduction	74	61	0.34	0.59
Instability:				
Varus	17	18	0.04	0.85
Valgus	14	21	1.22	0.27
Lachman	6	6	0.00	1.00
Abnormal Extension	38	54	3.10	**0.08**
Effusion:				
Sweep	37	52	2.94	**0.08**
Ballotment	19	41	7.30	**0.007**
Visible	28	33	0.49	0.48
4. RADIOGRAPHIC SIGNS				
3-Foot standing:				
Tibiofemoral varus (≤4°)	68	51	4.3	**0.04**
Medial joint space ≤ 2 mm	41	35	.44	0.63
Lateral joint space ≤ 3 mm	2	9	3.9	0.72
45° Posteroanterior weight bearing:				
Tibiofemoral varus (≤4°)	69	52	4.5	**0.03**
Medial joint space ≤ 2 mm	47	40	.68	0.41
Lateral joint space ≤ 2 mm	4	16	6.0	**0.01**

* Numbers given are mean ± standard deviation

** Student's *t* -test

Table 10. Variables considered in the logistic regression analysis for prediction of unstable medial meniscal tears in all patients.

Predictor variable	chi square p value	kappa
HISTORY		
Male	0.08	-
Swelling	0.01	0.33
Pain:		
At rest	0.07	0.16
PHYSICAL FINDINGS		
Special Tests:		
MacMurray	0.18	0.16
Effusion:		
Ballottable	0.007	0.19
RADIOGRAPHIC SIGNS		
3-Foot standing:		
Tibiofemoral varus (≤4°)	0.04	0.29
45° Posteroanterior weight bearing:		
Tibiofemoral varus (≤4°)	0.03	0.63
Lateral joint space ≤ 2 mm	0.01	0.80

Table 11. Odds ratios and 95 % confidence intervals for predictor variables remaining in logistic regression model predicting an unstable medial meniscal tear.

Variable	Odds Ratio	95 % Confidence Interval
History of swelling	.42	(0.19 , 0.96)
Ballottable effusion	.38	(0.16, 0.93)
Positive MacMurray test	2.21	(0.96 , 5.1)

Table 12. Hosmer and Lemeshow Goodness-of-Fit Test for logistic regression model predicting medial meniscal tears.

	No tear		Medial tear		
Group	Observed	Expected	Observed	Expected	Total
1	6	7	3	2	9
2	14	12	4	6	18
3	12	10	4	6	16
4	2	3	3	3	5
5	9	12	18	18	27
6	7	8	11	11	18
7	6	5	14	15	20

	chi-square	df	Significance
Goodness-of-fit test	3.73	5	.58

The probability of medial meniscal tear can be calculated for any patient by the formula:

$$\text{Prob (tear)} = 1 / (1 + e^{-(\alpha + \beta 1X1 + \beta 2X2 + \beta 3X3)})$$

where: α (0.33) is the value of the constant, $\beta 1$ (-0.87), $\beta 2$ (-0.97), $\beta 3$ (0.79) are the individual coefficients for those predictor variables remaining in the equation; and X_1, X_2 and X_3 are the specific values respectively of those variables for a particular case. We can determine the probability then of an individual who had no history of swelling, no ballottable effusion and a positive MacMurray sign by

$$\text{Prob (tear)} = 1 / (1 + e^{-(.33 - .97*0 + .79*1 - .87*0)}) = 1 /(1 + 0.29) = 0.77$$

The classification performance of the model can be observed for a probability of tear threshold of 0.5 (Table 13). It is apparent that at this threshold , this model has a greater sensitivity than specificity for predicting a medial meniscal tear. The ROC curve is given in Figure 7 and is discussed in Section 3.6.2.

Table 13. Classification performance of logistic regression model predicting medial meniscal tear using threshold of 0.5

		Observed Tear	
		Yes	No
Predicted Tear	Yes	46	24
	No	11	32

Sensitivity	**81%**
Specificity	**57%**
Accuracy	**69%**
Posit. Pred Value	**66%**
Negat. Pred. Value	**74%**

3.6 Subgroup analysis

3.6.1 Rationale

No radiographic variable remained in the multivariable model predicting unstable medial meniscal tears. We elected to split the data set into two sets of patients corresponding to different degrees of osteoarthritis defined by the chondral damage scored with the French Society of Arthroscopy (SFA) method. We felt that the preoperative radiographic information which is routinely obtained should be explored for further prediction assessments. A comparison of the three-foot standing and the 45° posteroanterior weight bearing view was done to determine predictive utility and other pertinent clinical factors.

a) Discrimination of medial compartment chondropathy: Both the three-foot and the 45° views showed similar predictions of severe chondropathy (defined as SFA Grades III and IV).

Using 2 mm of measured joint space as a cutpoint for chondropathy, the classification performances of both views were as follows:

	3 foot standing	45° posteroanterior weight bearing
Sensitivity (%)	82	82
Specificity (%)	60	63
Accuracy (%)	67	71

b) Discrimination of lateral compartment chondropathy: The 45° posteroanterior view was significantly more sensitive at predicting severe lateral chondropathy (defined as SFA Grades III and IV). Using 2mm of measured joint space as a cutpoint for chondropathy, the classification performances of both views were as follows:

	3 foot standing	45° posteroanterior weight bearing
Sensitivity (%)	50	100
Specificity (%)	98	96
Accuracy (%)	94	96

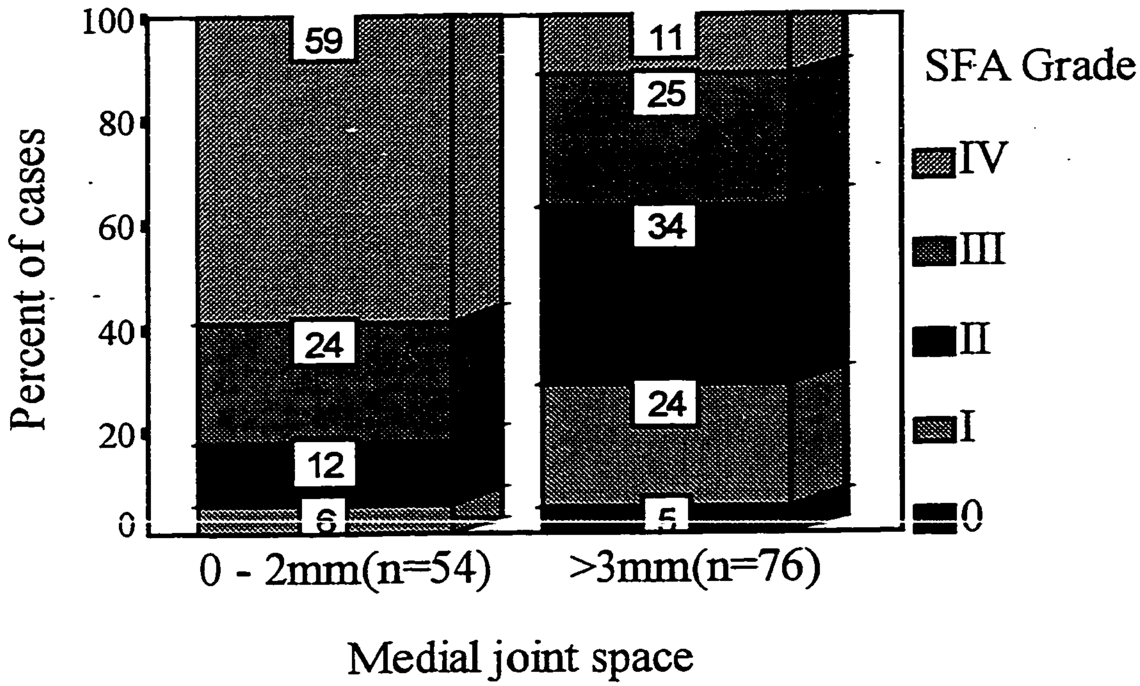
c)Reliability of measure: No significant difference was found between the kappa measures of interobserver agreement, displayed below with confidence intervals:

3 foot standing	45° posteroanterior weight bearing
0.63 (0.49, 0.77)	0.58 (0.44, 0.72)

d) Other: There are other issues which we believe favour the use of the 45° posteroanterior view. The technical and professional charge of the 45° view is $20.25CAN versus $30.23CAN for the three foot standing view according to the Ontario Ministry of Health's 1998 Schedule of Benefits. A regular sized viewbox is adequate for the 45° posteroanterior view allowing easier use. The presence of the other knee on the same projection allows for ease of comparison for the clinician and education for the patient to graphically depict asymmetries of joint space or the presence of osteophytes. This view can be duplicated in smaller Radiology departments of community hospitals and outpatient clinics where the majority of these patients are initially evaluated. The only disadvantage is the inability to calculate the mechanical weight bearing axis as with the 3 foot standing view. Nevertheless, it can be argued that the latter can be ordered on an individual basis and is more pertinent when more advanced procedures such as osteotomy and knee arthroplasty are being considered, usually in larger centres with full orthopaedic coverage and larger Radiology departments.

On the basis of the aforementioned comparisons , we elected to analyse the subgroup performance (for severity of medial compartment osteoarthritis) of the prediction rule by splitting the data set according to the medial joint space measured on the preoperative 45° posteroanterior weight bearing view. A cutpoint of 2mm was used after exploring the data and determining that this was the best discriminator of chondropathy. Using this cutpoint, 83% of patients (n = 56) with medial joint space of 2mm or less had arthroscopically proven chondropathy of SFA grades III and IV. Conversely, only 36% of patients (n = 69) had a similar degree of chondropathy and predominantly Grade III when medial joint space was 3mm or more. (Figure 6)

Figure 6. French Society of Arthroscopy (SFA) chondropathy scores vs. medial joint space height on preoperative 45° posteroanterior weight bearing view.



3.6.2 Statistical Analysis

Similar criteria were used in the selection of predictor variables for use in a logistic regression as was done for the entire derivation set. Those variables offered to the logistic regression are given in Table 14. Radiographic variables were omitted because the subgrouping was made according to radiographic evaluation of the medial compartment. A forward, stepwise methodology was used with entry and exit criteria set at a 0.10 significance level. Fifty four patients were entered into the model after listwise deletion. The model chi-square is 18.67 (df = 3, p = 0.0003) and the Hosmer and Lemeshow goodness of fit test chi-square was 1.01 (df = 4, p =

0.91). The model yielded similar variables as were found in the entire sample, but with larger coefficients (Table 15) and greater classification accuracy (Table 16).

Table 14. Variables considered in the logistic regression analysis for prediction of unstable medial meniscal tears in patients with severe medial joint space narrowing on 45° PA view.

Predictor variable	chi square p value	kappa
HISTORY		
Swelling	0.01	0.33
Locking	0.09	0.44
Pain:		
At rest	0.04	0.16
PHYSICAL FINDINGS		
Special Tests:		
MacMurray	0.2	0.16
Effusion:		
Ballotable	0.02	0.19
Visible	0.11	0.28

Table 15. Odds ratios and 95 % confidence intervals for predictor variables remaining in logistic regression model predicting medial meniscal tear in patients with medial joint space ≤ 2mm on 45°PA Xray.

Variable	Odds Ratio	95 % confidence intervals
History of swelling	.19	(0.04 , 0.84)
Ballottable effusion	.12	(0.02 , 0.60)
Positive MacMurray test	13.9	(2.3 , 85.3)

Table 16. Classification performance of logistic regression model predicting medial meniscal tear in patients with medial joint space ≤ 2mm on 45°PA Xray, using threshold of 0.5.

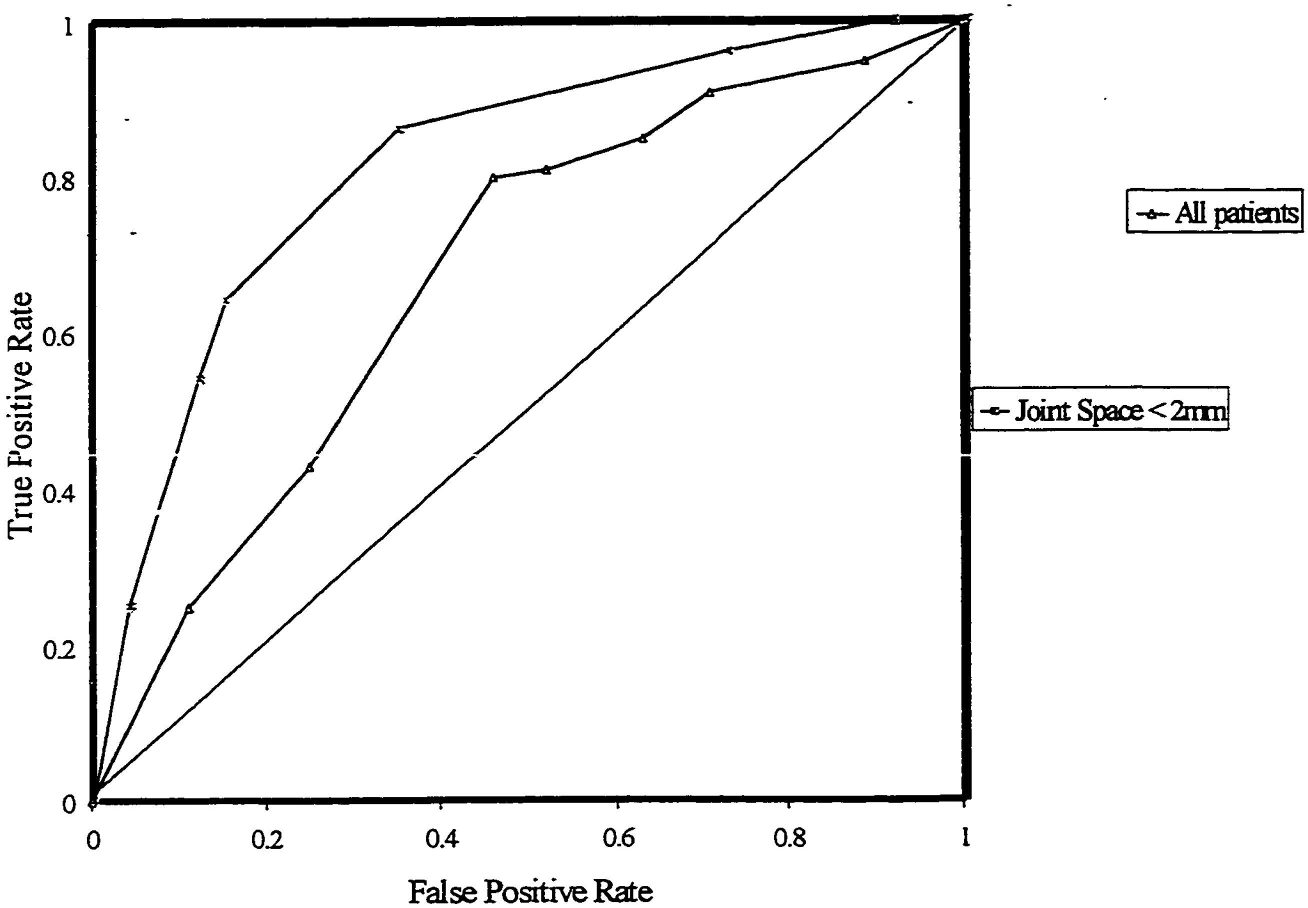
		Observed tear	
		Yes	No
Predicted tear	Yes	24	9
	No	4	17

Sensitivity	**86%**
Specificity	**65%**
Accuracy	**76%**
Posit. Pred. Value	**73%**
Negat. Pred. Value	**81%**

The probability thresholds were manipulated for the logistic regression model predicting unstable medial meniscal tears in the entire sample of patients and for the subset with severe radiographic joint space narrowing to create ROC curves for each. (Figure 7) Both curves subtend greater areas than chance and the preoperative predictions of fellows and staff. The area under the curve for the model in all patients was 0.67 (95% C.I.; 0.58, 0.76), significantly better than chance alone. For the subgroup of patients with severe joint space narrowing, the model's area under the curve (0.82 (95% C.I.; 0.71, 0.93)) was significantly greater than for all patients (p = 0.035).

When compared to staff's prediction, the model did not significantly improve the accuracy of prediction for all patients (p = 0.18). The model was however more accurate than the staff's prediction for the subgroup of patients with severe joint space narrowing (p = 0.005)

Figure 7. Receiver operating characteristic (ROC) curves for logistic regression model predicting medial meniscal tears in all patients and in those with severe medial joint space narrowing.



4. DISCUSSION

4.1 Derivation of model

The prediction model resulting from this derivation set of patients with osteoarthritis of the knee yielded a model with no history of swelling, no ballottable effusion, and a positive MacMurray test being significant predictor variables for an unstable medial meniscal tear as defined in this study. This is the first effort to our knowledge which attempted to predict meniscal tears in this group of carefully defined patients. Experienced physicians in this study could accurately classify only 60% of patients prior to surgery. The logistic regression model could only improve the overall accuracy to 69%. The model was considerably more beneficial in those patients with severe radiographic medial joint space narrowing of 2 mm or less, yielding an accuracy of 76%.

4.2 Methodological strengths and weaknesses

The principles for clinical prediction rules established by Wasson et al. and later modified by Laupacis et al. served as a blueprint for the methodology of this study.[45,46]

4.2.1 Derivation set

The cohort for this study was clearly defined though the clinical definition of osteoarthritis is still a matter of controversy. We chose to adhere to the criteria espoused by the consensus group of the American Rheumatological Association because of a lack of clear consistent definition in the literature and a desire to describe the subjects of this study and improve the generalizability of the results. The reliance on crepitus as a salient clinical feature remains troublesome. Though Bergquist[81] found that knee crepitus was very common in a cohort of elderly patients and not a reliable discriminatory sign for the presence of osteoarthritis, Cushnagan et al.[82] found substantial interobserver agreement ($\kappa = 0.64$) for tibio-femoral crepitus in 8 patients with established osteoarthritis of the knee. Ike [83] also reported that bony crepitus was transmitted in the more severely advanced cases if the compartments were stressed (varus and external rotation for medial ; valgus and internal rotation for lateral). They however did not determine the interobserver reliability for their techniques. Similarly, the presence of osteophytes on radiographs is not a specific sign of the disease. Though Altman et al[60] did find osteophytes to be a moderately specific sign of osteoarthritis compared to inflammatory arthritis, no correlation could be made between osteophyte size and arthroscopic damage. Herborg [84] showed osteophytes are not always symptomatic and are associated with ageing. We agree that the American Rheumatological Association definition of osteoarthritis is problematic and does exclude cases where there is clearly joint space narrowing without osteophytosis, indicating articular cartilage wear. We did however select this definition for inclusion into the study as we felt it was the best available means of clinically defining osteoarthritis in order to capture a spectrum of disease severity which we felt was important to the ultimate clinical relevance of our decision rule.

We felt that we captured the spectrum of disease severity both by objective and subjective measures of to minimize the impact of a referral filter bias[52]. The subjective range was revealed in the WOMAC scores and the objective variability was seen in the distribution of radiographic and articular wear at arthroscopy. Though a log was not kept of those patients for whom arthroscopy was not offered or who declined, the cohort for this study did not show an obvious referral filter bias. This is a critical issue for optimising the generalizability of the derived rule for most patients with osteoarthritis of the knee.

4.2.2 Interobserver reliability

Reliability of predictor variables should be established to enhance the reproducibility of the rule for different observers and in a different subset of patients. The kappa agreements in the present study were disappointingly small despite the use of common clinical variables, and the production of a study manual describing each clinical variable. For reasons of feasibility the kappa agreement in this study was tested between clinical fellows and attending staff, two groups of physicians with different levels of clinical experience yet both highly trained in assessment of knee disorders. This may have contributed to differences for more subjective tests such as the MacMurray test (kappa = 0.15) which showed significant differences in univariate association with unstable meniscal tear ($p = 0.51$ for fellows and $p = 0.08$ for staff).

Others have reported on the reliability of physical examination of the knee with variable results. Bergquist et al. [81] studied a cohort of ambulatory senior citizens who were not symptomatic with osteoarthritis and showed acceptable levels of agreement using the kappa statistic for tenderness (.56) and crepitus (.48), while patellar grind (.39) and bony enlargement (.27), fared less well. Stiell et al.[85] found moderate interobserver agreement for medial and lateral joint line tenderness ($\kappa = 0.5$ and 0.45 respectively) and visible effusion ($\kappa = 0.59$) in a cohort of acutely injured patients (mean age 36) assessed in an emergency department though the underlying prevalence of osteoarthritis was unknown. Jones et al [86] showed considerably less interobserver reliability for a number of physical signs from a randomly selected cohort of elderly inpatients.

They reported values for medial and lateral joint line tenderness ($\kappa = 0.48$ and 0.44 respectively) and effusion ($\kappa = 0.22$). This was despite the use of a training session to standardise the scoring.

To our knowledge, the only previously documented study of clinical sign reliability in osteoarthritis of the knee is that of Cushnagan et al. [82]. They studied physical signs in 8 patients with osteoarthritis and found substantial intraobserver agreement for several clinical signs and lower levels of interobserver agreement for joint tenderness ($\kappa = 0.4$) and effusion ($\kappa = 0.28$). We did not measure intraobserver agreement in this study because we were most concerned with interobserver agreement in formulating a decision rule which could be generalizable to other clinicians. It may also be that clinicians are less attentive to clinical signs in this study format where the medical decision has already been made for arthroscopy. The lack of reliable clinical variables was a limitation in the present study.

4.2.3 Outcomes

A prediction rule should have a clinically significant and clearly defined outcome measure. The choice of unstable meniscal tear as the primary dependent variable of interest merits further scrutiny. Many have described the results of partial menisectomy in older subjects and have differentiated outcome based on the presence of degenerative compartmental changes. Arthroscopic partial menisectomy is a well-tolerated and effective procedure in patients aged 40 or more without significant degenerative change[37-41]. The participating surgeons in the present study agreed with resection of all unstable meniscal tears in this cohort though we await the clinical consequences of this when we review our functional outcome results at minimum 2 year follow-up, the subject of a subsequent report. It is conceivable that successful patient outcomes are related to other variables whose prediction are more clinically important. Abrasion arthroplasty was not performed based on our interpretation of the literature suggesting suboptimal results with this procedure[28-30]. Consensus was sought to objectively define an unstable meniscal tear which was largely based on the experience of the participating surgeons.

4.2.4 Mathematical techniques

The mathematical techniques used for building the model were described in sufficient detail to be understood and reproducible. The model yielded only three variables which is quite simple for others to remember. Additional variables might improve the prediction but at the risk of introducing too many variables which diminishes the ease with which it would be used. Another risk is multicollinearity, which refers to the presence of predictor variables which are highly associated with one another. This phenomenon is manifest by increasing the magnitude of the coefficients and their standard errors, creating inaccuracies in the odds ratios which are of greatest use to clinicians. The addition of effusion and swelling would suggest a possibility of multicollinearity but this was not borne out by increasing coefficients, though they were significantly associated (chi-square = 7.1, p = 0.008). The accuracy of the model was tested by applying the goodness-of-fit statistic described by Hosmer and Lemeshow.[87] Their method involves dividing the data set into groups of increasing probability thresholds and calculates a chi-square statistic at each level. A significant value of the goodness-of-fit statistic indicates a significant discrepancy between the observed and expected frequencies of meniscal tears and a poor fit. In this study, the p value was 0.58 indicating a good fit of the model to the data set. Finally, the precision of the model can be evaluated by overall accuracy, the proportion of correct classifications in the entire derivation set at a defined probability threshold generally set at 0.5. Accuracy is dependent on prevalence and can give deceptively high values for inaccurate models in cases where prevalence of the predicted outcome variable is near 0 or 1. In the present study, the accuracy for prediction of medial meniscal tears was improved from the preoperative clinical prediction of 60% to the models for all patients (accuracy 69%) and for those with decreased medial joint space (accuracy 76%). The latter finding was a surprise and one which we believe has greatest potential clinical utility from the study. Receiver operating characteristic (ROC) curves provide a better description of the model's performance for a range of probability thresholds and their resulting sensitivities and specificities. The areas beneath the curve can be used to test hypotheses of significance comparing a priori and post model predictions of tears.

4.3 Clinical Relevance

4.3.1 Non - discriminatory variables

The prediction of meniscal tears in symptomatic arthritic knees is complicated by a few factors. First, several traditional key clinical signs are very common in this subgroup of patients who failed medical management and are not good discriminators. These symptoms and signs which would be otherwise helpful in patients without underlying articular cartilage damage, were pervasive in this sample of patients with established osteoarthritis. Pain with stairs (92%), arising from a chair (83%), medial joint line tenderness (90%) and pain on forced flexion (87%) are examples from the present study. Joint line tenderness in particular has been found to be commonly associated with meniscal tears in the non arthritic population.[58,59] We expected it would also perform well in discriminating unstable meniscal tears in this cohort which provided a spectrum of osteoarthritis severity. In fact, it appears to be less useful in osteoarthritis likely representing discomfort secondary to a number of possible factors including osteophytes, medial collateral ligament bursitis, synovitis or meniscal tear. The study results emphasize that many symptoms and signs which have been thought to be relevant for the diagnosis of medial meniscal tears were not clinically useful. In addition to jointline tenderness, symptoms of locking or giving way were also unhelpful. While this is somewhat disappointing, it does permit the clinician to pay particular attention to those signs which were more pertinent, and allow the assessment to be more efficient.

4.3.2 MacMurray test

The enhanced performance of the model in the subgroup of patients with severe medial joint space narrowing merits further consideration. The hypothesis that signs of meniscal irritation such as the MacMurray test would be less useful with advanced articular disease because of an overlap of symptoms from these two separate pathological entities proved to be false. We speculate that this is so because the MacMurray test is done near full flexion which is compressing the articular surfaces in a non weight bearing location on the femoral side which is generally spared of severe chondral damage until the later stages of osteoarthritis. This minimizes the contact of damaged articular surfaces while optimizing the accessibility to unstable meniscal tears which are compressed by the articular surfaces in full flexion. While this may explain the validity of the test in all patients with osteoarthritis it fails to explain why the test is superior in the more severe subgroup where its performance would be expectedly worse. In fact, the improved performance of the MacMurray test was modest; with equal specificities of only 45%, the test had a sensitivity of 70% in the severe group and 64% in the mild group.

4.3.3 Effusion

Ballotable effusion was not significantly associated with medial compartment articular chondropathy defined as SFA Grades III and IV in this study (O.R. = 1.12). It was however negatively associated with an unstable medial meniscal tears (O.R. = 0.32). The etiology of swelling in osteoarthritis of the knee remains speculative. If neither articular chondropathy nor unstable meniscal tears correlate with effusion and swelling , then what does? It may be that the temporal relationship of chondral breakdown is more important. Joint effusions are secondary to an imbalance of synovial fluid secretion versus resorption and are more common in the setting of active synovial irritation rather than quiescent or latent disease. This hypothesis was not formerly tested in the current study but remains speculative. It does appear however that clinicians should not ascribe swelling solely to meniscal tears on the basis of our findings. It must be noted that patients with effusions were persistently symptomatic despite the absence of unstable meniscal tears or severe chondropathy suggesting that synovial inflammation was indeed a definitive cause of subjective discomfort. The importance of effusion relative to arthroscopic intervention will hopefully be elucidated in the second phase of the study.

4.3.4 Swelling

The subjective sensation of swelling was strongly associated with the physical presence of a ballottable effusion (Chi - Square = 7.1, p = 0.008). Like effusion, it was associated with a normal meniscus. Despite its significance in the multivariable model, it remains a somewhat difficult variable to depend on. Many patients had difficulty answering the question "Does your knee ever swell ?" and initially responded with uncertainty. Some also pointed to localized areas of the knee as areas of swelling which is difficult to rationalize if we accept that the knee is one large compartment and that it should be perceived as generalized. Some also had swelling initially related to an injury but none subsequent to that making the recording of their answer problematic, though any affirmative answer was recorded as such. We feel the question should be refined to relate to ongoing, consistent activity related swelling though we await the clinical implication of the variable in relation to the outcome by quality of life.

4.4 Further study

Refinement of the rule and prospective validation must follow the initial derivation of a prediction rule. There are several areas in this study that must be improved upon. First is the interrator reliability which we hope can be improved by a training session to ensure standardisation. Second is the choice of outcome. Patient outcomes are most important and will form the basis of the refined rule perhaps with unstable meniscal tear as a surrogate measure if at all. The role of arthroscopy in osteoarthritis of the knee remains to be defined. Previous studies show a spectrum of responses to the procedure and it follows that better patient selection is required. We acknowledge that the rule is more accurate for cases where resecting meniscal tears has been felt to be less beneficial in previous studies, a finding that will be closely scrutinized when we evaluate the outcome data from quality of life scales. The current cohort of patients is being prospectively followed to determine the effectiveness of meniscal tear resection after a minimum of 2 years. Chondral flap was also recorded and will be studied as a separate entity. The clinical significance of our rule will ultimately be determined by the results of the procedure using patient relevant and clinically responsive outcome measures such as the WOMAC and SF-36 scores. We hope to answer this question in the subsequent phase of this study when all of the outcomes with minimum two-year follow up are reviewed (July 1999). We are especially intrigued with the role of swelling and effusion as they relate both to reported symptoms and postoperative change. It would appear that these findings, which are negatively associated with unstable medial meniscal tears, are related to subjective symptoms that incite patients to seek medical attention.

5.0 CONCLUSION

The methodological criteria for derivation of a prediction rule for unstable medial meniscal tears in osteoarthritic knees were defined and adhered to in the present study. Unfortunately, a standardized clinical assessment of patients with osteoarthritis of the knee did not yield sufficiently reliable variables to ensure a reproducible prediction rule for unstable meniscal tears. Several interesting points did emerge from univariate analysis. Mechanical symptoms such as locking or giving way and joint line tenderness did not discriminate for the presence of tears as has been traditionally espoused. Logistic regression did yield the following variables (with their odds ratios) for the prediction of unstable medial meniscal tears: history of swelling (0.44), ballottable effusion (0.39), and positive MacMurray test (2.02). The MacMurray test was particularly useful in cases of significant medial joint space narrowing. (odds ratio = 13.9). The clinical relevance of predicting and subsequently resecting an unstable meniscal tear will be verified in the second phase of this ongoing prospective evaluation of arthroscopy for symptomatic osteoarthritis of the knee. In the interim, the variables mentioned can be used by practising orthopaedic surgeons to enhance their diagnostic accuracy in this group of patients. Swelling and effusion in particular should be noted to be negative predictors for unstable tears in this population with articular wear. Caution should be exercised not to apply this rule to non arthritic patients where swelling and effusion may indeed positively predict meniscal tears. Ultimately, a rule tied to improved patient outcomes will be most relevant and attractive to both clinicians and patients.

REFERENCES

1.Hochberg MC, Altman RD, Brandt KD, et al. Guidelines for the medical management of osteoarthritis. *Arthritis Rheum* 1995;38(11):1541-1546.

2.Bradley JD, Brandt KD, Katz BP, Kalasinski LA, Ryan SI. Comparison of an anti-inflammatory dose of ibuprofen, an analgesic dose of ibuprofen, and acetaminophen in the treatment of patients with osteoarthritis of the knee. *N Engl J Med* 1991;325(2):87-91.

3.Bellamy N, Buchanan WW, Chalmers A, et al. A multicenter study of tenoxicam and diclofenac in patients with osteoarthritis of the knee. *Journal of Rheumatology* 1993;20(6):999-1004.

4.Feinberg J, Marzouk D, Sokolec C, Katz B, Bradley J, Brandt K. Effects of isometric versus range of motion exercise on joint pain and function in patients with knee osteoarthritis(abstract). *Arthritis Rheum* 1992;35 (suppl 5):R28

5.Dieppe PA, Sathapatayavongs B, Jones HE, Bacon PA, Ring EFJ. Intra-articular steroids in osteoarthritis. *Rheum Rehabil* 1980;19:212-217.

6.Jones AC, Doherty M. Intra-articular steroids are effective in osteoarthritis but there are no clinical predictors of response. *Ann Rheum Dis* 1996;55:829-832.

7.Peyron J. Intraarticular hyaluronan injections in the treatment of osteoarthritis: state-of-the-art review. *J Rheumatol* 1993;20(Supp39):10-15.

8.Listrat V, Ayral X, Patarnello F, et al. Arthroscopic evaluation of potential structure modifying activity of hyaluronan (Hyalgan) in osteoarthritis of the knee. *Osteoarthritis Cart* 1997;5(3):153-160.

9.Jones AC, Pattrick M, Doherty S, Doherty M. Intra-articular hyaluronic acid compared to intra-articular triamcinolone hexacetonide in inflammatory knee osteoarthritis. *Osteoarthritis Cart* 1995;3(4):269-273.

10.Coventry MB. Upper tibial osteotomy for osteoarthritis. *J Bone Joint Surg Am* 1985;67(7):1136-1140.

11.Coventry MB, Ilstrup DM, Wallrichs SL. Proximal tibial osteotomy. A critical long-term study of eighty-seven cases. *J Bone Joint Surg Am* 1993;75(2):196-201.

12.Hernigou P, Medevielle D, Debeyre J, Goutallier D. Proximal tibial osteotomy for osteoarthritis with varus deformity. A ten to thirteen-year follow-up study. *J Bone Joint Surg Am* 1987;69(3):332-354.

13.Insall JN, Joseph DM, Msika C. High tibial osteotomy for varus gonarthrosis. A long-term follow-up study. *J Bone Joint Surg Am* 1984;66(7):1040-1048.

14.Jackson JP, Waugh W. The technique and complications of upper tibial osteotomy. *J Bone Joint Surg Br* 1974;56(2):236-245.

15.Rubens F, Wellington JL, Bouchard AG. Popliteal artery injury after tibial osteotomy: report of two cases [see comments]. *Can J Surg* 1990;33(4):294-297.

16.Curley P, Eyres K, Brezinova V, Allen M, Chan R, Barnes M. Common peroneal nerve dysfunction after high tibial osteotomy. *J Bone Joint Surg Br* 1990;72(3):405-408.

17.Cameron HU, Welsh RP, Jung YB, Noftall F. Repair of nonunion of tibial osteotomy. *Clin Orthop* 1993;(287):167-169.

18.Malkani AL, Rand JA, Bryan RS, Wallrichs SL. Total knee arthroplasty with the kinematic condylar prosthesis. *J Bone Joint Surg Am* 1995;77-A(3):423-431.

19.Novak PJ, Bach BR. Selection criteria for knee arthroscopy in the osteoarthritic patient. *Orthopaedic Review* 1993;798-804.

20.Wright JG, Coyte P, Hawker G, et al. Variation in orthopaedic surgeons' perceptions of the indications for and outcomes of knee replacement. *Can Med Assoc J* 1995;152(5):687-697.

21. Personal Communication, Ontario Ministry of Health.

22.Ike RW, Arnold WJ, Rothschild EW, Shaw HL. Tidal irrigation versus conservative medical management in patients with osteoarthritis of the knee: a prospective randomized study. *J Rheumatol* 1992;19(5):772-779.

23.Livesley PJ. Articular debridement versus washout for degeneration of the medial femoral condyle [letter; comment]. *Journal of Bone & Joint Surgery - British Volume* 1996;78(5):854

24.Chang RW, Falconer J, Stulberg SD, Arnold WJ, Manheim LM, Dyer AR. A randomized, controlled trial of arthroscopic surgery versus closed-needle joint lavage for patients with osteoarthritis of the knee. *Arth Rheum* 1993;36(3):289-296.

25.Hubbard MJ. Articular debridement versus washout for degeneration of the medial femoral condyle. A five-year study [see comments]. *Journal of Bone & Joint Surgery - British Volume* 1996;78(2):217-219.

26.Johnson L. Arthroscopic abrasion arthroplasty historical and pathologic perspective: present status. *Arthroscopy* 1986;2(1):54-69.

27.Insall J. Intraarticular surgery for degenerative arthritis of the knee: a report of the work of the late K. H. Pridie. *J Bone Joint Surg Br* 1967;49-B:211-228.

28.Friedman M, Berasi C, Fox J, Del Pizzo W, Snyder S, Ferkel R. Preliminary results with abrasion arthroplasty in the osteoarthritic knee. *Clin Orthop* 1984;182:200-205.

29.Bert J, Maschka K. The arthroscopic treatment of unicompartmental gonarthrosis: A five year follow-up study of abrasion arthroplasty plus arthroscopic debridement and arthroscopic debridement alone. *Arthroscopy* 1989;5(1):25-32.

30.Rand J. Role of arthroscopy in osteoarthritis of the knee. *Arthroscopy* 1991;7(4):358-363.

31.Ghosh P, Taylor TK. The knee joint meniscus. A fibrocartilage of some distinction. *Clin Orthop* 1987;(224):52-63.

32.Fairbank TJ. Knee joint changes after meniscectomy. *J Bone Joint Surg Br* 1948;30:644-670.

33.Newman AP, Daniels AU, Burks RT. Principles and decision making in meniscal surgery. *Arthroscopy* 1993;9(1):33-51.

34.DeHaven KE. Decision-making factors in the treatment of meniscus lesions. *Clin Orthop* 1990;(252):49-54.

35.Hede A, Larsen E, Sandberg H. The long term outcome of open total and partial meniscectomy related to the quantity and site of the meniscus removed. *Int Orthop* 1992;16(2):122-125.

36.Fauno P, Nielsen AB. Arthroscopic partial meniscectomy: a long-term follow-up. *Arthroscopy* 1992;8(3):345-349.

37.Covall D, Wasilewski S. Roentgenographic changes after arthroscopic meniscectomy; Five - year follow-up in patients more than 45 years old. *Arthroscopy* 1992;8(2):242-246.

38.Matsusue Y, Thomson N. Arthroscopic partial menisectomy in patients over 40 years old: A 5- to 11-year follow-up study. *Arthroscopy* 1996;12(1):39-44.

39.McBride G, Constine R, Hofmann A, Carson R. Arthroscopic partial medial menisectomy in the older patient. *Journal of Bone & Joint Surgery - American Volume* 1984;66-A(4):547-551.

40.Wouters E, Bassett III F, Hardaker W, Garrett W. An algorithm for arthroscopy in the over-50 age group. *Am J Sports Med* 1992;20(2):141-145.

41.Yang S, Nisonson B. Arthroscopic surgery of the knee in the geriatric patient. *Clin Orthop* 1995;316:50-58.

42.Jones R, Smith E, Reisch J. Effects of medial menisectomy in patients older than forty years. *Journal of Bone & Joint Surgery - American Volume* 1978;60-A(6):783-786.

43.Lotke P, Lefkoe R, Ecker M. Late results following medial menisectomy in an older population. *Journal of Bone & Joint Surgery - American Volume* 1981;63-A(1):115-119.

44.Jackson R, Rouse D. The results of partial arthroscopic menisectomy in patients over 40 years of age. *Journal of Bone & Joint Surgery - British Volume* 1982;64-B(4):481-485.

45.Wasson JH, Sox HC, Neff RK, and Goldman L. Clinical prediction rules. Applications and methodological standards. N Engl J Med 1985;313(3):793-9.

46.Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. [Review] [56 refs]. *JAMA* 1997;277(6):488-494.

47.Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Annals of Emergency Medicine* 1992;21(4):384-390.

48.Stiell IG, Greenberg GH, McKnight RD, et al. Decision rules for the use of radiography in acute ankle injuries. Refinement and prospective validation [see comments]. *JAMA* 1993;269(9):1127-1132.

49.Stiell IG, McKnight RD, Greenberg GH, et al. Implementation of the Ottawa ankle rules [see comments]. *JAMA* 1994;271(11):827-832.

50.Stiell I, Wells G, Laupacis A, et al. Multicentre trial to introduce the Ottawa ankle rules for use of radiography in acute ankle injuries. Multicentre Ankle Rule Study Group. *BMJ* 1995;311(7005):594-597.

51.Verbeek PR, Stiell IG, Hebert G, Sellens C. Ankle radiograph utilization after learning a decision rule: a 12-month follow-up. *Academic Emergency Medicine* 1997;4(8):776-779.

52.Sackett DL. Biases in analytical research. *J Chron Dis* 1979;32:51-63.

53. Feinstein, A.R. Clinimetrics. First. New Haven and London: Yale University Press. 1987; p.36

54.Lachenbruch PA, Mickey RM. Estimation of error rates in discriminant analysis. *Technometrics* 1968;10:1-11.

55.Efron B. Bootstrap methods: another look at the jackknife. *Ann Stat* 1979;7:1-26.

56.Daniel D, Daniels E, Aronson D. The diagnosis of meniscal pathology. *Clin Orthop* 1982;163:218-224.

57.Terry GC, Tagert BE, Young MJ. Reliability of the clinical assessment in predicting the cause of internal derangements of the knee. *Arthroscopy* 1995;11(5):568-576.

58.Abdon P, Lindstrand A, Thorngren K-G. Statistical evaluation of the diagnostic criteria for meniscal tears. *Int Orthop* 1990;14:341-345.

59.Fowler PJ, Lubliner JA. The predictive value of five clinical signs in the evaluation of meniscal pathology. *Arthroscopy* 1989;5(3):184-186.

60.Altman R, Asch E, Bloch De. Development of criteria for the classification and reporting of osteoarthritis. *Arth Rheum* 1986;29(8):1039-1049.

61.Rosenberg T, Paulos L, Parker R, Coward D, Scott S. The forty-five-degree posteroanterior flexion weight-bearing radiograph of the knee. *J Bone Joint Surg Am* 1988;70-A(10):1479-1483.

62.Moreland JR, Bassett LW, Hanker GJ. Radiographic analysis of the axial alignment of the lower extremity. *J Bone Joint Surg Am* 1987;69-A:745-749.

63.Sackett DL, Haynes RB, Guyatt GH, et al. Anonymous *Clinical Epidemiology. A basic science for clinical medicine.* 2nd ed. Boston: Little, Brown and Co. 1991;p. 25-31.

64.Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174.

65.Ware JE and Sherbourne CD. The MOS Short-Form health survey(SF-36) 1. Conceptual framework and item selection. Medical Care 1992;30473-83.

66.Brazier JEe. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160-164.

67.Bellamy, N. Osteoarthritis - an evaluative index for clinical trials McMaster University; 1982;

68.Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *Journal of Rheumatology* 1988;15(12):1833-1840.

69.Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt L. Validation study of WOMAC: a health status instrument for measuring clinically-important patient-relevant outcomes following total hip or knee arthroplasty in osteoarthritis. *J Orthop Rheumat* 1988;1:95-108.

70.Outerbridge R. The etiology of chondromalacia patellae. *Journal of Bone & Joint Surgery - British Volume* 1961;43-B(4):752-757.

71.Bentley G, Dowd G. Current concepts of etiology and treatment of chondromalacia patellae. *Clin Orthop* 1984;189:209-228.

72.Casscells S. Gross pathologic changes in the knee joint of the aged individual: a study of 300 cases. *Clin Orthop* 1978;132:225-232.

73.Ficat R, Philippe J, Hungerford D. Chondromalacia patellae: A system of classification. *Clin Orthop* 1979;144:55-62.

74.Insall J, Faluo K, Wise D. Chondromalacia patellae - A prospective study. *Journal of Bone & Joint Surgery - American Volume* 1976;58-A:1-8.

75.Noyes F, Stabler C. A system for grading articular cartilage lesions at arthroscopy. *Am J Sports Med* 1989;17(4):505-513.

76.Dougados M, Ayral X, Listrat V, et al. The SFA system for assessing articular cartilage lesions at arthroscopy of the knee. *Arthroscopy* 1994;10(1):69-77.

77.Ayral X, Dougados M, Listrat V, Bonvarlet JP, Simonnet J, Amor B. Arthroscopic evaluation of chondropathy in osteoarthritis of the knee [see comments]. *Journal of Rheumatology* 1996;23(4):698-706.

78.Hanley J.A., McNeil BJ. The meaning and use of the area under a receiver operating characteristic curve (ROC) curve. *Radiology* 1982;143:29-36.

79.Thompson WD, Walter SD. A reappraisal of the Kappa coefficient. *J Clin Epidemiol* 1988;41:949-958.

80.Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. *Ann Intern Med* 1993;118(3):201-210.

81.Bergquist SR, Hannan MT, Neuhauser KB, Wieland DA, Felson DT. Lack of validity and reliability of physical exam components o the American College of Rheumatology osteoarthritis diagnostic criteria (abstract). *Arthritis Rheum* 1993;36((suppl 9)):S58

82.Cushnagan J, Cooper C, Dieppe P, Kirwan J, McAlindon T, McCrae F. Clinical assessment of osteoarthritis of the knee. *Ann Rheum Dis* 1990;49:768-770.

83.Ike RW, O'Rourke KS. Compartment-directed physical examination of the knee can predict articular cartilage abnormalities disclosed by needle arthroscopy. *Arth Rheum* 1995;38(7):917-925.

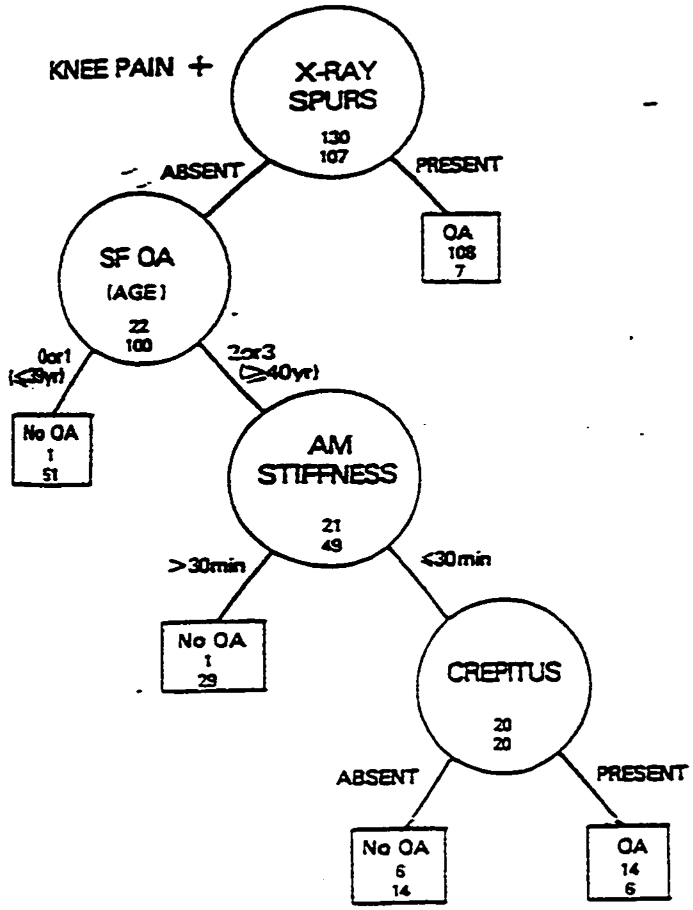
84.Hernborg J, Nilsson B. The relationship between osteophytes in the knee joint, osteoarthrosis and aging. *Acta Orthop Scand* 1973;44:69-74.

85.Stiell IG, Greenberg GH, Wells GA, McKnight RD, Cwinn AA, Cacciotti T, McDowell I, and Smith NA. Derivation of a decision rule for the use of radiography in acute knee injuries. Ann Emerg Med 1995;26(4):405-13.

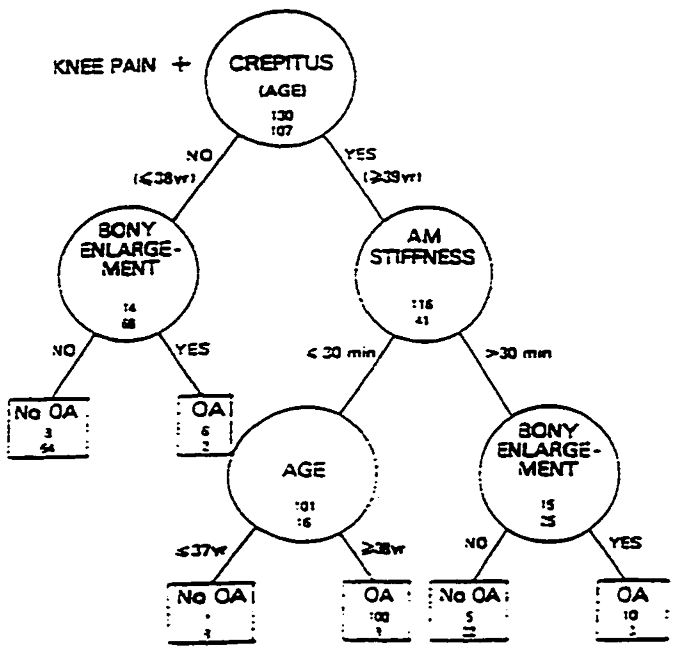
86.Jones AC, Hopkinson N, Pattrick M, Berman P, Doherty M. Evaluation of a method for clinically assessing osteoarthritis of the knee. *Ann Rheum Dis* 1992;51:243-245.

87.Hosmer DW, Lemeshow S. A goodness-of-fit test for the multiple logistic regression model. *Communication in Statistics* 1980;A10:1043-1069.

Appendix 1. American Rheumatological Association definition for the presence of osteoarthritis.



Osteoarthritis of the knee classification tree, for clinical, laboratory, and radiographic (x-ray) criteria. The overall sensitivity of this method is 94%; the overall specificity is 88%.



Osteoarthritis of the knee classification tree, for clinical criteria. The overall sensitivity of this method is 89%; the overall specificity is 88%.

Appendix 2. Patient consent form.

The Ottawa Hospital | L'Hôpital d'Ottawa

CONSENTEMENT DU PATIENT/ PROCURATION À DES RECHERCHES MÉDICALES

PATIENT/PROXY CONSENT FOR MEDICAL RESEARCH

Nom de l'étude / Name of research project

Étude d'arthroscopie du genou, Évaluation Prospective / A Study of Knee Arthroscopy, Prospective Study

Je confirme que j'ai expliqué la nature ainsi que les complications connues de ce projet de recherche/procuration au patient.

I confirm that I have explained the nature of and known complications of the research project to the patient/proxy.

Dr. G. Dervin, Dr. J.A. Brunet, Dr. J. Bouchard, Dr. J.P. Desjardins, Dr. R. Feibel, Dr. A. Giachino, Dr. G. Moreau

(chercheur) (researcher)

Je consens à la participation de : / I consent to the participation of:

Nom du patient / Patient's name	Nº de projet / Research no.

à l'étude précitée et j'autorise par la présente le(s) médecin(s) et le(s) chercheur(s) à procéder aux examens et/ou à dispenser les traitements suivants :

Questionnaires

in the above study and I hereby authorize the physician(s) and investigator(s) to proceed with the following questionnaire, examinations and/or to administer these treatments:

Study Questionnaires

J'ai lu la feuille d'explication détaillée approuvée par le Conseil d'éthique en recherches et je suis au courant des effets secondaires et des risques connus ayant trait aux examens et aux traitements.

J'ai également reçu une description de tous les avantages à attendre de ces examens et de ces traitements. On m'a fait connaître d'autres formes d'examen et de traitement.

J'ai eu l'occasion de poser des questions au sujet de ces examens et de ces traitements et on y a bien répondu.

On m'a dit que je pouvais retirer le consentement et suspendre la participation à l'étude à n'importe quel moment et pour quelque motif que ce soit et que cette action n'affectera pas la qualité des soins en cours et futurs.

Je reconnais qu'en agissant, comme procurateur, ce sera dans le meilleur intérêt du patient.

En toute connaissance de cause, je consens volontairement à ce que ______________________ participe à cette étude.

Ce protocole a été approuvé par le Conseil d'éthique en recherches de l'Hôpital d'Ottawa-site général. Ce conseil étudie les aspects éthiques de tous les projets de recherche sur les humains. Si je le désire, je peux contacter le président de ce comité.

I have read the detailed information sheet approved by the Research Ethics Board and am aware of the known side effects and risks related to the examinations and treatments.

I have also received a description of any benefits that may be expected from these examinations and treatments. As well, other forms of treatment and exams have been disclosed to me.

I have been given an opportunity to ask questions concerning the examinations and treatments involved and the questions which I have asked have been adequately answered.

I have been told that I can withdraw consent and stop participation in the study at any time and for any reason, and that such action will not affect the quality of ongoing and future care.

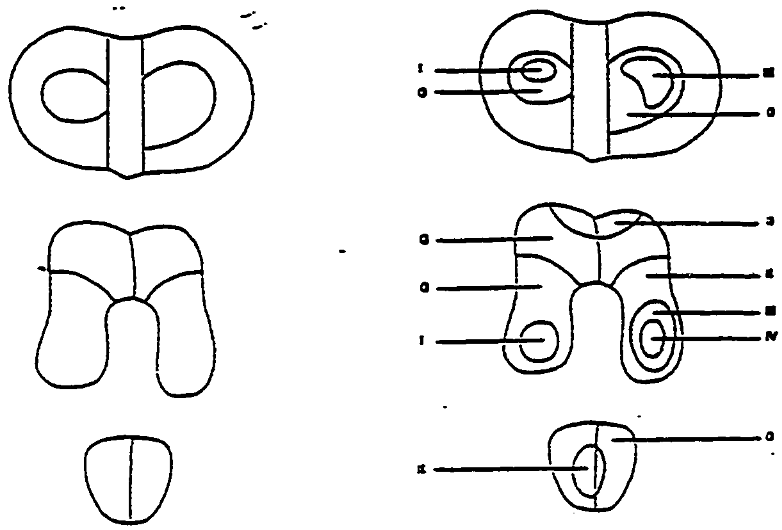
I acknowledge that if I am acting as a proxy that I am acting in the best interest of the patient.

With full knowledge of this, I voluntarily consent to the participation of ______________________ in the study.

This protocol has been approved by the Research Ethics Board of the Ottawa Hospital-General Site. This Board considers the ethical aspects of all hospital research projects using human subjects. If I wish, I may contact the chairperson of the Board with questions.

Patient ou nom de la personne légalement responsable (Procuration)-Patient or name of person legally responsible (Proxy)	Signature	Lien de parenté-Relationship
Nom du témoin-Name of witness	Signature	Date(ay-mm-jd)

Appendix 3. Example of SFA articular cartilage grading scheme.



Proposed diagram to report cartilage damage observed during chondroscopy (left side). An example is given on the right side where surface, grade, and localization of chondroplasty are reported.

With permission, *Journal of Arthroscopy*, Vol 10(1):71, 1994.

Appendix 4. Data collection sheet.

STUDY OF KNEE OSTEOARTHRITIS ARTHROSCOPY PROSPECTIVE EVALUATION (SKOAP)

Surname:__________ **Chart #:** __________ **Date:**__________ **Physician's Initials:**__________

DEMOGRAPHIC: Age:____yrs Sex: ○M ○F WCB Case: ○No ○Yes

NSAID/Analgesic Regular Use: ○No ○Yes BMI: ○<20 ○20-27 ○>27

HISTORY: Discrete event: ○No ○Yes

Pain:	Focal:	○None	○Mild	○Moderate	○Severe
	Generalized:	○None	○Mild	○Moderate	○Severe
	At Rest:	○None	○Mild	○Moderate	○Severe
	Arising From Chair:	○None	○Mild	○Moderate	○Severe
	Up/Down Stairs:	○None	○Mild	○Moderate	○Severe

Swelling:		○No	○Yes	
Giving Away:	Actual:	○No	○Yes	
	Apprehension:	○No	○Yes	
Locking:		○No	○Yes	if Yes, Duration in seconds:______

PHYSICAL	Effusion-	Visible:	○No	○Yes		
		Ballotment:	○No	○Yes		
		Sweep Test:	○No	○Yes		
Joint Line Tenderness		Medial:	○No	○Yes		
		Lateral:	○No	○Yes		
Joint Line Pain, Full Passive Flexion:			○None	○Mild	○Moderate	○Severe
if Yes:		Anteromedial:	○No	○Yes		
		Anterolateral:	○No	○Yes		
		Posteromedial:	○No	○Yes		
		Posterolateral:	○No	○Yes		
Full Passive Extension:			○None	○Mild	○Moderate	○Severe
if Yes:		Anteromedial:	○No	○Yes		
		Anterolateral:	○No	○Yes		
		Posteromedial:	○No	○Yes		
		Posterolateral:	○No	○Yes		

Ligamentous Stability:

Varus stress at 30° flexion:	○Stable	○Grade I	○Grade II	○Grade III
Valgus stress at 30° flexson:	○Stable	○Grade I	○Grade II	○Grade III
Lachman:	○Stable	○Grade I	○Grade II	○Grade III
Posterior drawer at 90° flexion:	○Stable	○Grade I	○Grade II	○Grade III
Passive Extension:	○Normal	○Abnormal (<0°)		

Special Tests:

MacMurray:	○Negative	○Positive
Steinman's:	○Negative	○Positive
Circumduction/Rotation (*full flexion*)	○Negative	○Positive

NOW PLEASE REVIEW 45° PA X-RAY TO PREDICT YOUR PROBABILITIES

PREDICT PROBABILITY

Of Degenerative Meniscal Tear: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Of 20% improvement in WOMAC pain score at 1 yr:

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%