Children's Hospital Orthopaedic Clinical Effectiveness Research Group

Mininder S. Kocher, MD, MPH

Department of Orthopaedic Surgery, Children's Hospital

INTRODUCTION

The evidence-based clinical practice and patient-derived outcomes assessment movements burst onto the scene of clinical medicine in the 1980's and 1990's as a result of contemporaneous technical, societal, and economic influences. Work by Wennberg and colleagues revealed large small-area variations in clinical practice, with some patients thirty times more likely to undergo an operative procedure than other patients with identical symptoms merely because of geographic location.[4-9] Further critical research suggested that up to 40% of some surgical procedures might be inappropriate and that up to 85% of common medical treatments were not rigorously validated.[1-3] Meanwhile, the costs of health care were rapidly rising to over two billion dollars per day, increasing from 5.2% of the gross domestic product in 1960 to 16.2% in 1997. As a result of the lack of evidence-based practice and the escalating costs of care, there has been an increased focus on the clinical effectiveness of care.

Clinical epidemiology provides the methodology to assess the clinical effectiveness of care. Epidemiology is the study of the causation and distribution of disease, using statistical and experimental methods. Epidemiologic methods focus on study design, data interpretation, and quantitative methods.

CHOCERG

We are currently developing the Children's Hospital Orthopaedic Clinical Effectiveness Research Group (CHOCERG). The research group's mission is to improve the effectiveness of management of musculoskeletal conditions in children and adolescents via applied epidemiologic research. The research group's goals are to provide epidemiologic and biostatistical support for pediatric orthopaedic clinical research, to apply advanced epidemiologic methods to pediatric orthopaedic clinical research, and to develop independent fields of inquiry into musculoskeletal conditions affecting public health.

Specifically, we are developing study sections in trauma, spinal disorders, hip disorders, upper extremity injuries, and sports medicine. We are developing a clinical research team

Mininder S. Kocher, MD, MPH is an Instructor in Orthopaedic Surgery at Harvard Medical School, Program in Clinical Effectiveness at Harvard School of Public Health, and Director of Children's Hospital Orthopaedic Clinical Effectiveness Research Group.

Address correspondence to:

Mininder S. Kocher, MD, MPH
Department of Orthopaedic Surgery
The Children's Hospital
300 Longwood Avenue
Boston, MA 02115

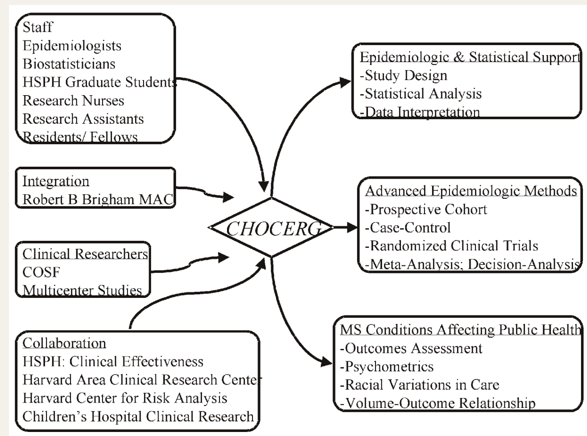

Figure 1: Children's Hospital Orthopaedic Clinical Effectiveness Research Group

Test \ Disease	+	-
+	A	B
-	C	D

A: True Positive
B: False Positive
C: False Negative
D: True Negative

Sensitivity: $P_{test+|Dz+} = \frac{TP}{Dz+} = \frac{A}{A+C}$

Specificity: $P_{test-|Dz-} = \frac{TN}{Dz-} = \frac{D}{B+D}$

Positive Predictive Value: $P_{Dz+|Test+} = \frac{p(D+)\,p(T+|D+)}{[p(D+)\,p(T+|D+)]+[(1-p(D+))\,p(T+|D-)]}$

Negative Predictive Value: $P_{Dz-|Test-} = \frac{(1-p(D+))\,(p(T-|D-))}{(1-p(D+))\,(p(T-|D-))+(p(D+)(p(T-|D+)}$

Figure 2: Diagnostic Performance

(Figure 1) and resources to apply advanced epidemiologic methodology including performing randomized clinical trials, prospective cohort studies, and quantitative synthesis methods such as meta-analysis, decision-analysis, and cost-effectiveness analysis. In addition, we intend to pursue fields of inquiry into musculoskeletal conditions affecting public health, including racial variations in the utilization of orthopaedic care, the volume-outcome relationship in orthopaedic care, and the development and validation of patient-derived outcome assessment instruments in pediatric orthopaedics.

EXAMPLES

Diagnostic Performance

The diagnostic performance of an ancillary test involves establishing the sensitivity, specificity, and predictive value of that test (Figure 2). In a clinical study to evaluate the

Figure 3: Diagnostic Performance of MRI of the Knee in Children and Adolescents
Reproduced with permission from the *American Journal of Sports Medicine*

Diagnosis	Sensitivity (%)			Specificity (%)			Positive Predictive Value (%)		Negative Predictive Value (%)	
	Clinical	MRI	P	Clinical	MRI	P	Clinical	MRI	Clinical	MRI
Anterior Cruciate Ligament Tear	81.3	75.0	0.55	90.6	94.1	0.39	49.0	58.6	97.8	97.1
Medial Meniscus Tear	62.1	79.3	0.15	80.7	92.0	0.03*	14.5	34.3	97.6	98.8
Lateral Meniscus Tear	50.0	66.7	0.24	89.2	82.8	0.21	34.0	30.1	94.1	95.7
Osteochondritis Dissecans	77.3	90.9	0.22	97.9	97.9	0.99	66.0	69.5	98.8	99.5
Lateral Discoid Meniscus	88.9	38.9	0.02*	98.0	100.0	0.15	31.0	100.0	99.9	99.4
OVERALL	71.2	72.0	0.89	91.5	93.5	0.26				

*p<0.05

diagnostic performance of MRI of the knee in children and adolescents, we studied a consecutive series (118 knees in 113 patients) of pediatric patients (≤16 years old) treated for intra-articular knee disorders who had an initial clinical diagnosis, followed by a MRI diagnosis, followed by arthroscopic findings.[10] There were 139 clinical lesions, 128 MRI lesions, and 135 arthroscopic lesions. There were no significant differences between clinical exam and MRI with respect to agreement with arthroscopic findings (clinical exam: 70%; MRI: 74%), overall sensitivity (clinical exam: 71%; MRI: 72%) and overall specificity (clinical exam: 92%; MRI: 94%). Stratified analysis by diagnosis revealed significant differences only for sensitivity of lateral discoid meniscus and specificity of medial meniscal tears (Figure 3). MRI had significantly lower sensitivity for younger children (<12 years old: 62%; 12-16 years old: 78%; p = 0.05). MRI also demonstrated lower specificity for younger children (<12 years old: 90%; 12-16 years old: 96%; p = 0.02). We concluded that selective MRI does not provide enhanced diagnostic utility over clinical exam, particularly in children, and should be used judiciously in cases where the clinical diagnosis is uncertain and MRI input will alter the treatment plan.

Figure 4: Clinical Prediction Algorithm for the Differentiation of Septic Arthritis and Transient Synovitis of the Hip in Children
Reproduced with permission from the *Journal of Bone and Joint Surgery*, Kocher MS, Zurakowski D, Kasser JR. Differentiating between septic arthritis and transient synovitis of the hip in children: An evidence-based clinical prediction algorithm. *Journal of Bone and Joint Surgery*, 1999, 81A:1662-1670

History of Fever	Non-Weight Bearing	Erythrocyte Sedimentation Rate ≥ 40 (mm/hr)	Serum White Blood Cell count > 12,000 (x1000/mm3)	Predicted Probability of Septic Arthritis
yes	yes	yes	yes	99.8%
yes	yes	yes	no	97.3%
yes	yes	no	yes	95.2%
yes	yes	no	no	57.8%
yes	no	yes	yes	95.5%
yes	no	yes	no	62.2%
yes	no	no	yes	44.8%
yes	no	no	no	5.3%
no	yes	yes	yes	93.0%
no	yes	yes	no	48.0%
no	yes	no	yes	33.8%
no	yes	no	no	3.4%
no	no	yes	yes	35.3%
no	no	yes	no	3.7%
no	no	no	yes	2.1%
no	no	no	no	1 in 700

Multivariable Methods: Regression

Regression analysis is a method of mathematical modeling used to determine independent associations of multiple independent variables with a binary categorical dependent variable (logistic regression), a continuous dependent variable (linear regression), an ordinal dependent variable (ordinal logistic regression), censored time-to-event data (Cox regression), or count data (Poisson regression). Regression is commonly used to predict outcomes, or to establish independent associations, controlling for confounding and colinearity.

We have used logistic regression to develop a clinical prediction rule to differentiate between septic arthritis and transient synovitis of the hip in children.[11] We reviewed children who presented to Children's Hospital from 1979 to 1996 with an acutely irritable hip. Diagnoses of true septic arthritis, presumed septic arthritis, and transient synovitis were explicitly defined based on joint fluid white blood cell count, joint fluid and blood cultures, and clinical course. We found that septic arthritis patients differed significantly (p<0.05) from transient synovitis patients with regard to ESR, serum WBC count and differential, weight-bearing status, history of fever, temperature, effusion on radiograph, history of chills, history of recent antibiotic use, hematocrit, and gender. Four independent multivariate clinical predictors were identified to differentiate between septic arthritis and transient synovitis: history of fever, non weight-bearing, ESR >40mm/hr, and serum WBC count >12,000/mm^3. The predicted probability of septic arthritis was determined for all sixteen combinations of these four predictors (Figure 4) and can be summarized as: <0.2% for 0 predictors, 3% for 1 predictor, 40% for 2 predictors, 93% for 3 predictors, and 99.8% for 4 predictors. The chi-square test for trend and the area under the receiver operating characteristic (ROC) curve indicated excellent diagnostic performance of this group of multivariate predictors in identifying septic arthritis.

We have also used logistic regression to determine independent associations, controlling for confounding. In a

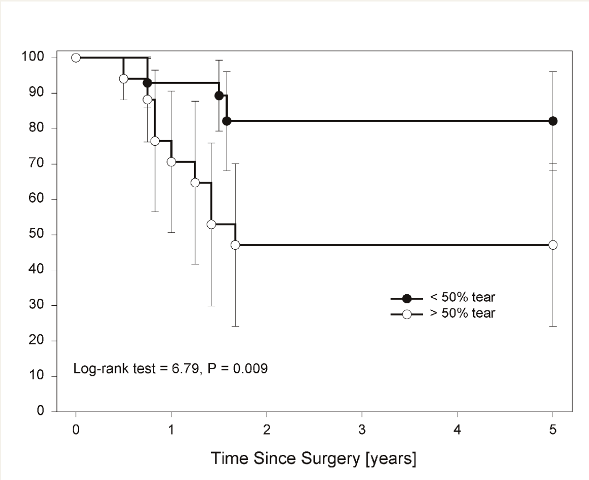

Figure 5: Partial Tears of the ACL in Children and Adolescents

prospective cohort study of 9,410 professional skiers who underwent preseason knee screening from 1992-1997, we studied the effect of functional knee bracing on subsequent knee injury in ACL deficient (ACLd) skiers.[12] ACL deficiency was operationally defined as an abnormal Lachman or pivot-shift examination with ≥ 5mm KT-1000 manual maximum side-to-side difference in a skier without prior ACL surgery in either knee. ACLd skiers self-selected the use of functional knee brace during skiing. Controlling for knee laxity, gender, age, ski occupation, Lachman and pivot-shift grades, we found that absence of bracing remained an independent, multivariate risk factor for knee injury. ACLd skiers without bracing were estimated to be over six times more likely to sustain a subsequent knee injury (OR = 6.6; 95% CI: 1.5-31.7). Because of the increased risk of subsequent knee injury in nonbraced skiers, we recommend functional bracing for ACL deficient skiers.

Survivorship Analysis

Survivorship analysis is used to analyze time-censored data, which poses unique challenges to univariate methods. We have used survivorship analysis to study the functional outcome of partial ACL injuries in children and adolescents treated without reconstruction.[13] A cohort of 45 skeletally mature and immature patients ≤17 years old who had an acute hemarthrosis, MRI signal changes within the ACL, an IKDC grade A or B Lachman and pivot-shift examination under anesthesia, an arthroscopically documented partial ACL injury, and minimum 2-year follow-up was studied prospectively. All patients were treated without ACL reconstruction and underwent a structured rehabilitation program. Fourteen patients (31%) underwent subsequent reconstruction. Using survivorship analysis, we found that >50% tear (p=0.01) (Figure 5), predominantly posterolateral tear (p=0.01), and older chronological age (p=0.01) were independent, time-dependent risk factors for subsequent reconstruction.

Decision Analysis

Expected-value decision analysis is a methodological tool, based in gaming theory, that allows for the quantitative

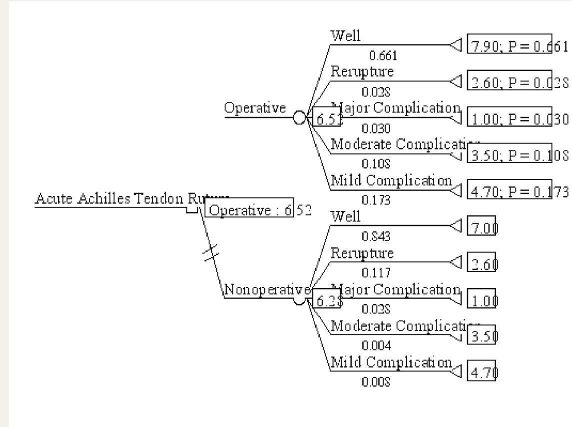

Figure 6: Decision Tree. Operative versus Nonoperative Management of Acute Achilles Tendon Ruptures: Expected Value Decision Analysis.

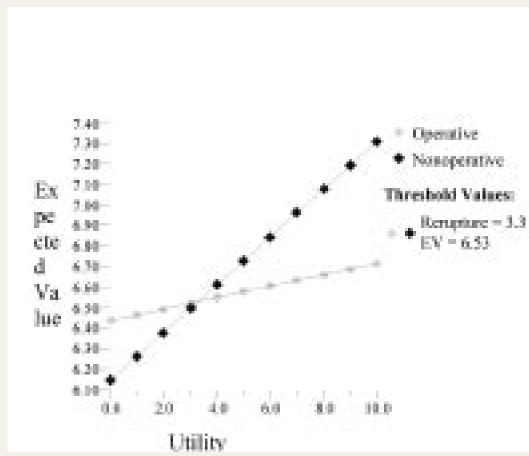

Figure 7: One-Way Sensitivity Analysis: Operative versus Nonoperative Management of Acute Achilles Tendon Ruptures: Expected Value Decision Analysis.

analysis of decision-making under conditions of uncertainty. The process of expected-value decision analysis involves the creation of a decision tree to structure the decision problem, roll-back analysis to calculate expected-value and determine the optimal decision-making strategy, and sensitivity analysis to determine the effect on decision-making of varying outcome probabilities and utilities. We used expected-value decision analysis to assess to determine the optimal management strategy, operative versus nonoperative, for acute Achilles tendon rupture.[14] Outcome probabilities were determined from a previously published quantitative synthesis of 83 published studies of acute Achilles tendon rupture from 1954-97. Patient-derived utility values were obtained from a visual analog scale questionnaire of 76 active males, age 30-50 years old without a history of Achilles tendon injury. Roll-back analysis revealed operative treatment to be the optimal management strategy. (EV: 6.52 vs 6.28; marginal value: 0.24) (Figure 6). Threshold values were determined for the probability of rerupture from nonoperative treatment (0.06), the probability of a moderate complication from operative treatment (0.16), the utility of rerupture (3.3) (Figure 7), and the utility of a moderate complication (1.3). We concluded that operative management was the optimal management strategy for acute Achilles tendon rupture given the outcome probabilities and patient utilities that we studied. Lowering rates of rerupture from nonoperative management or increasing rates of moderate complications from operative management would favor nonoperative management. Patients with a low utility for rerupture favor operative management, whereas patients with a low utility for complications favor nonoperative management. We advocate a model of doctor-patient shared decision-making in which both outcome probabilities and patient preferences are considered in order to optimize the decision process.

Cost-Effectiveness Analysis

Economic analysis of medical procedures includes cost-identification studies which determine costs and cost-evaluation studies which assess both costs and benefits. Cost evaluative study designs include cost-benefit analysis, cost-effectiveness analysis, and cost-utility analysis. In cost-benefit analysis, both costs and benefits are analyzed in terms of monetary units. In cost-effectiveness analysis, a ratio is determined of cost per health effect. In cost-utility analysis, this health effect is measured as quality-adjusted life years (QALY's). We studied the cost and effectiveness of routine pathological examination of surgical specimens from primary total hip and knee replacement for osteoarthritis by comparing clinical and pathological diagnoses in 1,234 consecutive cases of primary THR (n=471) and TKR (n=763) performed between 1992 and 1995 at the Brigham and Women's Hospital for the clinical diagnosis of osteoarthritis.[15] We found a low prevalence of discrepant (2.3%) and discordant (0.1%) diagnoses. The cost per discrepant diagnosis was $4,383 and the cost per discordant diagnosis was $122,728. We concluded that routine pathological examination of surgical specimens from primary THR and TKR for the clinical diagnosis of osteoarthritis had limited cost-effectiveness at our hospital due to the low prevalence of findings that altered patient management.

Outcomes Assessment

Health care outcomes assessment measures include generic measures, condition-specific measures, and measures of patient satisfaction. The legitimacy and importance of patient satisfaction as an outcome measure has burgeoned over the past decade because of its use for a variety of purposes such as indicating quality of care, assessing health care delivery, developing patient care models, impacting general health status, and allowing for continuous quality improvement. Furthermore, as the paradigm for health care has shifted toward a market model, patient satisfaction has become an outcome measure with great clinical and economic implications. We studied the determinants of patient satisfaction with outcome after ACL reconstruction in a cohort of 202 patients with minimum two year follow-up and complete demographic, surgical, subjective, and objective data.[16] We found that some specific objective variables were associated with satisfaction, however subjective variables of symptoms and function had the most robust associations with patient satisfaction. The seven independent multivariate determinants (adjusted $R^2=0.83$, $p<0.001$) of patient satisfaction included Lysholm score, overall subjective knee function, IKDC range-of-motion subscale, patella tenderness, full giving-way, flexion contracture, and swelling.

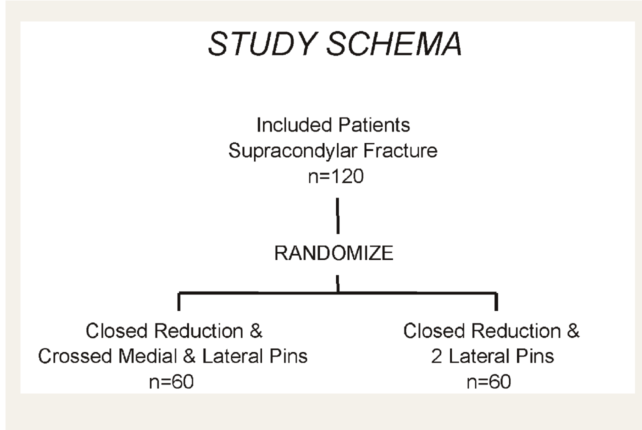

Figure 8: Randomized Clinical Trial study schema

Study Design

Associations between exposures and outcomes can be spurious due to chance, confounding, and bias. In hypothesis testing, the probability of a finding an association in a study by chance when, in reality, there is no association (type I error) is estimated by alpha, which is the p-value. The probability of finding no association in a study when, in reality, there is an association (type II error) is estimated by beta, which is 1 - power. Bias is nonrandom systematic error in design or conduct of a study. There are many forms of bias including selection bias, nonresponder bias, state of health bias, detection bias, performance bias, recall bias, acceptability bias, and publishing bias. A confounding variable has associations with both the exposure and the outcome and thus may distort their relationship. Frequent confounders include age, gender, socioeconomic status, and comorbidities.

Studies may be performed prospectively or retrospectively. Study designs may be observational, such as case reports, retrospective case series, case-control studies, prospective cohort studies, and cross-sectional studies, or they may be experimental, such as randomized clinical trials. Bias and confounding can be minimized through study design, with methods such as randomization, blinding, stratification, and striving for similar follow-up and assessment of groups.

The randomized clinical trial (RCT) is the highest level of clinical evidence because it minimizes bias and confounding. All variables between the groups are theoretically distributed equivalently, except the allocated study variable, because of randomization. Disadvantages of RCT's include difficulty, expense, length of time for study, ethical considerations, lack of acceptance by patients and clinicians, and lack of external generalizability. The design of a randomized clinical trial involves establishing explicit inclusion criteria, pre-hoc sample size calculation, patient accrual, blinded outcome assessment, interim analysis with stopping rules, and intent to treat analysis.

References

1. **Chassin MR et al.** Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures [see comments]. *JAMA*, 258:2533-7, 1987.
2. **Kahn KL, Kosecoff, J, Chassin MR, Flynn MF, Fink A, Pattaphongse N, Solomon DH, Brook RH.** Measuring the clinical appropriateness of the use of a procedure. Can we do it? *Med Care*, 26:415-22, 1988.
3. **Park RE, Fink A, Brook RH, Chassin MR, Kahn KL, Merrick NJ, Kosecoff J, Solomon DH.** Physician ratings of appropriate indications for three procedures: theoretical indications vs indications used in practice. *Am J Public Health*, 79:445-7, 1989.
4. **Wennberg JE, Gittelsohn A.** Small area variations in health care delivery. *Science*, 182:1102-8, 1973.
5. **Wennberg JE, Gittelsohn A.** Variations in medical care among small areas. *Sci Am*, 246:120-34, 1982.
6. **Wennberg JE.** Dealing with medical practice variations: a proposal for action. *Health Aff (Millwood)*, 3:6-32, 1984.
7. **Wennberg JE.** Outcomes research: the art of making the right decision. *Internist*, 3:26-8, 1990.
8. **Wennberg JE.** Practice variations: why all the fuss? *Internist*, 26:6-8, 1985.
9. **Wennberg JE, Bunker JP, Barnes B.** The need for assessing the outcome of common medical practices. *Annu Rev Public Health*, 1:277-95, 1980.
10. **Kocher MS, DiCanzio J, Zurakowski D, Micheli LJ.** Diagnostic performance of clinical examination and selective MRI in the evaluation of intra-articular knee disorders in children and adolescents. *Am J Sports Med*, in press. (Winner of the 2000 Clinical Research Essay Award of the Arthroscopy Association of North America.)
11. **Kocher MS, Zurakowski D, Kasser JR.** Differentiating between septic arthritis and transient synovitis of the hip in children: An evidence-based clinical prediction algorithm. *J Bone Joint Surg [Am]*, 81:1662-70, 1999. (Winner of the 12th Annual American Orthopaedic Association-Zimmer Resident Research Travel Award, the Richard Kilfoyle Award of the New England Orthopaedic Society, and the Vernon Thompson Research Award of the Western Orthopaedic Association).
12. **Kocher MS, Sterett WI, Briggs KK, Zurakowski D, Steadman JR.** Effect of functional bracing on knee injury in ACL-deficient professional skiers. Presented at the 68th Annual Meeting of the American Academy of Orthopaedic Surgeons. March 2001, San Francisco, CA.
13. **Kocher MS, Luke A, Zurakowski D, Micheli LJ.** Partial tears of the anterior cruciate ligament in children and adolescents. Presented at the 67th Annual Meeting of the American Academy of Orthopaedic Surgeons. March 2000, Orlando, FL.
14. **Kocher MS, Luke A, Marshall R, Hawkins RJ.** Operative vs. nonoperative management of acute achilles tendon ruptures: expected-value decision analysis. Presented at the 68th Annual Meeting of the American Academy of Orthopaedic Surgeons. March 2001, San Francisco, CA.
15. **Kocher MS, Erens G, Thornhill TS, Ready JE.** Cost-effectiveness of routine pathologic examination of surgical specimens from primary total joint replacement for osteoarthritis. *J Bone Joint Surg [Am]*, 82:1531-5, 2000.
16. **Kocher MS, Steadman JR, Zurakowski D, Briggs K, Sterett WI, Hawkins RJ.** Determinants of patient satisfaction after anterior cruciate ligament reconstruction. Presented at the American Orthopaedic Society for Sports Medicine Specialty Day Meeting. March 2001, San Francisco, CA. (Winner of the 2000 Clinical Research Essay Award of the Arthroscopy Association of North America.)