

The American Journal of Sports Medicine

<http://ajs.sagepub.com/>

Cost-effectiveness Analysis of the Diagnosis of Meniscus Tears

Richard C. Mather III, William E. Garrett, Brian J. Cole, Kristen Hussey, Michael P. Bolognesi, Tally Lassiter and Lori A. Orlando

Am J Sports Med 2015 43: 128 originally published online December 1, 2014

DOI: 10.1177/0363546514557937

The online version of this article can be found at:

<http://ajs.sagepub.com/content/43/1/128>

Published by:



<http://www.sagepublications.com>

On behalf of:

American Orthopaedic Society for Sports Medicine



Additional services and information for *The American Journal of Sports Medicine* can be found at:

Email Alerts: <http://ajs.sagepub.com/cgi/alerts>

Subscriptions: <http://ajs.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Dec 31, 2014

[OnlineFirst Version of Record](#) - Dec 1, 2014

[What is This?](#)

Cost-effectiveness Analysis of the Diagnosis of Meniscus Tears

Richard C. Mather III,^{*†} MD, William E. Garrett,[†] MD, Brian J. Cole,^{||} MD, MBA, Kristen Hussey,^{||} MS, Michael P. Bolognesi,[†] MD, Tally Lassiter,[‡] MD, MHA, and Lori A. Orlando,[§] MD, MHS

Investigation performed at the Department of Orthopaedic Surgery, Duke University Medical Center, Durham, North Carolina, USA

Background: Diagnostic imaging represents the fastest growing segment of costs in the US health system. This study investigated the cost-effectiveness of alternative diagnostic approaches to meniscus tears of the knee, a highly prevalent disease that traditionally relies on MRI as part of the diagnostic strategy.

Purpose: To identify the most efficient strategy for the diagnosis of meniscus tears.

Study Design: Economic and decision analysis; Level of evidence, 1.

Methods: A simple-decision model run as a cost-utility analysis was constructed to assess the value added by MRI in various combinations with patient history and physical examination (H&P). The model examined traumatic and degenerative tears in 2 distinct settings: primary care and orthopaedic sports medicine clinic. Strategies were compared using the incremental cost-effectiveness ratio (ICER).

Results: In both practice settings, H&P alone was widely preferred for degenerative meniscus tears. Performing MRI to confirm a positive H&P was preferred for traumatic tears in both practice settings, with a willingness to pay of less than US\$50,000 per quality-adjusted life-year. Performing an MRI for all patients was not preferred in any reasonable clinical scenario. The prevalence of a meniscus tear in a clinician's patient population was influential. For traumatic tears, MRI to confirm a positive H&P was preferred when prevalence was less than 46.7%, with H&P preferred above that. For degenerative tears, H&P was preferred until the prevalence reaches 74.2%, and then MRI to confirm a negative was the preferred strategy. In both settings, MRI to confirm positive physical examination led to more than a 10-fold lower rate of unnecessary surgeries than did any other strategy, while MRI to confirm negative physical examination led to a 2.08 and 2.26 higher rate than H&P alone in primary care and orthopaedic clinics, respectively.

Conclusion: For all practitioners, H&P is the preferred strategy for the suspected degenerative meniscus tear. An MRI to confirm a positive H&P is preferred for traumatic tears for all practitioners. Consideration should be given to implementing alternative diagnostic strategies as well as enhancing provider education in physical examination skills to improve the reliability of H&P as a diagnostic test.

Clinical Relevance: Alternative diagnostic strategies that do not include the use of MRI may result in decreased health care costs without harm to the patient and could possibly reduce unnecessary procedures.

Keywords: cost; cost-effectiveness analysis; decision analysis; economic analysis; health policy; meniscus tears; MRI; physical examination

*Address correspondence to Richard C. Mather III, MD, Department of Orthopaedic Surgery, Duke University Medical Center, Box 2887, Durham, NC 27710, USA (e-mail: richard.mather@dm.duke.edu).

[†]Department of Orthopaedic Surgery, Duke University Medical Center, Durham, North Carolina, USA.

[‡]Department of Orthopaedic Surgery, Bassett Health Care Network, Cooperstown, New York, USA.

[§]Department of Medicine and the Center for Applied Genomics & Precision Medicine, Duke University Medical Center, Durham, North Carolina, USA.

^{||}Department of Orthopaedic Surgery, Rush University Medical Center, Chicago, Illinois, USA.

The authors declared that they have no conflicts of interest in the authorship and publication of this contribution.

As diagnostic imaging is the fastest rising segment of costs in the US health care system (the Centers for Medicare and Medicaid Services estimated a 40% increase over the past decade),³¹ designing and implementing cost-effective diagnostic strategies is paramount to reducing waste and controlling costs. The accuracy of modern imaging is often marginally higher than that of patient history and physical examination.^{25,27,32,33} Formal analysis of the value added by specific aspects of the diagnostic algorithm is an important step in delivering high-value care.

In light of this information, we have proposed a framework for re-evaluating existing diagnostic standards meeting specific criteria by using cost-effectiveness analysis. Within this framework the added benefit attributable to diagnostic imaging can be judged in comparison with

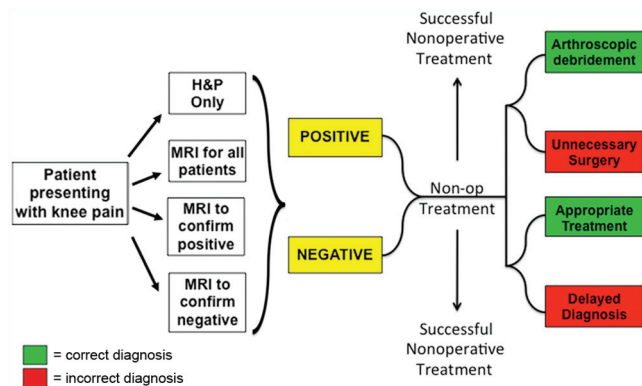


Figure 1. Health state diagram for the simple-chain model.

alternative strategies on both the basis of its financial cost and its benefits. In addition, if limited data are available, it can provide information on what operating characteristics are needed to optimize the strategy. To be appropriate for further study, diagnostic strategies must evaluate a disease that is prevalent, is not fatal in the setting of delayed diagnosis, includes modern imaging as part of the gold standard, and is well enough studied that data on alternative strategies exist.

In the current study, we used this framework to re-evaluate the diagnostic strategy for meniscus tears in adult patients. Meniscus tears meet the criteria outlined previously in that knee pain is one of the most commonly cited reasons for adults to visit their physicians and within this group meniscus injuries are highly prevalent (9%-33%).^{1,2,8,22,28,30} The gold standard diagnostic strategy is a history and physical examination with plain radiographs (H&P) followed by magnetic resonance imaging (MRI), and there are excellent data on the clinical validity of H&P and MRI.^{17,19,35,37,46} While this strategy has excellent performance, we thought it was important to investigate whether alternate strategies exist that can decrease costs while maintaining efficacy.

METHODS

General Model Overview

The model and analysis were constructed according to the guidelines set forth by the Panel on Cost-effectiveness in Health and Medicine by the US Public Health Service in 1993 and utilized the structure and approach of previous decision models by our group and others^{42,44,48} in orthopaedic surgery. A simple-chain decision model representing the diagnostic and therapeutic algorithm for a patient with knee pain (Figure 1) was constructed using a software program (TreeAge Pro 2009). We compared the cost-effectiveness of 4 potential diagnostic strategies: (1) H&P only (H&P), (2) H&P followed by MRI for all patients (MRI), (3) MRI to confirm a positive H&P (MRI+), and (4) MRI to confirm a negative H&P (MRI-). The primary effectiveness outcome was expressed in quality-adjusted life-years

TABLE 1 Assumptions in the Model Structure^a

Diagnosis	MRI not repeated if poor study
	Patients have isolated meniscus tear if positive
	Meniscus tear is not repaired—only meniscectomy is modeled
	No other diagnostic modalities utilized (CT/MRI hip or spine)
Treatment	Nonoperative treatment consists of physical therapy for 6 weeks and corticosteroid injection
	All patients presumed to have a degenerative meniscus tear undergo nonoperative treatment trial while traumatic tears do not
	False-negatives undergo diagnostic arthroscopy in 1 year
	True-positives undergo surgical treatment with success if they fail nonoperative treatment
	Improvement from treatment lasts a specified time
	Operative treatment consists of diagnostic and treatment arthroscopy
	There is a wait time for arthroscopy

^aCT, computed tomography; MRI, magnetic resonance imaging.

(QALYs), and costs were in 2012 \$US. Certain assumptions were made in the creation of the model and can be found in Table 1.

Decision Model

In our decision model, a patient with knee pain is assigned a pretest probability of a meniscus tear based on its prevalence in the presenting population. Afterward, all patients undergo evaluation with a history, physical examination, and radiographs (4 views: weightbearing anteroposterior in extension and posteroanterior 45° of flexion, sunrise/Merchant, and lateral). Additional tests may be performed depending on the strategy being evaluated (ie, in the MRI to confirm positive H&P strategy, those with a positive H&P will go on to have an MRI). Finally, all patients with suspected degenerative tears are treated nonoperatively with a corticosteroid injection and physical therapy for 6 weeks. They either succeed or fail this therapeutic trial based on their actual pathologic abnormalities, not the presumed diagnosis. Those with suspected traumatic tears proceed directly to surgery without a nonoperative treatment trial. Successful nonoperative treatment is regarded as improved utility equal to that of surgery and the patients' path in the model ends here. All patients who fail undergo arthroscopy.

Each diagnostic strategy had 4 potential outcomes: true-positive, false-negative, true-negative, and false-positive. Patients with true-positive results transition to the correct treatment and surgery 1 month after the 6-week nonoperative treatment trial. Patients with true-negative results, which consist of a number of diagnoses other than meniscus tears, are treated with surgery.

The false-negative patients are those not correctly identified at the 6-week mark but rather suffer from delayed diagnosis for 6 months, at which time they undergo arthroscopy. Patients with false-negative results not

improving with nonoperative treatment undergo an unnecessary surgery at the 6-week mark but are correctly diagnosed during the arthroscopy and patients are assumed to recover from their injury in 6 months. The relationships of these 4 outcomes are appropriate, with true-positives and true-negatives achieving the highest utility, followed by false-negatives and then false-positives. In other words, our model reflected a slightly greater penalty for an unnecessary surgery than for a delayed diagnosis.

As not all meniscus tears are identical, we ran the model separately for 2 cohorts based on the presence of a traumatic or degenerative meniscus tear.¹⁶ The time horizon for treatment benefit and level of improvement with treatment differs between groups. Traumatic tears, whether undergoing repair or partial meniscectomy, sustain their improvement over a 10-year period, while degenerative tears benefit only for 5 years.^{6,15,16,26,39} All probabilities, costs, and utilities were extracted from available literature after an exhaustive search or were estimated by expert opinion where necessary.

Model Parameters

Population. As mentioned previously, we ran the model under 2 scenarios, an expected traumatic tear or degenerative tear. This categorization is supported by several studies, including the prospective cohort by Englund et al,¹⁶ which found that of all factors studied, degenerative tears were by far the strongest predictors of response to meniscectomy. Fortunately, distinguishing clinically between degenerative and nondegenerative tears can be fairly straightforward using, for example, duration and onset of symptoms as well as plain radiographs. Traumatic tears required both of 2 findings: (1) acute onset of symptoms secondary to known injury and (2) Kellgren-Lawrence (K&L) grade 0. Degenerative tears were assumed to have unknown onset of symptoms with K&L grade 0 radiographs or any onset of symptoms (acute or insidious) with K&L grade 1 or 2. As arthroscopy is rarely indicated in knees with K&L grades of 3 or 4, these patients were excluded from the model population. Since age is not an accurate predictor of tear type (younger individuals can have degenerative tears and older individuals can have traumatic tears), both cohorts were run assuming a normal adult distribution.

To better understand the operating characteristics of the diagnostic strategies in 2 very distinct settings, primary care clinics and general orthopaedic clinics, we assigned a pretest probability for meniscus tear based on its prevalence in these 2 settings. Among patients with knee pain seen in primary care clinics, approximately 9% have meniscus tears,²⁰ while estimates for the orthopaedic and sports medicine physician are wider; they are generally higher.^{1,2,8,28} In our model we chose 30%, the median prevalence identified during a literature review,^{1,2,8,28} to reflect the population seen by the orthopaedic/sports medicine practitioner. Extensive sensitivity analyses were performed to identify the effect of disease prevalence on outcomes.

Diagnostic Test Accuracy. The sensitivity and specificity of MRI varies in the literature.^{17,19,35,37} Much evidence

suggests lateral meniscus tears are more difficult to accurately diagnose with imaging than are medial tears. Furthermore, the specificity appears to be higher in traumatic tears and lower in degenerative tears.^{4,14} Given this variability, the sensitivity of 93.3% and specificity of 88.3% in this study were taken from a large systematic review³⁷ that reported pooled estimates. Furthermore, we chose these numbers, higher than that in many other studies, to avoid any bias against MRI in the study.

The values used for sensitivity and specificity of H&P were based on a comprehensive knee examination reported in the *Journal of the American Medical Association* Rational Clinical Examination series rather than on individual specific maneuvers. We chose to use the comprehensive examination given its superior operating characteristics at a sensitivity of 77% (SD, $\pm 7\%$) and specificity of 93% (SD, $\pm 3\%$).⁴⁶ These values were derived from a meta-analysis of all available tests

Transition Probabilities. Two transition probabilities were used in the model, the success of nonoperative treatment for patients with and without meniscus tears. We thought it was important to allow for a period of nonoperative therapy, given that up to one-third of patients with meniscus tears,⁴¹ and an even higher percentage without meniscus tears, recover with nonoperative treatment. Several studies examining nonmeniscal causes for knee pain such as patellofemoral syndrome found 75% to 84% of patients improved with nonoperative treatment.^{10,11,13,23,40,49} In this module we used 75% as our point estimate since this is the value reported by the only available study with level 1 evidence.²³ The point estimates, their plausible ranges, and sources are found in Table 2.

Utilities. Utility values for each health state were assigned based on health-related quality of life and were graded on a scale of 0 to 1, with 1 being perfect health and 0 being death. The utilities used in this study are shown in Table 1. Utilities were calculated using the Short Form (SF)-6 dimension (SF-6D), a preference-based conversion of the SF-12, from a prospective cohort of 38 patients (mean age, 50 years [range, 19-69 years]; 13/38 female [34%]) undergoing arthroscopic meniscectomy at our institution (Rush University Medical Center). Preoperative values were used for knee pain, and 1-year follow-up or last value carried forward were used for utility after treatment of knee pain. The preoperative utility was 0.663 ± 0.115 , and postoperative utility was 0.78 ± 0.101 . These values are consistent with several studies, including other cost-effectiveness analyses, in the orthopaedic literature, where the utility of knee pain is measured at 0.6.^{5,18,21,38} Improvement in utility with treatment shows more variability, ranging from 0.1 to 0.3 across several nonoperative and treatment studies.^{3,7,12,21,47}

Surgical failures were not modeled separately but rather were factored into the average improvement for all patients undergoing surgery. In addition, since complication rates are exceedingly low and would not affect outcomes, they were not incorporated into the model.^{16,24,36,45}

"Disutility" values represent the short-term negative effect an intervention has on a patient's quality of life. With surgical procedures, this can include pain, immobility,

TABLE 2
Base Case Model Parameters^a

	Base Case	Sensitivity Analysis	Source
Health state utility, QALY			
Symptomatic meniscus tear	0.663	0.38-0.95	Study cohort
Successful treatment of meniscus tear	0.78	0.53-1.0	Study cohort
Disutility of diagnostic arthroscopy	0.01	0-0.10	Study cohort
Disutility of treatment arthroscopy	0.01	0-0.10	Study cohort
Time variables			
Length of treatment benefit (degenerative)	5 y	0-20 y	Reference 6
Length of treatment benefit (traumatic)	10 y	0-20 y	Reference 36
Delayed diagnosis for false-negative	6 mo	0-5 y	Expert opinion
Test accuracy, %			
Sensitivity of history and physical examination	77	1-99	Reference 45
Specificity of history and physical examination	93	1-99	Reference 45
Sensitivity of MRI	93.3	1-99	Reference 37
Specificity of MRI	88.4	1-99	Reference 37
Cost, 2012 US\$			
MRI	500	0-1500	Medicare
Diagnostic arthroscopy	2295	0-5000	Medicare
Diagnostic and treatment arthroscopy	2499	0-5000	Medicare
Nonoperative treatment	923	0-2000	Medicare
Prevalence of a symptomatic meniscus tear, %			
Primary care	9	1-99	Reference 20
Sports medicine	30	1-99	References 1, 2, 8, 28

^aMRI, magnetic resonance imaging; QALY, quality-adjusted life-year.

and nonlethal surgical complications in the postoperative and recovery periods. These transient periods of disutility are accounted for as a 1-time deduction from the health-related quality of life value during the year of procedure. We calculated a disutility value using the time to recovery and quality of life during recovery for both diagnostic and treatment arthroscopy.^{34,43} Most of our 38-patient cohort had 6-week values in addition to their final follow-up, and on average they reached 75% of their final outcome at 6 weeks. Using this recovery as a guide, we estimated the disutility of knee arthroscopy to be 0.01. This value is found in Table 2.

Costs. Since this model is performed from the societal perspective, as recommended by the Panel on Cost-effectiveness in Health and Medicine, the costs used are those that are accrued to society rather than an individual hospital or practice. Although this method may not account for all the itemized differences in cost seen by an individual hospital or surgical practice, it permits a more global comparison between the treatment strategies and avoids the pitfalls of regional cost differences. Therefore, the costs for MRI, nonoperative treatment, diagnostic arthroscopy, and combined diagnostic/therapeutic arthroscopy were estimated using the national average Medicare reimbursements for the procedures in 2012 \$US. The evaluation and management costs for H&P were excluded since they are common to each diagnostic strategy. Reimbursement for the diagnosis-related group (DRG) code and the professional fees for each procedure were summed to create the total reimbursement. Current Procedural Terminology (CPT) codes were used to estimate the professional costs. These costs are shown in Table 2.

Cost-effectiveness Analysis

Reported-outcome measures included average costs and effectiveness (QALYs) as well as the cost-effectiveness ratio for each strategy. QALYs combined the utility of a health state with the time spent in that health state. The incremental costs and effectiveness were also calculated and represented the relative difference between the 2 alternative strategies. The principle outcome measurement calculated was the incremental cost-effectiveness ratio (ICER), which is the ratio between the difference in costs and difference in QALY of each strategy. In terms of this model, 1 ICER would be expressed as $ICER = (Cost_{MRI} - Cost_{H\&P}) / (QALY_{MRI} - QALY_{H\&P})$. An ICER less than \$50,000 per QALY gained was considered to be cost-effective based on a willingness of the health care system to pay (WTP) value of \$50,000. In this cost-effectiveness analysis, the preferred treatment strategy was the most effective strategy, with an ICER < WTP. Additionally, we reported the likelihood of a true-positive, true-negative, false-positive, false-negative, unnecessary surgery, and delayed diagnosis with each diagnostic strategy (Table 3).

Sensitivity Analyses

Sensitivity analyses were used to validate the model and determine the effect of input variables across their range of reasonable values. These were performed to assess the effect that each variable has on the outcome of the model. If changing a variable changes the preferred strategy (ie, the most cost-effective strategy), then the model was deemed "sensitive" to that variable; if not, then it was

TABLE 3
Results of Analysis for Base Case^a

Scenario	Diagnostic Strategy	Cost, Average \pm SD, US\$ ^b	QALYs Gained, Average \pm SD	Average Cost Difference, US\$ ^b	Average Difference in QALYs Gained	Cost-effectiveness Ratio \pm SD, US\$/QALY ^b	Incremental Cost-effectiveness Ratio
Primary care							
Traumatic	H&P	1479 \pm 33	9.861 \pm 0.0013	—	—	150 \pm 3	—
	MRI	1978 \pm 32	9.861 \pm 0.0013	\$500	—	201 \pm 3	Dominated
	MRI+	1557 \pm 33	9.863 \pm 0.0012	\$78	0.002	158 \pm 3	\$40,565/QALY
Degenerative	MRI-	1897 \pm 32	9.859 \pm 0.0014	\$418	-0.002	192 \pm 3	Dominated
	H&P	1479 \pm 1	4.662 \pm 0.04	—	—	317 \pm 3	—
	MRI	1978 \pm 0	4.662 \pm 0.04	\$500	—	424 \pm 4	Dominated
Sports medicine	MRI+	1557 \pm 14	4.662 \pm 0.04	\$78	—	334 \pm 5	Dominated
	MRI-	1897 \pm 15	4.662 \pm 0.04	\$418	—	407 \pm 5	Dominated
	Traumatic						
Traumatic	H&P	2046 \pm 37	8.557 \pm 0.0017	—	—	176 \pm 4	—
	MRI	2620 \pm 37	8.556 \pm 0.0015	\$543	-0.001	227 \pm 4	Dominated
	MRI+	2077 \pm 39	8.600 \pm 0.0016	\$31	0.003	191 \pm 4	\$10,631/QALY
Degenerative	MRI-	2581 \pm 36	8.554 \pm 0.0015	\$504	-0.003	211 \pm 4	Dominated
	H&P	1735 \pm 37	4.659 \pm 0.0015	—	—	372 \pm 9	—
	MRI	2234 \pm 37	4.659 \pm 0.0013	\$500	—	479 \pm 9	Dominated
Sports medicine	MRI+	1884 \pm 39	4.659 \pm 0.0015	\$149	—	404 \pm 9	Dominated
	MRI-	2077 \pm 36	4.659 \pm 0.0013	\$342	—	446 \pm 9	Dominated
	Test Outcome, %						
		True-Positive	True-Negative	False-Positive	False-Negative	Delayed Diagnosis ^c	Unnecessary Surgery ^d
Primary care							
Traumatic	H&P	6.98	82.7	8.23	2.08	1.38	1.38
	MRI	8.49	80.4	10.4	0.63	0.42	0.42
	MRI+	6.53	90.0	0.9	2.58	1.72	1.72
Sports medicine	MRI-	8.97	73.2	17.6	0.14	0.09	0.09
	H&P	23.2	63.5	6.32	6.95	4.6	4.6
	MRI	28.1	61.7	8.12	2.09	1.37	1.37
Sports medicine	MRI+	21.6	69.1	0.71	8.6	5.68	5.68
	MRI-	29.7	56.1	13.7	0.46	0.30	0.30
	MRI	2.47	13.7	13.7	0.46	0.30	0.30

^aDiagnostic strategies: H&P, patient history and physical examination only; MRI, magnetic resonance imaging for all patients; MRI+, MRI only to confirm positive H&P; MRI-, MRI only to confirm negative H&P. All numbers reference H&P as the common baseline. The preferred strategy in each scenario is bolded and italicized. QALY, quality-adjusted life-year.

^b2012 \$US.

^cDelayed diagnosis, false-negatives not improving with nonoperative treatment. In other words, the percentage of patients suffering due to an incorrect diagnosis. The first number is for degenerative tears and the second for traumatic.

^dUnnecessary surgery, false-positives not recovering with nonoperative treatment; in other words, the percentage of patients. The first number is for degenerative tears and the second for traumatic.

“robust.” One-way sensitivity analyses, which change the value of a single variable at a time, were performed on the discount rate, age of the cohort, and all costs, utilities, and probabilities in the model. Any input variable capable of significantly influencing the ICER or altering the preferred treatment method was reported in the results. To further validate the base case findings, variables were subjected to multivariate sensitivity analyses by varying the value of more than 1 variable at a time.

RESULTS

Base Case

To fully assess the effect of the diagnostic strategies in different populations, 4 distinct clinical scenarios were examined for meniscus tears: traumatic tears evaluated in the primary care setting, traumatic tears evaluated in the orthopaedic/sports medicine setting, degenerative tears evaluated in the primary care setting, and degenerative

tears evaluated in the orthopaedic/sports medicine setting. Results of the analysis are shown in Table 3.

Regardless of the scenario, the least expensive strategy was always H&P, which cost \$1479 in primary care settings and \$1735 in specialist settings. The most expensive strategy was MRI, which cost \$500 more than H&P. The alternative MRI testing strategies, MRI+ and MRI-, were intermediate in cost between H&P and MRI, but MRI+ was the less expensive of the 2. In terms of efficacy, at the mean there was clinical difference in QALYs among the strategies. MRI- gained 0.001 QALYs over the other strategies, but this was clinically insignificant, with a minimal clinically important difference of the SF-6D of 0.03.

For degenerative tears in both settings, H&P was the most cost-effective strategy. In these scenarios the ICER for any use of MRI was greater than the WTP of \$50,000/QALY. For traumatic tears, MRI to confirm a positive history and physical examination was the preferred cost-effective strategy, with an ICER less than WTP (\$40,565/QALY in the primary care setting and \$10,631 in the orthopaedic/sports medicine setting). However, because the QALY

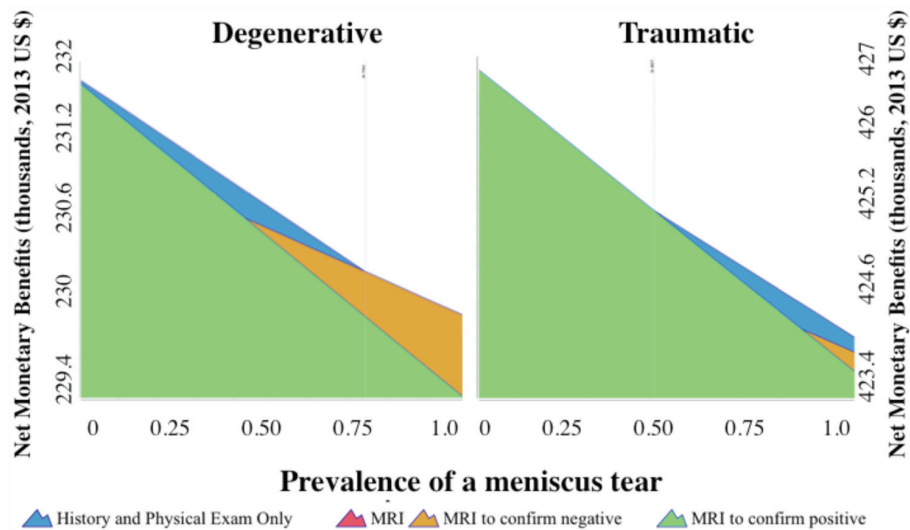


Figure 2. Strategy graph for the prevalence of meniscus tears.

gains were 10-fold lower than the minimally clinically important difference in the SF-6D, concern exists over this interpretation. The additional cost of an MRI to confirm a positive over H&P only was very small, at \$31 in the sports medicine clinic and \$78 in the primary care setting. As such, H&P and MRI to confirm positive were effectively equivalent strategies for the traumatic meniscus tear, both dominating MRI to confirm negative and MRI for all patients. These results are seen in Table 3.

While measuring monetary costs and effectiveness in QALYs are important, other measures are sometimes helpful in interpreting the effect of diagnostic tests. To this end we calculated the percentage probability of a true-positive, true-negative, false-positive, false-negative, unnecessary surgery, and delayed diagnosis for each strategy in the 2 different clinical settings, primary care and orthopaedic/sports medicine. In both settings MRI+ led to more than a 10-fold lower rate of unnecessary surgeries than did any other strategy, while MRI- led to a 2.08 and 2.26 higher rate than H&P alone in primary care and orthopaedic clinics, respectively. The opposite was true for delayed diagnosis; the MRI+ strategy resulted in the greatest likelihood of delayed diagnosis, while MRI- resulted in the lowest. These results are shown in Table 3.

Calculating the probability of superiority of each diagnostic strategy gives important information as to the likelihood that a strategy will be preferred, as opposed to the preferred strategy at the mean. For degenerative tears seen by a primary care physician, H&P was preferred in 80.7% of the trials and MRI+ in 19.3%. For degenerative tears seen by an orthopaedic or sports medicine physician, H&P was preferred in 96.2% of the trials and MRI+ in 3.8%. For traumatic tears seen by a primary care physician, H&P was preferred in 41.2% of the trials and MRI+ in 58.8%. For traumatic tears seen by an orthopaedic or sports medicine physician, H&P was preferred in

82% of the trials and MRI+ in 18%. MRI and MRI- were not preferred in any trial for any clinical scenario.

Sensitivity Analysis

One, 2, and 3-way sensitivity analyses were performed on all variables. The ones found to have an effect on the preferred strategy (ie, the sensitive variables) were prevalence of meniscus tears, accuracy of physical examination, percentage responding to nonoperative treatment in patients *without* a meniscus tear, and cost of an MRI. All remaining variables, including the accuracy of MRI, were robust. The results of the 1-way sensitivity analyses for all the variables along with their threshold values (the value above which the results of the base case analysis no longer held true and a different strategy became the most cost effective) are found in the Appendix (available online at <http://ajsm.sagepub.com/supplemental>).

Prevalence of Disease

The model outcomes were very sensitive to the prevalence of meniscus tears in the presenting population. This can also be thought of as the pretest probability of a meniscus tear before an evaluation by a health care provider. Figure 2 shows a preferred strategy graph across the range of possible values. For traumatic tears, MRI+ was preferred when prevalence is less than 46.7% and H&P was preferred when the prevalence is greater than 46.7%. For degenerative tears, H&P was preferred until the prevalence reaches 74.2%, and then MRI- was the preferred strategy.

Diagnostic Test Accuracy

Clearly the model's outcomes are highly dependent on the accuracy of the diagnostic tests. The sensitivity and

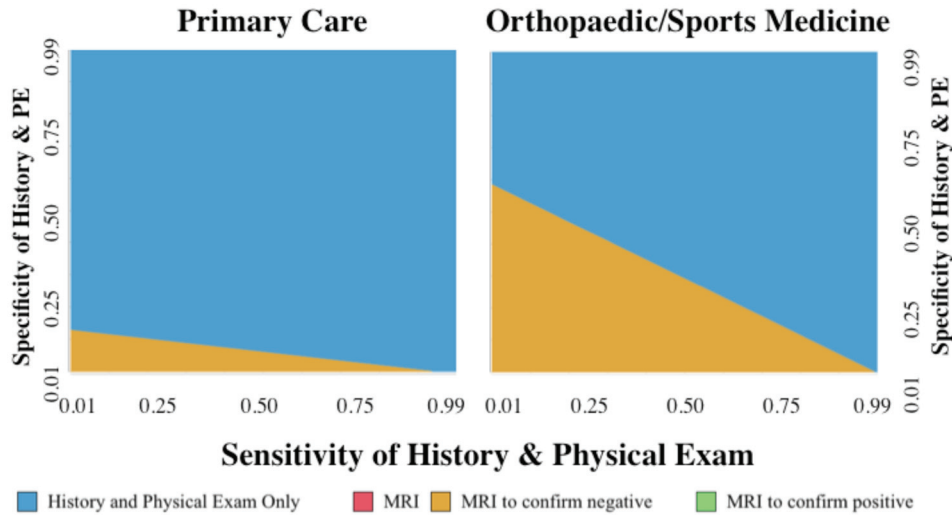


Figure 3. Two-way sensitivity analysis of the sensitivity and specificity of history and physical examination (PE) for the suspected degenerative meniscus tear.

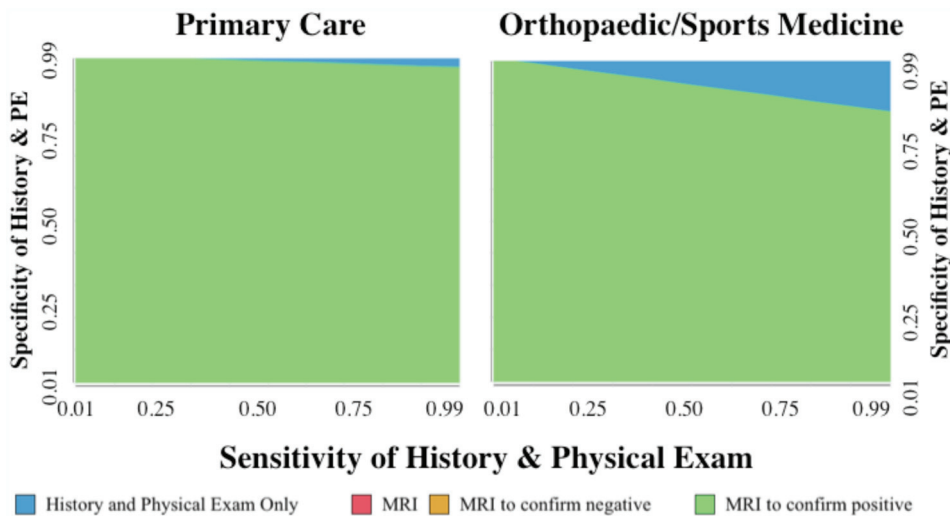


Figure 4. Two-way sensitivity analysis of the sensitivity and specificity of history and physical examination (PE) for the suspected traumatic meniscus tear.

specificity of MRI was robust in the model, and a 2-way sensitivity analysis varying sensitivity against specificity showed that test performance must be above 80% for both parameters for any diagnostic strategy utilizing an MRI to even be considered a viable strategy. The model, however, is sensitive to history and physical examination. This was depicted in 2-way sensitivity analyses varying sensitivity against specificity for the each of the clinical scenarios (Figures 3 and 4). Figure 3 shows the results for a degenerative tear in both the primary care and orthopaedic or sports medicine clinic, while Figure 4 shows the results for traumatic tears in both populations. For degenerative tears, the range of H&P accuracy over which H&P remains the preferred strategy was very wide. Only under

circumstances with unreasonably poor performance was H&P not preferred. For traumatic tears, the graph was upside-down, with MRI+ preferred over a wide range of values. However, the base case sensitivity and specificity values were close to the threshold line, where H&P was preferred.

DISCUSSION

The Centers for Medicare and Medicaid Services have shown that diagnostic imaging is the fastest rising segment of costs in the US health care system over the past decade, increasing an astounding 40%. Designing and

implementing cost-effective diagnostic strategies is paramount to meeting the sophisticated demands of health care delivery in the future. In this study, we proposed a framework for how to examine the value of diagnostic tests using the case of meniscus tears.

Knee pain is one of the most common presenting complaints to both primary care and orthopaedic or sports medicine physicians, and arthroscopic meniscectomy is the most commonly performed orthopaedic procedure in the US. It meets all of the proposed criteria for re-evaluation of diagnostic strategies, and others^{4,14} have already questioned the use of MRI in the diagnosis of symptomatic meniscus tears. This is the first study to examine the cost-effectiveness of the diagnosis of meniscus tears in a decision analysis framework.^{4,14} The results of this study offer opportunities for decreasing costs by utilizing history and physical examination alone for suspected degenerative meniscus tears.

The strategy for traumatic tears is less clear than that for degenerative tears, but it does appear to favor using MRI to confirm a positive history and physical examination. Our study found that in the primary care and orthopaedic or sports medicine clinic, MRI to confirm a positive H&P is the preferred strategy with an ICER below the WTP threshold. Because of the limitations of the available data (prevalence is not reported by tear type), we used the same prevalence for degenerative and traumatic meniscus tears in the model; however, the actual prevalence of traumatic tears is likely much lower than for degenerative. Sensitivity analysis on this variable demonstrates clearer superiority of MRI+, at low prevalence.

Furthermore, we did not model potential indirect applications of MRI such as using tear descriptions to guide repair and treatment of cartilage injuries. Since clinical decision making is complicated, we recognize that many physicians may order an MRI not simply to rule in or rule out a meniscus tear, but to identify other lesions such as cartilage injury. Fortunately, these concurrent pathologic abnormalities are nearly always addressable during surgery, but they do have distinctly different recoveries. If the recommendations of this study are to be adopted, consent before arthroscopic partial meniscectomy without an MRI will necessarily be broader; in cases where a patient is unwilling to undergo a procedure with extensive recovery such as microfracture, anterior cruciate ligament reconstruction, or meniscus repair, an MRI may be indicated for surgical decision making.

Lastly, we recognize that avoiding unnecessary surgeries is critical given the current medical-legal environment and the potential, but rare, catastrophic complications of surgery. The MRI+ strategy does achieve the lowest probability of unnecessary surgery at 0.12% to 1%. Given these considerations, the low incremental cost of MRI+ strategy, and the variability of the input data, we conclude that MRI to confirm a positive history and examination is the preferred diagnostic strategy for suspected traumatic meniscus tears.

A key policy effect of this study is the focus on H&P. It implies that a renewed focus should be given to teaching and utilizing effective physical examination skills. While

our study focused on meniscus injuries as an example, the published accuracy of patient history and physical examination is often at or near the levels of expensive and even invasive imaging tests.^{25,27,33,46} In fact, in the diagnosis of meniscus tears, H&P actually has a higher specificity than MRI, which drives the poor performance of the MRI strategy in our model. However, the virtually unlimited access to imaging tests and the ingrained culture of defensive medicine have permitted physical examination skills to deteriorate. The lack of thorough musculoskeletal education at US medical schools has also contributed to this problem.⁹ Educational policies that encourage teaching, performing, and deductive reasoning based on physical examination, such as adding a physical examination portion to the US Medical Licensing Examination Step II Clinical Skills, may dramatically improve the skill level of rising young physicians.

These findings also raise the specter of 2 other important issues in health care delivery: medical liability reform and payor requirements for exhaustive diagnostic tests before authorization of payment for procedures. While this study supports using H&P as the sole means of diagnosis for degenerative meniscus tears, practitioners in medically litigious states will be unlikely to adopt this strategy. Such a disincentive to efficient practice will need to change in order for the diagnostic algorithm set forth by this study to be adopted. We have highlighted the growing idea that less may be more for diagnostic tests and that in certain settings the use of expensive intermediate diagnostic tests to prevent unnecessary procedures is counterproductive and inefficient. We believe that this study is an excellent example of how careful evaluation of diagnostic strategies can help to prevent the devaluation of the clinical decision-making process and to permit a culture of medical decision makers free from dependence on high-cost imaging.

The weakness of this study lies primarily in the low level of evidence for some of the studies used to define model inputs. Variables in the model with medium to lower levels of evidence include treatment outcomes of partial meniscectomy and prevalence of meniscus tears. While these are weaknesses, one of the unique strengths of decision analysis is its ability to identify critical variables on which to focus future research. Both of these variables have important effects on the preferred diagnostic strategy and deserve to be further studied to refine their point estimates. The alternative outcomes such as delayed diagnosis and unnecessary surgery also give strength to the analysis for those with concerns regarding utility estimates. For example, if a particular policy preference were to minimize unnecessary surgeries, MRI+ would be preferred. To minimize delayed diagnoses, MRI- is preferred. However, MRI does result in a higher probability of performing an unnecessary surgery than does H&P. This last fact highlights a frequent diagnostic fallacy, that is, assuming additional imaging may help but not hurt the patient. Sophisticated imaging is not necessarily superior to traditional diagnostic methods, and reliance on imaging tests to make clinical decisions can be hazardous, as exemplified by the recent discovery of nephrogenic systemic sclerosis.²⁹

Another weakness is the complexity of the model. Decision models must be complex enough to faithfully represent the disease while still accurately answer the question at hand. We did not model specific diagnoses for knee pain other than meniscus tears, and we avoided detailed treatments such as meniscus repair. However, we believe that the outcomes and their relationship to each other reflect the contribution of the diagnostic tests on disease outcome without adding unneeded complex treatment modeling, which would dampen or dilute the reward of an accurate test. We recognize that other treatments may be necessary for the range of possible diagnoses, but modeling those introduces unnecessary complexity to the model and introduces a negative bias for accurately ruling out meniscus tears.

While the reported accuracy of H&P is high and comparable with MRI, it is likely that the variability in actual practice is greater than MRI since radiologists are subject to accreditation by oral and written boards but no standardization is applied to history and physical examination skills. This phenomenon makes a rigorous sensitivity analysis necessary. However, when examining the sensitivity analyses in Figures 3 and 4, the range of sensitivity and specificity over which the preferred strategy does not change is quite wide. Therefore, probable variation in H&P among practitioners does not pose significant weakness to the conclusions of this study.

Lastly, it is important to note that this study references *symptomatic meniscus tears*. Several studies have found the prevalence of asymptomatic meniscus tears in middle-aged or elderly patients is as high as 91%.^{4,14} In this case, MRI would have a lower specificity for degenerative tears. In this regard, our model is actually favorable to MRI since we use a high specificity. Further study is necessary to define the accuracy of MRI for degenerative tears of the meniscus.

Although cost-effectiveness studies are primarily a tool for population studies and policy decision makers, they can also be utilized by individual practitioners as well as the health care system. In the current health care climate, with policies such as benchmarking expected to become more prevalent, our results can help individual practitioners guide diagnostic strategy choice in their practice. To do this they would need to plot their own disease prevalence and H&P accuracy, as in Figures 2 to 4, to identify their preferred strategy. Additionally, doing so can aid in treatment choices for individual patients by comparing quality-of-life preferences for delayed diagnosis versus unnecessary surgery. For example, while in our model an unnecessary surgery is worse than a delayed diagnosis, the reverse may be true for an athlete, thereby altering the preferred strategy.

Cost-effectiveness analysis is established as a powerful tool for policymakers, insurers, and clinicians, and this study shows that using the proposed framework can successfully identify more efficient diagnostic strategies. In fact, if the results of this study were implemented, for every 1 million patients seen with knee pain, the US health system could save between \$500 million and \$5 billion dollars annually. Given the considerable potential for cost

savings, we recommend further level 1 evidence studies to better define sensitive variables and guide the diagnostic strategy recommendations for meniscus tears.

ACKNOWLEDGMENT

The authors thank Dustin Hambright, MD, for contributing to the literature review.

REFERENCES

1. Arendt E, Dick R. Knee injury patterns among men and women in collegiate basketball and soccer. NCAA data and review of literature. *Am J Sports Med*. 1995;23(6):694-701.
2. Baquie P, Brukner P. Injuries presenting to an Australian sports medicine centre: a 12-month study. *Clin J Sport Med*. 1997;7(1):28-31.
3. Barton GR, Sach TH, Jenkinson C, Doherty M, Avery AJ, Muir KR. Lifestyle interventions for knee pain in overweight and obese adults aged > or = 45: economic evaluation of randomised controlled trial. *BMJ*. 2009;339:b2273.
4. Bhattacharyya T, Gale D, Dewire P, et al. The clinical importance of meniscal tears demonstrated by magnetic resonance imaging in osteoarthritis of the knee. *J Bone Joint Surg Am*. 2003;85(1):4-9.
5. Black C, Clar C, Henderson R, et al. The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation. *Health Technol Assess*. 2009;13(52):1-148.
6. Chatain F, Robinson AH, Adeleine P, Chambat P, Neyret P. The natural history of the knee following arthroscopic medial meniscectomy. *Knee Surg Sports Traumatol Arthrosc*. 2001;9(1):15-18.
7. Coleman S, Briffa K, Conroy H, Prince R, Carroll G, McQuade J. Short and medium-term effects of an education self-management program for individuals with osteoarthritis of the knee, designed and delivered by health professionals: a quality assurance study. *BMC Musculoskelet Disord*. 2008;9:117.
8. Dallalana RJ, Brooks JHM, Kemp SPT, Williams AM. The epidemiology of knee injuries in English professional rugby union. *Am J Sports Med*. 2007;35(5):818-830.
9. Day CS, Yu YR, Yeh AC, Newman LR, Arky R, Roberts DH. Musculoskeletal preclinical medical school education: meeting an underserved need. *J Bone Joint Surg Am*. 2009;91(3):733-739.
10. Dehaven KE, Dolan WA, Mayer PJ. Chondromalacia patellae and the painful knee. *Am Fam Physician*. 1980;21(1):117-124.
11. Dehaven KE, Dolan WA, Mayer PJ. Chondromalacia patellae in athletes. Clinical presentation and conservative management. *Am J Sports Med*. 1979;7(1):5-11.
12. Dias RC, Dias JM, Ramos LR. Impact of an exercise and walking protocol on quality of life for elderly people with OA of the knee. *Physiother Res Int*. 2003;8(3):121-130.
13. Doucette SA, Goble EM. The effect of exercise on patellar tracking in lateral patellar compression syndrome. *Am J Sports Med*. 1992;20(4):434-440.
14. Englund M, Guermazi A, Gale D, et al. Incidental meniscal findings on knee MRI in middle-aged and elderly persons. *N Engl J Med*. 2008;359(11):1108-1115.
15. Englund M, Lohmander LS. Risk factors for symptomatic knee osteoarthritis fifteen to twenty-two years after meniscectomy. *Arthritis Rheum*. 2004;50(9):2811-2819.
16. Englund M, Roos EM, Roos HP, Lohmander LS. Patient-relevant outcomes fourteen years after meniscectomy: influence of type of meniscal tear and size of resection. *Rheumatology (Oxford)*. 2001;40(6):631-639.
17. Fischer SP, Fox JM, Del Pizzo W, Friedman MJ, Snyder SJ, Ferkel RD. Accuracy of diagnoses from magnetic resonance imaging of the knee. A multi-center analysis of one thousand and fourteen patients. *J Bone Joint Surg Am*. 1991;73(1):2-10.

18. Grootendorst P, Marshall D, Pericak D, Bellamy N, Feeny D, Torrance GW. A model to estimate health utilities index mark 3 utility scores from WOMAC index scores in patients with osteoarthritis of the knee. *J Rheumatol*. 2007;34(3):534-542.
19. Grossman JW, De Smet AA, Shinki K. Comparison of the accuracy rates of 3-T and 1.5-T MRI of the knee in the diagnosis of meniscal tear. *AJR Am J Roentgenol*. 2009;193(2):509-514.
20. Hing E, Hall MJ, Xu J. National Hospital Ambulatory Medical Care Survey: 2006 outpatient department summary. *Natl Health Stat Report*. 2008;4:1-31.
21. Hurley MV, Walsh NE, Mitchell HL, et al. Economic evaluation of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain. *Arthritis Rheum*. 2007;57(7):1220-1229.
22. Jackson JL, O'Malley PG, Kroenke K. Evaluation of acute knee pain in primary care. *Ann Intern Med*. 2003;139(7):575-588.
23. Kannus P, Natri A, Paakkala T, Jarvinen M. An outcome study of chronic patellofemoral pain syndrome. Seven-year follow-up of patients in a randomized, controlled trial. *J Bone Joint Surg Am*. 1999;81(3):355-363.
24. Katz JN, Brophy RH, Chaisson CE, et al. Surgery versus physical therapy for a meniscal tear and osteoarthritis. *New Engl J Med*. 2013;368(18):1675-1684.
25. Kocabey Y, Tetik O, Isbell WM, Atay OA, Johnson DL. The value of clinical examination versus magnetic resonance imaging in the diagnosis of meniscal tears and anterior cruciate ligament rupture. *Arthroscopy*. 2004;20(7):696-700.
26. Logan M, Watts M, Owen J, Myers P. Meniscal repair in the elite athlete: results of 45 repairs with a minimum 5-year follow-up. *Am J Sports Med*. 2009;37(6):1131-1134.
27. Lowery DJ, Farley TD, Wing DW, Sterett WI, Steadman JR. A clinical composite score accurately detects meniscal pathology. *Arthroscopy*. 2006;22(11):1174-1179.
28. Majewski M, Susanne H, Klaus S. Epidemiology of athletic knee injuries: a 10-year study. *Knee*. 2006;13(3):184-188.
29. Marckmann P, Skov L, Rossen K, et al. Nephrogenic systemic fibrosis: suspected causative role of gadodiamide used for contrast-enhanced magnetic resonance imaging. *J Am Soc Nephrol*. 2006;17(9):2359-2362.
30. McCaig LF, Burt CW. National Hospital Ambulatory Medical Care Survey: 1999 emergency department summary. *Adv Data*. 2001;320:1-34.
31. Medicare Payment Advisory Commission. *Health Care Spending and the Medicare Program: A Data Book*. Washington, DC: Medicare Payment Advisory Commission; 2010.
32. Meserve BB, Cleland JA, Boucher TR. A meta-analysis examining clinical test utilities for assessing meniscal injury. *Clin Rehabil*. 2008;22(2):143-161.
33. Miller GK. A prospective study comparing the accuracy of the clinical diagnosis of meniscus tear with magnetic resonance imaging and its effect on clinical outcome. *Arthroscopy*. 1996;12(4):406-413.
34. Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med*. 2002;347(2):81-88.
35. Muellner T, Weinstabl R, Schabus R, Vecsei V, Kainberger F. The diagnosis of meniscal tears in athletes. A comparison of clinical and magnetic resonance imaging investigations. *Am J Sports Med*. 1997;25(1):7-12.
36. Nawabi DH, Cro S, Hamid IP, Williams A. Return to play after lateral meniscectomy compared with medial meniscectomy in elite professional soccer players. *Am J Sports Med*. 2014;42(9):2193-2198.
37. Oei EH, Nikken JJ, Verstijnen AC, Ginai AZ, Myriam Hunink MG. MR imaging of the menisci and cruciate ligaments: a systematic review. *Radiology*. 2003;226(3):837-848.
38. Pavelka K, Gatterova J, Olejarova M, Machacek S, Giacobelli G, Rovati LC. Glucosamine sulfate use and delay of progression of knee osteoarthritis: a 3-year, randomized, placebo-controlled, double-blind study. *Arch Intern Med*. 2002;162(18):2113-2123.
39. Pearse EO, Craig DM. Partial meniscectomy in the presence of severe osteoarthritis does not hasten the symptomatic progression of osteoarthritis. *Arthroscopy*. 2003;19(9):963-968.
40. Post WR. Patellofemoral pain: results of nonoperative treatment. *Clin Orthop Relat Res*. 2005;436:55-59.
41. Pujol N, Beaufils P. Healing results of meniscal tears left in situ during anterior cruciate ligament reconstruction: a review of clinical studies. *Knee Surg Sports Traumatol Arthrosc*. 2009;17(4):396-401.
42. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC. The role of cost-effectiveness analysis in health and medicine. Panel on Cost-effectiveness in Health and Medicine. *JAMA*. 1996;276(14):1172-1177.
43. Servien E, Acquitter Y, Hulet C, Seil R. Lateral meniscus lesions on stable knee: a prospective multicenter study. *Orthop Traumatol Surg Res*. 2009;95(8 suppl 1):S60-S64.
44. Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-effectiveness in Health and Medicine. *JAMA*. 1996;276(16):1339-1341.
45. Sihvonen R, Paavola M, Malmivaara A, et al. Arthroscopic partial meniscectomy versus sham surgery for a degenerative meniscal tear. *N Engl J Med*. 2013;369(26):2515-2524.
46. Solomon DH, Simel DL, Bates DW, Katz JN, Schaffer JL. The rational clinical examination. Does this patient have a torn meniscus or ligament of the knee? Value of the physical examination. *JAMA*. 2001;286(13):1610-1620.
47. Underwood M, Ashby D, Carnes D, et al. Topical or oral ibuprofen for chronic knee pain in older people. The TOIB study. *Health Technol Assess*. 2008;12(22):iii-iv, ix-155.
48. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *JAMA*. 1996;276(15):1253-1258.
49. Whitelaw GP Jr, Rullo DJ, Markowitz HD, Marandola MS, DeWaele MJ. A conservative approach to anterior knee pain. *Clin Orthop Relat Res*. 1989;246:234-237.