**Washington State
Health Care Authority**

Health Technology Assessment

HTA Updated Final Report
Hyaluronic Acid/Viscosupplementation

Date: Thursday, April 26, 2010

Health Technology Assessment Program
676 Woodland Square Loop SE

P.O. Box 42712

Olympia, WA 98504-2712

http://www.hta.hca.wa.gov

MEDICAID EVIDENCE-BASED DECISIONS PROJECT (MED)

Summary of Findings

VISCOSUPPLEMENTATION FOR OSTEOARTHRITIS OF THE KNEE

April 15, 2010

Table of Contents

Tables

Appendixes

Glossary

References

Executive Summary

Background

Osteoarthritis (OA) is the most common form of chronic articular disease. OA affects approximately 27 million adults in the United States. The most commonly affected joint is the knee, with prevalence estimates ranging from 12% to 16%. To date, there is no known cure for OA nor is there a disease-modifying agent. Optimal management generally requires a combination of both nonpharmacological and pharmacological therapies, and joint replacement surgery or a joint salvage procedure may be considered for selected patients with severe symptomatic OA who have not obtained adequate pain relief and functional improvement from medical therapy. Pharmacological therapy generally begins with acetaminophen, followed by nonsteroidal antiinflammatory drugs (NSAIDs) if sufficient pain relief is not obtained. There is a small risk of systemic adverse effects with NSAIDs. Aspiration of fluid followed by intraarticular injection of a corticosteroid ameliorates pain in some patients, but duration of relief is usually limited to one to three weeks. Additionally, repeated intraarticular injections of corticosteroids have the potential to cause postinjection flare, infection, and progressive, long-term cartilage damage.

Recently, viscosupplementation with hyaluronan has been introduced as an alternative intraarticular injection therapy for OA. Hyaluronans are also known as sodium hyaluronate or hyaluronic acid (HA). HA is a normal component of synovial fluid and cartilage. The viscous nature of the compound allows it to act as a joint lubricant, whereas its elasticity allows it to act as a shock absorber. Hyaluronic products are characterized by their molecular weight, which varies according to the source of the compound and method of preparation. Five HA products are currently marketed in the United States: Euflexxa® (Ferring), Hyalgan® (Sanofi-Aventis), Orthovisc® (Anika Therapeutics), Supartz® (Seikagaku Corporation), and Synvisc® (Genzyme). Synvisc is a derivative of HA that consists of cross-linked polymers; the compound is referred to as Hylan G-F 20.

Hyaluronate preparations have been approved by the Food and Drug Administration (FDA) for treatment of pain associated with OA of the knee in patients who have not had an adequate response to nonpharmacological, conservative treatment and simple analgesics. Off-label use of HA in other joints has not been studied nearly as extensively as HA for OA of the knee. Recent systematic reviews have come to contradictory conclusions regarding the effectiveness of viscosupplementation, and national guidelines vary in their recommendations.

Methods

Studies reporting pain and function measures that are recognized as key measures by the OA research community were considered for inclusion. Surrogate measures of function, such as improvement in range of motion or changes in joint space, were not considered.

Literature searching began with the standard core sources that are part of the Medicaid Evidence-Based Decisions (MED) Project Rapid Review process. These sources were searched for systematic reviews, technology assessments, and guidelines published between January 2006 and the end of December 2009. MEDLINE and EMBASE searches were conducted to identify RCTs published after the last search dates of the most recent systematic review (September through December 2009). The following search terms were used: viscosupplementation or HA or hyaluronate or hyaluronan or hylan combined (and) with osteoarthritis or knee. The reference lists of the included systematic reviews and primary studies were manually searched. Additionally, a search was made for guidelines published in the last 10 years by the following organizations: American Association of Orthopedic Surgeons, American Pain Society, and American College of Rheumatology.

Three systematic reviews concerned primarily with the efficacy of viscosupplementation, a systematic review of trials comparing hylan with HA, and a systematic review of trials comparing HA or hylan with corticosteroids were selected for inclusion in this report. Four randomized controlled trials (RCTs) from the recently published literature were also selected; these included two placebo-controlled trials, a head-to-head comparison between hylan and non–cross-linked HA, and a head-to-head comparison between HA and exercise with placebo control. Cost and cost-effectiveness data were available various systematic reviews and in other publications.

Findings
There is consistent evidence demonstrating that viscosupplementation results in lower mean pain scores and improves mean function socres a few weeks after treatment. However, the magnitude of benefit may be too small to be clinically important. (Average change in pain score typically did not meet the threshold of minimal clinical importance, as defined by the OA research community.) There is a much greater volume of evidence regarding impact on pain than on function, and many studies did not follow patients beyond three months. Therefore, the impact of viscosupplementation on eventual recovery of function is uncertain. Compared with intraarticular corticosteroid injection, viscosupplementation appears to confer longer-lasting benefit, but the evidence was considered low quality. For comparisons with other treatments, there was insufficient evidence to allow any conclusion. Adverse events occur at a frequency of approximately 2% in single courses of treatment and are primarily transient local reactions; although rare, serious reactions are possible. The rate of adverse events per patient has been shown to increase with repeat courses of treatment, but the only available data were for hylan (high-molecular weight HA).

Evidence pertaining to issues other than efficacy and safety is of low quality:

- Available evidence suggests that viscosupplementation may be as effective as NSAIDs (four RCTs) and results in fewer systemic adverse events (two RCTs); in comparison with intraarticular corticosteroids, it has a delayed onset and longer-lasting benefit (nine RCTs plus meta-analysis).

- Hylan may have a superior benefit compared with that of non–cross-linked HA, but the magnitude of difference is very uncertain and hylan poses a small increase in the risk of adverse events.
- To date, there is no evidence of a difference in benefit between low and medium molecular weight HA.
- Younger age may be associated with greater efficacy; evidence pertaining to effectiveness by other patient characteristics and history is lacking.
- No definitive statement can be made regarding the cost-effectiveness of viscosupplementation.

Guidelines

Of the three high-quality guidelines, one made a weakly positive recommendation in favor of viscosupplementation, one concluded that no recommendation could be made due to the uncertainty in the evidence, and one recommended against viscosupplementation on the basis of a cost-effectiveness analysis that may not be applicable in the United States. The variation in guideline recommendations reflects a body of evidence that provides moderate-quality support of limited efficacy but leaves many important questions unanswered regarding clinical usefulness and cost-effectiveness.

Major Limitations of the Evidence

- Analyses suggesting that poorer-quality and smaller trials have inflated overall estimates of efficacy.
- Variation in methods across trials and across meta-analyses, which makes comparison of results difficult.
- Limited quantity of data pertaining to treatment effect in terms of proportion of patients who have clinically important improvement.
- No studies comparing viscosupplementation with glucosamine and/or chondroitin.
- No analysis of the synergistic effect of hyaluronans *combined* with specific therapies, and little or no relevant trial detail available in the selected reviews.
- No analysis of actual safety profile for any FDA-approved products other than hylan that is based on large, unbiased databases or registries.
- Few economic evaluations. Their estimates of clinical benefit cannot be directly compared with the conclusions of systematic reviews. The limited evidence that is available is primarily from non-U.S. healthcare systems.
- No controlled trials designed to test whether HA injections avert or delay total knee replacement (TKR), and a paucity of evidence concerning the efficacy and safety of different dosing regimens or repeat treatments. These issues were not key questions and thus were not targeted in the search for reviews and primary studies. However, it appears that little evidence on these issues is available.

Background

Clinical overview

Osteoarthritis (OA), the most common form of chronic articular disease, is characterized by degenerative loss of articular cartilage, subchondral sclerosis, joint deterioration, and biochemical and biomechanical changes of the extracellular matrix. According to the Centers for Disease Control and Prevention (CDC), OA affects approximately 27 million adults in the United States. The most commonly affected joint is the knee, with prevalence estimates ranging from 12% to 16%. The disease typically results in chronic pain, loss of joint function, and general disability. OA of the knee and hip has ranked high in terms of disability-adjusted life-years (DALYs) and years lived with disability (YLDs). Additionally, the economic burden of OA is substantial: costs associated with the disease have been reported to exceed $60 billion in the United States (Hayes, 2009).

To date, neither a known cure for OA nor a disease-modifying agent is available. Therefore, treatment is focused on reducing pain, maintaining and/or improving joint mobility, and limiting functional impairment. Optimal management generally requires a combination of both nonpharmacological and pharmacological therapies. Joint replacement surgery or a joint salvage procedure is considered for selected patients with severe symptomatic OA who have not obtained adequate pain relief and functional improvement from nonpharmacological and pharmacological therapies, and who experience progressive limitation in their activities of daily living (ADL). Nonpharmacological therapy generally includes education and support, physical therapy (including exercise), and occupational therapy. Pharmacological therapy may involve one or more of several options: (1) oral therapy with nonopioid analgesics (e.g., acetaminophen), nonsteroidal antiinflammatory drugs (NSAIDs), or opioid analgesics (e.g., propoxyphene, codeine, oxycodone); (2) topical therapy with analgesics such as capsaicin or methylsalicylate; (3) intraarticular injection with steroids; and (4) treatment with glucosamine and/or chondroitin sulfate (Hayes, 2009).

Acetaminophen is considered the preferred first-line pharmacological therapy for patients with symptomatic OA. NSAIDs are often used following insufficient pain relief with nonpharmacological treatments and acetaminophen. However, NSAIDs have been associated with a small risk of potentially serious systemic adverse effects. When oral and topical medications are inadequate, aspiration of fluid followed by intraarticular injection of a corticosteroid ameliorates pain in some patients, but relief is usually limited to one to three weeks. Additionally, repeated intraarticular injections of corticosteroids have the potential to cause postinjection flare, infection, and progressive, long-term cartilage damage (Hayes, 2009; Turajane et al., 2009).

Viscosupplementation with hyaluronan has been introduced as an alternative to NSAIDs or intraarticular injection therapy for OA. Hyaluronans are also known as sodium hyaluronate or hyaluronic acid (HA). HA is a glycosaminoglycan polymer and a normal

component of synovial fluid and cartilage. It plays a major role in the maintenance of the structural and functional characteristics of both the extracellular matrix of the cartilage and the synovial fluid. The viscous nature of the compound allows it to act as a joint lubricant, whereas its elasticity allows it to act as a shock absorber. The concentration and molecular weight of endogenous hyaluronan are reduced in patients with osteoarthritic joints, and hence, the joint is more susceptible to damage. HA injections not only replace lost HA but are thought to have a disease-modifying effect (Hayes, 2009; VA, 2008; Wang, Lin, Chang, Lin, & Hou, 2004).

Hyaluronic products can be characterized by varying molecular weight. Five HA products are currently marketed in the United States: Euflexxa® (Ferring), Hyalgan® (Sanofi-Aventis), Orthovisc® (Anika Therapeutics), Supartz® (Seikagaku Corporation), and Synvisc® (Genzyme). Synvisc is a derivative of HA that consists of cross-linked polymers; the compound is referred to as Hylan G-F 20. The manufacturer recommendations specify one to five injections per course of treatment (VA, 2008).

Policy context
Hyaluronate preparations have been approved by the Food and Drug Administration (FDA) for treatment of pain associated with OA of the knee in patients who have not had an adequate response to nonpharmacological, conservative treatment and simple analgesics. Off-label uses include OA of the hip, shoulder, and ankle; temporomandibular joint (TMJ) disorders; and rheumatoid arthritis of the knee. These nonapproved indications have not been studied nearly as extensively as HA for OA of the knee. Retreatment after a previous course of hyaluronan for any indication, including the knee, has not been studied as thoroughly as single treatments. Policymakers need to know whether treatment with viscosupplementation is effective in patients with the approved indication of knee OA; if so, in which patient subpopulations; and how viscosupplementation compares with other nonsurgical treatment options. Recent systematic reviews have come to contradictory conclusions regarding the effectiveness of viscosupplementation, and national guidelines vary in their recommendations. Some authors have suggested that the uncertainty surrounding the clinical effectiveness of this product relates to its original approval as a medical device and lack of the same preapproval investigation required of pharmaceutical FDA applications (Bannuru et al., 2009; VA, 2008).

Key Questions
1. **Key question 1. (a)** *What is the clinical effectiveness of viscosupplementation for treatment of OA of the knee?* **(b)** *Do different viscosupplementation products vary in effectiveness?*
2. **Key question 2.** *What are the adverse effects associated with viscosupplementation in patients with OA of the knee?*
3. **Key question 3.** *Does the effectiveness of viscosupplementation vary by subpopulation defined by these factors: age, race/ethnicity, gender,*

primary versus secondary OA, disease severity and duration, weight (body mass index), and prior treatments?

4. **Key question 4.** *What are the cost implications and cost-effectiveness of this type of product?*

Methods

Inclusion criteria

<u>Patient group</u>: Adults with OA of the knee.
<u>Intervention(s)</u>: Viscosupplementation (hyaluronic acid injection – Hyalgan, Synvisc, Supartz, Orthovisc, Euflexxa).
<u>Comparator(s)</u>: NSAIDs, corticosteroid injection, physical therapy, oral pain medications, placebo, arthroscopic lavage and/or debridement.
<u>Outcome(s)</u>: Pain, function, quality of life, adverse effects.

The primary outcomes for this Rapid Review were reduction in pain and improvement in function. The OA research community recognizes the following outcome measures (Bannuru et al., 2009; VA, 2008):

- Pain (visual analogue scale [VAS] or Likert scale) in the index joint—at rest, during walking, or during activities other than walking.
- Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index. The WOMAC scale is specific to the knee or hip and measures symptoms related to pain, stiffness, and physical function. A higher score represents *less* severe symptoms; scores range from 0 to 96.
- Lequesne Index. The Lequesne scale is specific to the knee and assesses pain with walking and with activities of daily living (ADL). Higher scores on this scale signify *greater* pain-induced impairment; scores range from 0 to 24 (VA, 2008).

Studies reporting these measures were considered for inclusion as evidence pertaining to Key Question #1. Surrogate measures of function, such as improvement in range of motion or changes in joint space, were not considered.

These additional inclusion criteria were applied: (1) English language, (2) randomized controlled trials (RCTs) (including trials where other active treatment served as the control) or systematic reviews of RCTs, and (3) uncontrolled studies cited by selected systematic reviews for safety data.

Exclusion criteria
Reviews and primary studies published more than 10 years ago; studies in which a nonstandard treatment regimen (e.g., < three injections unless the product was approved for single injection) was used; studies with less than one month of follow-up.

Search strategy

The following core sources were searched for systematic reviews, technology assessments, and guidelines published between January 2000 through December 2009: BMJ Clinical Evidence; Hayes, Inc.; Cochrane Library; UK National Library for Health (NLH), including National Institute for Health and Clinical Excellence (NICE); Canadian Agency for Drugs and Technologies in Health (CADTH); Institute for Clinical Systems Improvement (ICSI); Agency for Healthcare Research and Quality (AHRQ); Veterans Affairs/Department of Defense (VA/DoD); Washington State Health Technology Assessment (HTA) program; and the Blue Cross/Blue Shield HTA program. The following search terms were used: viscosupplementation or hyaluronan or hyaluronate.

MEDLINE and EMBASE searches were conducted to identify RCTs published after the last search dates of the latest systematic review (September through December 2009). The full search strategy is included in Appendix A. The reference lists of included systematic reviews and primary studies were manually searched.

A search of the following sources helped identify additional guidelines published in the last 10 years: American Association of Orthopedic Surgeons, American Pain Society, American College of Rheumatology, and the National Guidelines Clearinghouse. Search terms were *viscosupplementation or hyaluronan or hyaluronate*.

Quality assessment

The quality of the guidelines was assessed using a modified and adapted Appraisal of Guidelines Research & Evaluation (AGREE) instrument. The methodological quality of the included studies (systematic reviews and RCTs) was assessed using a modified NICE and Scottish Intercollegiate Guidelines Network (SIGN) tool. A summary judgment for the overall quality of the body of evidence (low, moderate, or high) was assigned to each key question using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system.

The GRADE system defines these categories for assessing the quality of bodies of evidence for therapeutic interventions (Brozek et al., 2009):

High: Further research is very unlikely to change our confidence in the estimate of effect. Typical sets of studies would be RCTs without serious limitations or well-performed observational studies with very large effects (or other qualifying factors).

Moderate: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. Typical sets of studies would be RCTs with serious limitations or well-performed observational studies yielding large effects.

*Low: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. Typical

sets of studies would include RCTs with very serious limitations or observational studies with important limitations or without special strengths.

*Very low: Any estimate of effect is very uncertain. Typical sets of studies would include RCTs with very serious limitations and inconsistent results, observational studies with serious limitations, or unsystematic clinical observations (e.g. case series or case reports).

*The MED Project has chosen to collapse these two categories because they usually have the same policy implications. Other organizations are also adopting this practice (Guyatt et al., 2008).

Quality assessment of individual guidelines, systematic reviews, and primary RCTs appears in Appendix B.

Search Results

Three systematic reviews concerned primarily with the efficacy of viscosupplementation, (Bellamy et al., 2006; Hayes, 2009; Samson et al., 2007), a systematic review of trials comparing hylan with HA (Reichenbach et al., 2007), and a systematic review of trials comparing HA or hylan with corticosteroids (Bannuru et al., 2009) were selected. The general reviews covered disparate sets of RCTs, not only because of different publication dates but also because of differences in selection criteria: Hayes included only RCTs published in full and with sample sizes ≥ 100; the meta-analyses reviewed by Samson et al., including the Bellamy meta-analysis, included RCTs published as abstracts as well as in full; Bellamy et al. also solicited information on unpublished trials; Samson et al. reviewed only data pertaining to placebo comparisons; and there were variations in the databases searched. A few very early trials were missing from the Hayes review. The Reichenbach special-issue review (hylan versus HA) review included some trials missed in the most comprehensive general review (Bellamy et al., 2006). The Bannuru special-issue review (HA versus corticosteroid) included one study that was not included in the other reviews.

The literature search yielded four RCTs published later than the last search date in the systematic reviews. These included two placebo-controlled trials (Altman, Rosen, Bloch, Hatoum, & Korner, 2009; Baltzer, Moser, Jansen, & Krauspe, 2009), a head-to-head comparison between hylan and non-cross-linked HA (Chou, Lue, Lee, Lin, & Lu, 2009), and a head-to-head comparison between HA and exercise with placebo control (Kawasaki et al., 2009).

Cost and cost-effectiveness data were available in three systematic reviews (Hayes, 2009; VA, 2008; Waddell, 2007), and an additional two primary economic studies were selected from the National Health Service (NHS) Economic Evaluation Database (EED) (Jacobs, Kane, & Clarke, 2008; Turajane, Labpiboonpong, & Maungsiri, 2007). However, the article by Jacobs et al. was subsequently excluded. The article calculates

the lost revenue for an orthopedic clinic in the UK due to the unavailability of reimbursement for HA injections in addition to office visit reimbursement. This situation does not exist in the United States (see Appendix C for CPT codes). Lastly, data from a cost-effectiveness analysis was abstracted from one of the selected guidelines (NICE, 2008).

Findings

Key Question #1a: *What is the clinical effectiveness of viscosupplementation for treatment of OA of the knee?*

Systematic reviews and technology assessments, Key Question 1a
Efficacy (versus placebo)
Quantitative synthesis of the evidence for efficacy came primarily from six published meta-analyses that were summarized and critically appraised in an Agency for Healthcare Research and Quality (AHRQ) Technology Assessment (Samson et al., 2007). (See Table 3.) A total of 5843 patients and 42 placebo-controlled RCTs are represented in the Samson review of meta-analyses. In addition, Samson et al. performed several sensitivity analyses, primarily with data abstracted by one of the reviewed meta-analyses, which was a Cochrane Review (Bellamy et al., 2006). The Bellamy review and meta-analysis emphasized by-product, by-comparator, and by-outcome results. However, Samson et al. chose to focus on by-class results because of two considerations: previous evidence that molecular weight is not associated with outcomes, except perhaps in the case of hylan versus non–cross-linked hyaluronans; and the preponderance of very small sets of studies (often only one) in the by-product and by-comparator analyses conducted by the Bellamy review. The outcomes targeted by the Samson review were pain, physical function, QOL, and adverse events. The reviewed meta-analyses addressed all of these outcomes except QOL. In some of the selected meta-analyses, various measures of the same outcome were pooled for a standardized estimate.

Each of the six meta-analyses calculated pooled estimates for multiple follow-up intervals. The estimates were sometimes calculated separately for pain in different situations (e.g., at rest or walking). In every analysis, after-treatment pain scores were lower for HA groups than for placebo groups, or improvement in pain was greater in HA groups than in placebo groups. The differences were sometimes nonsignificant for shorter follow-up intervals (especially when ≤ four weeks) or when data were available from four or fewer RCTs. In summary, after controlling for placebo effect, an actual treatment effect on pain was consistently observed, but was often delayed and peaked at two months or more.

These pain effects were generally small and thus, their clinical importance is unclear. For example, the statistically significant weighted mean differences (WMDs) fell in the range of 1.0 to 22.5 on a 100-point VAS scale. WMD refers to the between-group difference in pain score at follow-up or difference in pain improvement, depending on

the particular meta-analysis. According to the Samson review, the authors of the meta-analyses offered no definition of clinical importance. Samson and colleagues cite a source suggesting that a 20- to 40-point improvement in WOMAC pain (100-point scale) is considered a positive response. Other authors have considered a 20-point improvement on 100-point pain scales to be clinically important (Altman, Rosen, Bloch, Hatoum, & Korner, 2009; Kahan et al., 2003; Torrance et al., 2002; Yen et al., 2004). The Bellamy analysis reported some WMDs ≥ 20 for VAS pain, but only in the analysis of trials of hylan versus placebo. Significant and favorable standardized mean differences (SMDs), also referred to as effect sizes, were reported for pain outcomes by the Bellamy review and exceeded 0.8 for any type of HA. According to the Samson and Bellamy reviews, standardized effect sizes may be interpreted as follows: 0.2 or 0.3 = small, 0.5 = moderate (i.e., clinically recognizable), and 0.8 = large. (NOTE: Effect size categories are generic and do not necessarily translate to clinical importance for particular health problems, but the cutoffs suggested by Bellamy et al. are consistent with general convention.) The SMDs reported in other meta-analyses were very small (0.0 to 0.32).

A small mean effect does not convey whether only a few patients or a substantial proportion of patients experienced clinical improvement. The Samson review found that almost all placebo comparisons in individual trials failed to report the results in useful terms such as the proportion of patients in each arm who experienced clinically meaningful improvement. No results were provided in the Samson review for trials that did report in these terms, except for the one RCT omitted from the meta-analyses. In this study, 44% of patients treated with hylan and 30% of the placebo arm (*P*=0.048) were symptom free at 26 weeks; 43% of patients treated with non-hylan HA were symptom free, and differences between the two active treatment arms were nonsignificant (Rolf et al., 2005).

Samson and colleagues reported that some number-needed-to-treat (NNT) calculations were available, but NNT estimates were conflicting and only a few were tied to assumptions concerning clinical importance.

There were fewer meta-analyses of functional outcomes than of pain outcomes. Of 15 analyses reported in the Samson review, 9 were significant and favorable, and again, those were for the longer follow-up periods. Effect sizes for function outcomes ranged from 0.16 at best in one meta-analysis to 0.32 in another meta-analysis to ≥ 0.8 in the Bellamy review. One meta-analysis of functional outcomes reported positive results but these were difficult to understand in practical terms because of an unusual calculation. Analyses of composite outcomes measured with the Lequesne Index were significant and favorable, but the effects were small, i.e., differences at follow-up in the range of approximately one point on a 0 to 24 scale. The Samson review cites a source suggesting that 20% is the minimum clinically important improvement for the Lequesne Index.

The authors of the five study-level meta-analyses covered in the Samson review came to a variety of conclusions regarding the efficacy of viscosupplementation. These ranged from negative to moderately positive to strongly positive. The authors of the Samson review considered only one meta-analysis to have reported data and analysis that fully supported the meta-analysis authors' conclusion. This was also the meta-analysis with a negative conclusion—that the *clinical* effectiveness of viscosupplementation has not been proven and that viscosupplementation may be associated with a higher risk of adverse events. The primary flaws that Samson and colleagues reported for the other meta-analyses were failure to search EMBASE and use of language restrictions. The conclusion of Samson and colleagues was that a clinical benefit has not been clearly demonstrated for viscosupplementation as a treatment for OA of the knee and that rigorous, multicenter RCTs are needed.

The qualitative review by Hayes (2009) reported positive effects on pain, function, and/or QOL in six placebo-controlled RCTs (n=2140 evaluable patients) and no effect in four placebo-controlled RCTs (n=807); conflicting efficacy results were also seen in the subset of recent trials that were not included in the earlier systematic reviews. (See Tables 1 and 2.) The Hayes review did not discuss the magnitude of benefit or relationship between benefit and follow-up interval. However, the review did identify two studies that reported the proportion of patients who improved ≥ 20 points on a 100-point scale. The results favored HA over placebo and were statistically significant:

- 56% versus 41% in per-protocol analysis (*P*=0.031); 36% versus 28% in intention-to-treat (ITT) analysis (nonsignificant) (Altman & Moskowitz, 1998).
- 76% versus 62% (*P*=0.0346); < 10% dropout rate (Neustadt, Caldwell, Bell, Wade, & Gimbel, 2005).

Appendix D provides an overview of the placebo-controlled trials included in the various reviews and meta-analyses.

Comparisons with other therapies
For nonplacebo comparators, only one to four RCTs per comparator met the Hayes review's selection criteria of sample size ≥ 100. (See Tables 1 and 2.) In comparisons with NSAIDs, appropriate care only, exercise, and intraarticular corticosteroids, the results were either conflicting or available from a single trial.

RCTs comparing viscosupplementation with other therapies were included in the Bellamy review. (See Table 4.) These authors did not pool the results for different products compared with the same alternative but did convert the results to common measures. The review included six RCTs comparing a single viscosupplementation product with NSAIDs. The review authors concluded that the two treatments had generally comparable efficacy (two of the six RCTs assessed safety only). Considering seven comparisons with corticosteroid injection (three trials in one of the comparisons), they concluded that HA/hylan appeared to confer longer-term benefits. Trials comparing

HA or hylan with treatments other than NSAIDs or intraarticular corticosteroids were too few in number to allow the authors to reach conclusions.

A meta-analysis of seven RCTs (n=606) comparing viscosupplementation with corticosteroid injection (Bannuru et al., 2009) reported results consistent with the conclusion reached in the Bellamy review. Pooled effect sizes favored corticosteroids in the first few weeks but then began to increasingly favor HA/hylan with time. The greatest effect (at 17 to 26 weeks) was still modest (0.39; 95% CI, 0.18 to 0.59). The analysis was based on a hierarchy of pain measures. No functional outcomes were assessed. (See Tables 3 and 4.)

The Hayes review identified two comparisons with conventional treatment in which the proportion of patients who improved by ≥ 20 points on a 100-point scale were reported:

- 69% versus 40% (*P*=0.0001) for WOMAC pain and 62% versus 35% (*P*=0.0001) for WOMAC pain and either stiffness or physical functioning; ITT analysis (Raynauld et al., 2002).
- 65% versus 40% (*P*<0.0001); ITT analysis (Kahan, Lleu, & Salin, 2003).

RCTs, Key Question #1a
Efficacy versus placebo
One double-blind, placebo-controlled RCT of good quality demonstrated efficacy at 26 weeks (Altman, Rosen, Bloch, Hatoum, & Korner, 2009). No treatment effect was observed at 12 weeks. The effect on primary outcome met the authors' definition of a minimally important clinical benefit: the odds of a ≥ 20-point absolute improvement in pain on a 100-point VAS were 1.7 times (95% confidence interval 1.2 to 2.4) greater in the HA group. The effect on physical functional status (SF-36® Health Survey Physical Component Score [QualityMetric Inc.]) was also strong. No placebo-controlled effect was observed in the other selected RCT (see following discussion).

Comparison with other therapies
A double-blind randomized trial of good quality compared autologous conditioned serum (ACS) with HA and with saline placebo (Baltzer, Moser, Jansen, & Krauspe, 2009). ACS was found to have a substantial effect on function, pain, and quality of life (QOL) at 7, 13, and 26 weeks, compared with both HA and with placebo. The differences between HA and placebo were very small and nonsignificant. Similar results were observed in the approximately 60% of patients who could be traced at two years. In a comparison of HA with home exercise, there was no difference between groups in reduction of pain or improvement of function (Kawasaki et al., 2009); this trial was of fair quality.

Overall summary of evidence, Key Question #1a
Efficacy versus placebo
A large body of evidence, including approximately 50 RCTs comparing viscosupplementation with placebo, finds a small improvement in pain and possibly function, although benefits are generally not seen until after a few weeks. The

magnitude of benefit from HA alone is likely too small to be clinically important. Meta-analyses estimated that the difference in improvement between HA and placebo groups was < 20 points on a 100-point VAS pain scale, except in hylan trials, and reported variable standardized effect sizes (large in one meta-analysis but < 0.40 in two others) for function outcomes. There is a much greater volume of evidence regarding impact on pain than on function, and many studies did not follow patients beyond three months. Therefore, the impact of viscosupplementation on eventual recovery of function is uncertain.

Comparison with other therapies

Evidence on viscosupplementation compared with NSAIDs is limited, but an analysis of four randomized comparator trials showed generally nonsignificant differences in pain and function scores at follow-up. Comparison of HA with intraarticular corticosteroids in two systematic reviews, including one meta-analysis of nine randomized comparator trials, suggests that HA offers less immediate relief compared with intraarticular corticosteroids but greater benefits after the first few weeks. Two RCTs suggest that in comparison with conventional care, viscosupplementation leads to a greater number of patients experiencing clinically meaningful improvement.

Key Question #1b: *Do different viscosupplementation products vary in effectiveness?*

Systematic reviews and technology assessments, Key Question #1b
The Hayes review noted the paucity of data pertaining to the comparative effectiveness of different products. The review identified only three RCTs with sample size > 100 that addressed this question (see Table 2); these were not included in any other systematic reviews because of their publication dates. One single-blind trial with a large dropout rate found no significant differences in WOMAC or QOL outcomes comparing high (hylan) with two non–cross-linked products. Outcomes were similar for the low and medium molecular weight arms, but these were not analyzed statistically. An unblinded trial resulted in superior outcomes in the hylan arm at six and 12 months but no differences in QOL outcomes, but a double-blind trial found no difference between hylan and other HA.

Four of the meta-analyses covered in the Samson review provided evidence that hylan has a superior effect to that of non-hylan products. Such evidence was generated by eliminating outliers, separately analyzing hylan trials, or performing meta-regression, and so the magnitude of the difference in effect was not directly assessed (see following paragraphs). A fifth meta-analysis detected no association between molecular weight and effect size in meta-regression but did not specifically analyze according to hylan versus non-hylan. These analyses constitute indirect comparisons because no head-to-head comparator trials were included.

Sensitivity analyses conducted by the Samson review shed some doubt on previous indirect comparisons of hylan with non-hylan HA. In the Samson review's model, hylan

versus non-hylan was associated with larger effect size. However, because of significant heterogeneity, the authors of the Samson review split the six hylan (versus placebo) trials into two groups, two trials with larger effects (both poor quality) and four trials with smaller effects. When analyzed in this fashion, heterogeneity was minimal, and the confidence interval for the pooled effects of the poorer quality trials did not cross the confidence interval for the pooled effect of all six trials. See following discussion of another sensitivity analysis (Reichenbach et al., 2007) of indirect comparisons.

Some head-to-head comparator trials were included in the overall Bellamy review, but the authors concluded that they were too few in number to allow conclusions about the relative value of hylan over non-hylan HA or of any hyaluronic product compared with another.

One systematic review on the issue of hylan versus non-hylan was available (Reichenbach et al., 2007). The reviewers analyzed the effect on pain in 13 RCTs and quasi-randomized comparator trials (n=2085). A small absolute effect below the authors' definition of minimally important clinical improvement was detected, and this disappeared when two outlier trials were removed from analysis. Furthermore, meta-regression analysis showed no association between molecular weight (continuous variable) and effect size. The authors also conducted an indirect comparison based on the 31 placebo-controlled trials contributing to three of the meta-analyses covered in the Samson review. The SMD of pain scores at follow-up was approximately three times that calculated for the comparator trials. Further analysis suggested that this discrepancy could be due to the relatively small size of the hylan versus placebo trials, the large effects that they reported, and an association across all 31 trials of smaller sample size with larger effect.

RCTs, Key Question #1b
A very small randomized comparator trial of poor quality enrolled patients with bilateral OA (Chou, Lue, Lee, Lin, & Lu, 2009). Clinicians injected one knee with hylan and the other with non-hylan HA. Greater improvement in WOMAC pain and in VAS pain was observed in the hylan knees. For example, at 26 weeks WOMAC pain score on a 10-point scale was 1.2 in hylan knees and 1.7 in non-hylan knees; the respective VAS pain scores were 45 and 55. Differences were statistically significant after adjustment for very small baseline differences. However, differences in functional outcomes were small and nonsignificant.

Overall summary of evidence, Key Question #1b
Hylan may have a superior benefit compared with that of non–cross-linked HA, but the magnitude of difference is unlikely to be clinically significant. To date, there is no evidence of a difference in benefit between low and medium molecular weight HA.

Strength and limitations of the evidence, Key Questions #1a and #1b

The selected systematic reviews were of fair to good quality. The Bellamy and Hayes reviews came to positive conclusions regarding the overall efficacy of viscosupplementation, whereas the Samson review concluded that efficacy remains uncertain. A key reason for this discrepancy was the weight attached in the Samson review to the lack of demonstrated clinically important benefit in most studies and meta-analyses.

There was high consistency across meta-analyses of positive, though not always statistically significant, results in favor of a benefit compared with placebo; evidence was derived from RCTs, most of them considered fair to good quality; and outcome measures were directly related to Key Question #1. However, nearly all trials failed to report the results in useful terms such as the proportion of patients in each arm who experienced clinically meaningful improvement. In addition, the overall body of evidence is characterized by a lack of directness. First, all meta-analyses reported high heterogeneity. The authors were often able to explain heterogeneity in terms of one or more of the following factors and their association with larger effects:

- Poorer trial quality. The trials included in the systematic reviews were typically of fair quality (see details from Samson review in Table 1), but many were of poor quality. The limitations included absence of clear allocation concealment, lack of ITT analysis, and high dropout rates.
- Smaller sample size.
- Outlier trials.
- Protocol allowing use of escape medication.
- Patient age < 65 years.

Secondly, methods were inconsistent across trials, meta-analyses, and systematic reviews. As noted in the critique by Campbell, Bellamy, & Gee (2007) (see Table 1), trials have differed in whether they allow rescue medication, whether baseline is considered to be pretreatment or time of last injection, and whether change in score or final score was reported. None of the systematic reviews analyzed the results by these factors or provided this type of trial-specific data.

Limitations other than indirectness have also been explored. According to the Samson review, 55% of placebo-controlled trials were funded by industry, but the quantitative evidence of bias attributable to industry sponsorship was mixed. Several analyses of publication bias also led to conflicting conclusions; the Samson review, itself, and three of the meta-analyses included in Samson suggested publication bias, while two others did not. Publication bias occurs when negative or disappointing results are not submitted or accepted for publication. Some meta-analyses included trials that were unreported or published only as abstracts. Thus, some of the remaining existing publication bias was overcome in these meta-analyses. In summary, there are some strengths in the evidence but there are also some hindrances in comparing the

calculated treatment effects; thus, the overall body of evidence regarding the efficacy of viscosupplementation versus placebo is of moderate quality.

The quality of evidence pertaining to other issues can be summarized as follows:

- Rates of clinically important improvement: Clinical importance could be more fully assessed if RCTs reported success rates in addition to mean improvement. In other words, it would be useful to know not only the mean difference in pain score before and after treatment, but also the proportion of patients who improved to a degree that is predefined as clinically important. Only four out of more than 50 trials reported this statistic: two RCTs with placebo comparison groups and two RCTs with conventional treatment comparison groups. ITT analysis in three of the trials showed statistically significant differences in the number of patients who achieved clinically meaningful improvement. It is not possible to know whether other trials included this type of calculation or if they did, why the authors did not report results.
- Synergistic effect of viscosupplementation combined with other therapies: Insufficient information to allow a judgment regarding the specific combination(s) of concurrent therapy to which the addition of viscosupplementation provides the greatest benefit. The systematic reviews provided no analysis of this issue and little or no relevant detail describing the selected RCTs.
- Comparative effectiveness of viscosupplementation versus NSAIDs: Consistent results in several trials but no analysis of trial quality or heterogeneity across trials (low quality).
- Comparative effectiveness versus corticosteroids: Generally poor-quality trials, according to Bannuru et al. (2009) (low quality).
- Comparative effectiveness versus glucosamine and/or chondroitin: No studies.
- Differential effect, hylan versus non-hylan: Indirect nature of pooled estimates favoring hylan, conflicting results among comparator trials, and poor quality of both placebo-controlled and comparator trials (low quality).
- Differential effect according to molecular weight among non-hylan products: Evidence came primarily from metaregression analysis of placebo-controlled trials (low quality).

Key Question #2: *What are the adverse effects associated with viscosupplementation in patients with OA of the knee?*

Systematic reviews and technology assessments, Key Question #2
The Hayes and Bellamy reviews described adverse events as occurring at very low rates in RCTs. However, it is understood that most RCTs are underpowered to detect significant adverse events, and these reviews did not include any observational studies or findings from safety databases. The Samson review, on the other hand, described minor adverse events as "common" and cited event rates from large case series. In two case series (n=3931 and n=4253 for first treatment) cited by the Samson review,

adverse events occurred at frequencies of 2.1% of injections or 5.3% of patients in a single course of treatment and 8.5% of injections or approximately 10% of patients in repeat courses of treatment. In both of the studies with repeat treatment data, hylan, rather than medium or low molecular weight HA, was used. Adverse events were generally local reactions (injection site pain, injection site infection, and local joint pain or swelling). See Table 1 for additional details.

According to Samson and colleagues, trial data suggest that *severe* adverse events are *not* common. In one meta-analysis reviewed by Samson et al., major adverse events occurred in three of 1002 knees (0.30%) treated with non-hylan and one of 139 knees (0.72%) treated with hylan. Serious adverse events included swelling, vasculitis, and hypersensitivity reaction. The Samson review reported that the FDA's Manufacturer and User Facility Device Experience (MAUDE) database suggests that rare serious events, e.g., pseudosepsis, are possible with Hyalgan, Euflexxa, and Synvisc (hylan). In 85 of 236 reports between January 2005 and January 2007, patients were hospitalized. Nine of the 236 reports mentioned pseudosepsis or pseudoseptic reaction (Hyalgan, four; Euflexxa, one; Synvisc, four). Pseudosepsis is a noninfectious reaction that mimics sepsis; pseudosepsis following knee injection consists of severe joint inflammation with pain and effusion 24 to 72 hours after injection.

While it is understood that intraarticular injections carry some inherent risk, the extent to which adverse events are associated with the HA compound itself, as opposed to the injection procedure, is unclear. One of the meta-analyses covered by the Samson review reported a pooled relative risk (RR) of 1.08 for any adverse event, whereas another reported a pooled RR of 1.20 for minor events. The patient-level meta-analysis included in the Samson review reported a lower rate of adverse events in HA arms (1.8%) than in the placebo arms (3.2%). The Bellamy review found no significant differences in placebo trials for 12 measures of adverse events, but reported a high pooled RR for pain at the injection site (1.7 to 1.9).

The Reichenbach review of hylan/non-hylan comparator trials observed approximately a twofold increase in the risk of any adverse events associated with hylan and low statistical heterogeneity across trials, even though definitions and reporting varied considerably. The absolute rate of any adverse event in the hylan arms ranged from 0.05% to 18%. The same review estimated that 14 patients would need to be treated with hylan rather than non-hylan HA in order for one patient to suffer an adverse event. No long-term complications were reported. The frequency of adverse events in a very large case series (n=1537) of hylan treatments was 2.7% per injection, as cited in the Samson review.

The Bellamy review observed that, in trials comparing viscosupplementation with systemic treatment, e.g., NSAIDs, there were more local reactions but fewer systemic adverse effects such as gastrointestinal problems. The Hayes report included an RCT in which fewer patients in the HA arm (52%) had had adverse events at one year compared with the appropriate care group (68%; *P*=0.0116).

The long-term safety of viscosupplementation is unknown. The Hayes review noted that the safety of repeated courses of viscosupplementation has not been well studied. As already noted, the Samson review identified two case series showing adverse events to increase with repeat courses of treatment. None of the reviews described concerns regarding long-term after-effects from injection of HA, but they also did not identify any studies with follow-up longer than six months, and most studies had follow-up periods of ≤ three months.

RCTs, Key Question #2
The frequency of treatment-related adverse events was nearly identical (10% to 11%) between groups receiving injections of HA and saline in a large (n=588) RCT (Altman, Rosen, Bloch, Hatoum, & Korner, 2009). In the comparison of ACS and HA (n=376), the HA group experienced local adverse events more frequently (38%) than the ACS group (23%) or saline group (28%) (Baltzer et al., 2009). The nature of the events reported in these trials was similar to those described in the systematic reviews and meta-analyses, but the event rates were larger than those suggested by the reviews. These might be considered outlier trials; there was no readily apparent reason for the discrepancy.

Overall summary of evidence, Key Question #2
As with any intraarticular injection, there is a small risk of local, transient reactions with the administration of viscosupplementation (in the range of 2% of patients in a single course of treatment); serious adverse events are rare (less than 1%). It is unclear whether the hyaluron product adds to the risk associated with the injection procedure. Use of hylan, compared with non-crosslinked HA, poses a small absolute increase in the risk of reactions overall and has a higher potential to produce serious adverse effects. Direct evidence pertaining to the safety of viscosupplementation versus other treatments is too sparse to allow conclusions. There is some evidence that repeat courses of treatment result in increased risk (in the range of 8% of patients) of adverse events, at least with the use of hylan. Long-term studies of the safety of viscosupplementation are lacking.

Strength and limitations of the evidence, Key Question #2
As noted in Bellamy et al., the sample sizes typical of RCTs diminish the value of meta-analyses of RCTs as sources of safety data. However, a few thousand patients treated with viscosupplementation were the basis of reviewers' assessments that serious adverse events are uncommon. In addition, the Samson review cited three case series involving 1500 to more than 4000 patients each. It was not clear whether these case series were systematically selected. No number needed to harm (NNH) calculations were reported for viscosupplementation versus placebo or for other types of treatment; this would have provided an indication of the clinical importance of risks associated with viscosupplementation.

Key Question #3: *Does the effectiveness of viscosupplementation vary by subpopulation defined by these factors: age, race/ethnicity, gender, primary*

versus secondary OA, disease severity and duration, weight (body mass index), and prior treatments?

Systematic reviews and technology assessments, Key Question #3
The Samson review included a key question regarding subpopulation effects. The authors reported the following:

- A trial (also described in the Hayes review) comparing intraarticular HA with placebo found no overall treatment effect but did observe a significant effect in a subgroup of patients who were > 60 years of age and had more severe OA (Lequesne Index scores > 10). This finding was not replicated in a confirmatory study.
- Two RCTs failed to detect a differential effect according to age, sex, or body mass index (BMI)/weight.
- One of these two trials also failed to detect a differential effect by disease severity.
- However, another trial observed substantial pain improvement in patients with Kellgren-Lawrence grade 3 to 4 disease, whereas no effect was observed in patients with grade 2 disease or in the overall study group. (The Kellgren-Lawrence scheme classifies severity, i.e., progression of OA, on a 0 to 4 scale according to the presence of several radiographic changes. Grades 3 and 4 represent moderate and severe OA, respectively [Kellgren & Lawrence, 1957].)
- There were no trials that enrolled only patients with secondary OA or that evaluated outcomes by primary versus secondary disease.
- There were no trials examining race/ethnicity, disease duration, or prior treatment.

The conclusion of the Samson review was that the available evidence does not demonstrate a differential effect by subpopulation, but that the quantity of evidence is limited. No comment was made on the quality of these trials.

A meta-analysis of 20 trials (Wang et al., 2005) included in the Samson review assessed the influence of patient factors on the treatment effect of HA (versus placebo). Using meta-regression and subgroup analysis, the authors found greater mean patient age to be associated with smaller treatment effect. The Samson review did not take this into account when addressing their key question regarding subpopulation effects.

A very small before-and-after study (n=32) cited by Hayes (2009) reported that a higher HA concentration in the synovial fluid predicted greater clinical response in patients who were treated with injections of hylan (60% sensitivity and 77% specificity at an optimal cutoff) (Anandacoomarasamy et al., 2008).

RCTs, Key Question #3

In a comparison of HA with home exercise, less severe OA at baseline was an independent predictor of better outcomes after adjustment for age and BMI, which were not independent predictors (Kawasaki et al., 2009). Severity was measured by a continuous measure (joint space width) of progression in joint deformity. Less severe disease was prognostic of better composite pain and function outcome in both HA and home exercise groups, and both groups had similar outcomes. In an earlier trial cited in the Hayes review (see Table 2), outcomes were significantly better in the HA group than the exercise group, perhaps because that trial included only patients with advanced OA.

Overall summary of evidence, Key Question #3

No strong conclusions can be drawn regarding the differential effectiveness of viscosupplementation by age, race/ethnicity, gender, primary versus secondary OA, disease severity and duration, weight (BMI), and prior treatments because of a paucity of data. Individual trial evidence regarding the influence of age and disease severity has been conflicting, but a meta-regression and subgroup analysis of 20 trials suggested that younger age predicts greater response. Factors other than age or disease severity have either not been studied or have been shown by one or two studies to be unrelated to treatment effect.

Strength and limitations of the evidence, Key Question #3
The evidence is of low quality. There are very few data. Most subgroup analyses were based on post hoc subgroup analysis.

Key Question #4: *What are the cost implications and cost-effectiveness of this type of product?*

Unless otherwise noted, dollar amounts refer to US dollars.

Hayes (2009) cited the following cost information, obtained from the website of a supplier (Axon Medical Supplies):

- Hyalgan: $69 for one 2.0-mL syringe; 10 syringes for $570
- Orthovisc: $706.27 for one 2.0-mL syringe; three syringes or 10 ampules for $1950
- Supartz: $318.99 for five 2.5-mL syringes

A review completed for the Veterans Administration and Department of Defense included these cost estimates, from the perspective of a payer/healthcare system (VA, 2008):

- Euflexxa®: $87/injection, $260/course of treatment (three injections)
- Hyalgan®: $65/injection, $195 to $325 /course of treatment (three to five injections)

- Orthovisc®: $198/injection, $595 to $793/course of treatment (three to five injections)
- Supartz®: $68/injection, $205 to $341/course of treatment (three to five injections)
- Synvisc®: $142/injection, $426/course of treatment (three to five injections)

A study cited by the Waddell review reported costs of $100 to $200 per HA injection, not including the cost of an office visit (Lo, LaValley, McAlindon, & Felson, 2003), but Lo et al. did not provide the source of this information.

The Waddell review was based on a search of the MEDLINE database through October 2006. The author concluded that the bulk of evidence suggests a favorable cost-effectiveness profile for HA, and that this evidence justifies its use (Waddell, 2007). The review presented the following results from economic evaluations, two of which studied *hylan* and none of which were conducted in the United States:

- An incremental cost-utility ratio of CAD $10,000 (1999 costs) per quality-adjusted life-year (QALY), comparing the use of hylan plus appropriate care with appropriate care alone. The trial (n=255), which was randomized but not blinded, was not included in any of the selected systematic reviews. The Bellamy review listed this as one of the studies excluded because it was available only as an abstract, and the Samson review excluded it because it was not an efficacy trial. This cost-effectiveness analysis is based on one of the clinical trials (Raynauld et al., 2002) included in the Hayes review. "Appropriate care" was defined as consistent with guidelines published by the American College of Rheumatology, with instructions to clinicians to treat conservatively. Utilities for calculating QALYs were derived from patient responses to a generic, validated health status questionnaire. The trial reported that an additional 29% (69% versus 40%) of patients were improved at one year in the hylan group compared with the appropriate care group. Improvement was defined as a 20-point reduction in WOMAC score (100-point scale). The authors reported an incremental cost-effectiveness ratio (ICER) of $2505 per patient improved over the one-year time frame. This study was from the societal perspective (including not only work time lost but also time away from usual activity). The cost-utility ratio was described as falling under the suggested Canadian adoption threshold (Torrance et al., 2002). NOTE: ACR guidelines recommend use of nonpharmacologic treatment, e.g., exercise or physical therapy, to minimize reliance on NSAIDs and acetaminophen (ACR, 2000). It is difficult to assess the representativeness of this trial's results since event rates were generally not available in the evidence selected for this Rapid Review.

- Similar per-patient medical and sick leave costs (public payer perspective in France) for patients treated for 9 months with hylan or with conventional treatment, and greater effectiveness in the hylan group (Kahan et al., 2003). This

randomized trial was included in the Hayes and Bellamy reviews but not in the Samson review because of the lack of a placebo control. The observed difference in improvement was 11 to 13 points greater in the hylan arm on 100-point WOMAC and VAS scales. Although differences in mean improvement were not clinically significant according to the previously described 20-point threshold, event rates showed clinically important improvement to be more likely in the HA arms. More patients in the HA arm (65%) than in the conventional treatment arm (40%; $P<0.0001$) experienced ≥ 20% improvement in Lequesne Index score. The proportion of patients who experienced a 20-point or greater decrease in pain with walking (100-point scale) was 88% in the hylan group and 68% in the conventional care group. QALYs were not calculated. NOTE: Conventional care was not defined, and it was not clear whether patients in the hylan arm continued to receive conventional care. However, all patients in the study had failed at least two courses of NSAID therapy. See note regarding event rates in the study by Torrance et al. (2002).

- Inferior cost-effectiveness of HA compared with celecoxib in patients who have a poor global knee assessment and who have declined surgery (Yen et al., 2004). This modeling study was conducted from a societal perspective (costs to public payer plus productivity losses) in Taiwan; the time frame was 26 weeks. HA was assumed to be both more expensive and more effective, in terms of QALYs gained, than either of the two NSAIDs. QALY estimates took into account the probability of gastrointestinal complications from the NSAIDs and related mortality, as well as injection pain from HA. The incremental cost-utility ratio reported for celecoxib versus naproxen was $21,226 per QALY gained. (The authors did *not* report the corresponding ratio for HA versus naproxen, which is $33,148/QALY according to data supplied in the article.) The authors reported an incremental cost-utility ratio of $42,000 for HA versus *celecoxib* and concluded that celecoxib had reasonable cost-effectiveness, while HA might not be economically feasible in Taiwan. NOTE: The conventional limit for cost-effectiveness in the United States is $50,000/QALY, so conclusions might be different for a U.S. setting. The utility values for translating the clinical effect into QALYs were derived from a panel of physicians, rather than patients with knee osteoarthritis; again, representativeness is unknown. The estimate for QALYs gained was based on a single trial in which 36% of patients in the HA arm achieved clinical success (20-point improvement in pain with walking on a 100-point VAS) at 26 weeks (Altman et al., 1998). For reasons already noted, the representativeness of the event rate is unknown. However, the adjusted mean difference in change in pain score at 26 weeks in this trial was 8.8, which is similar to weighted mean differences reported by meta-analyses of placebo-controlled trials: 7.3 at 22 to 30 weeks (Arrich et al., 2005) and 9.0 at 14 to 26 weeks (Bellamy et al., 2006). Sensitivity analysis showed that cost utility was very sensitive to estimates of both the cost and effectiveness of HA. As noted in a review by NICE (NICE, 2008), the cost-effectiveness study did not take into

account the possibility of cardiovascular events associated with NSAIDs, which created a bias in favor of the NSAIDs.

NOTE: These articles were retrieved so that details not reported by Waddell et al. could be included.

One of the RCTs selected from the primary literature included a cost-effectiveness analysis to test the hypothesis that hylan is cost-effective compared with non-hylan (Chou et al., 2009). However, the analysis was incorrectly structured to answer the question. Separate ICERs for hylan and for non-hylan, rather than a single ICER showing the cost-effectiveness of hylan *versus* non-hylan, were reported.

A retrospective cost analysis study from the payer perspective in Thailand was designed to demonstrate potential cost savings from the use of HA (product unidentified) in patients who had failed all other conservative treatment after at least six months (Turajane, Labpiboonpong, & Maungsiri, 2007). The authors found that when HA was ineffective and surgery was necessary, HA contributed only 6% of the total direct medical costs of treatment. They also demonstrated that the cost of viscosupplementation was much less than the cost of surgery. However, there was no analysis of whether the cost of surgery could be avoided altogether or how long it could be delayed in patients who responded to HA.

An economic analysis was conducted in development of the NICE guidelines on osteoarthritis (NICE, 2008). To informally estimate the cost-effectiveness of HA versus placebo, the health economist first constructed a cost consequence table showing health benefits at 26 weeks and costs (non-trial sources) for each of two trials assessing viscosupplementation for OA of the knee. An ICER was then calculated for each trial; one ICER exceeded the National Health Service cost-effectiveness threshold and the other analysis showed placebo to be both more effective and less expensive. The generalizability of these findings is limited for these reasons: the estimate of clinical benefit was based on only two trials; the two trials excluded patients with severe OA; both trials used products that are not available in the United States; the cost estimates assume that injections are always delivered by physicians; and the comparison was with placebo, not standard care.

Washington State Agency Data

The following data is provided by the Washington State agencies on their utilization and cost information.

Washington State Health Care Authority

Table 1   Payment Summary by Agency for 2006-2008

UMP, L&I, DSHS DATA 2006-2008

2006-2008	Patient count	Procedure Count	Total Cost
Uniform Medical Plan	1,969	8,424	$1,201,323
Labor & Industry	934	2,917	$850,330
Dept of Social and Health Services	848	2,780	$461,353
All Agencies	3,571	14,121	$2,513,006

Table 2   Payment Detail by Patient and Procedure, 2006-2008, Combined Agencies

Combined Agency Data, 2006-2008

2006	Pt Ct	Inj Ct	Avg Injs/Pt	Avg Cost/Pt	Avg Cost/Inj	Total Cost
Procedure						
HCPCS 20610 (Injection Procedure)	692	2709	3.9	$ 209	$ 53	$ 144,671
Injectables						
5 Injection Injectables (Supartz 3-5 injs)*						
Hyalgan/Supartz (J7321/Q4083/J7317)	181	814	4.5	$ 392	$ 87	$ 71,005
3 Injection Injectables						
Synvisc (J7322/Q4084/J7320)	517	1718	3.3	$ 510	$ 154	$ 263,860
Euflexxa (J7323/Q4085)	0	0	0	$ -	$ -	$ -
Orthovisc (J7324/Q4086)	0	0	0.0	$ -	$ -	$ -
3 Injection Injectable Totals	517	1718	3.3	$ 510	$ 154	$ 263,860
All Injectables totals	698	2532	3.6	$ 480	$ 132	$ 334,865
2006 Totals/Avg Total Costs			3.6	$ 693	$ 177	$ 479,536

2007	Pt Ct	Inj Ct	Avg Injs/Pt	Avg Cost/Pt	Avg Cost/Inj	Total Cost
Procedure						
HCPCS 20610 (Injection Procedure)	1271	4738	3.7	$ 216	$ 58	$ 274,785
Injectables						
5 Injection Injectables (Supartz 3-5 injs)*						
Hyalgan/Supartz (J7321/Q4083/J7317)	449	1867	4.2	$ 388	$ 93	$ 174,087
3 Injection Injectables						
Synvisc (J7322/Q4084/J7320)	530	1568	3.0	$ 486	$ 164	$ 257,493
Euflexxa (J7323/Q4085)	45	122	2.7	$ 174	$ 64	$ 7,836
Orthovisc (J7324/Q4086)	286	861	3.0	$ 449	$ 149	$ 128,335
3 Injection Injectable Totals	861	2551	3.0	$ 457	$ 154	$ 393,664
All Injectables totals	1310	4418	3.4	$ 433	$ 129	$ 567,751
2007 Totals/Avg Total Costs			3.4	$ 663	$ 178	$ 842,536

2008	Pt Ct	Inj Ct	Avg Injs/Pt	Avg Cost/Pt	Avg Cost/Inj	Total Cost
Procedure						
HCPCS 20610 (Injection Procedure)	1788	6674	3.7	$ 232	$ 105	$ 414,725
Injectables						
5 Injection Injectables (Supartz 3-5 injs)*						
Hyalgan/Supartz (J7321/Q4083/J7317)	596	2294	3.8	$ 363	$ 55	$ 216,080
3 Injection Injectables						
Synvisc (J7322/Q4084/J7320)	686	2127	3.1	$ 463	$ 81	$ 317,600
Euflexxa (J7323/Q4085)	163	472	2.9	$ 284	$ 12	$ 46,292
Orthovisc (J7324/Q4086)	411	1256	3.1	$ 477	$ 50	$ 196,237
3 Injection Injectable Totals	1260	3855	3.1	$ 445	$ 47	$ 560,129
All Injectables totals	1856	6149	3.3	$ 418	$ 49	$ 776,209
2008 Totals/Avg Total Costs			3.3	$ 666	$ 303	$ 1,190,934

*FDA approved Supartz label to change from 5 injections to 3 to 5 injections in January 2006

Table 3. 2008 Injection Series Evaluation, All Agencies

COMBINED AGENCY DATA, 2008

2008 Injection Series	Hyalgan/ Supartz*	Synvisc	Euflexxa	Orthovisc	All Injection Types
Background info					
Total Patients 2008	596	686	163	411	1,856
FDA Injection Counts per Procedure	5(3-5)*	3	3	3	
Series Completions					
Patients completing at least 1 series of injections	39%	72%	64%	71%	61%
Patients completing 2 series or more	4%	12%	7%	10%	
Series Incompletions					
Patients who did not complete any injection series	61%	27%	29%	24%	37%
Patients with a single injection only	16%	12%	15%	12%	
Patients with two injections only	11%	15%	15%	12%	
Three injection incomplete series (Hyalgan/Supartz only)	24%				
Four injection incomplete series (Hyalgan/Supartz only)	11%				

Hyalgan (5 inj) and Supartz (3-5 inj) are combined due to a shared billing code. Completion data may be skewed by the proportion of each drug prescribed, the speed of adoption of FDA approved label changes in practice, and the severity of the patients' condition.

Table 4. UMP Only, HA Procedure Overlap with Knee Surgeries*

UMP DATA, 2006-2008

Year	HA Pt Counts by Year	General Knee Surgery Comparisons				Total Knee Replacement Comparisons			
		Counts		Percentages		Counts		Percentages	
		All KS Pts	HA Pts with KS	HA Pts with KS	KS Pts with HA	All TKA Pts	HA Pts with TKA	HA Pts with TKA	TKA Pts with HA
2006	376	567	149	40%	26%	43	12	3%	28%
2007	647	579	176	27%	30%	46	19	3%	41%
2008	942	1,478	191	20%	13%	56	19	2%	34%

KS = Knee Surgery   TKA = Total Knee Arthroplasty (Knee replacement)

*Notes:
- UMP data is presented due to inability to link patients and claims for other agencies
- Analysis constrained to 2006-2008 due to approval of HA injections in 2006, and incomplete annual data available after 2008.
- Short time frames for all procedures (HA, KS, TKA) reduces our ability to form linkages between events
- General estimated rate of turnover for plan beneficiaries is 30% annually
- Small populations for procedures may skew results

Overall summary of evidence, Key Question #4
The cost of viscosupplementation, from a U.S. payer perspective, has been reported to be in the range of $65 to $195 per injection; one to five injections are required for a single course of treatment.

No definitive statement can be made regarding the cost-effectiveness of viscosupplementation. Only two pragmatic studies (societal perspective, Canada and France) reported either an acceptable one-year cost-utility ratio for the addition of viscosupplementation to appropriate care or similar cost and improved effectiveness when hylan was compared with conventional care. It is difficult to assess the representativeness of these results because of the paucity of other trials comparing HA with appropriate or conventional care or reporting the results in terms of the proportion of patients with clinically important improvement. Also, the results should be interpreted in light of the fact that comparisons of HA with placebo have generally shown less than clinically significant treatment effects. Only one economic evaluation (Taiwan) compared viscosupplementation with NSAIDs. The study concluded that celecoxib was more cost-effective than HA as an alternative to naproxen in patients who have declined total knee replacement. That evaluation was a modeling study, and sensitivity analysis showed that the results were dependent on assumptions of cost and effectiveness.

Strength and limitations of the evidence
Cost-effectiveness studies do not typically include placebo control since their intent is to compare the technology of interest with real-world alternatives. The results of cost-effectiveness analyses that express effectiveness in terms of QALYs thus do not demonstrate the absolute efficacy of a new technology, but rather how its balance of benefits and harms compare with those of standard or alternative treatments, some of which also have placebo effects. Cost-effectiveness analyses also reflect the findings or assumptions that some proportion of patients *do* experience clinically important effects, even if mean improvement is not clinically important.

Evidence pertaining to the cost-effectiveness of viscosupplementation has several deficiencies:

- Time frames were short (six months to one year).
- The number of cost analyses and cost-effectiveness studies is very small and estimates of clinical benefit cannot be assessed due to the paucity of comparable data.
- There were no cost data or cost-effectiveness data specific to single-injection treatments, now possible for at least one product (FDA, 2010).
- The full economic evaluations were not conducted in the United States; the results may not apply to the U.S. due to differences in prices, reimbursement policies, standards of care, and definitions of cost-effectiveness limits.

- There was no cost-effectiveness analysis of HA versus intraarticular corticosteroid injection.

Guidelines

The search of the core sources and relevant specialty groups identified six publications from within the past 10 years that addressed viscosupplementation for OA of the knee (AAOS, 2008; ACR, 2000; APS, 2002; NICE, 2008; VA, 2008; Zhang et al., 2007, 2008).

Summary of guidelines and quality assessment
Three guidelines of good quality made varying recommendations with regard to viscosupplementation for OA of the knee (AAOS, 2008; NICE, 2008; Zhang et al., 2007, 2008):

Osteoarthritis Research Society International (OARSI) published a guideline for the management of hip and knee OA (Zhang et al., 2007, 2008). The OARSI guideline is a critical evaluation of existing treatment guidelines (published between 1945 and October 2005) and a systematic review of research evidence from recent studies (up to January 2006). One specific recommendation pertaining to viscosupplementation was issued. OARSI recommends that *injections of intraarticular hyaluronate may be useful in patients with OA of the knee (level of evidence Ia, strength of recommendation 64% on a 100-point VAS)*. The authors note that these injections are characterized by delayed onset, but prolonged duration, of treatment benefit compared with intraarticular injections of corticosteroids.

The American Academy of Orthopaedic Surgeons (AAOS) published a guideline on the treatment for OA of the knee that was rated as good quality (AAOS, 2008). The physician workgroup responsible for development of the guideline used an Agency for Healthcare Research and Quality (AHRQ) technology assessment (Samson et al., 2007) as the evidence base for the recommendation pertaining to the use of intraarticular HA for treatment of OA of the knee. The authors of the guideline concluded that they *could not recommend for or against the use of intraarticular HA as treatment for OA of the knee*. This inconclusive rating was due to conflicting evidence in pooled effects from poor-quality trials relative to higher-quality trials, as well as unclear clinical significance of the results. The AHRQ report did not consider viscosupplementation versus conventional care or cost-effectiveness.

The National Institute for Clinical Health and Excellence (NICE) issued a guideline for the care and management of OA in adults (NICE, 2008). The quality of this guideline was rated as good. The authors note that the evidence suggests that intraarticular hyaluronans may provide a treatment benefit for pain reduction up to three months after a series of three to five injections, but with a generally

small effect size. A limited cost-effectiveness analysis led to the conclusion that *hyaluronans are not within the realm of affordability*. The guidance from NICE states that intraarticular hyaluronan injections are not recommended for the treatment of OA.

Three guidelines of poor quality supported the use of viscosupplementation (ACR, 2000; APS, 2000; VA, 2008). The Pharmacy Benefits Management Service-Medical Advisory Panel (PBM-MAP) division of the Veterans Health Administration issued a Drug Class Review stating that the evidence supports the use of intraarticular HA for OA of the knee, but does not support the use of one product over another (VA, 2008). The American College of Rheumatology (ACR) guideline stated that intraarticular hyaluronan therapy is indicated for use in patients who have not responded to a program of nonpharmacological therapy and simple analgesics (ACR, 2000). The guideline also indicates that intraarticular hyaluronan injections may be especially advantageous for patients in whom nonselective NSAIDs and COX-2–specific inhibitors are contraindicated, or in whom these drugs have been associated with either a lack of efficacy or adverse events. The American Pain Society (APS) has published comprehensive guidelines on pain management in OA, rheumatoid arthritis, and juvenile chronic arthritis (APS, 2000). A small section on HA concludes by stating that HA supplements "may be considered" in persons with OA and knee pain who are unresponsive to acetaminophen, unresponsive to nonselective and COX-2 NSAIDS, or unable to take these medications.

Comparison of guidelines and evidence summary

Of the three high-quality guidelines, one made a weakly positive recommendation in favor of viscosupplementation, one concluded that no recommendation could be made due to the uncertainty in the evidence, and one recommended against viscosupplementation on the basis of a cost-effectiveness analysis that may not be applicable in the United States. The variation in guideline recommendations reflects a body of evidence that provides moderate-quality support of limited efficacy but leaves many important questions unanswered regarding clinical usefulness and cost-effectiveness.

Summary

General conclusions

There is consistent evidence demonstrating that viscosupplementation results in lower mean pain scores and improves mean function scores a few weeks after treatment. However, the magnitude of benefit of HA alone may be too small to be clinically important. (Average change in pain score typically did not meet the threshold of minimal clinical importance, as defined by the OA research community.) There is a much greater volume of evidence regarding impact on pain than on function, and many studies did not follow patients beyond three months. Therefore, the impact of viscosupplementation on the eventual recovery of function is uncertain. Compared with intraarticular corticosteroid injection, viscosupplementation appears to confer longer-lasting benefit,

but the evidence was considered low quality. For comparisons with other treatments, there was insufficient evidence to allow any conclusion. Adverse events occur at a frequency of approximately 2% in single courses of treatment and are primarily transient local reactions; although rare, serious adverse reactions are possible. The rate of adverse events per patient has been shown to increase with repeat courses of treatment, but the only available data were for hylan (high molecular weight HA).

Well-designed, adequately powered RCTs with minimal loss to follow-up are needed to establish the magnitude of benefit that can be expected from viscosupplementation. Such trials should adopt standard approaches to factors such as rescue analgesics, definition of follow-up intervals, type of placebo, and definitions of treatment-related adverse events. To allow a full assessment of clinical relevance, the results should be reported in terms of both mean change and success rates (proportion of patients achieving clinically important improvement according to a standard definition). Future research should not only compare viscosupplementation with other active pharmacological and nonpharmacological treatments, but also provide sufficient detail regarding concomitant treatment to allow an assessment of the synergistic effect of viscosupplementation with other treatments. Trials with preplanned subgroup analyses are needed to determine the patient and disease characteristics that are associated with clinically important benefit. Lastly, studies assessing the overall utility of viscosupplementation to patients could provide perspective for assessing the clinically measured impact on pain and function.

Evidence pertaining to issues other than effectiveness and safety is of low quality:

- Available evidence suggests that viscosupplementation may be as effective as NSAIDs (four RCTs) and results in fewer systemic adverse events (two RCTs); in comparison with intraarticular corticosteroids, it has a delayed onset and longer-lasting benefit (nine RCTs plus meta-analysis).
- Hylan may have a superior benefit compared with that of non–cross-linked HA, but the magnitude of difference is very uncertain and hylan poses a small increase in the risk of adverse events.
- To date, there is no evidence of a difference in benefit between low and medium molecular weight HA.
- Younger age may be associated with greater efficacy; evidence pertaining to effectiveness by other patient characteristics and history is lacking.
- No definitive statement can be made regarding the cost-effectiveness of viscosupplementation.

Limitations of the evidence
- Analyses suggesting that poorer-quality and smaller trials have inflated overall estimates of efficacy.

- Variation in methods across trials and across meta-analyses, which makes comparison of results difficult.
- Limited quantity of data pertaining to treatment effect in terms of proportions of patients who have clinically important improvement.
- No studies comparing viscosupplementation with glucosamine and/or chondroitin.
- No analysis of the synergistic effect of hyaluronans *combined* with specific therapies, and little or no relevant trial detail available in the selected reviews.
- No analysis of actual safety profile for any FDA-approved products other than hylan that is based on large, unbiased databases or registries.
- Few economic evaluations. Their estimates of clinical benefit cannot be directly compared with the conclusions of systematic reviews. The limited evidence that is available is primarily from non-U.S. healthcare systems.
- No controlled trials designed to test whether HA injections avert or delay total knee replacement (TKR), and a paucity of evidence concerning efficacy and safety of different dosing regimens or repeat treatments. These issues were not key questions and thus were not targeted in the search for reviews and primary studies. However, it appears that little evidence on these issues is available.

Table 1. Overview of Systematic Reviews, Key Questions #1 and #2

Reference	Study Design, Search Dates	Studies, Patients	Outcomes and Comparisons Evaluated	Main Findings/Authors' Conclusions	Comments and Quality of Evidence*
Bannuru 2009 See Table 3 for additional details.	SR with MA Searched MEDLINE, EMBASE, CINAHL, BIOSIS, and Cochrane Controlled Trial Register to February 2009. No language restrictions. Conference proceedings, including: American College of Rheumatology, British Society for Rheumatology, Osteoarthritis Research Society International (1990 to February 2009), additional unpublished studies solicited from experts and manufacturers	7 RCTs (n=606 participants, 610 knees; mean age 49-72 yrs; 53%-100% women); all trials were published in full	Studies had to report ≥1 of a hierarchy of outcome measures recommended for OA clinical trials (WOMAC, OA Index Pain Subscale*, knee pain when walking*, knee pain during activities other than walking*, spontaneous joint pain*) *VAS or Likert HA vs corticosteroids	Efficacy/effectiveness: Pooled effect sizes favored corticosteroid up until 3 to 6 wks and significantly favored HA by 11-16 wks. Effect size reached 0.39 at 17-26 wks. Several types of analysis ruled out any influence of covariance between outcomes and time points, trial quality, baseline differences, or type of HA product and corticosteroid. Safety: Not assessed. Authors' conclusions: Corticosteroids are more effective than HA in the short term (up to approximately 4 wks), whereas HA is more effective in the long term (4-26 wks).	*Quality of included RCTs*: No formal assessment tool; 1 trial clearly reported allocation concealment; 3 were open label, 3 single-blind, and 1 double-blind; 5 had industry sponsorship and 2 were unclear; 8%-30% withdrawal rates. 5 trials judged to be of "low quality", 2 of "higher quality." *Quality of SR*: Overall rating fair to good. Rationale for not including Lequesne Index as an outcome measure not explained. Authors were unclear whether all reported outcomes were extracted from each study or only the outcome highest in the hierarchy.
Hayes 2009	SR	19 RCTs reviewed in detail (n=4969 pts with	Pain	Efficacy/effectiveness: HA can provide statistically	A **Hayes Rating** of **B** assigned for single course of treatment.

See Table 2 for details of included RCTs and findings.	1996 – September 2009 (MEDLINE and EMBASE) English language; sample size ≥100; RCTs published only as abstracts were excluded	moderate to severe radiographically confirmed OA)	Clinical function Disease-specific disability/severity (e.g., WOMAC and Lequesne Index) Generic functional status (e.g., SF-36©, ADL) QOL (e.g., Health Utilities Index) Adverse effects All comparisons	significant pain relief and improvement in function compared with placebo, but duration and magnitude of benefit vary (10 RCTs showing positive effect; 4 RCTs showing no effect). HA was more effective than NSAIDs (1 RCT) or was equivalent (1 RCT). Some comparisons with intraarticular corticosteroids reported superiority of HA (3 RCTs); no difference in 1 study; long-term relief was generally considered superior with HA. Safety: No serious HA-related complications reported, but data lacking with respect to repeated treatments. Conclusions: Moderate to strong evidence from placebo-controlled trials suggests that HA can relieve pain and increase functional activity if conservative therapy has failed or cannot be tolerated. An effect comparable with that of corticosteroids suggests that HA could be an option for pts who have failed steroid therapy or in whom steroid therapy is contraindicated. Promising but scarce data suggest that HA has an effect similar to that of conventional exercise.	Quality of included RCTs: Review described individual studies as well designed, but noted frequent deficiencies: relatively high dropout rates; omission of ITT analysis; financial support from manufacturers; confounding by concurrent therapy. Quality of SR: Fair to good Analyses in other reviews suggest that the Hayes exclusion of trials with <100 pts resulted in a body of evidence reflecting smaller, and perhaps more representative, treatment effects.

| Campbell 2007 (not considered as evidence; only as comment on other SRs) | An analysis of MAs | 6 previously published MAs (the 5 study-level MAs reviewed by Samson et al. plus one other) | 6 potential sources of discordance were examined. | Clinical question: Similar across MAs.

Study selection and inclusion: 47 placebo-controlled trials total; 1 common to all; 9 common to 5; 6 common to 4; 31 common to ≤3 MAs. Reasons included search strategies (language, dates, sources), selection criteria, application of selection criteria, inclusion/exclusion of abstracts and unpublished trials, inclusion/exclusion of trials using nonstandard products or regimens.

Data extraction: Differences over whether certain trials' data were adequately reported for inclusion in MA; different baselines (pretreatment, first injection, last injection); various endpoints; change vs final scores; different choices concerning pain measures; data by knees vs by pts.

Assessment of study quality: 4 MAs used different formal tools; sensitivity analysis of certain factors (e.g., ITT analysis) in 1 MA; no specific assessment in 1 MA.

Assessment of ability to combine studies: All

Statistical methods: WMD vs SMD vs MD; 1 MA factored in | MA (Medina 2005) omitted by Samson et al. selected only RCTs reporting WOMAC and Lequesne outcomes; all 7 included RCTs were covered by other MAs; authors concluded that HA may improve function for up to 6 mos.

Using GRADE criteria, Campbell et al. judged the overall evidence to be of moderate quality (downgraded from high due to numerous study differences and inconsistencies in MA conclusions).

No quality assigned to this review; SR criteria do not entirely apply. |

				interval between measurement times; different methods for eliminating or explaining heterogeneity; only 3 MAs explored publication bias and reported different results; 5 statistical packages.	
				Conclusions of Campbell and colleagues: "Probably" use HA.	
Reichenbach 2007 See Table 3 for additional details.	SR with MA of RCTs or quasi-randomized comparator trials (hylan vs HA) with MA plus indirect comparison using results from previous MAs MEDLINE, EMBASE, and CINAHL to November 2006 Unpublished trials from these sources: textbooks, FDA advisory panel proceedings, and additional studies solicited from industry representatives	13 RCTs (n=2085) or quasi-randomized trials published in full (11 RCTS) or as abstracts (2 RCTs); no unpublished trials selected Pt characteristics by trial (mean age 54-71 yrs, median 61; mean duration symptoms 4-7.7 yrs, median 5)	Pain (global, with walking, WOMAC, Lequesne, *or* with activities other than walking, in order of decreasing preference) AEs (flares, effusions, any) Hylan vs HA	Efficacy/effectiveness: Absolute effect had CI near null, did not meet authors' definition of clinical importance, and was characterized by high heterogeneity. Stratified analysis suggested poor quality of some trials inflated the overall effect and was largely responsible for the high heterogeneity. Metaregression of placebo-controlled trials included in other MAs showed inverse association between trial size and effect size. Safety: Robust evidence of an approximately 2-fold increase in risk of local adverse events associated with use of hylan as opposed to HA. Authors' conclusions: (a) Because of lack of clear superior effectiveness over that of HAs, and increased risk of local adverse events, use of	Quality of RCTs: Generally poor quality and/or incomplete reporting. 2 clearly reported allocation concealment; 6 clearly reported pt blinding; 1 clearly reported therapist blinding; 4 were clearly *not* supported by industry. Quality of SR): Good

				hylan should be discouraged in research and clinical practice. (b) Heterogeneity appears to be due more to trial quality than to different HA comparators, as assumed by the Bellamy review. (c) Large effect size in formal, indirect comparison of placebo-controlled trials in comparison with estimate based on direct comparison trials suggests that previous implicit indirect comparisons were misleading.	
Samson 2007 (AHRQ) See Table 3 for results of 6 MAs and other details.	SR of MAs with supplemental analyses (a) MEDLINE (through March 2007) (b) EMBASE (November 2006), and (c) Cochrane Controlled Trials Register (November 2006) Conference proceedings, including: American Association of Orthopedic Surgeons, American College of Rheumatology, and Osteoarthritis Research Society International (2004-2006)	6 MAs (41 trials) plus 1 additional RCT (42 RCTs total; n=5843 pts); mean age 54 yrs, predominately men, early-stage OA 5 of 42 RCTs published as abstracts only 5 MAs analyzed study-level data; 1 MA, pt-level analysis Pt characteristics by trial: mean age 45-72 yrs; 28%-100% women; mean baseline pain (100-point VAS) with movement was 44-79 in hyaluronan arms and 42-80 in placebo arms Additionally, 5 case	Overall review targeted pain, function, AEs, and QOL. Individual MAs analyzed pain (5 MAs), physical function (3 MAs), pt global assessment (1 MA), composite Lequesne Index (1 MA) Placebo comparisons only.	Efficacy/effectiveness: MA authors conclusions: (a) effective at 5 wks and beyond; (b) comparable with other treatments; (c) no proof of clinical effectiveness and possible increased risk of AEs; (d) moderately effective at 5-7 and 8-10 wks; (e) effective and safe but questions remain about differential effectiveness; and (f) small effect but caution about potential publication bias. Safety: AE profiles were not consistent across trials, but when reported, were generally similar in frequency between HA and placebo arms. Most common events included injection site pain, injection site infection, and local joint pain/swelling. MAUDE data suggested rare serious AEs associated with	Quality of RCTs in study-level MAs: Quality ratings for 37 evaluable RCTs: good (9 RCTs), fair (16 RCTs), poor (12 RCTs). ITT results reported in 17 RCTs; 9 RCTs reported ≥20% loss to f/u; double blinding reported by 35 RCTs. Supplemental analysis: (a) Smaller effect sizes associated with higher-quality trials, use of non-hylan vs hylan, and larger sample size (>100). Further analysis added uncertainty to conclusions regarding differences associated with hylan. (b) Positive, underpowered studies more likely than negative studies to be published. 15.5% pts in unreported studies and 9.7% pts in abstracts only. Suggests publication bias in overall body of research. Industry involvement: funding (23 RCTs [55%]), statistical analysis (8 RCTs), coauthor (8 RCTs).

		series (articles or abstracts) and MAUDE (FDA) reviewed for AE data.		Halgan®, Euflexxa®, and hylan.	Quality of MAs:
				Survey of rheumatologists: Pseudoseptic arthritis may not be as rare as thought (no denominator, low survey response rate, not published).	Quality ratings for MAs: Major flaws (3 MAs); minor flaws (2 MAs). Primary flaws were failure to search EMBASE and language restrictions. Impact of language restrictions was minimal; impact of omission of EMBASE was not elucidated). Conclusions considered to be supported: fully (1 MA with negative conclusions), partially (3 MAs), not supported (1 MA). NOTE: No quality assessment of pt-level MA due to lack of a validated instrument; also no assessment of validity of the MA's conclusion.
				Case series, *hylan only*: 2.1% (82/3931) AE rate per injection; 1% (34/3367) for single course, and 8.5% (48/564) for second course.	
				Case series: 5.3% pts (n=4253), most commonly arthropathy; 1 severe AE (large effusion and synovitis); 2-fold increase in subgroup with previous HA treatment.	Industry sponsorship: 30% RCTs (1 MA), 65%-73% RCTs (3 MAs), not reported (1 MA).
				Case series, *hylan only*: 2.7% injections (n=1537) and 8.3% pts (n=336).	Quality of SR: Good
				Subgroup analysis: No evidence of differential effect by age, sex, primary/secondary OA, BMI/weight, or disease severity (see Findings, Key Question #3, for more detail).	
				Conclusions of Samson and colleagues: Evidence does not clearly demonstrate clinical benefit. Variations in the approaches and characteristics of the 5 study-level MAs provide multiple perspectives that permit broad synthesis of evidence.	

| Bellamy 2006 (Cochrane) See Tables 3 and 4 for more detail. | SR of RCTs with MA

MEDLINE through mid-July 2003, EMBASE through week 29 2003, Current Contents to mid-September 2000, and Cochrane Central Register of Controlled Trials

Published conference proceedings through 2005 and additional studies solicited from industry representatives and investigators | 76 blinded RCTs published in full, published as abstracts, or unpublished; 32 RCTs were placebo-controlled; 30 RCTs included in pooled analyses.
Blinding was not a selection criterion.

Studies selected if pts were diagnosed with OA according to ACR criteria, a published algorithm, or detailed clinical or radiographic information.

Pt characteristics not summarized. | Pain

Physical function

Pt global assessment

Joint imaging

(Selected studies had to assess ≥1 of first 3 outcomes.)

All comparisons | Efficacy/effectiveness: Authors concluded that HA is effective, especially at the 5- to 13-wk postinjection period. They note that analyses suggest differential efficacy for different products on different variables and at different time points.

At 5-13 wks, *relative* pain difference ranged from 28% to 54% (favoring HA), and relative function, from 9% to 32% (favoring HA).

Effect size in placebo comparisons was moderate to large for some products on some variables.
In comparison with corticosteroids, HA/hylan may have more prolonged effects.

Safety:
No major safety issues; in some analyses HA/hylan was comparable in efficacy to systemic forms of active intervention with more local reactions but fewer systemic AEs. Almost all AEs were relatively transient. | Quality of included RCTs: Mean quality of RCTs was 3.6 on Jadad scale (1-5), median 3.

This report was an update of an earlier report that had received manufacturer funding.

Quality of SR: Good

NOTE: Although the methodological quality of this review was considered to be good, the narrative synthesis of the large number of pooled estimates was lacking. |

* The methodological quality of systematic reviews was assessed using a modified National Institute for Clinical Health and Excellence (NICE) and Scottish Intercollegiate Guidelines Network (SIGN) tool: Good, Fair, Poor.

Key: ACR, American College of Rheumatology; ADL, activities of daily living; AE, adverse event; FDA, Food and Drug Administration; GRADE, Grading of Recommendations Assessment, Development and Evaluation; HA, hyaluronic acid; ITT, Intention-to-treat; MA, meta-analysis; MAUDE, Manufacturer and User Facility Device Experience; MD, (unstandardized) mean difference; OA, osteoarthritis; pt(s), patient(s); QOL, quality of life; RCT, randomized controlled trial; SMD, standardized mean difference; SR, systematic review; VAS, visual analogue scale; WMD, weighted mean difference; WOMAC, Western Ontario and McMaster Universities (Index)

Table 2. Summary of Narrative Findings in Primary Studies Reviewed by Hayes 2009, Key Questions #1 and #2

NOTES: Abstracted from Table 1 in Hayes 2009 with supplementation of group rates, where effect was positive, from study articles. In almost all studies, a series of HA injections were delivered.

Study Details by Comparator or Analysis		Efficacy and Safety Outcomes
HA vs placebo	Efficacy: 10 RCTs (n=2140 evaluable pts); f/u 1 mo to 1 yr	Conflicting efficacy results
	Positive efficacy results: 6 RCTs (n=1333 evaluable pts); f/u 1 mo to 1 yr	Positive effect on pain, function, and/or QOL (Dougados 1993, Altman 1998, Petrella 2002, Jubb 2003, Neustadt 2005, Kotevoglu 2006; n=78). NOTE: 2 studies by Raynauld may have overlapping pt groups.
	Negative efficacy results: 4 RCTs (n=807 evaluable pts); f/u 20 wks to 1 yr	No effect (Lohmander 1996, Brandt 2001, Karlsson 2002, Lundsgaard 2008), except for marginally better long-term relief when data for 2 different HA arms were pooled (Karlsson 2002), in a subgroup of pts >60 yrs with more severe disease (Lohmander 1996), or on investigator global assessment of pt condition (Lundsgaard 2008). Although between-group differences were nonsignificant, a substantially greater proportion of HA pts (30%) than placebo pts (17%) had ≥7-point improvement in WOMAC score in Brandt 2001 study.
	Safety: 1 RCT (n=255 pts)	Substantially and significantly fewer pts had side effects in HA group, compared with placebo group (Raynauld 2002).
HA vs NSAIDS	Comparative effectiveness: 2 RCTs (n=441 evaluable pts); f/u 3-6 mos	Conflicting results:
	Positive results: 1 RCT (n=333); f/u 6 mos	Substantially better pain outcomes: 48% pts in HA group and 33% in naproxen group (P=0.022) had no or slight pain (Altman 1998).
	Negative results: 1 RCT (n=108 evaluable pts); f/u 3 mos	No difference in symptoms or function (Petrella 2002).
	Safety: 1 RCT (n=333 pts)	Fewer adverse effects associated with HA (Altman 1998).

HA + appropriate care vs appropriate care only	1 RCT (n=255 evaluable pts); f/u 1 yr Comparative effectiveness: Comparative safety:	Positive effect on pain, function, and QOL (Raynauld 2002). 52% adverse event rate in HA arm vs 68% in appropriate care arm (*P*=0.0116)
HA vs conventional care	Comparative effectiveness: 1 RCT (n=506); f/u≤1 yr.	Positive effect on pain, function, and QOL (Kahan 2003)
HA vs exercise	Comparative effectiveness: 1 RCT (n=84 evaluable pts); f/u 6 mos	Positive effect on pain and physical function at 3 mos but loss of effect at 6 mos (Karatosun 2005).
HA vs intraarticular corticosteroid	Comparative effectiveness: 3 RCTs (n=395 evaluable pts); f/u 6-9 mos Positive results: 2 RCTs (n=223 evaluable pts); f/u 26 wks to 6 mo Negative results: 1 RCT (n=100); f/u up to 6 mos	Conflicting results. Positive effect on pain and/or physical function (Frizziero 2002 [at 6 mos although not at 1 mo], Caborn 2004). No effect on pain or function (Leopold 2003).
HA vs other HA or vs hylan	Comparative effectiveness: 3 RCTs (n=1298 evaluable pts); f/u 6 mos to 1 yr Difference detected: 1 RCT (n=392); f/u up to 1 yr No difference: 2 RCTs (n=906); f/u 6 mos to 1 yr Safety: 1 RCT (n=660 pts)	Conflicting results: Hylan G-F superior to HA for pain, physical function but no effect on QOL (Raman 2008). No significant differences between high, medium, and low molecular weight HA (Juni 2007) or between HA and hylan (Karlsson 2002). Trend toward significantly more frequent adverse events in high molecular weight group, compared with medium or low molecular weight groups (Juni 2007).
Single vs multiple courses of treatment	Efficacy: 1 RCT (n=231 evaluable pts); f/u 3 mos Safety: 1 RCT (n=100 evaluable pts)	Analysis by subgroup (first-time and repeat-course of treatment). Treatment effect in first-time subgroup but not in repeat subgroup with respect to use of NSAIDs; treatment effect in both subgroups with respect to all 3 WOMAC scores (Raynauld 2005). In a comparison of RCT data for single-course treatment with chart data for multiple treatments, acute local reactions were significantly greater in multiple-treatment pts (Leopold 2003).

Key: f/u, follow-up; HA, hyaluronic acid; NSAID, nonsteroidal antiinflammatory drug; pt(s), patient(s); QOL, quality of life; RCT, randomized controlled trial; WOMAC, Western Ontario and McMaster Universities (Index)

Table 3. Key Results of Meta-analyses, Key Questions #1 and #2

Details pertaining to individual meta-analyses of HA/hylan versus placebo, including those conducted by Bellamy et al. (2006), were derived primarily from descriptions provided by Samson et al. (2007); a few details were confirmed or supplemented by referring to the original articles. See Table 4 for a description of other analyses conducted by Bellamy et al., i.e., by-product and nonplacebo comparisons. See Table 2 for overviews of Bannuru 2009, Bellamy 2006, Hayes 2009, Reichenbach 2007, and Samson 2007. Estimates in this table with confidence intervals that do not cross the null value are **bolded**. Different outcomes (pain, function, pain/function, and adverse events) have been color coded for easier tracking. Analyses specific to hylan (as opposed to non-cross-linked HA) are denoted by ** **. For most trials, difference at follow-up had to be used as a proxy for difference in change from baseline.

Reference (type of evidence)	Compar-ator	Outcome	No. of Participants (no. of trials)	F/u	Relative ((RR) or Absolute (WMD, SMD, NNT/H) Effect (95% CI)	Comments (Including review authors' quality assessment)
Bannuru 2009 (study-level MA)	Cortico-steroid	Hierarchy of outcome measures (WOMAC, OA Index Pain Subscale*, knee pain when walking*, knee pain during activities other than walking*, spontaneous joint pain*) *VAS or Likert	606 (610 knees) (7 RCTs, published in full)	1-2 wks 3-6 wks 7-10 wks 11-16 wks 17-29 wks	Hedges' g statistic as the effect size for each study; represents score change corrected for small samples. (Hedges' g is a particular formula for calculating standardized differences between means. The article provides a textbook reference but no further explanation.) Pooled g statistics *(positive values favor HA)*: **−0.39 (−0.65 to −0.12)** −0.01 (−0.23 to 0.21) 0.22 (−0.05 to 0.49) **0.35 (0.03 to 0.66)** **0.39 (0.18 to 0.59)**	Very similar rates were observed in multivariate analyses, adjusting for within- and between-study covariance (among outcomes and between time points). The same was true in sensitivity analysis, pooling results only for the 5 trials using ITT analysis, for the 4 trials with blinding, for the 4 trials comparing Hyalgan with methylprednisolone acetate. Metaregression revealed no significant interactions with blinding or ITT status. 2 trials had significant baseline differences, but MA of change scores did not reveal these trials to differ from pooled data.

Reichenbach 2007 (MA)	**Hylan vs HA**		2085 (13 RCTs or quasi-randomized trials, publicized or unpublicized; abstracts or in full)			Unclear whether this was study-level or pt-level.
	Hylan vs HA	Pain (global, with walking, WOMAC, Lequesne, *or* with activities other than walking, in order of decreasing preference)	2085 (13 RCTs)	Data extracted from last f/u or at 6 mos following last injection, whichever came first.	SMD or, if value at last f/u not available, then SMD of change. Negative values favor hylan. **SMD –0.27 (–0.55 to 0.01**); P<0.001; I^2=88% SMD –0.10 (–0.26-0.06; I^2=48% when 2 outlier trials were removed.	RCT effect sizes: CI crossed null (7 trials), favoring hylan (–1.21 to –0.44; 4 trials); favoring HA (0.18 to 0.26; 2 trials). A priori determination of –0.30 as threshold for clinical importance. Funnel-plot analysis and univariate metaregression revealed no association between trial size and treatment effect. Cutoff value in metaregression was 200. Univariate metaregression produced effect sizes near null in trials with adequate allocation concealment (2 RCTs and 46% pts), blinding of pts (6 RCTs and 72% pts), ITT analysis. CIs for these estimates did not include the minimally clinically relevant effect size of –0.30, and statistical heterogeneity was reduced to low-moderate range. However, tests for interaction did not reveal statistically significant effect modification, possibly because of small # of trials. Longer f/u resulted in a clinically important effect size, but test for interaction was again nonsignificant. Metaregression with f/u as continuous variable showed no association. No association between molecular weight (continuous variable in metaregression) and effect size.

	Hylan vs HA	AEs Flares Effusions Any	1067 (4 RCTs) 981 (2 RCTs) 1540 (6 RCTs)		RR 7.27 (0.39-134) RR **2.40 (1.21-4.76)** RR **1.91 (1.04-3.49)** NNH 14 (5 to 324) for 1 additional AE.	Variable definitions and reporting detracts from validity of pooled estimate, but RR measure compensates for between-trial differences. Risk increase observed consistently in individual trials. Low statistical heterogeneity across trials. Unclear basis of NNH calculation.
	Hylan vs HA	Indirect comparison of pain effect	31 trials contributing to 3 MAs (Arrich 2005, Bellamy 2006, Lo 2003)		SMD **–0.64 (–1.25 to –0.02)**; I^2=72%	3 of the 31 trials compared hylan with placebo, were small, and reported large benefits. The other trials compared non-hylan HA with placebo. Metaregression showed a nonsignificant effect size of 0.23 (favoring HAs) in trials ≥200 and a significant effect size of –1.19 (favoring hylan) in trials with <200 pts; test for interaction between effect size and trial size was significant.
Samson 2007 (Study-level MA using data abstracted by Bellamy 2006) **Hylan trials only**	Placebo	VAS pain (weight bearing or WOMAC)	6 RCTs	5-13 wks post-injection	Negative WMD favors HA. All trials: **–20.2 (CI, –29.5 to –10.9); I^2=82%; Egger test NS** 2 trials with larger effects*: **–34 (–37 to –30)** 4 trials with smaller effects*: **–12 (–14 to –10)** *Estimated from forest plot. Note that the 2 CIs do not overlap with CI for all 6.	3 poor-quality trials; 3 fair-quality trails (on basis of 7 defined criteria); 24%-29% dropout rates in 2 trials; 1 trial unblinded, double-blinding in others. The 2 trials with larger effects were pooled with others in 4 of the study-level MAs. Samson and colleagues concluded that the pooled effect for hylan should be considered more uncertain than the CI would suggest.

From Samson 2007 Bellamy 2006 (study-level MA; Cochrane Review)	Placebo or other compar-ator (*only results for placebo compar-ator present-ed here*)	Pain* Physical function* Patient global assessmen t Joint imaging *SMDs calculated when different measures were used.	Not available (32 placebo-controlled RCTs published in full, published as abstracts, or unpublished;)	Varied from last day of injection to 18 mos.	Authors concluded that HA is effective, especially at the 5- to 13-wk postinjection period. No major safety issues, but a review of RCTs are not the best source of AE rates. Negative values for WMD at f/u and SMD at f/u favor HA. Positive values for NNT favor HA; negative values favor placebo.	No language restrictions. 30% all included RCTs had industry sponsorship. Mean RCT quality was 3.5 (Jadad scale, 5 maximum), median 3. No specific analysis of publication bias, but Egger test results consistent with publication bias. Of >850 forest plots, only 38 provide pooled estimates for >3 trials. Analysis by RevMan 4.2.8 software. The following results, except where noted, are presented as comparison 50 in Bellamy 2006. *Quality (according to Samson and colleagues):* Minor flaws; conclusions partially justified by data/analysis. 25% pts and trials were unreported or reported only as abstracts.

	Placebo	Pain at rest (100-point VAS)	9 RCTs	1-4 wks	WMD –3.5 (–9.2 to 2.1); I^2=80%	Comparison 50 in Bellamy 2006 reports –5.37 (–9.90 to –0.85).
	Placebo	Weight-bearing pain (100-point VAS)	20 RCTs 16 RCTs 8 RCTs 3 RCTs	1-4 wks 5-13 wks 14-26 wks 45-52 wks	WMD **–7.7 (–11.3 to –4.1)**; I^2=80% WMD **–13.0 (–17.8 to –8.2)**; I^2=82% WMD **–9.0 (–14.8 to –3.2)**; I^2=77% WMD –2.6 (–7.4 to 2.2); I^2=0	Comparison 50 in Bellamy 2006 reports –11.00 for 5-13 wks (same CI).
	Placebo	WOMAC pain	6 RCTs 6 RCTs 3 RCTs	1-4 wks 5-13 wks 14-26 wks	SMD **–1.2 (–1.9 to –0.5)**; I^2=88% SMD **–1.0 (–1.6 to –0.5)**; I^2=88% SMD **–1.0 (–1.8 to –0.3)**; I^2=80%	
	Placebo	Pain on weight bearing (100-point VAS), **hylan trials**	6 RCTs 5 RCTs 4 RCTs	1-4 wks 5-13 wks 14-26 wks	WMD **–12.54 (–20.39 to –4.69)** WMD **–22.5 (–35.2 to –9.7)**; I^2=83% WMD –20.7 (–35.56 to 5.83)	Presented as Comparison 20 in Bellamy 2006. Samson review reported only for 5-13 wks and provided the I statistic.
	Placebo	WOMAC function	6 RCTs 6 RCTs 3 RCTs	1-4 wks 5-13 wks 14-26 wks	SMD **–1.0 (–1.6 to –0.4)**; I^2=85% SMD **–0.9 (–1.3 to –0.4)**; I^2=84% SMD **–0.8 (–1.4 to –0.2)**; I^2=70%	
	Placebo	Lequesne Index (pain and function, 0-24)	5 RCTs 4 RCTs 3 RCTs 1 RCT	1-4 wks 5-13 wks 14-26 wks 45-52 wks	WMD **–0.8 (–1.4 to –0.2)**; I^2=44% WMD **–1.4 (–2.0 to –0.7)**; I^2=16% WMD **–0.1 (–0.8 to 0.9)**; I^2=6% WMD **–1.1 (–2.7 to 0.5)**; I^2=not applicable	

	Placebo	Patient global improvement	5, 6, 4, and 2 RCTs	1-4, 5-13 14-26, 45-52 wks	RR 1.0-1.1 (NS) at all f/u intervals; I^2=30% (45-52 wks) to 70% (14-26 wks)	
	Placebo	# pts improved	3 RCTs 2 RCTs 2 RCTs	1-4 wks 5-13 wks 14-26 wks	NNT 5, 11, 100 NNT 10, infinity NNT 7.1, 20	Samson et al. note that these calculations are not tied to a definition of clinically important improvement.
	Placebo	# pt clinical failures	1 RCT	14-26 wks 45-52 wks	NNT 11 NNT 6.7	Samson et al. note that these calculations are not tied to a definition of clinically important improvement.
	Placebo	40% relative or 5-point absolute (20-point scale) improvement in WOMAC pain	1 RCT	1-4 wks 5-13 wks 14-26 wks	NNT 14 NNT −33 (favoring placebo) NNT −33 (favoring placebo)	Considered by Samson et al. to be related to a definition of clinically important improvement.
	Placebo	5-point absolute (20-point scale) improvement in WOMAC pain	1 RCT (different trial from previous row)	14-26 wks	NNT 5.9	Considered by Samson et al. to possibly be related to a definition of clinically important improvement.
	Placebo	Improvement global assessment	7 RCTs	1-4, 5-13, 14-26 wks	Generally negative NNT values	Samson et al. note that these calculations are not tied to a definition of clinically important improvement.
	Placebo	Injection site pain	Not available		RR 1.7 (1.19-2.44; *P*=0.004)	No other significant differences in AE occurrence, e.g., discontinuance of study drug or GI complaint, at any f/u interval.

	Placebo	AEs, *non-hylan trials*	5 RCTs		RR 1.6 (0.54-5.6)	
	Placebo	AEs, **hylan** *trials*	5 RCTs		RR 1.9 (0.51-7.3)	

From Samson 2007
 	
Strand 2006
 	
(pt-level MA)
	Placebo (IA saline)	Lequesne Index (pain and function, 0-24)	1155 (5 double-blind RCTs; 3 published and 2 unpublished; selected from 18 trials included in PMA application)	5 and 13 wks (all trials); 9 wks (4); 17, 20, and/or 25 wks (3); integrated analysis tested for effects of treatment over time	Authors concluded that HA is comparable with other treatments, given the magnitude of improvement in HA arms and the significance of HA-saline differences.
					
					Negative estimate favors HA.
					
					Group mean change (treatment vs placebo):
					Fixed-effects model: –2.74 vs –2.16
					Random-effects model: –2.68 vs –2.00
					
					Translates to a *difference* in mean *change* (treatment minus placebo):
					    Fixed-effects model: **–0.58 (–0.95 to –0.20)**
					    Random-effects model: **–0.68 (–0.79 to –0.56)**
						Pooled estimate of difference in change (–0.58 or –0.68) is very small compared with magnitude of the scale (0-24).
						
						Quality: No formal evaluation of quality (lack of validated instrument MAs of pt-level data); no deficiencies noted other than 10% dropout rate in treatment arm and 15% in placebo arm.

	Placebo	AEs	Not available		1.8%, HA; 3.2%, placebo	
From Samson 2007 Arrich 2005 (study-level MA)	Placebo		Sample sizes 38-330 (22 single-/double-blind RCTS, published in full)		Authors concluded that HA has not been proven clinically effective and may be associated with greater risk of AEs. Negative WMD/SMD at f/u favors HA.	RCTs with English or German abstracts included. In general, no evidence of publication bias. *Quality (according to Samson and colleagues):* Minor flaws; conclusions fully supported by data/analysis. No explanation of why some trials, e.g., the 2 hylan trials with large effects (Scale 1994, Wobig 1998), could not be used although other MAs used them.

	Placebo	Pain at rest (100-point VAS)	468 (8 RCTs)	2-6 wks	WMD **–8.7 (–17.2 to –0.2)**; I^2=94%	Pooled estimates for trials that did not use ITT analysis, did not clearly report allocation concealment, or were unblinded showed a greater treatment effect.
	Placebo	Pain during/after exercise (100-point VAS)	941 (9 RCTs)	2-6 wks 10-14 wks 22-30 wks	WMD –3.8 (–9.1 to 1.4); I^2=81% (9 trials) WMD –4.3 (–7.6 to –0.0); I^2=0 (5 trials) **WMD –7.3 (–11.8 to –2.4); I^2=0 (4 trials)**	Omitting an outlier trial in which pain *increased* among those with more advanced disease resulted in a WMD for 2-6 wks of –4.2 and I^2=20%.
	Placebo	Function (multiple measures)	994 (9 RCTs)	2-6 wks 10-14 wks 22-30 wks	SMD 0 (–0.23 to 0.23); I^2=66% (9 trials) SMD –0.11 (–0.31 to 0.09); I^2=59% (7 trials) **SMD –0.16 (–0.16 to –0.13)**; I^2=62% (5 trials)	Pooled estimates for trials that did not clearly report allocation concealment showed a greater treatment effect for first 2 time periods.
	Placebo	AEs	2019 (15)		**RR 1.08 (1.01-1.15)**	AEs were typically minor. Trials reporting AEs were more frequently published than trials not reporting AEs.

From Samson 2007 Modowal 2005 (study-level MA)	Placebo	Pain (100-point VAS)	(9 double-blind RCTs, published in full)	1 wk 5-7 wks 8-12 wks 15-22 wks	Negative WMD in *change* favors HA. WMD –4.4 (–7.2 to –1.1); I^2=92% (9 RCTs) WMD **–17.6 (–28.0 to –7.5)**; I^2=92% (6 RCTs) WMD **–18.1 (–29.9 to –6.3)**; I^2=95% (6 RCTs) WMD –4.4 (–24.1 to 15.3); I^2=94% (3 RCTs) Authors concluded that HA is moderately effective at 5-7 and 8-10 wks.	Trials reporting pain as part of WOMAC score were excluded. 73% of RCTs had industry sponsorship. 4 RCTs were considered low quality (score ≤0.75, maximum 1.0); excluding them lowered pooled estimates considerably. In metaregression, the relationships between trial quality and outcomes varied by f/u interval. No publication bias detected (tendency toward significant Eggers test, *P*=0.096). **Metaregression showed hylan to be associated with significantly better outcomes at 5 wks and beyond.** Few studies relative to the literature; no justification for excluding WOMAC pain as an outcome measure. *Quality (according to Samson and colleagues):* Major flaws; conclusions not supported by data/analysis.

| From Samson 2007 Wang 2004 (study-level MA) | IA injection of placebo | Pain (with or without activities) Function AEs | 1647 knees (pain/function outcomes); 2252 knees (AEs) (20 single-/double-blind RCTs published in full or as abstracts) | Integrated analysis | SPID% and SFID% are overall measures of efficacy that standardize different outcome measures, different evaluation time points, and different trial durations across studies; expresses cumulative response. ASPID% and ASFID% are adjusted for baseline values. Peak PID% and peak FID% reflect maximum efficacy observed in each trial as a percentage of the maximum possible on the scale used). Authors concluded that MA confirmed therapeutic efficacy and safety; additional studies needed to resolve uncertainty regarding differential effect by product, clinical situation, and pt population. Positive estimates favor HA. | English-only. Mean RCT quality score was 19 points (maximum 28); allocation concealment unclear in all; 65% RCTs had industry sponsorship. MA reported no evidence of publication bias (no funnel plot asymmetry, using sample size as ordinate); funnel plots constructed by Samson and colleagues using precision at the ordinate showed asymmetry. No explanation of how the efficacy measures relate to clinical assessment. *Quality (according to Samson and colleagues):* Major flaws; conclusions partially supported by data/analysis. |

	IA injection of placebo	Pain with activities, *non-hylan trials*	Not available (17 RCTs)	Integrated analysis	SPID%, **7.90% (4.10-11.70)**; I^2=84% (17 trials) ASPID%, **13.4% (5.5-21.3)**; I^2=83% (15 trials) Peak PID%, **9.9% (4.8-15.0)**; I^2=91% (16 trials)	Metaregression and/or subgroup analysis showed trial quality, sample size, allowing escape analgesics, and age to have no association with pain outcomes. Evidence of the influence of industry sponsorship was mixed.
	IA injection of placebo	Pain with activities, **hylan** *trials*	Not available (3 RCTs)	Integrated analysis	SPID%, 23.6%; ASPID%, 34.8%; peak PID%, 27.1% (no CIs)	In contrast to non-hylan trials, no heterogeneity. Greater treatment-placebo differences than those reported for non-hylan trials.
	IA injection of placebo	Pain without activities	Not available (10 RCTs)	Integrated analysis	SPID%, **6.0% (0.7 to 11.2)**; significant heterogeneity (10 trials) ASPID%, 11.0% (−3.7 to 25.7); significant heterogeneity (9 trials) Peak PID%, 7.0% (−1.8 to 15.7); significant heterogeneity (9 trials)	Significant heterogeneity for each overall calculation. NOTE: Results for this outcome were not reviewed by Samson and colleagues; data taken from MA article.
	IA injection of placebo	Function (any of multiple measures), *non-hylan trials*	Not available (17 RCTs)	Integrated analysis	SFID%, **5.3% (2.1-8.5)**; no heterogeneity AFPID%, **11.7% (6.3-16.2)** in favor of HA; no heterogeneity Peak FID%, **8.2% (3.8-12.6)** in favor of HA; significant heterogeneity	Metaregression and/or subgroup analysis showed trial quality, sample size, allowing escape analgesics, and age >65 yrs to be associated with *smaller* effect on function.

	IA injection of placebo	Function (any of multiple measures), **hylan** trials	Not available (3 RCTs)	Integrated analysis	SFID%, 21.9%; AFPID%, 38.3%; peak FID%, 26.8% (no CIs)	Greater treatment-placebo differences than those reported for non-hylan trials.
	IA injection of placebo	AEs	2252 knees (20 RCTs)		RR of *minor* AE, 1.2 (1.01-1.41)	*Major* AEs occurred in 3/1002 knees in *non-hylan* trials (severe swelling, vasculitis, hypersensitivity reaction); 1/139 knees (acute painful local reaction) in h*ylan* trials.
From Samson 2007 Lo 2003 (study-level MA)	IA placebo	Pain (global, with walking, WOMAC, Lequesne, *or* with activities other than walking, in order of decreasing preference)	2949 knees, 2927 pts (22 single-/double-blind RCTs published in full or as abstracts)	1-4 mos (prefer-ence given to measure-ment at 2-3 mos)	Authors concluded small effect, but publication bias may overestimate. SMD in change from baseline at: –0.32 (–0.47 to –0.17); significant heterogeneity **SMD diminished to –0.19 (–0.27 to –0.10) with no heterogeneity when 3 RCTs of hylan were excluded. Authors considered 2 (Scale 1994, Wobig 1998) of 3 trials to be outliers.**	Special RCT inclusion criteria: ≥3 injections, dropout <50%. 77% of RCTs had industry sponsorship; 7 reported ITT data, provided raw data for ITT analysis, or had no dropouts; overall dropout rate 12.4%. Evidence of publication bias (funnel plot asymmetry based on sample size of published trials; very small pooled effect in unpublished trials). In a subset of 8 trials that reported change from baseline, the difference in pooled change in each arm suggested that a placebo effect accounted for 79% of the improvement in HA arms (f/u intervals not reported). *Quality (according to Samson and colleagues):* Major flaws; conclusions partially supported by data/analysis.

Key: AE, adverse event (or effect); ASPID/ASFID, adjusted (for baseline pain/function intensity) SPID/SFID; CI, confidence interval; f/u, follow-up; GI, gastrointestinal; HA, hyaluronic acid; I, I index (statistical measure of heterogeneity); IA, intraarticular; ITT, intention-to-treat; MA, meta-analysis; NNT, number needed to treat (harm) (in order for one patient to experience benefit or harm according to related outcome measure and unit); OA, osteoarthritis; PID/FID, pain/function intensity difference; PMA, Premarket Approval (FDA); RCT, randomized controlled trial; RR, relative risk; SMD, standard mean difference; SPID/SFID, sum of pain/function intensity differences; SR, systematic review; WMD, weighted mean difference; WOMAC, Western Ontario and McMaster Universities (Index)

Table 4. Summary of Findings from Bellamy 2006, Key Questions #1 and #2, Nonplacebo Comparisons

NOTES: Except for the trials comparing viscosupplementation with placebo, Bellamy and colleagues did not pool results for different viscosupplementation products. For head-to-head comparator trials, they also did not pool results across trials using different corticosteroid products or different forms of conventional therapy as comparators. Each referenced comparison (total 53) represents a single analysis by the authors. The HA/hylan versus placebo analysis (Comparison 50) is summarized in Table 3. Other comparisons are excluded from the table below for one of these reasons: HA was not the sole difference between treatment arms, outcomes of interest to this review were not reported, or Bellamy and colleagues could not calculate standardized estimates. Standardized adverse event rates were often not available. CIs that do not cross the null value are **bolded**.

Comparator*	Comparison Number Assigned by Bellamy 2006	Outcome	No. of Participants (no. of trials)	F/u	Relative ((RR) or Absolute (WMD) Effect (95% CI) (Negative WMD favors HA/hylan)	Key Findings as summarized by Bellamy 2006
HA or hylan vs exercise, physical therapy, appropriate care, or trigger point injection	9, 16, 17, 25, 26, 29, 27, 28, 39, 40	Various	982 (11 RCTs)	Various	Mostly nonsignificant differences for outcomes of interest. Largest significant effect: WMD –13.20 (–17.02 to –9.38), WOMAC function (100-point VAS) at 36 wks, hylan vs appropriate care	No comment.
HA vs other HA	1, 7, 43	Various	504 (3 RCTs)	Various	No significant differences were observed in available comparator studies for key outcomes.	Paucity of head-to-head comparisons of different HA products warrants caution in drawing conclusions regarding relative value. On the other hand, class estimates may overestimate or underestimate the effect of single products.
Hylan vs HA	3, 5, 18, 30, 41. 42	Various	1012 (10 RCTs)	Various	Mostly nonsignificant differences for outcomes of interest. Largest significant effect: –7.07 ([-13.41, -0.73]), pain at night (100-point VAS) at 1-4 wks RR of 2.91 (0.47 to 17.86) was reported for frequency of AEs (2 RCTs, 146 pts).	

NSAIDs	10, 21, 22, 51 (3 trials included in comparison 51)	Various	891 (6 RCTs; 2 of the 6 RCTs assessed safety only)	Various	Mostly nonsignificant differences for outcomes of interest. Exception: In 3 trials of hylan, significant WMDs in the range of –12 to –19 were observed, but CIs were wide.	In general, similar efficacy compared with NSAIDs. Few AEs; compared with systemic interventions, may result in more local reactions but fewer systemic AEs.

HA or hylan vs intraarticular corticosteroid (9 RCTs).
Larger differences were observed in these comparisons than in others described in this table. Statistically significant differences (bolded) often favored HA or hylan at follow-up intervals exceeding 4 wks. NOTE: In their "Main Findings", the authors refer to 10 trials comparing HA or hylan with corticosteroids, but only 9 trials in seven comparisons could be accounted for.

Methylprednisol one acetate	11 (Hyalgan®)	Spontane-ous pain intensity (100-point VAS)	Not available (3 RCTs)	Not available	**WMD –7.73 (95% CI, –12.81 to –2.64)**	In general, HA appeared to confer longer-term benefits compared with corticosteroid injections.
		Local or systemic reaction			RR 3.0 (0.13 to 71.74)	
6-methylprednisol one acetate	36 (Orthovisc®)	Pain on weight bearing (100-point VAS)	55 (1 RCT)	1-4 wks 5-13 wks 14-26 wks	WMD 5.03 (–4.95 to 15.00) **WMD –15.64 (–24.51 to –6.77)** **WMD –15.40 (–25–91 to –4.89)**	
		Pain at rest (100-point VAS)		1-4 wks 5-13 wks 14-26 wks	WMD 3.53 (–2.09 to 9.15) **WMD –7.70 (–13.50 to –1.90)** WMD –2.90 (–9.47 to 3.67)	
		Pain on walking (100-point VAS)		1-4 wks 5-13 wks 14-26 wks	WMD –0.40 (–11.46 to 10.66) **WMD –18.43 (–29.19 to –7.67)** WMD –14.90 (–25.91 to 3.89)	
		Lequesne Index (0-24)		1-4 wks 5-13 wks 14-26 wks	WMD –0.10 (–0.91 to 0.71) **WMD –1.40 (–2.13 to –0.67)** **WMD –1.14 (–2.16 to –0.12)**	

Triamcinolone hexacetonide	12 (Hyalgan)	Pain on nominal activity	56 (1 RCT)	4 wks 14-26 wks	WMD –0.20 (–17.39 to 16.99) WMD –10.0 (–31.83 to 11.83)
		Pain at rest		4 wks 14-26 wks	WMD –0.70 (–18.17 to 16.77) WMD –20.40 (–43.92 to 3.12)
Triamcinolone hexacetonide	24 (hylan)	WOMAC pain	215 (1 RCT)	5-13 wks 14-26 wks	**WMD –0.40 (–0.65 to –0.15)** **WMD 0.40 (–0.68 to –0.12)**
		WOMAC function		5-13 wks 14 to 26 wks	**WMD –5.0 (–8.86 to –1.14)** **WMD –5.20 (–9.10 to –1.30)**
		WOMAC total score		5-13 wks 14 to 26 wks	**WMD –7.40 (–12.74 to –2.06)** **WMD –7.30 (–12.76 to –1.84)**
Mucopolysaccharide polysulfuric acid ester	13 (Hyalgan)	Pain *change*	59 (1 RCT)	6 wks	WMD 4.0 (0.98 to 7.02)
		Function *change*		6 wks	WMD 0.60 (–1.95 to 3.15)
Betamethasone	23 (hylan)	No assessment of pain, function, or frequency of AEs	# pts not available, 1 RCT		
Betamethasone	35 (Orthovisc®)	WOMAC function	40 (1 RCT)	1-4 wks 5-13 wks	WMD 3.00 (–2.39 to 8.39) **WMD –9.0 (–14.15 to –3.85)**

*In some comparisons, the HA/hylan arm also received the comparison treatment.

Key: AE, adverse event; GI, gastrointestinal; HA, hyaluronic acid; WMD, weighted mean difference; WOMAC, Western Ontario and McMaster Universities (Index); NSAID, nonsteroidal antiinflammatory drug; PT, physical therapy; RCT, randomized controlled trial; RR, relative risk; VAS, visual analog scale

Table 5. Summary of Findings, RCTs

Reference (type of evidence)	Outcomes	No. of Partici-pants	Participant characteristics	Interven-tions	F/u	Main Findings	Quality*	Comments
Altman 2009 (36-site, double-blind, placebo-controlled RCT; superiority trial)	Primary: Difference at 26 wks in LSM change score on 100-point VAS for pain following 50-foot walk Secondary: High response according to OARSI Index (composite), change in WOMAC (100-point scale), # tablets rescue analgesic, change in SF-36 (Acute Form Health Survey), GPA (pain at resting at time of evaluation)	588	OA according to ACR criteria, mean age 61-62 yrs; 63% women; similar prior treatment; moderate baseline pain (mean 55-56 according to primary outcome measure) Exclusions: knee surgery within 12 mos (target knee) or 6 mos (contralateral)	1 injection/wk x 3 wks HA (Euflexxa®, Ferring) (n=291) vs saline (n=295)	12 and 26 wks from baseline	Efficacy: Results apply to evaluation at 26 wks, unless otherwise noted. All effect measures favor HA. 95% CIs in parentheses. Absolute differences are differences in LSM change score. *Primary outcome (100-point VAS, after 50-foot walk):* –6.6 (–10.8 to –2.5; *P*=0.002) Significant differences first noted at 18 wks. *Pts with ≥20-point improvement, same measure:* OR 1.7 (1.2-2.4) *Subgroups:* Significant difference in pts with Kellgren-Lawrence grade 2 OA; no treatment effect in pts with grade 3 *Significant differences in secondary outcomes:* High OARSI response: OR 1.4 (1.0-2.1; *P*=0.047) WOMAC physical function: –4.3 (–7.9 to –0.7; *P*=0.019) in physical function GPA: –4.5 (–8.6 to –0.3; *P*=0.035) SF-36 PCS score: OR 1.609 (0.245-2.973; *P*=0.021 for greater improvement in HA group over time). All effects were NS at 12 wks. Safety: *Treatment-related AEs:* 29 (10%), HA; 32 (11%), saline. Most common AE in both groups was arthralgia. No effusions in HA group.	Good	Stable pharmacological (except for oral NSAIDs) and nonpharmacological therapy could continue during trial if no changes made. Acetaminophen as rescue analgesic. 90% power calculations for sample size. ITT population for efficacy analysis defined as all who received ≥1 injection and had ≥1 postbaseline evaluation. All randomized pts included in safety analysis. 12% overall dropout rate in each group; 2 pts in HA not included in ITT analysis because no postbaseline evaluation. 20-point improvement

						Efficacy/effectiveness		

								in primary outcome represents threshold for clinical importance, according to OARSI standards cited in article. High OARSI responders defined as ≥50% or ≥20-point improvement in pain *or* function *or* any of 3 other defined combinations of relative and absolute improvement.
Baltzer 2009 (Multi-center, randomized comparator trial with placebo control and double blinding)	WOMAC, pain (100-point VAS), QOL (SF-8 HRQL), AEs GPA defined as pt satisfaction (not presented here)	376	Primary OA of knee for ≥3 mos, mean age 54-60 yrs, 55% women, mean baseline pain 66-70 on 100-point VAS Exclusions: Grade IV OA, knee surgery within previous 3 mos	2 injections/wk x 3 wks ACS (n=134) vs HA (HYA-Ject, Ormed) (n=135) vs saline (n=107)	7, 13, and 26 wks from baseline (ITT population) 2 yrs ([a] per-protocol population, i.e., no additional therapy n=188); [b] fully traceable from original ITT population, n=310);	Efficacy/effectiveness: Compared with either HA or saline, ACS had: (1) better WOMAC and VAS outcomes for all evaluation times, e.g., 2.42 overall WOMAC at 26 wks vs 3.75 (HA) and 3.93 (saline) and 29.5 VAS at 26 wks vs 49.3 (HA) and 48.2 (saline); (2) more pts who experienced >50% VAS improvement at 26 wks: 67% ACS pts vs 32% (HA) and 33% (saline); (3) greater improvement in SF-8 HRQL dimensions and component scores. (P<0.001 for each comparison). Significant differences remained between ACS and the HA/saline groups at 2 yrs, both in fully per-protocol population and in population traceable at 2 yrs (analyzed with LOCF). No significant differences in f/u scores between HA and saline groups at 2 yrs. Safety: % pts with local AE: ACS, 23%; HA,	Good	Sample size based on power calculations and an assumption of differences between ACS and HA; ITT analysis; correction for repeated measures (for each outcome measure but not across outcome measures). 3-wk washout period starting with first injection; no NSAIDs during trial; paracetamol as rescue medication. 8.2% dropout rate at 26 wks; 41% at 2 yrs (66 pts lost to f/u and 122 received subsequent therapy).

					new, blinded observer	38%; saline, 28% % pts requiring rescue analgesic: ACS, 17%; HA, 26%; saline, 31%. 3.7% pts in HA group and 2% in saline group withdrew due to acute reaction after an injection; 1 pt in HA group withdrew due to an allergic skin reaction. No pts in ACS group withdrew due to an AE.		
Chou 2009 Ran-domized comparator trial	Pain (100-point VAS), WOMAC (presumably 10-point), Lequesne Index, HSS knee score, cost-effectiveness	37	Mild-moderate, *bilateral* OA	After last injection and then 8, 12, 16, 20, and 26 wks after first injection	Hylan (Synvisc ®; 1/wk x 3) vs HA (Artz®; (1/wk x 3)	Baseline differences very small. Greater improvement in hylan knees in WOMAC pain and in VAS pain. For example, f/u scores at 26 after adjustment for baseline differences were 1.2 vs 1.7 (WOMAC pain, P=0.024) and 45 vs 55 (VAS pain; P=0.003). NS change in WOMAC function, Lequesne Index, HSS score. Synvisc was a cost-effective option from both pt and national payer perspectives.	Poor	
Kawasaki 2009 (Evaluator-blind, randomized comparator trial)	Pain (VAS), Japanese Knee Osteoarthritis Measure, modified OARSI criteria (composite, pain and function)	102	OA of knee, mean age 70 yrs, 100% women	HA (42) vs home exercise (n=45)	24 wks	Only very small, nonsignificant differences between grps for change in all outcome measures. Multiple regression showed that less advanced disease at baseline (according to radiographic joint space) was a positive, independent predictor of better OARSI outcome; age, BMI, treatment group, baseline swelling, and baseline ROM were not significant predictors.	Fair	Responders defined as in Altman 2009.

* The methodological quality of randomized controlled trials was assessed using a modified National Institute for Clinical Health and Excellence (NICE) and Scottish Intercollegiate Guidelines Network (SIGN) tool: Good, Fair, Poor.

Key: ACR, American College of Rheumatology; ACS, autologous conditioned serus; AE, adverse event; BMI, body mass index; CI, confidence interval; f/u, follow-up; GPA, global patient assessment; HA, hyaluronic acid; HSS, Hospital for Special Surgery; ITT, intention-to-treat; LOCF, last observation carried forward; LSM, least squares mean; OA, osteoarthritis; OARSI, Osteoarthritis Research Society International; NSAID, nonsteroidal antiinflammatory drug; pt(s), patient(s); QOL, quality of life; ROM, range of motion; SF-8 HRQL, Short Form-8 Health-Related Quality of Life; VAS, visual analogue scale; WOMAC, Western Ontario and McMaster Universities (Index)

Table 6. Summary of Guidelines

Recommending Body, Year Published	Guideline(s)	Evidence Base	Overall Quality
ACR 2000	Intraarticular hyaluronan therapy is indicated for use in patients who have not responded to a program of nonpharmacological therapy and simple analgesics. Intraarticular hyaluronan injections may be especially advantageous in patients in whom nonselective NSAIDs and COX-2-specific inhibitors are contraindicated, or in whom they have been associated either with a lack of efficacy or with adverse events.	Evidence used in guideline, but no clear methodology provided.	Poor
APS 2002	"The injection of HA supplements into the knee may be considered in persons with OA and knee pain who are unresponsive to acetaminophen, nonselective and COX-2 selective NSAIDS, or who cannot take these medications."	None described.	Poor
AAOS 2008	AAOS concluded that they could not recommend for or against the use of intraarticular hyaluronic acid for patients with mild to moderate symptomatic OA of the knee (level of evidence I and II; grade of recommendation inconclusive).	AHRQ (2007) evidence report served as the basis for this recommendation; the systematic review in the AHRQ (2007) report included 6 meta-analyses (41 RCTs) and 1 additional RCT.	Good
NICE 2008	Intraarticular hyaluronan injections are not recommended for the treatment of OA of the knee, or any other joint.	Evidence from 1 Cochrane systematic review with meta-analysis in patients with OA of the knee (40 RCTs) and 3 additional RCTs was basis for the recommendation.	Good
VA 2008	Evidence supports the use of intraarticular hyaluronan or hylan injections for OA of the knee.	MEDLINE literature search with unclear methodology; 7 systematic reviews with meta-analyses included as evidence.	Poor
Zhang 2008 (OARSI)	Injections of intraarticular hyaluronate may be useful in patients with knee OA (level of evidence Ia; strength of recommendation 64% (95% CI, 43-85). They are characterized by delayed onset, but prolonged duration, of symptomatic benefit when compared with intraarticular injections of corticosteroids.	Systematic search of MEDLINE, EMBASE, CINAHL, AMED, and Science Citation Index identified 9 guidelines and 6 systematic reviews pertaining to viscosupplementation (23 guidelines and 40 studies total for the whole guideline).	Good

Appendix A. Updated Search Strategy

The following search was conducted in PubMed to find systematic reviews published later than the reviews selected from the MED Core Sources.

> Search terms: *viscosupplementation or HA or hyaluronate or hyaluronan or hylan* combined (with AND) with *osteoarthritis or knee*
> Limited to Humans; Publication dates 2006 – 2009
> > Limited to Practice Types: meta-analysis, practice guideline, consensus development conference, NIH, guideline
> > Limited to Journal Groups: systematic review
> > Limited to "systematic review" in Title/Abstract
> > These 3 searches combined with OR.

The following search was conducted in OVID (MEDLINE and EMBASE databases) to find primary studies published later than the search conducted by the latest selected systematic review.

1. (viscosupplementation or HA or hyaluronate or hyaluronan or hylan).mp.
[mp=ti, ab, sh, hw, tn, ot, dm, mf, nm, ui]

2. (osteoarthritis or knee).mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, nm, ui]

3. 1 and 2

4. limit 3 to English language

5. limit 4 to human

6. limit 5 to yr="September 2009—December 2009"

7. limit 6 to humans

Appendix B. Quality Assessment of Guidelines, Systematic Reviews and Primary Studies

Quality of Guidelines

Key Recommendations	ACR 2000	APS 2002	AAOS 2008	NICE 2008	Zhang 2008 (OARSI)	VA, 2008
Ending date of literature search	NR	2001	February 2008	April 2007	January 2006	October 2007
Section 1: Primary Criteria						
Rigor of Development: Evidence	Poor	Poor	Good	Good	Good	Poor
Rigor of Development: Recommendations	Poor	Poor	Good	Good	Good	Poor
Editorial Independence	Good	Poor	Good	Good	Good	Poor
Section 2: Secondary Criteria						
Scope and Purpose	Poor	Poor	Good	Good	Fair	Fair
Stakeholder Involvement	Poor		Fair	Fair	Good	Poor
Clarity and Presentation	Poor	Good	Fair	Fair	Good	Poor
Applicability	Poor	Good	Poor	Poor	Fair	Poor
Section 3: Overall Assessment of the Guideline						
How well done is this guideline?	Poor	Poor	Good	Good	Good	Poor
Other Comments:		*				

*The HA recommendation is made in the context of a comprehensive set of guidelines on pain management for OA, rheumatoid arthritis, and juvenile chronic arthritis. The publication refers to nine systematic reviews for numerous.

(Appendix B continued)
Methodology Checklist: Systematic Reviews and Meta-Analyses (Adapted from NICE)

KEY: Y= Yes   N = No   U = Unclear   N/A = Not applicable

QUALITY ASSESSMENT CRITERIA	Bellamy 2006	Samson 2007	Reichenbach 2007	Bannuru 2009	Hayes 2009
1.1 The study addresses an appropriate and clearly focused question.	N[1]	Y	Y	Y	Y
1.2 An adequate description of the methodology used is included, and the methods used are appropriate to the question.	Y	Y	Y	Y	Y
1.3 The literature search is sufficiently rigorous to identify all the relevant studies.	Y	Y	Y	Y	Y
1.4 The criteria used to select articles for inclusion is appropriate	Y[2]	Y	Y	U	Y
1.5 Study quality is assessed and taken into account.	Y	Y	Y	Y	N[3]
1.6 There are enough similarities between the studies selected to make combining them reasonable.[2]	Y	Y	Y	Y	Y
1.7 There is a conflict of interest statement.	Y[4]	Y	Y	N	Yes (website)
1.8 There is a description of source(s) of funding.	Y	Y	N	N	N/A
2.1 How well was the study done to minimize bias?	Fair-Good	Good	Good	Good	Fair-Good
2.2 If coded as fair or poor, what is the likely direction in which bias might affect the study results?	Unknown				In favor of a treatment effect.[5]
2.2 Are the results of this study directly applicable to the patient group targeted by this topic?	Y	Y	Y		Y

2.4 Other reviewer comments:

[1]The scope of outcomes and comparators considered in the review by Bellamy et al. was very broad; there were no clear clinical questions in mind. The authors' justification for focusing on by-product, by-outcome, by-comparator analysis was to allow practitioners to assess questions of interest to them, but the narrative synthesis of data did not provide sufficient detail on patterns in the results. The review is, however, an excellent source of information.

[2]The reviews that included meta-analysis or reviewed other authors' meta-analyses all reported high heterogeneity, but also attempted to explain the sources of heterogeneity, so this was not seen as a deficiency.

[3]The Hayes review cited the same study weaknesses that were reported by others but did not attempt to rate the quality of studies or assess how study quality might be related to reported results.

[4]The review by Bellamy et al. was an update of a previous Cochrane Review, and the earlier review had industry sponsorship (Genzyme).

[5]Lack of an assessment of patterns suggesting differences in findings according to study quality might have biased conclusions, given the findings of analyses in other reviews. Omission of abstracts and unpublished studies could have subjected the review to publication bias; however, this is uncertain, given other reviewers' mixed results in formal assessments of publication bias in this literature.

(Appendix B continued)
Methodology Checklist: Randomized Controlled Trials (Adapted from NICE)
KEY: Y= Yes   N = No   U = Unclear   N/A = Not applicable

(Table continues on next page.)

QUALITY ASSESSMENT CRITERIA	Altman 2009	Baltzer 2009	Chou 2009	Kawasaki 2009
1.1 An appropriate method of randomization was used to allocate participants to intervention groups.	Y	Y	U	Y
1.2 An adequate concealment method was used such that investigators, clinicians, and participants could not influence enrolment or intervention allocation.	N	Y	U	U
1.3 The intervention and control groups are similar at the start of the trial. (The only difference between groups is the treatment under investigation.)	Y	Y	Y	Y
1.4 Investigators, participants, and clinicians were kept 'blind' about treatment allocation and other important confounding/prognostic factors. If the answer is no, describe any bias that might have occurred.	Y	Y	N (not possible)	N (not possible)
1.5 The intervention and control groups received the same care apart from the intervention(s) studied.	Y	Y	Y	N
1.6 The study had an appropriate length of follow-up.	Y	Y	Y	Y
1.7 All groups were followed up for an equal length of time (or the analysis was adjusted to allow for differences in length of follow-up).	Y	Y	Y	Y
1.8. What percentage of the individuals or clusters recruited into each group of the study dropped out before the study was completed? What percentage did not complete the intervention(s)?	12% each group	8.2%, 26 wks; 41%, 2 yrs	10%	5% w/in 3 mos; 22% after 3 mos
1.9 All the subjects were analyzed in the groups to which they were randomly allocated (often referred to as intention to treat analysis).	Y	Y	n	Y (except 5% dropping out before 3 mos)
1.10 All relevant outcomes are measured in a standard, valid and reliable way.	Y	Y	Y	Y
1.11 The study reported only on surrogate outcomes. (If so, please comment on the strength of the evidence associating the surrogate with the important clinical outcome for this topic.)	N	N	N	N

1.12 The study uses a composite (vs. single) outcome as the primary outcome. If so, please comment on the appropriateness of the composite and whether any single outcome strongly influenced the composite.	N	Not applic able	Y (standard measure used in many OA studies)	Y (standard measure used in many OA studies)
1.13 There is a conflict of interest statement.	N (but one author's affiliation was manufacturer)	Y	N	Y
1.14 There is a description of source(s) of funding.	N* (see above)	Y	N	N
2.1 How well was the study done to minimize bias?	Good	Good	Poor	Fair
2.2 If coded as fair or poor, what is the likely direction in which bias might affect the study results?			Unknown	Unknown
2.3 Are the results of this study directly applicable to the patient group targeted by this topic?	Y	Y	Y	Y
Other reviewer comments: In the Kawasaki trial, the exercise group was not only given specific exercises to perform but encouraged to walk as much as possible during ADLs; would seem more reasonable to have given the advice to walk during ADLs to the HA group as well since this advice is likely given regardless of other treatment.				

Appendix C. Relevant ICD and CPT Codes

CODES	DESCRIPTION
ICD-9	
715	Osteoarthrosis and allied disorders Note: Localized, in the subcategories below, includes bilateral involvement of the same site. Includes: arthritis or polyarthritis: degenerative hypertrophic degenerative joint disease osteoarthritis
715.16	Osteoarthrosis localized primary involving lower leg
715.26	Osteoarthrosis localized secondary involving lower leg
715.36	Osteoarthrosis localized not specified whether primary or secondary involving lower leg
715.96	Osteoarthrosis unspecified whether generalized or localized involving lower leg
717	Internal derangement of knee Includes: degeneration of articular cartilage or meniscus of knee rupture, old of articular cartilage or meniscus of knee tear, old of articular cartilage or meniscus of knee
ICD-10	
M15	Polyarthrosis Includes: arthrosis with mention of more than one site Excludes: bilateral involvement of single joint (M16-M19)
M15.0	Primary generalized (osteo)arthrosis
M15.3	Secondary multiple arthrosis
M15.4	Erosive (osteo)arthrosis
M15.8	Other polyarthrosis
M15.9	Polyarthrosis, unspecified
M17	Gonarthrosis (arthrosis of knee)
M17.0	Primary gonarthrosis, bilateral
M17.1	Other primary gonarthrosis
M17.2	Post-traumatic gonarthrosis, bilateral
M17.3	Other post-traumatic gonarthrosis
M17.4	Other secondary gonarthrosis, bilateral
M17.5	Other secondary gonarthrosis
M17.9	Gonarthrosis, unspecified
M19	Other arthrosis
CPT codes applicable to viscosupplementation	
20610	Arthrocentesis, aspiration, and/or injection; major joint or bursa (e.g. shoulder, hip, knee joint)
CPT codes applicable to total knee replacement (TKR)	
27440	Arthroplasty, knee tibial plateau
27441	Arthroplasty, knee tibial plateau; with debridement and partial synovectomy

27442	Arthroplasty, femoral condyles, or tibial plateau(s) knee
27443	Arthroplasty, femoral condyles, or tibial plateau(s) knee; with debridement and partial synovectomy
27445	Arthroplasty, knee, hinge prosthesis (e.g., Walldius type)
27446	Arthroplasty, knee condyle and plateau; medial or lateral compartment
27437	Arthroplasty, patella; without prosthesis
27438	Arthroplasty, patella; with prosthesis
27447	Arthroplasty, knee condyle and plateau; medial and lateral compartments with or without patella resurfacing (total knee arthroplasty)
HCPCS Level II codes for viscosupplementation	
J7321	Hyaluronan or derivative, Hyalgan or Supartz, for intraarticular injection, per dose
J7323	Hyaluronan or derivative, Euflexxa, for intraarticular injection, per dose
J7324	Hyaluronan or derivative, Orthovisc, for intraarticular injection
J7325	Hyaluronan or derivative, Synvisc or Synvisc-One, for intraarticular injection, 1 mg
HCPCS Level II codes for intraarticular cortisone injection	
J0702	Injection betamethasone acetate 3 mg and betamethasone sodium phosphate, 3 mg
J0704	Injection, betamethasone sodium phosphate per 4 mg
J1020	Injection, methylprednisone acetate, 20 mg
J1030	Injection, methylprednisone acetate, 40 mg
J1040	Injection, methylprednisone acetate, 80 mg
J1094	Injection, dexamethasone acetate, 1 mg
J1100	Injection, dexamethasone sodium phosphate, 1 mg
J1700	Injection, hydrocortisone acetate, up to 25 mg
J1710	Injection, hydrocortisone sodium phosphate, up to 50 mg
J1720	Injection, hydrocortisone sodium succinate, up to 100 mg
J2650	Injection, prednisolone acetate, up to 1 mL
J2920	Injection methylprednisone sodium succinate up to 40 mg
J2930	Injection methylprednisone sodium succinate up to 125 mg
J3302	Injection triamcinolone diacetate, per 5 mg
J3303	Injection triamcinolone hexacetonide, per 5 mg

Appendix D. Placebo-Controlled Trials

NOTE: This table lists the placebo-controlled trials included in the six meta-analyses reviewed in the report by Samson et al. (2007) and in the systematic review by Hayes (2009). (Table continues on next page.)

Trial	Lo 2003 (study-level MA)	Wang 2004 (study-level MA)	Arrich 2005 (study-level MA)	Modowal 2005 (study-level MA)	Bellamy[1] 2006 (study-level MA)	Strand 2006 (patient-level MA)	Hayes 2009 (only trials with n>100)
Shichikawa 1983a					X		
Shichikawa 1983b					X		
Bragantini 1987		X	X		X		
Grecomoro 1987		X	X	X	X		
Dixon 1988	X	X	*		X		
Russell 1992	X		X				
Dougados 1993	X	X	X		X		X
Moreland 1993					X		
Puhl 1993	X	X	X	X	X	X	
Cohen 1994	X	X					
Creamer 1994	X	X			X		
Dahlberg 1994	X		X				
Henderson 1994	X	X	X	X	X		
Scale 1994	X	X	**	X	X		
Carrabba 1995	X	X	X		X		
Corrado 1995	X	X	X		X		
Formiguera 1995	X	X	*		X		
France 1995						X	
Guler 1996					X		

Lohmander 1996	X	X	**	X	X	X	X
U.K. 1996						X	
Wu 1997		X	*		X		
Altman 1998	X	X	X	X	X		X
Dickson 1998[2]		X					
Wobig 1998	X	X	*	X	X		
Hizmetli 1999					X		
Huskisson 1999	X	X	X	X	X		
Brandt 2001	X	X	X		***		X
Bunyaratavej 2001			X		***		
Dickson 2001[2]					X		
Tamir 2001	X	X	*		X		
Karlsson 2002	X		X		X		X
Petralla 2002	X		X	X	X		X
Jubb 2003	X		X		X		X
Pham 2003[3]	X						
Tsai 2003					X		
Altman 2004					***		
Cubukcu 2004					X		
Day 2004			X		X	X	
Pham 2004[3]					X		
Neustadt 2005					***		X
Sezgin 2005					X		
Rolf 2005							
Kotevoglu 2006					X		X (n=78)
Lundsgaard 2008							X

Raman 2008							X
Altman 2009							
Baltzer 2009							

Adapted from Table 6 in: Treatment of primary and secondary osteoarthritis of the knee. AHRQ Evidence Report/Technology Assessment No. 157. AHRQ Publication No. 107-E012. Evidence Report/Technology Assessment, (157), 1-157.

Published as abstracts only.

Unpublished.

Not included in any meta-analysis. One study (Rolf 2005) was identified by Samson et al. in addition to the studies identified by the meta-analyses authors. Other placebo-controlled trials were published later than the search time frame for the latest meta-analysis.

[1] The studies included in the Bellamy 2006 review were the ones used for supplemental analyses by Samson et al.

[2] Published first as abstract and later in full.

[3] Published first as abstract and later in full.

*Included for adverse events, but not in any pooled efficacy result.

**Identified in search, but data "could not be used" for any outcome other than adverse events.

***Included in systematic review, but data not used in a pooled by-class result.

MED REPORT GLOSSARY[1]

Absolute Risk

The probability or chance that a person will have a medical event. Absolute risk is expressed as a percentage. It is the ratio of the number of people who have a medical event divided by all of the people who could have the event because of their medical condition.

For example, research studies have found that among 10,000 people age 75 and over who take a drug like ibuprofen for osteoarthritis pain, 15 of them will die from stomach bleeding. The *absolute risk* of dying from stomach bleeding is 15 out of 10,000, or 0.15 percent of people taking ibuprofen.

Association

A relationship. In research studies, association means that two characteristics (sometimes also called variables or factors) are related so that if one changes, the other changes in a predictable way. An association does not necessarily mean that one variable causes the other.

Bias

Any factor, recognized or not, that distorts the findings of a study. In research studies, bias can influence the observations, results, and conclusions of the study and make them less accurate or believable.

For example, in studies of new drugs, it often was customary to record adverse events only when they occurred in more than 5 or 10 percent of the people taking the drugs. Since information about rare adverse events was not reported, this led to conclusions that the drugs caused fewer side effects than was actually the case. This led to a *bias* in the number of reported side effects.

Blinding (Masking)

A way of making sure that the people involved in a research study — participants, clinicians, or researchers — do not know which participants are assigned to each study group. Blinding usually is used in research studies that compare two or more types of treatment for an illness. Blinding is used to make sure that knowing the type of treatment does not affect a participant's response to the treatment, a health care provider's behavior, or assessment of the treatment effects.

For example, *blinding* is usually done in a type of study known as a randomized controlled trial. The participants are considered "blinded" if they do not know whether

---

[1] Portions of the Glossary were adapted from AHRQ Effective Healthcare Program at
http://effectivehealthcare.ahrq.gov/tools.cfm?tooltype=glossary&report=full

they are taking the drug being researched or a placebo. When neither the participants nor the researchers know who is taking the drug, the study is called "double-blinded."

Clinical Research
The branch of medical science devoted to finding information that improves people's health. It includes research studies that examine the safety and effectiveness of medications, medical devices, diagnostic tests, and treatment regimens intended for human use. Usually, more than one person with the same disease is studied.

For example, the Women's Health Initiative (WHI) is a national *clinical research* study composed of several study components. The WHI includes:

a. A randomized controlled clinical trial of promising but unproven approaches to prevention;
b. An observational study to identify predictors of disease; and
c. A study of community approaches to developing healthful behaviors.

Clinical Trial

A carefully conducted research study that compares the effects of drugs, treatments, or diagnostic tests.

For example, in a randomized controlled *clinical trial* to understand whether calcium tablets work to prevent broken bones in women with low bone density, women with low bone density in one group are randomly assigned to receive calcium and women with low bone density in another group are randomly assigned to the control group and receive a placebo (inactive substance). The numbers of women who suffer fractures in each group are compared to find out whether calcium works.

Cohort Study (Prospective Observational Study)
A clinical research study in which people who presently have a certain condition or receive a particular treatment are followed over time and compared with another group of people who are not affected by the condition.

In a cohort study (also known as a prospective observational study), the researchers take measurements of the people who belong to a cohort at several points in time. The measurements can be symptoms, blood tests, X-rays, or whether the disease has caused the person to die.

For example, the Women's Health Initiative is a *cohort study* that collects information from a group of older women who are followed over several years.

Comparative Effectiveness
A type of health care research that compares the results of one approach for managing a disease to the results of other approaches. Comparative effectiveness usually

compares two or more types of treatment, such as different drugs, for the same disease. Comparative effectiveness also can compare types of surgery or other kinds of medical procedures and tests. The results often are summarized in a systematic review.

The kinds of results that are studied to compare drugs or procedures include relief of symptoms, length of life, or whether people need to go to the hospital. These results are called outcomes. Many other kinds of outcomes can also be compared.

Researchers examined the *comparative effectiveness* of drugs used to treat depression. They examined all the studies about using drugs known as antidepressants. The studies looked at how well people's symptoms improved after taking an antidepressant and also examined the occurrence of side effects. The researchers summarized their findings in a systematic review.

Confidence Interval

The range in which a particular result (such as a laboratory test) is likely to occur for everyone who has a disease. "Likely" usually means 95 percent of the time.

Clinical research studies are conducted on only a certain number of people with a disease rather than all the people who have the disease. The study's results are true for the people who were in the study but not necessarily for everyone who has the disease.

The confidence interval is a statistical estimate of how much the study findings would vary if other different people participated in the study. A confidence interval is defined by two numbers, one lower than the result found in the study and the other higher than the study's result. The size of the confidence interval is the difference between these two numbers.

For example, a study shows that the risk of heart attack from a drug is 3 percent (0.03). The *confidence interval* is shown as "95% CI: 0.015, 0.04." This means that if you conduct this study on 100 different samples of people, the risk of heart attack in 95 of the samples will fall between 1.5 percent and 4 percent. We are 95 percent confident that the true risk is between .015 and .04.

Controlled Clinical Trial

A type of clinical trial comparing the effectiveness of one medication or treatment with the effectiveness of another medication or treatment. In many controlled trials, the other treatment is a placebo (inactive substance) and is considered the "control."

An example of a *controlled clinical trial* is one in which people who took glucosamine were compared with people who did not take glucosamine to determine its effectiveness in relieving pain and improving function for people with osteoarthritis.

Effect Size (Treatment Effect)

The amount of change in a condition or symptom because of a treatment (compared to not receiving the treatment). It is commonly expressed as a risk ratio (relative risk), odds ratio, or difference in risk.

For example, in studies about side effects of antipsychotic drugs, researchers found that these drugs were three times more likely to cause drowsiness than a placebo. The odds ratio was 3.0, meaning the *effect size* was "three times as likely."

Effectiveness
Whether a drug or other treatment works in real life. Effectiveness studies of drugs look at whether they work when they are used the way that most people take them. Effectiveness means that most people who have the disease would improve if they used the treatment.
For example, antidepressant drugs are considered to be *effective* for the treatment of depression. These drugs have been examined in many clinical trials and other types of research studies.

Efficacy
Whether a drug or other treatment works under the best possible conditions. In a research study about efficacy, the study participants are carefully selected, and the researchers can make sure the drug is taken properly and stored properly. The study participants may differ from other people in the general public who have the disease. A treatment that has efficacy under the best conditions may not work as well in a different group of people with the same disease.

For example, a recent clinical trial compared people treated with insulin to people treated with oral medicine for diabetes. Only people with no other medical problems were enrolled in the study, and most were under age 65. The people treated with insulin had better improvement in their blood glucose than the people treated with oral medicines. This study is considered an *efficacy* study, because only younger people without any other health problems were included. Many people who have diabetes are over age 65 and have other problems such as heart disease. It is not known whether the same results would be found in these people.

External Validity
The extent to which clinical research studies apply to broader populations. A research study has external validity if its results can be generalized to the larger population.

For example, researchers analyzed a group of studies to determine the effectiveness of diagnostic tests for breast cancer. In general, the prevalence of breast cancer for women who undergo these tests is 20 percent. The prevalence of breast cancer in most of the studies the researchers analyzed was 50 percent or higher. These high prevalences suggested that the women studied were not typical of the general population. Therefore the studies lacked *external validity*.

Heterogeneity

Differences among research studies. Heterogeneity can apply to either the way the studies were conducted, the methodologies used in the studies, or differences in the way people respond to the treatment. Research reports may describe different types of heterogeneity:

- Statistical Heterogeneity — Differences in the effects of the treatment or intervention.
- Methodological Heterogeneity — Differences in study design.
- Clinical Heterogeneity — Differences in the characteristics of the participants, interventions, or outcome measures.

For example, if three clinical trials of a new drug were performed and only one of the trials found that the drug had efficacy, the results would show *heterogeneity*. Careful review of the three studies would show whether the heterogeneity was probably caused by statistical heterogeneity, methodological heterogeneity, or clinical heterogeneity.

ICER

Incremental Cost Effectiveness Ratio. Calculated by dividing the cost or marginal cost of an intervention by the number of units of outcomes gained (usually QALYs) by providing the intervention. A cost of $50,000 - $100,000 per QALY is a commonly accepted threshold for cost-effectiveness.

I.C.E.R.

The Institute for Clinical and Economic Review. ICER's mission is to lead innovation in comparative effectiveness research through methods that integrate considerations of clinical benefit and economic value. Through a unique collaboration with patients, clinicians, manufacturers, insurers and other healthcare stakeholders, ICER develops tools to support patient decisions and medical policy that share the goal of achieving maximum value for every healthcare dollar.

Internal Validity

The extent to which the results of a clinical research study are not biased. Several characteristics of a study affect its internal validity. Are the two groups of people being compared similar in all the important characteristics that may affect the measurements of data? Are the data collected being measured using accurate methods?

For example, a study was performed comparing people receiving a new type of surgical treatment to people who had received a different treatment in an earlier year. The researchers concluded that the people who received the new treatment had less pain. However, the method used to measure how much pain they had was different for the earlier year than for the new treatment. This study was felt to have poor *internal validity* because of the difference in the method of measuring pain.

Likelihood Ratio

A measure of the accuracy of a diagnostic test. It is used to determine how likely it is that a person has a specific disease based on test results. When the test result is positive, the likelihood ratio is known as a positive likelihood ratio (LR+). When the test result is negative, the likelihood ratio is known as a negative likelihood ratio (LR-). The likelihood ratio is a way of comparing the probability that the test result would occur in people with the disease as opposed to occurring in people without the disease.

A positive likelihood ratio greater than 10 (>10) or a negative likelihood ratio less than 0.1 (<0.1) would be considered clinically useful in helping guide health care decision making.

For example, a diagnostic test called a large core needle biopsy, when used to diagnose breast cancer, has a *positive likelihood ratio* (LR+) of 16.2 and a *negative likelihood ratio* (LR-) of 0.03.

Masking (Blinding)
A way of making sure that the people involved in a research study — participants, clinicians, or researchers — do not know which participants are assigned to each study group. Masking usually is used in research studies that compare two or more types of treatment for an illness. Masking is used to make sure that knowing the type of treatment does not affect a participant's response to the treatment, a health care provider's behavior, or assessment of the treatment effects.

For example, *masking* is usually done in a type of study known as a randomized controlled trial. The participants are considered "blinded" if they do not know whether they are taking the drug being researched or a placebo. When neither the participants nor the researchers know who is taking the drug, the study is called "double-blinded."

Meta-Analysis
A way of combining data from many different research studies. A meta-analysis is a statistical process that combines the findings from individual studies.

For example, researchers wanted to know about the risk of stomach bleeding in people taking aspirin. They did a *meta-analysis* of data from 24 clinical trials with nearly 66,000 participants and found that the risk of stomach bleeding was 2.47 percent with aspirin compared to 1.42 percent with placebo (inactive substance).

Negative Predictive Value
Indicates the likelihood that people with a negative test result would not have a condition.
The higher the value of the negative predictive value (for example, 99 percent would usually be considered a high value), the more useful the test is for predicting that people do not have the condition.

For example, the negative predictive value (PV-) of a normal Prostate Specific Antigen (PSA) test in prostate cancer screening is about 98 percent. It is very unlikely that men with a normal PSA test result on routine screening have the disease.

Number Needed to Harm
The number of people who would need to be treated over a specific period of time before one bad outcome of the treatment will occur. The number needed to harm (NNH) for a treatment can be known only if clinical trials of the treatment have been performed.

For example, a meta-analysis looked at the risk of death for people with Alzheimer's disease who were taking atypical antipsychotic medication. The researchers found that for every 100 people with Alzheimer's disease using atypical antipsychotic medication, there was one additional death compared to 100 people with Alzheimer's disease using placebo (Number Needed to Harm=100).

Number Needed to Treat
The number of people who need to be treated over a specific period of time to promote one additional good outcome (or prevent one additional bad outcome). The number needed to treat (NNT) for a treatment can be known only if clinical trials of the treatment have been performed.

For example, a clinical research study compared two anti-ulcer drugs for treating gastroesophageal reflux disease (GERD). It found that for every 10 people treated with drug A, 9 had relief of their symptoms. For every 10 people treated with drug B, 8 had relief of their symptoms. The study concluded that, if 10 people were switched from drug B to drug A, the number who had relief of symptoms would increase from 8 to 9. This means that one more person would obtain relief for every 10 who had the medicine switched (Number Needed to Treat = 10).

Odds Ratio
The chance of an event occurring in one group compared to the chance of it occurring in another group. The odds ratio (OR) is a measure of effect size and is commonly used to compare results in clinical trials.

For example, a research study compared two groups of women who developed diabetes during their pregnancies. One group was treated with metformin, and the other group was treated with insulin. The researchers recorded how many of the mothers delivered their babies earlier than expected (less than 37 weeks after becoming pregnant). When they calculated the odds of an early delivery, the *odds ratio* (OR) for metformin was 1.06. This means that the women taking metformin had a small increase (1.06 times) in the odds of having an early delivery compared to the women taking insulin.

Outcome

The end result of health care practices. There are many kinds of outcomes. How long people live following a health care treatment is one kind of outcome, known as survival. Other outcomes measure the effects a treatment has on people's lives, such as changes in their ability to function or changes in their quality of life.

Outcomes also include undesirable events such as side effects of drugs. Another type of outcome is whether people needed to change to another kind of treatment.

Researchers studied the *outcomes* of treatment for coronary artery disease. They examined how long people lived after the treatments. They also examined how many people had chest pain or heart attacks after the treatments.

Placebo Controlled Study

A study in which the effect of a drug is compared with the effect of a placebo (an inactive substance designed to resemble the drug). In *placebo controlled clinical trials*, participants receive either the drug being studied or a placebo. The results of the drug and placebo groups are then compared to see if the drug is more effective in treating the condition than the placebo is.

For example, two *placebo controlled studies* have compared the drug risperidone with a placebo for treating behavioral problems in children having the diagnosis of autism. Both studies found that the children given risperidone had fewer behavioral problems than the children given a placebo.

Pooled Odds Ratio

When the data on odds ratios from multiple studies are combined, the result is a pooled odds ratio (POR). An odds ratio (OR) is the comparison of the chance of an event occurring in one group to the chance of it occurring in another group. The odds ratio is a measure of effect size and is commonly used to compare results in clinical trials.

For example, researchers looked at the results of five different studies that compared using a particular drug for treating depression with using a placebo (inactive substance) to treat depression. They looked at the amount of weight gain in the people taking the drug compared to the people taking a placebo. When they calculated the odds of weight gain from each of the studies, the *pooled odds ratio* (POR) for this particular drug was 11.16. This means that the people taking the drug had more than 11 times the odds of gaining weight compared to the people who were taking a placebo.

Positive Predictive Value

Indicates the likelihood that a person with a positive test result would actually have the condition for which the test is used. The higher the value of the positive predictive value (for example, 90 percent would be considered a high value), the more useful the test is for predicting that the person has the condition.

For example, the *positive predictive value* (PV+) of mammography in breast cancer screening has been estimated to be less than 30 percent. For every 100 people who have something discovered on their mammogram that looks like cancer, it will turn out actually to be cancer in less than 30 of those people.

Prevalence

How often or how frequently a disease or condition occurs in a group of people. Prevalence is calculated by dividing the number of people who have the disease or condition by the total number of people in the group.

For example, health leaders in a community were concerned that some women were not receiving health care and therefore did not know they had serious problems. The health leaders began a breast cancer screening program. Women were encouraged to come into a community clinic and have breast examinations. In this example, out of 1,000 women, 20 had breast cancer. Therefore, the *prevalence* of breast cancer in women who undergo screening for breast cancer is 20/1000, or 2 percent.

Prospective Observational Study

A clinical research study in which people who presently have a certain condition or receive a particular treatment are followed over time and compared with another group of people who are not affected by the condition.

In a prospective observational study (also called a cohort study), the researchers take measurements of the people who belong to a cohort at several points in time. The measurements can be symptoms, blood tests, X-rays, or whether the disease has caused the person to die.

For example, the Women's Health Initiative is a *prospective observational study* that collects information from a group of older women who are followed over several years.

Publication Bias

The tendency of researchers to publish experimental findings that have a positive result, while not publishing the findings when the results are negative or inconclusive. The effect of publication bias is that published studies may be misleading. When information that differs from that of the published study is not known, people are able to draw conclusions using only information from the published studies.

For example, two research studies of a new drug are being conducted. One study finds that people with a certain disease improve while taking the drug. The second study finds that people with the same disease do not improve while taking the drug. If the first study is published but the second is not, then *publication bias* has occurred. If both studies are published, then *publication bias* has not occurred.

QALY

Quality Adjusted Life Year: a unit used to measure health status by combining quality of life and survival duration. QALYs are calculated by multiplying the time spent in a particular disease or illness state by the relative desirability of that state (with 1 = benchmark of perfect health, and 0 = death); the result represents the equivalent number of years of full health.

Randomization
A method of assigning participants in clinical trials into two or more groups randomly (by chance). One group receives the treatment or drug being researched, and one group receives either no treatment, a placebo (inactive substance), or another drug. Participants are assigned to a group by various methods.

For example, researchers wanted to use *randomization* in a new study. The researchers decided to flip a coin for each new study participant and assign the person to the first group if the coin is heads and to the second group if the coin is tails. Researchers usually use other techniques than a coin flip. A method called a random numbers table is often used.

Randomized Controlled Trial
A controlled clinical trial that randomly (by chance) assigns participants to two or more groups. There are various methods to randomize study participants to their groups. An example is a *randomized controlled trial* (RCT) to understand whether calcium tablets work to prevent broken bones in women with low bone density. Women with low bone density are randomly assigned to one of two groups. One group receives calcium and the control group receives a placebo (inactive substance). The numbers of women who suffer fractures in each group are compared to find out whether calcium works.

Relative Risk
A comparison of the risk of a particular event for different groups of people. Relative risk (RR) is usually used to estimate exposure to something that could affect health. In a clinical research study, the experimental group is exposed to a particular drug or treatment. The control group is not. The number of events in each group is compared to determine relative risk.

For example, researchers analyzed information from a group of different studies that looked at the number of people who had stomach bleeding when taking the drug ibuprofen. They found that people who took ibuprofen had stomach bleeds 2½ times more often than those who did not take ibuprofen. The *relative risk* of having stomach bleeding while taking ibuprofen was 2.5 (RR 2.5) in these studies.

Risk
A way of expressing the chance that something will happen. It is a measure of the association between exposure to something and what happens (the outcome). Risk is the same as probability, but it usually is used to describe the probability of an adverse

event. It is the rate of events (such as breast cancer) in the total population of people who could have the event (such as women of a certain age).

For example, the *risk* of a woman developing invasive breast cancer at some time in her life is about 1 in 8, or 13 percent.

Risk/Benefit Ratio

A method for comparing a treatment's benefits and risks, such as curing a disease (benefit) versus having a serious side effect from the treatment (risk). The risk/benefit ratio of a treatment is different depending on the disease or condition being treated.

For example, some types of pneumonia often are fatal if not treated but can be cured with antibiotic medications. Antibiotics have a low rate of adverse events. The *risk/benefit ratio* of antibiotic treatment for serious pneumonia is low. This means that the risk of an adverse event is low compared to the probability of improvement from the treatment.

Screening

Using tests or other methods of diagnosis to find out whether or not a person has a specific disease or condition before it causes any symptoms. For many diseases (for example, cancers), starting treatment earlier leads to better results. The purpose of screening is to find the disease so that treatment can be started as early as possible.

For example, a breast exam and a mammogram are both *screening* tests used to find small breast cancers.

Selection Bias

A type of bias caused by an error in the way people are assigned to groups in a clinical research study. This can occur when the study and control groups are chosen so that they differ from each other in ways that may affect the outcome of the study.

For example, a research study compared rates of side effects in men who received surgical removal of the prostate with rates in men who received radiation therapy for prostate cancer. The men in the two groups differed in their age and the rates of other medical problems. Because of this *selection bias*, differences in the side effect rates may not be due just to the effects of the type of treatment.

Sensitivity (True-Positive Rate)

The ability of a test to identify correctly people with a condition. A test with high sensitivity will nearly always be positive for people who have the condition (the test has a low rate of false-negative results). Sensitivity is also known as the true-positive rate.

For example, researchers looked at 10 studies that evaluated the sensitivity of MRI (magnetic resonance imaging) for diagnosing cancer in people who had suspicious breast lumps. Out of every 100 people whose breast lumps were eventually found to be

cancerous, 92 had positive MRI tests. In these studies, the sensitivity of MRI was 92 percent.

Side Effects

Any effects of a drug or treatment that are not wanted. Side effects may be temporary and go away when the drug is stopped. Sometimes they continue for a longer time, even when the drug is no longer being taken.

For example, headache, nausea, hair loss, and skin irritation are *side effects* that commonly occur with drugs.

Specificity (True-Negative Rate)

The ability of a test to identify correctly people without a condition. A test with high specificity will rarely be wrong about who does NOT have the condition (the test has a low rate of false-positive results). Specificity is also known as the true-negative rate.

For example, researchers looked at 10 studies that evaluated the specificity of MRI (magnetic resonance imaging) for diagnosing cancer in people who had suspicious breast lumps. Out of every 100 people whose breast lumps were eventually found not to cancerous, 72 had negative MRI tests. In these studies, the *specificity* of MRI was 72 percent.

Standard Treatment

The treatment or procedure that is most commonly used to treat a disease or condition. In clinical trials, new or experimental treatments sometimes are compared to standard treatments to measure whether the new treatment is better.

For example, the *standard treatment* for anemia (low blood iron) is iron pills.

Statistical Significance

A mathematical technique to measure whether the results of a study are likely to be true. *Statistical significance* is calculated as the probability that an effect observed in a research study is occurring because of chance. Statistical significance is usually expressed as a P-value. The smaller the P-value, the less likely it is that the results are due to chance (and more likely that the results are true). Researchers generally believe the results are probably true if the statistical significance is a P-value less than 0.05 (p<.05).

For example, results from a research study indicated that people who had dementia with agitation had a slightly lower rate of blood pressure problems when they took Drug A compared to when they took Drug B. In the study analysis, these results were not considered to be *statistically significant* because p=0.2. The probability that the results were due to chance was high enough to conclude that the two drugs probably did not differ in causing blood pressure problems.

Systematic Review

A summary of the clinical literature. A systematic review is a critical assessment and evaluation of all research studies that address a particular clinical issue. The researchers use an organized method of locating, assembling, and evaluating a body of literature on a particular topic using a set of specific criteria. A systematic review typically includes a description of the findings of the collection of research studies. The systematic review may also include a quantitative pooling of data, called a meta-analysis.

A comparative effectiveness review is a type of systematic review in which all the available evidence about particular treatments for a disease is reviewed and compared. Scientists collected all the published studies that compared types of treatment for prostate cancer that had not spread beyond the prostate gland. They compiled the results of these studies in a comparative effectiveness review, which is a type of *systematic review*.

Treatment Effect (Effect Size)

The amount of change in a condition or symptom because of a treatment (compared to not receiving the treatment). It is commonly expressed as a risk ratio (relative risk), odds ratio, or difference in risk.

For example, in studies about side effects for antipsychotic drugs, researchers found that these drugs were three times more likely to cause drowsiness than a placebo. The odds ratio was 3.0, meaning the *treatment effect* was "three times as likely."

True-Negative Rate (Specificity)

The ability of a test to identify correctly people without the condition. A test with a high true-negative rate will rarely be wrong about who does NOT have the condition (the test has a low rate of false-positive results). The true-negative rate is also known as specificity .

For example, researchers looked at 10 studies that evaluated the specificity of MRI (magnetic resonance imaging) for diagnosing cancer in people who had suspicious breast lumps. Out of every 100 people whose breast lumps were eventually found not to cancerous, 72 had negative MRI tests. In these studies, the *true-negative rate* of MRI was 72 percent.

True-Positive Rate (Sensitivity)

The ability of a test to identify correctly people with a condition. A test with a high true-positive rate will nearly always be positive for people who have the condition (the test has a low rate of false-negative results). The true-positive rate is also known as sensitivity.

For example, researchers looked at 10 studies that evaluated the sensitivity of MRI (magnetic resonance imaging) for diagnosing cancer in people who had suspicious

breast lumps. Out of every 100 people whose breast lumps were eventually found to be cancerous, 92 had positive MRI tests. In these studies, the *true-positive rate* of MRI was 92 percent.

Validity

Whether a test or technique actually measures what it is intended to measure. Validity can refer to an individual measurement or to the design and approach taken in a clinical research study. When referring to a single measurement, validity means the accuracy of the measurement.

For example, obtaining a blood pressure measurement has *validity* if the blood pressure device works correctly and the person using the device knows how to use it properly.

Variable

Any characteristic that can be measured in different individuals. A variable is also any factor that can affect the outcome of an experiment or study.

Research studies have both independent and dependent *variables*. A dependent *variable* is the change or outcome that results from an independent *variable*. Independent *variables*, such as receiving an experimental treatment, can be changed by the researchers.

For example, a clinical trial found that a nonsteroidal anti-inflammatory (NSAID) drug is better than acetaminophen for osteoarthritis pain relief. In this study, whether a study participant received the NSAID drug or acetaminophen is an independent *variable*, and the amount of pain relief is a dependent *variable*.

Weighted Mean Difference

The mean difference is the average of difference between start and finish values. These differences are then weighted by the variances, for example, number in a study (sample size) or precision of estimate of effect. The purpose is to give more weight to the studies that give more information about the treatment effect and have larger sample sizes.

References

Altman, R. D., Rosen, J. E., Bloch, D. A., Hatoum, H. T., & Korner, P. (2009). A double-blind, randomized, saline-controlled study of the efficacy and safety of EUFLEXXA for treatment of painful osteoarthritis of the knee, with an open-label safety extension (the FLEXX trial). *Seminars in Arthritis and Rheumatism, 39*(1), 1-9.

Altman, R. D., & Moskowitz, R. (1998). Intraarticular sodium hyaluronate (Hyalgan) in the treatment of patients with osteoarthritis of the knee: a randomized clinical trial. Hyalgan Study Group. *The Journal of Rheumatology, 25*(11), 2203-2212.

American Academy of Orthopaedic Surgeons (AAOS). (2008, December 6, 2008). *Clinical Practice Guideline. Treatment of osteoarthritis of the knee (non-arthroplasty).* Retrieved December 18, 2009, from http://www.aaos.org/research/guidelines/OAKguideline.pdf

American College of Rheumatology (ACR). (2000). *Practice Guidelines. Recommendations for the Medical Management of Osteoarthritis of the Hip and Knee*. Retrieved December 16, 2009, from http://www.rheumatology.org/publications/guidelines/oa-mgmt/oa-mgmt.asp?aud=mem

American Pain Society (APS). (2002). *Guideline for the Management of Pain in Osteoarthritis, Rheumatoid Arthritis and Juvenile Chronic Arthritis*. 2nd Edition. Glenview, IL: American Pain Society.

Anandacoomarasamy, A., Bagga, H., Ding, C., Burkhardt, D., Sambrook, P. N., & March, L. M. (2008). Predictors of clinical response to intraarticular Hylan injections -- a prospective study using synovial fluid measures, clinical outcomes, and magnetic resonance imaging. *The Journal of Rheumatology*, 35(4), 685-690.

Arrich, J., Piribauer, F., Mad, P., Schmid, D., Klaushofer, K., & Mullner, M. (2005). Intraarticular hyaluronic acid for the treatment of osteoarthritis of the knee: systematic review and meta-analysis. *Canadian Medical Association Journal, 172*(8), 1039-1043.

Baltzer, A. W., Moser, C., Jansen, S. A., & Krauspe, R. (2009). Autologous conditioned serum (Orthokine) is an effective treatment for knee osteoarthritis. *Osteoarthritis and Cartilage, 17*(2), 152-160.

Bannuru, R. R., Natov, N. S., Obadan, I. E., Price, L. L., Schmid, C. H., & McAlindon, T. E. (2009). Therapeutic trajectory of hyaluronic acid versus corticosteroids in the treatment of knee osteoarthritis: a systematic review and meta-analysis. *Arthritis and Rheumatism, 61*(12), 1704-1711.

Bellamy, N., Campbell, J., Robinson, V., Gee, T., Bourne, R., & Wells, G. (2006, April 19). Viscosupplementation for the treatment of osteoarthritis of the knee. *Cochrane Database of Systematic Reviews, 2006* (2), DOI: 10.1002/14651858. CD005321.pub2. Retrieved December 28, from The Cochrane Library Database.

Brozek, J. L., Akl, E. A., Alonso-Coello, P., Lang, D., Jaeschke, R., Williams, J. W., et al. (2009). Grading quality of evidence and strength of recommendations in clinical practice guidelines. Part 1 of 3. An overview of the GRADE approach and grading quality of evidence about interventions. *Allergy, 64*(5):669-677.

Campbell, J., Bellamy, N., & Gee, T. (2007). Differences between systematic reviews/meta-analyses of hyaluronic acid/hyaluronan/hylan in osteoarthritis of the knee. *Osteoarthritis and Cartilage, 15*(12), 1424-1436.

Chou, C. W., Lue, K. H., Lee, H. S., Lin, R. C., & Lu, K. H. (2009). Hylan G-F 20 has better pain relief and cost-effectiveness than sodium hyaluronate in treating early osteoarthritic knees in Taiwan. *Journal of the Formosan Medical Association, 108*(8), 663-672.

Food and Drug Administration (FDA). (2009). Center for Devices and Radiological Health (CDRH). *Premarket Approval (PMA) Database* [search: P940015]. Retrieved April 12, 2010 from http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfTopic/pma/pma.cfm?num=P940015S012.

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., et al. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal, 336*(7650):924-926.

Hayes, Inc. (2009). *Hayes Medical Technology Directory*. *Sodium hyaluronate for osteoarthritis*. Lansdale, PA: Hayes, Inc.

Jacobs, N., Kane, T., & Clarke, H. (2008). Cost of outpatient viscosupplementation. *British Journal of Healthcare Management, 14*(12), 535-538.

Kahan, A., Lleu, P. L., & Salin, L. (2003). Prospective randomized study comparing the medicoeconomic benefits of Hylan GF-20 vs. conventional treatment in knee osteoarthritis. *Joint, Bone, Spine: Revue du Rhumatisme, 70*(4), 276-281.

Kawasaki, T., Kurosawa, H., Ikeda, H., Takazawa, Y., Ishijima, M., Kubota, M., et al. (2009). Therapeutic home exercise versus intraarticular hyaluronate injection for osteoarthritis of the knee: 6-month prospective randomized open-labeled trial. *Journal of Orthopaedic Science, 14*(2), 182-191.

Kellgren, J. H. & Lawrence, J. S. (1957). Radiological assessment of osteo-arthrosis. *Annals of the Rheumatic Diseases, 16*(4), 494-502.

Lo, G. H., LaValley, M., McAlindon, T., & Felson, D. T. (2003). Intraarticular hyaluronic acid in treatment of knee osteoarthritis: a meta-analysis. *Journal of the American Medical Association, 290*(23), 3115-3121.

Modawal, A., Ferrer, M., Choi, H. K., & Castle, J. A. (2005). Hyaluronic acid injections relieve knee pain. *The Journal of Family Practice, 54*(9), 758-767.

National Institute for Health and Clinical Excellence (NICE). (2008). *National Collaborating Centre for Chronic Conditions. Osteoarthritis: National Clinical Guideline for Care and Management in Adults*. Retrieved December 16, 2009, from http://www.nice.org.uk/nicemedia/pdf/CG059FullGuideline.pdf

Neustadt, D., Caldwell, J., Bell, M., Wade, J., & Gimbel, J. (2005). Clinical effects of intraarticular injection of high molecular weight hyaluronan (Orthovisc) in osteoarthritis of the knee: a randomized, controlled, multicenter trial. *The Journal of Rheumatology, 32*(10), 1928-1936.

Raynauld, J. P., Torrance, G. W., Band, P. A., Goldsmith, C. H., Tugwell, P., Walker, V., et al. (2002). A prospective, randomized, pragmatic, health outcomes trial evaluating the incorporation of hylan G-F 20 into the treatment paradigm for

patients with knee osteoarthritis (Part 1 of 2): clinical results. *Osteoarthritis and Cartilage, 10*(7), 506-517.

Reichenbach, S., Blank, S., Rutjes, A. W., Shang, A., King, E. A., Dieppe, P. A., et al. (2007). Hylan versus hyaluronic acid for osteoarthritis of the knee: a systematic review and meta-analysis. *Arthritis and Rheumatism, 57*(8), 1410-1418.

Rolf, C. G.., Engstrom, B., Ohrvik, J., Valentin, A., Lilja, V., & Levine, D. W. (2005). A comparative study of the efficacy and safety of hyaluronan viscosupplements and placebo in patients with symptomatic and arthroscopy-verified cartilage pathology. *Journal of Clinical Research, 8*, 15-32.

Samson, D. J., Grant, M. D., Ratko, T. A., Bonnell, C. J., Ziegler, K. M., & Aronson, N. (2007). Treatment of primary and secondary osteoarthritis of the knee. AHRQ Evidence Report/Technology Assessment No. 157. AHRQ Publication No. 107-E012. *Evidence Report/Technology Assessment,* (157), 1-157.

Strand, V., Conaghan, P. G., Lohmander, L. S., Koutsoukos, A. D., Hurley, F. L., Bird, H., et al. (2006). An integrated analysis of five double-blind, randomized controlled trials evaluating the safety and efficacy of a hyaluronan product for intraarticular injection in osteoarthritis of the knee. *Osteoarthritis and Cartilage, 14*(9), 859-866.

Torrance, G. W., Raynauld, J. P., Walker, V., Goldsmith, C. H., Bellamy, N., Band, P. A., et al. (2002). A prospective, randomized, pragmatic, health outcomes trial evaluating the incorporation of hylan G-F 20 into the treatment paradigm for patients with knee osteoarthritis (Part 2 of 2): economic results. *Osteoarthritis and Cartilage, 10*(7), 518-527.

Turajane, T., Labpiboonpong, V., & Maungsiri, S. (2007). Cost analysis of intraarticular sodium hyaluronate treatment in knee osteoarthritis patients who failed conservative treatment. *Journal of the Medical Association of Thailand, 90*(9), 1839-1844.

Turajane, T., Wongbunnak, R., Patcharatrakul, T., Ratansumawong, K., Poigampetch, Y., & Songpatanasilp, T. (2009). Gastrointestinal and cardiovascular risk of non-selective NSAIDs and COX-2 inhibitors in elderly patients with knee osteoarthritis. *Journal of the Medical Association of Thailand, 92*(Suppl 6), S19-S26.

Veterans Administration (VA). (2008). *Clinical Guidance. Drug Class Reviews. VISCOSUPPLEMENTATION for Osteoarthritis of the Knee: Intraarticular Administration of Hyaluronan (Hyaluronic Acid) and Hylan G-F 20 Products.* Retrieved December 31, 2009, from http://www.pbm.va.gov/Clinical%20Guidance/Drug%20Class%20Reviews/Hyaluronic%20acid-Hylan%20G-F%2020%20for%20Osteoarthritis%20of%20the%20Knee-Final%20Review%20(May%202008).doc

Waddell, D. D. (2007). Viscosupplementation with hyaluronans for osteoarthritis of the knee: clinical efficacy and economic implications. *Drugs and Aging, 24*(8), 629-642.

Wang, C. T., Lin, J., Chang, C. J., Lin, Y. T., & Hou, S. M. (2004). Therapeutic effects of hyaluronic acid on osteoarthritis of the knee. A meta-analysis of randomized

controlled trials. *The Journal of Bone and Joint Surgery, American Volume, 86-A*(3), 538-545.

Yen, Z. S., Lai, M. S., Wang, C. T., Chen, L. S., Chen, S. C., Chen, W. J., et al. (2004). Cost-effectiveness of treatment strategies for osteoarthritis of the knee in Taiwan. *The Journal of Rheumatology, 31*(9), 1797-1803.

Zhang, W., Moskowitz, R. W., Nuki, G., Abramson, S., Altman, R. D., Arden, N., et al. (2007). OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. *Osteoarthritis and Cartilage, 15*(9), 981-1000.

Zhang, W., Moskowitz, R. W., Nuki, G., Abramson, S., Altman, R. D., Arden, N., et al. (2008). OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis and Cartilage, 16*(2), 137-162.