Osteoarthritis and Cartilage (2007) **15,** 1424–1436

doi:10.1016/j.joca.2007.01.022

Osteoarthritis and Cartilage

International Cartilage Repair Society

OARSI OSTEOARTHRITIS RESEARCH SOCIETY INTERNATIONAL

Differences between systematic reviews/meta-analyses of hyaluronic acid/hyaluronan/hylan in osteoarthritis of the knee

J. Campbell B.A., Research Associate†, N. Bellamy M.D., D.Sc., Professor and Director‡* and T. Gee Ph.D., Biostatistician‡

† *Division of Orthopaedic Surgery, The University of Western Ontario, London, Ontario, Canada*

‡ *Centre of National Research on Disability and Rehabilitation Medicine (CONROD), The University of Queensland, Brisbane, Queensland, Australia*

Summary

Objective: To explore reasons for discrepant results between systematic reviews (SR)/meta-analyses (MA) of the efficacy and safety of hyaluronic acid/hyaluronan/hylan (HA) therapy in the treatment of osteoarthritis (OA) of the knee.

Methods: A decision algorithm was utilised to identify reasons for discordance among six SR. Sources of discordance such as clinical question, trial selection and inclusion, data extraction, assessment of study quality, assessment of the ability to combine trials, and statistical methods for data synthesis were examined.

Results: A similar question was asked in all six SR. Different trials were selected for inclusion in the reviews mainly because of differences in the search strategies and selection criteria. Although similar methods for data extraction were utilised, differences were found both in the outcome measures and time-points selected for extraction. Methodological quality was not always formally assessed. Different statistical methods for data synthesis resulted in conflicting estimates of therapeutic effect.

Conclusions: Reasons for the inconsistency of results reported in the six SR were identified. Using the principles of the GRADE approach for estimating the therapuetic effect of HA in the treatment of OA of the knee, there is moderate evidence suggesting that further research is unlikely to change our confidence in the estimate of the effect. In the balance of benefit to harm, the trade-off is probable benefit with respect to pain reduction and physical function improvement with low risk of harm.

Key words: Osteoarthritis, Knee, Systematic review, Meta-analysis, Hyaluronic acid, Hyaluronan, Hylan.

Introduction

Level I evidence from systematic reviews (SR) is widely regarded as high level evidence likely to influence evidence-based decision-making in routine clinical care[1]. SR synthesize the results of all relevant studies by using strategies that limit bias and random error[2]. As the number of SR addressing the same therapeutic question increases, the possibility for conflict among reviews increases. The extent to which different methodological approaches and analytic strategies may influence the outcome of meta-analyses in SR has been previously reported[3]. The authors proposed a decision algorithm to help decision-makers understand discordance among reviews[3]. This algorithm has been utilised to explore reasons for discrepant results of SR in complementary medicine and in *Helicobacter pylori* eradication therapy in non-ulcer dyspepsia[4,5].

The current debate over the efficacy and safety of HA in the treatment in knee OA has not been resolved by the publication of several discordant SR. Six SR have been published within the last four years with differing conclusions (Table I)[6–11]. Although other SR and critical appraisals of HA have been reported, these were not assessed in this publication because they did not include a meta-analysis.

We compared the methodology used in the six SR and examined potential explanations for discordance in order to compare the strength of the evidence for the therapeutic efficacy of HA compared to placebo in the treatment of OA of the knee.

Method

The decision algorithm proposed by Jadad *et al.* to help decision-makers select between discordant reviews was applied to six reported SR[3]. Potential sources of discordance identified to be examined were: the clinical question, study selection and inclusion, data extraction, assessment of study quality, assessment of the ability to combine studies, and the statistical methods for data synthesis. In order to grade the quality of the evidence and make a recommendation about the therapeutic effect of HA compared to placebo in the treatment of knee OA, the four elements of the GRADE (Grades of Recommendation Assessment, Development and Evaluation) system, study design, study quality (high, moderate, low, very low), consistency, and directness were assessed to answer the question, "Should HA be used for the treatment of knee OA?"[12].

*Address correspondence and reprint requests to: Nicholas Bellamy, M.D., D.Sc., Faculty of Health Sciences (CONROD), The University of Queensland, Level 3, Mayne Medical School, Herston Road, Brisbane, Queensland 4006, Australia. Tel: 61-7-3365-5558; Fax: 61-7-3346-4603; E-mail: n.bellamy@uq.edu.au

Received 18 August 2006; revision accepted 28 January 2007.

Table I
What were the results of the six reviews?

Arrich *et al.*[6]	"According to the currently available evidence, intra-articular HA has not been proven clinically effective and may be associated with a greater risk of adverse events." WMD -3.8, -4.3, -7.1 and -0.5 for pain during movement from 2-6, 10-14, 22-30 and 44-60 weeks, respectively.
Bellamy *et al.*[7]	"The analyses support the contention that the HA class of products is superior to placebo. There is considerable between-product, between-variable and time-dependent variability in the clinical response. …within the constraints of the trial designs employed, no major safety issues were detected. Overall, the…analyses support the use of the HA class of products in the treatment of knee OA." WMD -8, -13, -9 and -3 for pain on weight bearing from 1-4, 5-13, 14-26 and 45-52 weeks, respectively.
Lo *et al.*[8]	"Intra-articular HA has a small effect when compared with an intra-articular placbo. The presence of publication bias suggests even this effect may be overestimated." Pooled effect size 0.32 (= nonsteroidal anti-inflammatory drugs over acetaminophen).
Medina *et al.*[9]	"HA injection may provide short-term relief of pain and improved functionality for patients with OA of the knee, but benefits do not last beyond 6 months."
Modawal *et al.*[10]	"Intra-articular viscosupplementation was moderately effective in relieving knee pain in patients with OA at 5 to 7 and 8 to 10 weeks after the last injection but not at 15 to 22 weeks." WMD 4.4, 17.7, 18.1 and 4.4 for pain VAS from 1, 5-7, 8-12 and 15-22 weeks, respectively.
Wang *et al.*[11]	"This meta-analysis confirmed the therapeutic efficacy and safety of intra-articular injection of HA for the treatment of OA of the knee." Adjusted sum of the pain intensity differences percentage 13.4.

WMD: weighted mean difference; HA: hyaluronic acid/hylaluronan/hylan; OA: osteoarthritis.

Table II
Do the reviews ask the same question?

Arrich *et al.*[6]	"We performed a systematic review and meta-analysis of randomized controlled trials to assess the effectiveness of intra-articular HA for the treatment of OA of the knee."
Bellamy *et al.*[7]	"To assess the effects of viscosupplementation in the treatment of OA of the knee."
Lo *et al.*[8]	"To evaluate whether intra-articular HA is efficacious in treating knee OA."
Medina *et al.*[9]	"Should your patient opt for HA injection?"
Modawal *et al.*[10]	"To evaluate the efficacy of intra-articular viscosupplementation therapy with HA for pain relief of knee OA, we conducted a meta-analysis of randomized, double-blind, placebo-controlled trials."
Wang *et al.*[11]	"To elucidate the therapeutic efficacy and safety of intra-articular injection of HA in the treatment of OA of the knee by conducting a meta-analysis of randomized controlled trials."

HA: hyaluronic acid; OA: osteoarthritis.

Results

The first step in the algorithm is to determine if the reviews asked the same clinical question. It appears that all six reviews asked a similar question (Table II). Since the reviews addressed the same question, the next step was to establish whether the reviews included the same primary trials[13–63] (Table III). One of the 47 trials was common to all six reviews: Huskisson 1999[36]. Nine trials were common to five reviews: Altman 1998[14], St. J. Dixon 1988[58], Dougados 1993[27], Henderson 1994[34], Lohmander 1996[46], Petrella 2002[49], Puhl 1993[52], Scale 1994[54], and Wobig 1998[61]. Six trials were common to four reviews: Brandt 2001[17], Carrabba 1995[19], Grecomoro 1987[30], La Sala 1995[28], Tamir 2001[59], and Wu 1997[63]. Arrich *et al.* identified 24 randomized controlled trials (RCT) and completed analyses on 22 RCT (excluding Lohmander 1996[46] and Scale 1994[54])[6]. Bellamy *et al.* identified 76 RCT of which 40 were placebo-controlled trials[7]. Although these reviewers reported results by single HA product, a class-based analysis was completed in which 22 RCT for the pain on weight-bearing outcome measure were included. Lo *et al.* identified 22 RCT and completed an intention-to-treat meta-analysis on seven RCT (Carrabba 1995[19], Creamer 1994[22], Dahlberg 1994[24], Jubb 2003[40], La Sala 1995[28], Pham 2003[50], and Wobig 1998[61])[8]. Medina *et al.* identified 35 potentially eligible studies and selected seven for inclusion in their meta-analysis (Altman 2004[15], Day 2004[25], Dougados 1993[27], Huskisson 1999[36], Karlsson 2002[41], Petrella 2002[49], and Pham 2004[51]).[9] Modawal *et al.* identified 17 RCT and completed analysis on nine RCT (Altman 1998[14], Grecomoro 1987[30], Henderson 1994[34], Huskisson 1999[36], Lohmander 1996[46], Petrella 2002[49], Puhl 1993[52], Scale 1994[54], and Wobig 1998[61])[10]. Wang *et al.* identified 25 RCT and analysed 20 (excluding abstracts which did not provide quantitative data: Isdale 1993[37], Jubb 2001[39], Karlsson 1999[41], Moreland 1993[47], and Russell 1992[53])[11].

Differences were detected due to different strategies to search the literature, different criteria for selecting trials for inclusion (for example, publication status and language of publication), and application of the selection criteria (Table IV). Different electronic databases were searched to identify trials. Search dates varied but this was mainly attributable to publication dates of the reviews. Hand-searching of specialized journals, journal supplements, and proceedings of conferences was utilised in four of the reviews[7,8,10,11]; and of reference lists of retrieved studies in the Medina *et al.* review[9], while not utilised in the Arrich *et al.* review[6]. Some review authors attempted to contact authors of trials or industry for unpublished data and/or manuscripts. Three review groups restricted trials to those published only in the English language[9-11]. Bellamy *et al.* included seven published RCT which were not included in any of the other reviews[7]: Cubukcu 2005[23], Kotevoglu 2006[43], Neustadt 2005[48], Sezgin 2005[55], Shichikawa 1983[56], Shichikawa1983a[57], and Wobig 1999[62]. The Shichikawa trials were originally published in Japanese but an English translation was available. Abstracts were included in three reviews[7,8,11] but not in the other three[6,9,10]. Bellamy *et al.* included three abstracts (Groppa 2001[32], Guler 1996[33], and Tsai 2003[60]) that were not included in any of the other five reviews[7]. Unpublished manuscripts and data from unpublished manuscripts were included in some reviews but not in others. Arrich *et al.*[6] included an unpublished article from Russell 1992[53]. It was classified as a high quality trial (i.e. reported blinding, allocation concealment and intention-to-treat analysis). Bellamy *et al.* included three unpublished manuscripts (Hizmetli 1999[35], Lin 2004[45], and Moreland 1993[47]) the first two which had been submitted for publication[7]. Lo *et al.* received data for