Accepted Manuscript

Title: Serum Cartilage Oligomeric Matrix Protein (sCOMP) is Elevated in Patients with Knee Osteoarthritis: A Systematic Review and Meta-Analysis

Authors: Johanna M. Hoch, MS, ATC Carl G. Mattacola, PhD, ATC, FNATA Jennifer M. Medina McKeon, PhD, ATC, CSCS Jennifer S. Howard, MS, ATC Christian Lattermann, MD

PII: S1063-4584(11)00281-0
DOI: 10.1016/j.joca.2011.09.005
Reference: YJOCA 2513

To appear in: Osteoarthritis and Cartilage

Received Date: 11 March 2011
Accepted Date: 23 September 2011

Please cite this article as: Hoch JM, Mattacola CG, Medina McKeon JM, Howard JS, Lattermann C. Serum Cartilage Oligomeric Matrix Protein (sCOMP) is Elevated in Patients with Knee Osteoarthritis: A Systematic Review and Meta-Analysis, Osteoarthritis and Cartilage (2011), doi: 10.1016/j.joca.2011.09.005

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Serum Cartilage Oligomeric Matrix Protein (sCOMP) is Elevated in Patients with Knee Osteoarthritis: A Systematic Review and Meta-Analysis

Corresponding Author:
Johanna M. Hoch, MS, ATC
Division of Athletic Training & Rehabilitation Sciences Doctoral Program
University of Kentucky
College of Health Sciences
Wethington Building
900 South Limestone
Lexington, KY 40536-0200
859-323-1100 x80839
johanna.clark@uky.edu

Carl G. Mattacola, PhD, ATC, FNATA
Division of Athletic Training & Rehabilitation Sciences Doctoral Program
University of Kentucky
College of Health Sciences
210E Wethington Building
900 South Limestone
Lexington, KY 40536-0200
859-323-1100 x80860
carlmattacola@uky.edu

Jennifer M. Medina McKeon, PhD, ATC, CSCS
Division of Athletic Training & Rehabilitation Sciences Doctoral Program
University of Kentucky
College of Health Sciences
206A Wethington Building
900 South Limestone
Lexington, KY 40536-0200
859-323-1100 x80594
Jmedi2@email.uky.edu

Jennifer S. Howard, MS, ATC
Division of Athletic Training & Rehabilitation Sciences Doctoral Program
University of Kentucky
College of Health Sciences
Wethington Building
900 South Limestone
Lexington, KY 40536-0200
859-323-1100 x80839
Christian Lattermann, MD
Assistant Professor of Orthopaedics and Sports Medicine
Director: Center for Cartilage Repair and Restoration
University of Kentucky
College of Medicine
Lexington, KY

SERUM COMP AND OSTEOARTHITIS SYSTEMATIC REVIEW
ABSTRACT

Objective: To be used in diagnostic studies, it must be demonstrated that biomarkers can differentiate between diseased and non-diseased patients. Therefore, the purpose of this study was to answer the following questions: (1) Is serum cartilage oligomeric matrix protein (sCOMP) elevated in patients with radiographically diagnosed knee osteoarthritis (OA) compared to controls? (2) Are there differences in sCOMP levels when comparing differing radiographic OA severities to controls? Methods: Systematic review and meta-analysis. Data Sources: A systematic search of CINAHL, PEDro, Medline, and Sports Discus was completed in March 2010. Keywords: knee, osteoarthritis, sCOMP, radiography. Study Inclusion Criteria: Studies were written in English, compared healthy adults with knee OA patients, used the Kellgren Lawrence (K/L) classification, measured sCOMP, and reported means and standard deviations for sCOMP. Results: For question 1, seven studies were included resulting in 7 comparisons. A moderate overall effect size (ES) indicated sCOMP was consistently elevated in those with radiographically diagnosed knee OA when compared to controls (ES=0.60, p<0.001). For question 2, 4 studies were included resulting in 13 comparisons between radiographic OA severity levels and controls. Strong ES were calculated for K/L-1 (ES=1.43, p=0.28), K/L-3 (ES=1.05, p=0.04), and K/L-4 (ES=1.40, p=0.003). A moderate ES was calculated for K/L-2 (ES=0.60, p=0.01). Conclusions: These results indicate sCOMP is elevated in patients with knee OA and is sensitive to OA disease progression. Future research studies with a higher level of evidence should be conducted to investigate the use of this biomarker as an indicator for OA development and progression.

Key terms: biomarkers, Kellgren Lawrence, radiography, degenerative joint disease
INTRODUCTION

Characterized by irreversible joint destruction such as cartilage degradation, osteophyte development and joint space narrowing, osteoarthritis (OA) affects millions of individuals each year. Knee OA, either affecting the patellofemoral and or the tibiofemoral joint, is the most common cause of disability in the United States, causing pain and loss of function. Currently, there are few diagnostic tools used to identify individuals with knee OA. The diagnosis of OA is based on patient reports of pain and stiffness, and the presence of osteophytes and joint space narrowing as viewed on radiographs. Although many patients will demonstrate both symptomatic and visual indicators of OA, there is not a direct correlation between clinical indicators and actual joint damage. Given the limitation of current diagnostic tools and that early osteoarthritic changes such as articular cartilage abnormalities are silent, OA is often unrecognized until it has reached an irremediable and disabling level. The ability to develop intervention strategies with the hope of delaying irreversible joint damage remains difficult due to the lack of sensitive and valid pre-radiographic diagnostic tools. Identification of sensitive diagnostic tools to recognize pre-radiographic OA are necessary in order to develop and implement intervention strategies aimed at delaying irreversible joint damage.

Several serum and/or synovial fluid biomarkers have been identified in the literature to diagnose pre-radiographic OA. For a biomarker to be useful in diagnosing early joint damage, it must be sensitive to differences between healthy individuals and those with OA, and also among varying degrees of severity of joint disease. Examples of these biomarkers include keratan sulfate and pentosidine, both which tend to be elevated in patients with OA. Another biomarker that is theorized to have significant diagnostic value for beginning OA, is serum cartilage oligomeric matrix protein (sCOMP).
Serum COMP is a non-collagen biomarker for cartilage degradation present in articular cartilage, and other tissues such as ligament, meniscus, synovial membrane, and tendon\textsuperscript{1,17-21}. Numerous studies have investigated the relationship of sCOMP in patients with and without knee OA\textsuperscript{3, 4, 12, 14, 22, 23}. Validation of this relationship will provide scientists and physicians with a prospective pre-radiographic diagnostic indicator that may be clinically applicable and may assist in the development of treatment interventions for early stage OA.

The purpose of this systematic review was to answer the following questions: (1) Is sCOMP elevated in patients with radiographically diagnosed knee OA compared to controls? (2) Are there differences in sCOMP levels when comparing differing radiographic OA severities to controls?

**METHODS**

*Search Strategy*

A computerized literature search was completed in March of 2010 utilizing: CINAHL (from 1981), PEDro (from 1929), Medline (from 1966), and SportDiscus (from 1985). The search terms used were, knee, osteoarthritis, sCOMP, and radiography. All abstracts from the search results were reviewed. If the abstracts did not contain enough information to include or exclude the study from the review, the study was reviewed in its entirety. In addition, all reference lists were cross-referenced for relevant studies not included in the original searches.

*Criteria for Study Selection*

The inclusion criteria for the studies used in this systematic review were:
• Subjects with radiographically diagnosed knee OA and disease free control groups.

• Studies using the Kellgren Lawrence (K/L) scale to classify knee OA.

• Studies that measured sCOMP or used sCOMP as an outcome.

• Studies reporting means and associated measures of variability.

• Studies using human adults (18+ years or older).

• Studies published in the English language.

**Assessment of Publication Bias**

A funnel plot was used to provide an illustrative assessment of publication bias. In addition, Duval and Tweedie’s Trim and Fill method and Orwin’s Fail-Safe N were used to further interpret possible publication bias. The Duval and Tweedie’s Trim and Fill method looks for missing studies on the left side of the mean effect using a fixed effects model. The asymmetric studies from the right hand side of the mean effect are trimmed, the unbiased effect is located, then the studies to left of the mean are then filled in. This method results in an adjusted cumulative effect, and provides a conservative estimate of the total number of studies that are “missing”. Orwin’s Fail-Safe N test was employed to assess the robustness of the observed overall effects of the moderators on sCOMP.

**Sensitivity Analysis**

The “1-study removed method” was used to test the stability of the cumulative effect across the included studies by determining if the results of one particular study substantially influenced the overall effect. The analysis systematically removes each
study and replaces it so that the influence of each study can be individually evaluated. If the removal of any given study results in little change, it can be concluded that the pooled result is robust\textsuperscript{24}. For the second question, we performed an additional sensitivity analysis to determine the influence of sample size on the overall effect for each of the individual K/L comparisons. Study comparisons were dichotomized into “large (>10 subjects per group) or “small” (<10 subjects per group). As a group, “large” studies and then “small” studies were selectively removed in order to assess for changes in the overall result based on sample size.

Assessment of Study Quality

The study quality was assessed independently by two authors using a quality index for non-randomized studies\textsuperscript{26}. This index was adapted from a previously published version by Downs and Black\textsuperscript{27}. Based on the study designs for the included studies, the quality index assessment tool\textsuperscript{26} was selected in order to compare case-control and retrospective-cohort studies. A total of 16 items were used to assess study quality for each study. The quality index assessment tool addressed areas such as: clarity of objectives, main outcomes, subject characteristics and main findings, as well as, external validity and internal validity concerning bias and confounding\textsuperscript{26}. Any discrepancies in scores between authors were discussed and a mutual score was reached. Using previously published criteria\textsuperscript{26}, those studies achieving ≥75% of the criteria were considered high quality, 60-74% were considered moderate quality, and ≤60% were considered low quality.

Data Extraction and Statistical Analysis

The variable of interest for this study was sCOMP. The reported unit of measure is typically ng/ml, but sCOMP levels have been reported using µg/ml and U/L. For meta-analysis, all sCOMP units of measure were used for data extraction and statistical
analysis. Furthermore, in some cases sCOMP levels are not normally distributed. Recognition of this will allow for the data to be transformed using a logarithmic transformation, assuring the assumptions of the general linear model\textsuperscript{23, 28}. For the purposes of this meta-analysis, we recognize that a normal distribution might not have been present before data analysis; however, we did not or could not modify the data to control for this.

For this systematic review of the literature, K/L severity classification system for OA was used as an inclusion criterion. This classification system was chosen as it is a common classification system used to grade OA\textsuperscript{29}. Studies using other forms of OA severity classification systems were excluded to ensure consistent comparisons across all studies.

To determine if sCOMP was elevated in patients with radiographically diagnosed knee OA compared to controls, bias-corrected Hedges’s $g$\textsuperscript{30} effect sizes (ES) were calculated to estimate differences between OA and control groups and 95% confidence intervals (CI) to assess the uncertainty in these estimates. Hedge’s $g$ adjusts for sample size, which is often referred to as an unbiased estimate\textsuperscript{30}. When applicable, separate ESs were calculated for individual K/L classifications compared to controls. All ESs, 95% CIs, and p-values were calculated in Comprehensive Meta Analysis (Comprehensive Meta Analysis Version 2.0, Biostat, Englewood, NJ).

The calculated Hedges’ $g$ ES is a measure of a population’s mean effect estimated under the assumption that the individual variability has a Gaussian distribution. A positive ES indicated elevated sCOMP levels in OA patients as compared to controls. ESs $\leq 0.40$ were interpreted as representing a weak effect, ESs between 0.41 and 0.69 were interpreted as representing a moderate effect, and ESs $\geq 0.70$ were interpreted as representing a strong effect\textsuperscript{31}. An ES was interpreted as statistically significant if $< 0.05$ (in which the corresponding CI would exclude the nil value).
Level of Evidence

The level of evidence for the included studies was assessed using the Oxford Centre for Evidence-Based Medicine (CEBM)-Levels of Evidence\textsuperscript{32}. The levels of evidence range from 1a to 5, with 1a representing a systematic review of prospective cohort studies, and level 5 representing expert opinion without critical appraisal, bench research or “first principles”. The CEBM strength-of-recommendation grades are A, B, C, and D. Grade A represents consistent level 1 studies; grade B represents consistent level 2 or 3 studies or extrapolations from level 1 studies; grade C represents level 4 studies or extrapolations from level 2 or 3 studies; and grade D represents level 5 evidence, troublingly consistent or inconclusive studies of any level\textsuperscript{32}.

RESULTS

Study Selection

Computerized and hand searches yielded 57 studies that were included in the initial review (Figure 1). Based on the inclusion criteria and presentation of necessary data, a total of seven studies were included in this review (Table 1)\textsuperscript{3, 4, 11, 12, 14, 23, 33}. The reason(s) for exclusion for the remaining 50 studies can be found in the online Appendix Table 1.

Is sCOMP elevated in patients with radiographic knee OA compared to controls?

Seven studies met the inclusion criteria to answer this question (Table 1)\textsuperscript{3, 4, 11, 12, 14, 23, 33}. The mean quality index assessment for the included studies was 59% (25%-87.5%). Three studies\textsuperscript{4, 14, 23} were considered high quality, one\textsuperscript{3} was considered moderate...
quality, and three studies\textsuperscript{11, 12, 33} were considered low quality. Level C evidence exists that sCOMP is elevated in patients with knee OA. This recommendation was reached based on consistent level 4 studies with extrapolations from level 2 or 3 studies\textsuperscript{32}.

A total of 7 ESs and 95% CIs were calculated (Figure 2). Calculated ESs ranged from 0.06 ((CI) -0.47-0.60) to 2.70 (1.85-3.54). A total of 4 ESs were weak, 1 ES was moderate and 2 ESs were strong. The results of the random effects meta-analysis revealed a moderate overall ES of 0.60 (0.25-0.94, p=0.001), indicating sCOMP was consistently elevated in patients with radiographically diagnosed knee OA compared to controls.

\textit{Publication Bias}

The trim and fill\textsuperscript{24} analysis indicated one study is missing, and the addition of this study would result in an insignificant weak ES of 0.39 (-0.03-0.81, Figure 3). In addition, the Orwin’s Fail-Safe N indicated a range of 20-50 additional studies (based on the trivial ES range of 0.05 to 0.10) would be needed to nullify the overall effect. Therefore, the effect of publication bias introduced across the studies is likely trivial. If all relevant studies beyond those analyzed in this meta-analysis were included, the ES would probably remain unchanged.

\textit{Sensitivity Analysis}

Following the 1-study removed method, the ESs ranged from 0.37 to 0.71. The lowest lower confidence limit was 0.17 and the highest upper confidence limit was 1.20. All p-values were p<0.01. This indicated that there was not one particular study which substantially influenced the overall effect.
Are there differences in sCOMP levels when comparing differing radiographic OA severities to controls?

To be included in data analysis, a study must have presented data for the K/L severities in order to make control comparisons. Using this additional inclusion criteria, four studies\(^4,11,12,23\) were included and 13 comparisons were made (Table 1). The mean quality index assessment for the included studies used to answer this question was 56\% (25\%\text{-}87.5\%). Two studies\(^4,23\) were considered high quality and two studies\(^11,12\) were considered low quality. Level C evidence exists that sCOMP levels are elevated according to OA severity when compared to healthy controls. This recommendation was reached based on consistent level 4 studies with extrapolations from level 2 or 3 studies\(^32\).

A total of 13 comparisons were used, with a calculated overall effect of 1.00 (0.65\text{-}1.35, \(p<0.001\)). To answer question 2, subgroup ESs were calculated for each of the OA severities (K/L-1,-2,etc.) compared to controls. Strong ESs were observed for K/L-1, K/L-3, and K/L-4 while a moderate ES was observed for K/L-2 (Figure 4). The subgroup ES for the K/L-1 comparison\(^11,12\) was 1.43 (-1.15\text{-}4.02, \(p=0.28\)). The subgroup ES for the K/L-2 comparison\(^4,11,12,23\) was 0.60 (0.13\text{-}1.06, \(p=0.01\)). The subgroup ES for the K/L-3 comparison\(^11,12,23\) was 1.05 (0.06\text{-}2.03, \(p=0.04\)). The subgroup ES for the K/L-4 comparison\(^4,11,12,23\) was 1.40 (0.47\text{-}2.36, \(p=0.003\))\(^4,11,12,23\).

Our results indicate a significant moderate effect for K/L-2, and a significant strong effect for K/L-3 and K/L-4 when compared to controls. Therefore, the subgroup meta-analysis revealed strong trends in elevated sCOMP levels as OA severity levels increase. However, the CI for the ESs of each subgroup do overlap, therefore caution must be used when interpreting these results.

<insert Figure 4 near this paragraph>
Publication Bias

A publication bias assessment using the trim and fill method\textsuperscript{24} was performed for all 13 comparisons. The results indicated publication bias is likely and there are 5 comparisons missing (Figure 5). The addition of these 5 comparisons to the left of the overall mean effect would result in an insignificant weak overall effect of 0.39 (-0.002-0.79), with the CIs crossing zero. Additionally, the results of Orwin’s Fail-Safe N indicated a range of 44-100 additional studies (based on the trivial ES range of 0.05 to 0.10) would be needed to nullify the overall effect. Therefore, the effect of publication bias introduced across the studies is trivial. If all relevant studies beyond those analyzed in this meta-analysis were included, the ES would probably remain unchanged. <insert Figure 5 near this paragraph>

Sensitivity Analysis

Following the 1-study removed method; the ESs remained strong and ranged from 0.78 to 1.30. The lowest lower confidence limit was 0.48, and the highest upper confidence limit was 1.60. All p-values were p<0.001. This indicated that there was not one particular comparison that substantially influenced the overall effect calculated for the question.

A sense of caution must be employed when interpreting the K/L-1 vs. control ES, as the two comparisons used to calculate this ES had a combined total of 9 subjects in the group, and the CI encompassed zero. Given the recent results of a meta-epidemiological study\textsuperscript{34}, small comparisons such as those exhibited in the K/L-1 comparison, can inflate the overall effect (see K/L-1 diamond, Figure 4). In order to determine the effect of the smaller subgroups (n<10) on the overall ESs, we performed a sensitivity analysis comparing the overall effect of large vs. small subgroups. The results of this analysis (Figure 6) indicated small groups had a much larger overall
effect (ES=2.33, CI:1.09-3.60), compared to larger groups (ES=0.41, CI:0.25-0.60), demonstrating small groups significantly
influenced the overall effect for each subgroup. However, if we were to remove the comparisons with < 10 subjects per group, no
comparisons would be available for the K/L-1 severity. <insert Figure 6 near this paragraph>

DISCUSSION

To be a useful biomarker for the diagnosis of pre-radiographic knee OA, the biomarker must have a strong correlation with the
disease, thereby having levels that are distinguishable between patients with and without the condition\textsuperscript{4,15,16}. Based on the results of
this systematic review, we concluded that sCOMP is consistently elevated in patients with radiographic knee OA compared to healthy
controls (Figure 2). Furthermore, higher levels of sCOMP are associated with a trend toward greater radiographic OA severity when
compared to controls, as indicated by the ESs (Figure 3). Further investigation is warranted to determine if this marker can be utilized
to assess the presence of pre-radiographic OA, and in evaluating OA interventions in order to delay permanent joint degradation.

Study Quality Assessment

The mean quality index assessment for these studies was 59%. Three\textsuperscript{4,14,23} of the studies were considered high quality, one was
considered moderate quality\textsuperscript{3} and three of the studies\textsuperscript{11,12,33} were considered low quality. For the studies that were considered low
quality, a majority of the criterion examining external and internal bias were either unable to be determined or not represented in the
manuscript. These criterion included subject recruitment and representation of the population being sampled, blinding of investigators,
appropriateness of statistical tests used and validity and reliability of the measurement tools used. Also, missing were criterion that
were specific to internal validity. For example, a majority of the internal validity (confounding) criterion were not represented in most
manuscripts considered moderate and low quality. These criteria addressed whether subjects were recruited from the same population, during the same period of time and if there was adequate adjustment for confounding in the analyses.

**Kellgren/Lawrence Classification and Assessment**

As part of the inclusion criteria, the K/L classification for knee OA must have been included and reported. This classification system was chosen because it is the most commonly used radiological classification system used to grade OA. Two studies reviewed did not report the number of radiologists used to determine K/L classification. Senolt et al. reported the use of two radiologists for the review of subject radiographs and K/L classification assignment. One of the studies reported using a blinded rheumatologist to assign K/L classifications. Cibere et al. reported two blinded investigators with good interrater reliability (ICC=0.79) read the radiographs and independently classified the patients. Consensus was reached between readers if they disagreed on K/L classification for the patients. Finally, two studies reported the use of a single radiologist assessing the radiographs for K/L classification assignment and reported the inter- and intrarater reliability. It is important to report the number of individuals that are involved in reviewing radiographs, as this information will allow the readers to understand that bias was adequately controlled for. Also, reporting the inter- and intrarater reliability allows for interpretation of the generalizability of these measures among different clinicians.

Only three studies used in this review provided definitions for each of the K/L classifications used to assess OA severity. A recent review of K/L classifications used in current published research, specifically in reference to a grade of 2, reported grade 2...
definitions are different throughout the reported literature\textsuperscript{29}. Therefore it is essential the K/L grades be defined in each of the studies\textsuperscript{35}. Knowing which definition the authors used to identify K/L-2 subjects will allow better comparisons to be made across the studies.

\textit{Serum COMP ELISAs and associated coefficients of variation}

It has been previously reported that certain ELISAs are more appropriate for detecting human sCOMP than others\textsuperscript{36}. For the purposes of this systematic review we did not exclude studies based on the ELISA used to detect human sCOMP levels. Therefore, multiple ELISAs were used to detect human sCOMP for the studies that were included in our review (Table 2). Three studies used an ELISA manufactured by AnaMar Medical,\textsuperscript{3,14,33} three studies\textsuperscript{4,11,23} used an in house ELISA as previously reported by Vilim et al.\textsuperscript{37}, and one study used an ELISA manufactured by Kamiya Biomedical Company\textsuperscript{12}. The study using the Kamiya Biomedical Company ELISA\textsuperscript{12} did have the largest calculated ES. However, the authors cannot conclude whether or not this ELISA was more sensitive than the others used in this review. It must be noted that Jordan et al.\textsuperscript{23} and Clark et al.\textsuperscript{4} did report modifications to the Vilim et al.\textsuperscript{37} ELISA protocol in that an alkaline phosphatase-conjugated avidin was used rather than the previously reported peroxidase.

Previous research reported a weak correlation (R\textsuperscript{2} =0.210) between the AnaMar Medical ELISA and the in house ELISA reported by Vilim et al.\textsuperscript{36,37} for measuring sCOMP levels in the same subjects\textsuperscript{36}. For the purposes of this review, readers must be aware that there are differences in detecting sCOMP levels for each of the different ELISA manufactures. The only way to ascertain which ELISA kit is the best at detecting sCOMP levels can only be accomplished by comparing each technique with serum samples from the same subjects, which cannot be done using the available data for this meta-analysis. One advantage of the data contained
within this systematic review and meta-analysis are that the ESs provide a unitless measure of the magnitude of change between
groups which allow standardization for comparison between studies.

Coefficients of variation (CVs) are reported in order to determine the assay variability, and the interassay and intraassay CVs
are most commonly reported\(^{38}\). Four studies did not report the inter- or intraassay CVs for the ELISAs used to determine human
sCOMP\(^3,4,11,12\). Three studies did report inter- and intraassay CVs; with Mundermann et al.\(^{33}\) reporting the lowest CVs (< 1.9% for
interassay, < 2.7% for intraassay). Table 2 reports the inter- and intraassay CVs for each of the studies used in this systematic review
and meta-analysis. The acceptable range of reportable CVs are often laboratory specific, therefore it is important that this information
be presented in order to ascertain the variability of the reported sCOMP levels in each of the ELISAs performed.  \(<\text{Insert table 2 near this paragraph}>\)

*Sensitivity Analysis and Publication Bias Assessments*

We conducted two separate sensitivity analyses. For the first question, the lowest pooled ES was interpreted as weak (ES=0.37,
CI: 0.18-0.60). For the second analysis, the lowest pooled ES was interpreted as strong (ES=0.78, CI: 0.50-1.01). Based on these
findings, we concluded that our reported results were stable, and that no individual study substantially influenced the overall pooled
effects calculated for either meta-analysis.

For the second question, we performed a sensitivity analysis investigating the effects of the small groups (n<10) on the overall
effects for each of the sub groups (Figure 6). The results of this sensitivity analysis revealed the small groups have a much larger
overall effect (ES=2.33, CI: 1.1-3.6) compared to the large groups (ES=0.41, CI: 0.25-0.60). However, the smaller groups also
demonstrated greater imprecision, indicating potentially inflated results of the smaller groups which will directly influence the overall effect.

We also assessed publication bias using the Orwin’s fail-safe N\(^{25}\) and Duval and Tweedie’s Trim and Fill method\(^{24}\). Although the initial funnel plot indicated there was a possibility of publication bias for each of our questions (Figures 3 and 5), our secondary publication bias analysis indicated that more than 20 studies for the first question and 44 studies for the second question, with non-significant results, would need to be included in the analysis in order to make the pooled results become trivial and nullify our overall effects.

**Limitations**

This review is not without limitations. First, it must be noted that only studies using the K/L classification scale for OA were included. This classification system was chosen as it is a common classification system used to grade OA\(^{29}\). There are other systems for classifying OA; however, we chose to use only the K/L classification scale in order to ensure consistent comparisons across studies.

Second, a recent meta-epidemiological study associated with randomized control trials for OA treatment determined smaller studies (n<100) can have a deleterious effect on the interpretation of the overall meta-analysis results\(^{34}\). For the purposes of our meta-analysis, we considered a group small if they contained <10 subjects. For question 2, we had 6 groups that were considered small. In order to determine whether or not these groups had a harmful effect on the overall ESs for each subgroup, we conducted a secondary analysis of small vs. large comparisons (Figure 6). Based on this analysis, we determined that the small group comparisons may
inflate the overall ESs. Therefore caution is necessary when interpreting our results for the K/L-1 comparison as this subgroup only had 9 subjects. In addition, it was not possible to determine if several of the studies were prospective or retrospective in nature. Inclusion of this information in the future will allow more accurate synthesis of the available data.

Conclusions

The usefulness of sCOMP as a diagnostic biomarker for identification of knee OA prior to the occurrence of permanent joint degradation is contingent upon the sensitivity to detect differences in patients with and without the disease and between differing severities of the disease\(^4,15,16\). For both questions posed in this review, level C evidence is currently available in the literature. This recommendation was reached based on consistent level 4 studies or extrapolations from level 2 studies\(^32\). To strengthen this recommendation, future rigorous prospective investigations should be conducted. However, based on the available reported data, our results indicate sCOMP is elevated in patients with diagnosed knee OA compared to controls (Figure 2). Furthermore, clear trends were identified based on disease severity with larger ESs for K/L-3 and K/L-4 comparisons when compared to K/L-2 comparisons (Figure 4).

Author Contributions

All authors made substantial contributions to all three sections: 1) the conception and design of the study, or acquisition of data, or analysis and interpretation of the data, 2) the drafting of the article or revising it critically for important intellectual content and, 3) final approval of the version submitted.

Acknowledgements
There are no sources of funding for this study.

Competing Interest Statement

There are no competing interests to disclose for any of the authors on this manuscript.
References


Legends of Figures

Figure 1: Summary of search history and included studies.

Figure 2: A Forest plot depicting the calculated effect sizes of serum cartilage oligomeric matrix protein in patients with radiographically diagnosed knee osteoarthritis when compared to controls. The diamond at the bottom of the plot represents the overall effect size.

Figure 3: A funnel plot of standard error by Hedges’s g using Duval and Tweedie’s Trim and Fill method to assess publication bias of the studies used to determine if serum cartilage oligomeric matrix protein is elevated in patients with radiographically diagnosed knee osteoarthritis when compared to controls.

Figure 4: A Forest plot depicting the calculated effect sizes of serum cartilage oligomeric matrix protein exist when comparing patients with differing radiographic knee osteoarthritis severities when compared to controls. Each of the diamonds represents the overall effect size for each of the comparisons made for each Kellgren Lawrence classification.

Figure 5: A funnel plot of standard error by Hedges’s g using Duval and Tweedie’s Trim and Fill method to assess publication bias of the comparisons used to determine if differences in serum cartilage oligomeric matrix protein levels exist when comparing differing knee osteoarthritis severities when compared to healthy controls.

Figure 6: A Forest plot depicting the calculated effect sizes for large groups (n>10) and small groups (n<10) used to answer the second question.
Figure 1

Searched: CINAHL, Medline, PEDro, Sports Discus and hand search.
57 Studies identified.

18 studies excluded based on abstract content.

39 studies included.

32 studies excluded based on exclusion criteria.

7 studies included. 3, 4, 11, 12, 14, 23, 33
Figure 2

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Subgroup within study</th>
<th>Statistics for each study</th>
<th>Hedge's g</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cibere et al. (2009)</td>
<td>Control vs. KL 2,3,4</td>
<td>0.061</td>
<td>-0.472</td>
<td>0.930</td>
<td>0.822</td>
<td></td>
</tr>
<tr>
<td>Clark et al. (1999)</td>
<td>Control vs. KL 2,3,4</td>
<td>0.304</td>
<td>0.074</td>
<td>0.536</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td>Fernandes et al. (2007)</td>
<td>Control vs. KL 2,3,4</td>
<td>1.075</td>
<td>0.042</td>
<td>1.659</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Jordan et al. (2003)</td>
<td>Control vs. KL 2,3,4</td>
<td>0.375</td>
<td>0.229</td>
<td>0.520</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Mundermann et al. (2009)</td>
<td>Control vs. KL 1,2,3,4</td>
<td>0.061</td>
<td>-0.365</td>
<td>0.488</td>
<td>0.778</td>
<td></td>
</tr>
<tr>
<td>Senolt et al. (2004)</td>
<td>Controls vs. KL 1,2,3,4</td>
<td>0.586</td>
<td>0.131</td>
<td>1.041</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>Wakani et al. (2007)</td>
<td>Controls vs. KL 1,2,3,4</td>
<td>2.397</td>
<td>1.853</td>
<td>3.541</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

sCOMP lower sCOMP higher

-4.00 -2.00 0.00 2.00 4.00
K/L = Kellgren Lawrence Classification
The observed funnel plot suggests a publication bias towards studies demonstrating larger effect sizes, with an asymmetrical distribution of the studies to the bottom of the funnel. The black dot represents the missing study. The black diamond represents the adjusted overall effect size if the missing study was present.
Figure 4
<table>
<thead>
<tr>
<th>Group by</th>
<th>Study name</th>
<th>Subgroup within study</th>
<th>Hedges’s g</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control vs. K/L 1</td>
<td>Senolt et al.</td>
<td>Control vs. K/L 1</td>
<td>0.070</td>
<td>-1.332</td>
<td>1.464</td>
<td>0.921</td>
</tr>
<tr>
<td>Control vs. K/L 1</td>
<td>Wakiyama</td>
<td>Control vs. K/L 1</td>
<td>2.708</td>
<td>1.646</td>
<td>3.769</td>
<td>0.000</td>
</tr>
<tr>
<td>Control vs. K/L 1</td>
<td></td>
<td>Control vs. K/L 1</td>
<td>1.429</td>
<td>-1.541</td>
<td>4.012</td>
<td>0.278</td>
</tr>
<tr>
<td>Control vs. K/L 2</td>
<td>Clark et al.</td>
<td>Control vs. K/L 2</td>
<td>0.256</td>
<td>0.007</td>
<td>0.503</td>
<td>0.044</td>
</tr>
<tr>
<td>Control vs. K/L 2</td>
<td>Jordan et al.</td>
<td>Control vs. K/L 2</td>
<td>0.275</td>
<td>0.116</td>
<td>0.433</td>
<td>0.001</td>
</tr>
<tr>
<td>Control vs. K/L 2</td>
<td>Senolt et al.</td>
<td>Control vs. K/L 2</td>
<td>0.476</td>
<td>-0.084</td>
<td>1.036</td>
<td>0.096</td>
</tr>
<tr>
<td>Control vs. K/L 2</td>
<td>Wakiyama</td>
<td>Control vs. K/L 2</td>
<td>3.410</td>
<td>2.049</td>
<td>4.772</td>
<td>0.000</td>
</tr>
<tr>
<td>Control vs. K/L 2</td>
<td></td>
<td>Control vs. K/L 2</td>
<td>0.997</td>
<td>0.127</td>
<td>1.866</td>
<td>0.013</td>
</tr>
<tr>
<td>Control vs. K/L 2</td>
<td>Jordan et al.</td>
<td>Control vs. K/L 2</td>
<td>0.274</td>
<td>0.056</td>
<td>0.493</td>
<td>0.014</td>
</tr>
<tr>
<td>Control vs. K/L 3</td>
<td>Senolt et al.</td>
<td>Control vs. K/L 3</td>
<td>0.828</td>
<td>0.204</td>
<td>1.453</td>
<td>0.009</td>
</tr>
<tr>
<td>Control vs. K/L 3</td>
<td>Wakiyama</td>
<td>Control vs. K/L 3</td>
<td>2.426</td>
<td>1.380</td>
<td>3.491</td>
<td>0.000</td>
</tr>
<tr>
<td>Control vs. K/L 3</td>
<td></td>
<td>Control vs. K/L 3</td>
<td>1.045</td>
<td>0.061</td>
<td>2.030</td>
<td>0.037</td>
</tr>
<tr>
<td>Control vs. K/L 4</td>
<td>Clark et al.</td>
<td>Control vs. K/L 4</td>
<td>0.492</td>
<td>0.118</td>
<td>0.867</td>
<td>0.010</td>
</tr>
<tr>
<td>Control vs. K/L 4</td>
<td>Jordan et al.</td>
<td>Control vs. K/L 4</td>
<td>0.777</td>
<td>0.496</td>
<td>1.098</td>
<td>0.000</td>
</tr>
<tr>
<td>Control vs. K/L 4</td>
<td>Senolt et al.</td>
<td>Control vs. K/L 4</td>
<td>0.785</td>
<td>-0.239</td>
<td>1.810</td>
<td>0.133</td>
</tr>
<tr>
<td>Control vs. K/L 4</td>
<td>Wakiyama</td>
<td>Control vs. K/L 4</td>
<td>4.723</td>
<td>3.290</td>
<td>6.156</td>
<td>0.000</td>
</tr>
<tr>
<td>Control vs. K/L 4</td>
<td></td>
<td>Control vs. K/L 4</td>
<td>1.398</td>
<td>0.472</td>
<td>2.325</td>
<td>0.003</td>
</tr>
<tr>
<td>Control vs. K/L 4</td>
<td></td>
<td>Control vs. K/L 4</td>
<td>0.437</td>
<td>1.199</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Hedges’s g and 95% CI

Meta Analysis

sCOMP lower: sCOMP higher

Favours B
K/L = Kellgren Lawrence Classification

*: Clark et al.\textsuperscript{4} reported combined data for the K/L 3 and 4 subjects. For the purposes of this subgroup analysis, we placed this group of subjects into the K/L 4 group.
The observed funnel plot suggests a publication bias towards studies demonstrating larger effect sizes, with an asymmetrical distribution of the studies to the bottom of the funnel. The black dots represent the missing comparisons. The black diamond represents the adjusted overall effect size if the missing comparisons were present.
Table 1: Studies systematically reviewed to determine if sCOMP is elevated in patients with radiographically diagnosed knee osteoarthritis.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Level of Quality</th>
<th>Study</th>
<th>OA Group</th>
<th>Inclusion</th>
<th>No.</th>
<th>No. OA</th>
<th>Dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6
<table>
<thead>
<tr>
<th>Evidence (CEBM)</th>
<th>Index Score (%)</th>
<th>Design</th>
<th>Criteria</th>
<th>control patients</th>
<th>patients</th>
<th>variables measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cibere et al. (2009)</td>
<td>2b</td>
<td>81.25%</td>
<td>Exploratory cohort</td>
<td>Aged 40-79 years; pain, aching, or discomfort in or around the knee on most days of the month at any time in the past; any pain, aching, or discomfort in or around the knee in the past 12 months.</td>
<td>16</td>
<td>Pre-ROA= 105 ROA= 80† Total= 185</td>
</tr>
<tr>
<td>Clark et al. (2007)</td>
<td>2b</td>
<td>81.25%</td>
<td>Retrospective cohort</td>
<td>Caucasian, knee OA K/L ≥ 2.</td>
<td>148</td>
<td>K/L 2 = 109 K/L 3,4= 34 Total= 143</td>
</tr>
<tr>
<td>Fernandes et al. (2007)</td>
<td>4</td>
<td>68.75%</td>
<td>Case control</td>
<td>Aged 40-70 years, mechanical pain in one or both knees for minimum 3 months, and knee crepitus upon clinical evaluation.</td>
<td>86</td>
<td>SOA= 75± Pain= 11 NOA= 18± Total= 104</td>
</tr>
<tr>
<td>Jordan et al. (2003)</td>
<td>2b</td>
<td>87.5%</td>
<td>Retrospective cohort</td>
<td>Radiographs and serum COMP samples in database.</td>
<td>302</td>
<td>K/L 2= 313 K/L 3= 110 K/L 4= 44 Total= 467</td>
</tr>
<tr>
<td>Mundermann et al. (2009)</td>
<td>4</td>
<td>43.75%</td>
<td>Case control</td>
<td>Definite osteophyte presence in the medial or lateral tibiofemoral compartment; a narrowest point inter-bone distance of the medial compartment less than the</td>
<td>41</td>
<td>K/L 1= 11 K/L 2= 7 K/L 3= 12 K/L 4= 12 Total= 42</td>
</tr>
</tbody>
</table>
lateral compartment; pain in and around at least one knee for most of the days in the past months; at least some difficulty with two or more items in the WOMAC physical function scale.

Senolt et al. (2004)\textsuperscript{11} 

<table>
<thead>
<tr>
<th>Study Details</th>
<th>N</th>
<th>%</th>
<th>Study Design</th>
<th>Sample Description</th>
<th>No. With K/L</th>
<th>K/L</th>
<th>Articular Cartilage Assessment, keratan sulfate, CS6, CS846, HA and sCOMP and S. Pentosidin.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senolt et al. (2004)\textsuperscript{11}</td>
<td>4</td>
<td>25%</td>
<td>Case control</td>
<td>Undergoing arthrocentesis, no history of renal disease of diabetes mellitus and had normal levels of creatinine.</td>
<td>38</td>
<td>K/L 1= 2, K/L 2= 18, K/L 3= 14, K/L 4= 4, Total= 38</td>
<td>sCOMP and Pentosidine.</td>
</tr>
</tbody>
</table>

Wakitani et al. (2007)\textsuperscript{12} #

<table>
<thead>
<tr>
<th>Study Details</th>
<th>N</th>
<th>%</th>
<th>Study Design</th>
<th>Sample Description</th>
<th>No. with K/L</th>
<th>K/L</th>
<th>Articular Cartilage Assessment, keratan sulfate, CS6, CS846, HA and sCOMP and S. Pentosidin.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wakitani et al. (2007)\textsuperscript{12} #</td>
<td>4</td>
<td>25%</td>
<td>Case control</td>
<td>Patients undergoing knee surgery.</td>
<td>24</td>
<td>K/L 1= 7, K/L 2= 4, K/L 3= 6, K/L 4= 7, Total= 24</td>
<td>Articular cartilage assessment, keratan sulfate, CS6, CS846, HA and sCOMP and S. Pentosidin.</td>
</tr>
</tbody>
</table>

K/L= Kellgren Lawrence;  
* pre-ROA= pre-radiographic osteoarthritis (K/L <2); ROA= radiographic osteoarthritis (K/L >2); C2C= Type II collagen cleavage neopeptide; C1,2C= Types I and II collagen cleavage neopitope; CS846= cartilage proteoglycan aggrecan turnover epitope; NTX-I= N-telopeptide of type I collagen; CTX-II= C-telopeptide of type II collagen; HA= hyaluronan acid; WOMAC= Western Ontario McMaster University Index.  
† Only the ROA group was used for this review.
SOA (symptomatic osteoarthritis) = K/L grades 2,3,4; NOA (non-symptomatic osteoarthritis) = K/L 2 or higher; Pain = K/L 0 or 1;

VAS = Visual Analogue Scale

± Only the SOA and NOA groups were used for this review.

# C6S = chondroitin 6 sulfate;
Table 2. ELISA manufacture information and associated CVs for each of the studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>ELISA Manufacture Information</th>
<th>Type of ELISA</th>
<th>Specific antibodies used</th>
<th>Interassay CV</th>
<th>Intraassay CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cibere et al.</td>
<td>AnaMar Medical, Lund, Sweden</td>
<td>Sandwich</td>
<td>Two monoclonal antibodies (not specified)</td>
<td>&lt; 5%</td>
<td>&lt; 5%</td>
</tr>
<tr>
<td>Clark et al.</td>
<td>In house method *</td>
<td>Competitive</td>
<td>Monoclonal antibody 17-C10</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Fernandes et al.</td>
<td>Anamar Medical, Uppsala, Sweden</td>
<td>Sandwich</td>
<td>Two monoclonal antibodies (not specified)</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Jordan et al.</td>
<td>In house method *</td>
<td>Sandwich</td>
<td>Monoclonal antibodies 16-F12 and 17-C10</td>
<td>9.7% for “normal” controls</td>
<td>5.8% for “normal” controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.7% for “high” controls</td>
<td>6.6% for “high” controls</td>
</tr>
<tr>
<td>Mundermann et al.</td>
<td>AnaMar Medical AB, Lund, Sweden</td>
<td>Sandwich</td>
<td>Mouse monoclonal antibodies (not specified)</td>
<td>&lt; 1.9%</td>
<td>&lt; 2.7%</td>
</tr>
<tr>
<td>Senolt et al.</td>
<td>In house method *</td>
<td>Sandwich</td>
<td>Monoclonal antibodies 17-C10 and 16-F12</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Wakitani et al.</td>
<td>Kamiya Biomedical Company, Seattle, WA.</td>
<td>Sandwich</td>
<td>Not specified</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
</tbody>
</table>

Note. CV= coefficient of variation.

*: The same in house method was used for each of these studies, and the information regarding this method can be found in Vilim V, Voburka Z, Vytasek R, Senolt L, Tchetverikoc I, Kraus VB et al. Monoclonal antibodies to human cartilage oligomeric matrix protein:
epitope mapping and characterization of sandwich ELISA. *Clin Chim Acta* 2003;328:59-69. However, it must be noted that Clark et al.\textsuperscript{4} and Jordan et al.\textsuperscript{23} reported slight modifications to the in house method used for their studies.