Using ordered values of subregional cartilage thickness change increases sensitivity in detecting risk factors for osteoarthritis progression

R.J. Buck †*, B.T. Wyman ‡, M.-P. Hellio Le Graverand §, D. Hunter §§, E. Vignon ||, W. Wirth ¶#, F. Eckstein ¶#

StatAnswers Consulting LLC, San Diego, CA, USA
Pfizer Inc., Groton, CT, USA
Northern Clinical School, Sydney Medical School, University of Sydney, Australia
Claude Bernard University Lyon 1, Lyon, France
Institute of Anatomy & Musculoskeletal Research, Paracelsus Medical University (PMU), Salzburg, Austria
Chondrometrics GmbH, Ainring, Germany

SUMMARY

Objective: To examine whether ordered values of (sub)regional femorotibial cartilage thickness change are superior to region-based approaches in detecting risk factors for cartilage loss in osteoarthritis (OA).

Methods: 58 women with knee OA had 3 Tesla MR images acquired at baseline and 24 months. Changes in cartilage thickness (ΔThCtAB) were determined in eight medial femorotibial subregions. An ascending sort of individual ΔThCtAB measurements was done to create “ordered values”. Risk factors for cartilage loss considered were: age, BMI, anatomical knee axis (AAA), minimal (medial) joint space width (mJSW), and percent of medial tibial plateau covered by the meniscus (percent cover). All change metrics were tested for association with the risk factors using Kendall’s τ and relative sensitivity of multiple tests of subregions and ordered values were compared with single metrics of change from plate and compartment summaries and the first ordered value.

Results: The associations between subregion ΔThCtAB and AAA (P = 0.0002), mJSW (P = 0.016), and age (P = 0.011) were significant, but only AAA (at a = 0.05) and age (at a = 0.1) remained significant after adjusting for multiple subregions. In contrast, cMFTC had P-values < 0.05 for AAA (P = 0.0001), mJSW (P = 0.016), and meniscus subluxation (0.04). The first ordered value had significant associations with AAA (P = 0.0004), mJSW (P = 0.003), meniscus subluxation (P = 0.02) and percent cover (P = 0.031) of which were significant at a = 0.05 after adjusting for tests on multiple risk factors.

Conclusion: Ordered values of ΔThCtAB were more sensitive in detecting risk factors of cartilage loss than subregional ΔThCtAB. Sensitivity was further enhanced by considering the minimum ordered value as a single test, thus not requiring adjustment for multiple tests. Using ordered values there was a significant association between ΔThCtAB and baseline AAA, mJSW, meniscus subluxation and meniscus percent cover. This study provides an important step in validating ordered values of cartilage change.

© 2010 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Introduction

Magnetic resonance imaging (MRI) is increasingly used to determine cartilage loss quantitatively in osteoarthritis (OA) and to explore the relationship between cartilage loss and potential risk factors. Such risk factors include malalignment, i.e., the deviation from the normal knee axis, high body mass index (BMI), meniscus alterations, small minimal joint space width (mJSW), subchondral bone sclerosis, or other signs of advanced disease, including presence of cartilage lesions, denuded bone area, or low cartilage thickness at baseline.

Recent interest has been in measuring cartilage change not only over entire compartments or cartilage plates, but in distinct femorotibial subregions. Because (sub)regions with the greatest changes appear to vary between subjects, a recent paper proposed the ordered values of subregional change as a more efficient metric of cartilage thickness change. The ordered values focus on the magnitude/direction of the change alone and are therefore independent of the location, i.e., subregion, of the maximal change in each participant. This approach was found to be superior in detecting differences in cartilage loss between OA participants (Kellgren–Lawrence Grade [KLG] 3) and a healthy reference group, whereas the differences (in cartilage loss) were...
not statistically significant between these two groups when using total plate or (sub)regional cartilage thickness changes\(^1\).

The ordered value approach also has potential for exploring relationships between cartilage loss and risk factors of OA progression with greater sensitivity. When change within a specific subregion is assessed across subjects, the link between cartilage change and a risk factor may be weak, because few participants show progression in that particular subregion. The ordered values approach, in contrast, provides opportunity to capture cartilage thickness changes across subjects regardless of spatial location, without the potential dilution of the signal when using plate or compartment summaries. Thus, the ability to detect significant associations between cartilage loss and risk factors of OA progression may be stronger when using ordered values than when using region-specific measures.

In this study we assess whether ordered values of subregional femorotibial cartilage thickness change\(^1\) display greater sensitivity in revealing a relationship between risk factors and progression of knee OA (cartilage loss) than conventional strategies that summarize cartilage thickness change across anatomical (sub)regions\(^14\). The number of risk factors was not intended to be exhaustive but meant to include key risk factors, as established by previously published data.

### Methods

The longitudinal observational study on which this analysis was based included 180 women (70 with radiographic OA, 110 healthy controls), of which 152 (age 56.7 ± 8.6 years, 58 with radiographic OA, 94 healthy controls) completed the baseline and month 24 visits\(^14,17\). This particular study evaluated the 58 OA subjects who completed both visits. Inclusion criteria for OA participants were frequent symptoms, mild to moderate radiographic OA in the medial femorotibial compartment in conventional weight-bearing, extended anterior—posterior (AP) radiographs, less or equal joint space narrowing (JSN) in the medial than in the lateral femorotibial compartment, a BMI of ≥30, and a medial femorotibial mJSW of ≥2 mm in modified Lyon Schuss (mLS) radiographs\(^18,19\). The study was conducted in compliance with the ethical principles derived from the Declaration of Helsinki, the local Institutional Review Boards, informed consent regulations, and International Conference on Harmonization Good Clinical Practices Guidelines.

KLG1 first scored from AP radiographs at the imaging site, was then reassessed by a single experienced reader after enrollment was completed, and was eventually adjudicated by a third reader, if the first two disagreed.

Previously validated\(^20,21\), double oblique coronal water excitation spoiled gradient echo MRI sequences were acquired at 3 Tesla, at (cMFTC = controls), of which 152 (age 56.7 ± 8.6 years, 58 with radiographic OA, 94 healthy controls) completed the baseline and month 24 visits\(^14,17\). This particular study evaluated the 58 OA subjects who completed both visits. Inclusion criteria for OA participants were frequent symptoms, mild to moderate radiographic OA in the medial femorotibial compartment in conventional weight-bearing, extended anterior—posterior (AP) radiographs, less or equal joint space narrowing (JSN) in the medial than in the lateral femorotibial compartment, a BMI of ≥30, and a medial femorotibial mJSW of ≥2 mm in modified Lyon Schuss (mLS) radiographs\(^18,19\). The study was conducted in compliance with the ethical principles derived from the Declaration of Helsinki, the local Institutional Review Boards, informed consent regulations, and International Conference on Harmonization Good Clinical Practices Guidelines.

KLG1 first scored from AP radiographs at the imaging site, was then reassessed by a single experienced reader after enrollment was completed, and was eventually adjudicated by a third reader, if the first two disagreed.

Previously validated\(^20,21\), double oblique coronal water excitation spoiled gradient echo MRI sequences were acquired at 3 Tesla, at baseline and 24 months\(^22,23\). Segmentation of femorotibial cartilages (medial tibia = MT, lateral tibia = LT, weight-bearing medial femoral condyle = cMF, and weight-bearing lateral femoral condyle = cLF) was performed by seven readers with formal training and at least 3 years experience in cartilage segmentation, using dedicated software (Chondrometrics GmbH, Germany)\(^22\). The images were read in pairs with readers blinded to order of acquisition and segmentation quality control was performed for all data sets by one person (FE). The precision and stability of the measurements in this multi-center study have been described previously\(^22\).

Mean cartilage thickness over the total area of subchondral bone (ThCBAB)\(^22,24\) was used to measure cartilage loss. ThCBAB was computed for total cartilage plates (MT, cMF, LT, and cLF), five tibial subregions (central, external, internal, anterior, and posterior), and three femoral subregions (central, external, internal), the medial and lateral femorotibial compartments (MFTC = MT + cMF, LFTC = LT + cLF), and central portion of compartments (cMFTC = cMT + ccMF, cLFTC = cLT + ccLF). Subregion borders within each cartilage plate were determined by an automated algorithm\(^8,16\). Only medial compartments, plates, and subregions were considered for analysis, as all OA study participants had medial femorotibial disease.

The correlation of risk factors measured at baseline with ΔThCBAB was explored in the subset of subjects with KLG2 (n = 30) or KLG3 (n = 28), since healthy volunteers were not expected to have cartilage loss. The risk factors examined were: age, BMI, anatomical knee axis alignment (AAA) as measured from LS radiographs\(^22,25\), mJSW measured from mLS radiographs\(^19,26\) and medial meniscal subluxation and percent cover measured from coronal T1-weighted WE 3D\(^10,27\).

AAA was defined as the angle between the tibial and femoral axes, i.e., the lines joining the midpoints between the inner and outer margins of the bone and the central point between the tibial spines; varus alignment corresponding to positive angles. The inner and outer margins of the femur and of the tibia were detected at least 10 cm from the central point between the tibial spines. AAA was calculated using LS films, digitized image analysis software (Holy’s software, UCLB, Lyon, France), and manual identification of the tibial spines by an experienced investigator (EV) who was blinded to visit.

Medial meniscal position measures of subluxation and the percent of the tibial plateau covered by the meniscus were calculated using eFilm Workstation software as described previously\(^10,27\). Percent of meniscus covering of the medial tibia was calculated as “meniscus covering of the medial tibia” divided by the sum of “meniscal covering” and “meniscal uncovering” of the medial tibia.

The mJSW of the medial femorotibial compartment was measured using semi-automatic digitized image analysis software (Holy’s software, UCLB, Lyon, France), permitting the contours of the medial JSW to be detected automatically by an edge-based algorithm\(^26\). The external limits were determined by an experienced observer who excluded marginal osteophytes.

To generate the ordered values, the ΔThCBAB (mm/y) values in the eight medial subregions (five tibial, three femoral) were sorted, and the smallest value (most negative or least positive change) in each subject was assigned to ordered value one. This was repeated with the next smallest values, in order for remaining subregions. These ordered values are denoted as C (Δchange)\(_1\), C\(_2\), C\(_3\), C\(_4\), C\(_5\), C\(_6\), C\(_7\), C\(_8\), where the subscript represents the rank of observed ΔThCBAB relative to other rates of change in the same subject and compartment (Table 1).

The associations between potential risk factors and measures of change, either compartment, plate, subregional, or ordered values, were assessed using Kendall’s τ\(^28\). Kendall’s τ has the same characteristics as other measures of correlation, notably ±1 is perfect association and 0 is no association. Kendall’s τ is a measure of concordance and summarizes whether two measurements (in this case a risk factor and cartilage loss) are both larger (or smaller) in one individual compared to another over all pairs of individuals. A

### Table 1

<table>
<thead>
<tr>
<th>Subject 1:</th>
<th>(\Delta)ThCBAB (mm/y) values for eight subregions in an individual are presented along with the ordered values names C(_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subregion</td>
<td>cMT</td>
</tr>
<tr>
<td>Subject 1:</td>
<td>(\Delta)ThCBAB</td>
</tr>
<tr>
<td>C(_1) = 1</td>
<td>1*</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Subject 2:</th>
<th>(\Delta)ThCBAB (mm/y) values for eight subregions in an individual are presented along with the ordered values names C(_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subregion</td>
<td>cMT</td>
</tr>
<tr>
<td>Subject 2:</td>
<td>(\Delta)ThCBAB</td>
</tr>
<tr>
<td>C(_1) = 1</td>
<td>1*</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>Subject 3:</th>
<th>(\Delta)ThCBAB (mm/y) values for eight subregions in an individual are presented along with the ordered values names C(_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subregion</td>
<td>cMT</td>
</tr>
<tr>
<td>Subject 3:</td>
<td>(\Delta)ThCBAB</td>
</tr>
<tr>
<td>C(_1) = 1</td>
<td>2</td>
</tr>
</tbody>
</table>

*The average change for C\(_1\) is (\(-0.45\)–\(0.25\)–\(0.55\))/3 = \(-0.42\) mm
The measures of change were considered in four groups: compartment or central compartment and plates; subregions; ordered values; and C(1). The minimum ordered value, C(1), can be viewed as the first in a series of sequential tests with testing continuing only as long as test results are statistically significant. Based on this approach, tests between risk factors and the minimum ordered value C(1) can be considered a univariate test for a general test of association. This approach increases test sensitivity in detecting associations between thickness change and risk factors by dropping the multiple comparison restrictions.

The reason for splitting the measures into these four groups was that each was considered a potential primary endpoint that may be used in future clinical trials. The evaluation of tests of subregions or ordered values of change requires adjustment for multiple tests on the eight subregions, while plates or compartments have a limited number of tests (two tests for plates and a single test for MFTC or cMFTC). Hence, direct comparison of P-values between subregions or ordered values of ΔThCtAB on one hand, and plates or compartments on the other hand is not appropriate.

Comparisons between methods and metrics are based on P-values, and on whether significance levels are met. If no association exists, then P-values reflect random variation and relative magnitudes are not meaningful. For this reason the focus was on results with significant P-values, particularly for examples of single test conditions using plates, compartments or C(1). Since multiple risk factors were considered, the false discovery rate (FDR) methods described below were used to help confirm that the observed results were significant, and not a reflection of multiple tests.

The comparison of the sensitivity of Kendall’s τ as a measure of association between a given risk factor and ΔThCtAB of subregions/ordered values was evaluated by the magnitude of P-values < 0.05, using ΔThCtAB of subregions/ordered values, respectively. Comparison of sensitivity between compartments, central compartments, plates and C(1) was based on the magnitude of P-values for tests of association. Although there were tests for two plates vs single tests for compartments or C(1), it was felt the effect of multiple comparisons was negligible enough to permit direct comparisons in this particular case.

Comparisons between conditions where multiple tests (subregions/ordered values) and single tests (compartments/plates/C(1)) were done by comparing significance after adjusting for multiple tests. Multiple tests of Kendall’s τ over the eight different medial regions or ordered values were handled using the FDR(m) methods at two different significance criteria, α = 0.05 and α = 0.1, to create conditions considered “significant” and “moderately significant”.

### Results

Tests of association, Kendall’s τ, between risk factors and medial compartment (MFTC) ΔThCtAB (Table II) had P-values < 0.05 for AAA ($τ = 0.339, P = 0.0003$) and mJSW ($τ = 0.196, P = 0.03$). Tests of association between ΔThCtAB in MFTC and risk factors (Table II) had P-values < 0.05 for AAA ($τ = 0.37, P = 0.0001$), meniscus subluxation ($τ = -0.181, P = 0.04$), and mJSW ($τ = 0.212, P = 0.016$). Tests of association between risk factors and plate (cMF, MT) ΔThCtAB had P-values < 0.05 for AAA (cMF $τ = 0.345, P = 0.0002$, MT $τ = 0.22, P = 0.0187$). The tests of association between meniscus subluxation and MT ΔThCtAB ($P = 0.055$) and mJSW and cMF ΔThCtAB ($P = 0.06$) had the next smallest P-values. AAA was significant at α = 0.05 for MFTC, cMFTC, and MFTC ΔThCtAB after adjusting for multiple risk factors, while mJSW was significant for MFTC ΔThCtAB ($a = 0.1$) and cMFTC ΔThCtAB ($a = 0.05$), and meniscus subluxation was significant ($a = 0.1$) for cMFTC ΔThCtAB.

Individual tests of association, Kendall’s τ, between demographic risk factors and ΔThCtAB of subregions (Table III) resulted in P-values of < 0.05 for age in icMF ($τ = 0.228, P = 0.011$). After adjusting for the
Discussion

In this study we assessed whether ordered values of subregional femorotibial cartilage change display greater sensitivity in revealing a relationship between established risk factors and progression of knee OA (cartilage loss) than conventional strategies that summarize cartilage thickness change across anatomical (sub) regions. Tests of association were conducted for ΔThCtAB at 12 femorotibial anatomical regions (including compartment, plate and subregions measures), and ordered values of eight subregion ΔThCtAB values against six risk factors. Compartments, plates and subregions were viewed as summaries of cartilage change that create a single metric and hence only a single test of association was needed for each risk factor. Considering C(1) as a single metric is based on viewing the minimal ordered value as the first in a series of sequential tests that is sufficient to gauge the association between cartilage change and risk factors. Tests for subregion ΔThCtAB and ordered values require multiple tests to assess the relationship of cartilage loss and a risk factor. Interest was in comparing the sensitivity of single tests based on plates, compartments, or C(1) with multiple tests for subregions and ordered values of those subregions.

The risk factors most frequently found with P-values < 0.1 were AAA, mJSW, meniscus subluxation, and percent meniscal cover of the medial tibia. The ordered values approach was significantly associated with more risk factors (AAA and mJSW at α = 0.05) and meniscus subluxation at α = 0.1, compared to subregion approach (AAA at α = 0.05, and age at α = 0.1). This indicates that using ordered values when viewed as multiple tests is more sensitive in detecting relationships between risk factors and progression of knee OA than using anatomical regions (or subregions).

Risk factors showing significant associations with ΔThCtAB in the current study have been previously identified as relevant risk factors of structural progression of OA, namely malalignment, low radiographic mJSW, and meniscus changes. There are several reasons why certain risk factors, such as weight and BMI, did not have significant associations with ΔThCtAB in this study: The study may have lacked adequate sample size to detect the association or, because only participants with a BMI > 30 were included in the study, the variability in BMI may have been too low for detecting a significant association. Meniscal position and coverage were measured using MRI scans that are not generally considered optimal for meniscus delineation but provided sufficient spatial resolution for performing quantitative measurements. This may have affected the strength of the relationship found with cartilage thickness change, but it needs to be stressed that the study was not designed to confirm known or identify novel risk factors of cartilage change, but to evaluate the relative performance of the ordered value vs region-specific approach.

When a study is designed to identify risk factors for OA progression, interest is mainly in observing the strength and direction of the association; understanding the spatial magnitude and location of change is generally of secondary importance. The presumption when using ordered values is that the extreme ordered values (minimum, maximum) will be populated with real changes, if they occur, from any subregion of an individual. Hence if a relationship exists between cartilage loss and a risk factor, the evidence would most likely be observed in the smallest ordered values (i.e., the region of greatest cartilage loss). This implies that a single test of the minimum ordered value, C(1), should be sufficient to test for association.

### Table IV

Kendall’s τ and P-values for tests of association between change in specified ordered values of subregion change and selected risk factors. P-values are unadjusted for multiple comparisons. Here C(1) is viewed as one of multiple tests covering all ordered values.

<table>
<thead>
<tr>
<th>Region</th>
<th>Age</th>
<th>BMI</th>
<th>AAA</th>
<th>Meniscus Subluxation</th>
<th>Percent Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(1)</td>
<td>-0.002</td>
<td>0.984</td>
<td>-0.045</td>
<td>0.615</td>
<td>0.333</td>
</tr>
<tr>
<td>C(2)</td>
<td>-0.002</td>
<td>0.984</td>
<td>0.024</td>
<td>0.794</td>
<td>0.363</td>
</tr>
<tr>
<td>C(3)</td>
<td>-0.003</td>
<td>0.973</td>
<td>0.026</td>
<td>0.773</td>
<td>0.305</td>
</tr>
<tr>
<td>C(4)</td>
<td>0.067</td>
<td>0.456</td>
<td>0.018</td>
<td>0.846</td>
<td>0.340</td>
</tr>
<tr>
<td>C(5)</td>
<td>0.071</td>
<td>0.432</td>
<td>-0.030</td>
<td>0.742</td>
<td>0.291</td>
</tr>
<tr>
<td>C(6)</td>
<td>0.120</td>
<td>0.181</td>
<td>0.064</td>
<td>0.481</td>
<td>0.191</td>
</tr>
<tr>
<td>C(7)</td>
<td>0.111</td>
<td>0.219</td>
<td>0.030</td>
<td>0.742</td>
<td>0.216</td>
</tr>
<tr>
<td>C(8)</td>
<td>0.025</td>
<td>0.783</td>
<td>-0.049</td>
<td>0.587</td>
<td>0.181</td>
</tr>
</tbody>
</table>

Significance for each risk factor accounting for multiple comparisons across all ordered values was assessed by FDR at two levels denoted by ** α = 0.05; * α = 0.1.
Generally, extreme values, e.g., $C_{1}$, are highly variable compared to intermediate ordered values. This statistical property may diminish the benefit of reducing the problem to a single test, as $C_{2}$ may reduce variability while maintaining differences between groups leading to a larger standardized response mean. The current study shows that the ordered value $C_{1}$ was always amongst the ordered values with the smallest $P$-values for risk factor associations. Further, a previous study comparing change between healthy reference subjects and OA participants showed that $C_{1}$ was the ordered value with the strongest evidence (smallest $P$-value) for differences in the rate of change in ThCTAB between two populations with different KLG; so it appears that using $C_{1}$ that as a single metric for testing is a reasonable choice.

Amongst compartment, plate and subregion measures of $\Delta$ThCTAB, tests of association between risk factors and cMFTC were the most sensitive. The use of an aggregate measure of subregion thickness changes, e.g., compartment, plate or cMFTC $\Delta$ThCTAB, has both benefits and drawbacks. If change occurs in several subregions, which may vary between individuals, all these changes may be encompassed into the aggregate measure. On the other hand, if changes are localized, the changes observed may be diluted by including areas with minimal change. Another benefit of using a single metric per patient, rather than subregions is that multiple comparisons across subregions are no longer needed. The benefit of choosing a single test a priori over considering multiple regions to evaluate the relationships between risk factors and cartilage loss is substantial, as can easily be appreciated by the effect of the commonly used Bonferroni correction, $a/n$, where $n$ is the number of tests. It is impossible to avoid multiple testing conditions when considering multiple subregions, except through summation, e.g., compartments, or finding a sequential strategy for testing that puts the primary endpoint as the first test. Both the use of the minimum ordered value, $C_{1}$, and cMFTC can be viewed as approaches to providing a single test condition for assessing the association of ThCTAB with risk factors. Tests for $C_{1}$ were found to have smaller $P$-values than cMFTC for mJSW, meniscal subluxation and percent meniscal cover of the medial tibia, while tests for cMFTC had a smaller $P$-value for AAA, but both measures were highly significant ($P < 0.0005$). The magnitude of the differences in $P$-values is highlighted by the fact that these four risk factors were still significantly associated with $C_{1}$ after adjusting for tests over multiple risk factors, while only AAA and mJSW were significantly associated with cMFTC after adjusting for multiple tests over all risk factors. These results indicate that $C_{1}$ was more sensitive than the cMFTC for detecting a relationship between established risk factors and progression of knee OA (i.e., cartilage loss).

While $C_{1}$ is derived from multiple observations on each individual, only one test for differences between groups is carried out and hence is free of multiple comparison issues. The effect of using $C_{1}$ is wrapped up in the distribution of $C_{1}$ and this is one of the reasons the use of non-parametric methods was considered, as the nature of this distribution is unclear. It is important to distinguish this situation from the ordering of test results, which clearly does not reflect a single test situation.

The smallest $P$-values for tests of association between subregional $\Delta$ThCTAB and the risk factors AAA, mJSW, subluxation and percent meniscal cover of the medial tibia generally occurred in the same three subregions: ecMF, eMT, and cMT. These subregions also have been shown to be amongst the subregions showing the greatest mean rates of change in the medial femorotibial compartment across different studies. However, these subregion test results were generally only marginally significant ($0.05 < P < 0.1$). This provides indirect evidence that $\Delta$ThCTAB is associated with these risk factors, but that the changes are differentially distributed anatomically across individuals, which leads to the association with each subregion being diluted. Since $C_{1}$ captures the largest cartilage loss of each individual, this may explain why $C_{1}$ was generally more sensitive to showing risk factor associations than compartment, plate or subregional changes.

The ordered values approach can also include both medial and lateral femorotibial subregions. Considering ordered values over medial and lateral subregions in populations with medial or lateral or bi-compartmental disease (or with varus and valgus malalignment) may be particularly useful when analyzing non-compartment-specific risk factors (such as BMI). Another potential advantage of the ordered value approach is that it is not restricted to identifying risk factors of cartilage loss. A recent paper suggested that OA may not be a one way road of cartilage loss, but that at certain (early) stages of the disease, cartilage thickness is increasing (due to swelling or hypertrophy) rather than diminishing. Whereas these effects (swelling/hypertrophy vs loss) may be cancelling each other when looking at means of change in larger cohorts, analysis of the maximum ordered value $C_{8}$ (the ordered value with greatest increase in cartilage thickness) may provide opportunity to identify risk factors of cartilage swelling/hypertrophy at the earlier stages of the disease.

The statistical properties of ordered values in the context of examining the association between risk factors and OA progression are optimal when measurement error is the same for all subregions. If heterogeneous measurement error was an issue, it is unclear how it could bias the tests to create unwarranted significant results. However, these properties should still hold as long as individual OA-related changes are large relative to measurement error, and past reports indicate that this is a reasonable assumption.

In conclusion, this is the first study to suggest that ordered values of (sub)regional femorotibial cartilage thickness change derived from MRI are superior to region-based approaches (i.e., more sensitive) in identifying risk factors for cartilage loss in OA. Sensitivity is further improved by considering tests using ordered values sequentially. $C_{1}$ was found to be more sensitive than cMFTC (the most sensitive region measure) in tests of association with selected risk factors, with both providing a basis for a single test for association. The current conclusions are based on empirical evidence and should be confirmed in other (larger) studies and also including other risk factors. However, the significant association of ordered values with risk factors provides an important step in validating that ordered values of cartilage change, particularly $C_{1}$, captures OA-related change, as statistically significant associations with risk factors would not regularly occur if $C_{1}$ simply reflected random measurement error.

Contributions

RB, the corresponding author, takes responsibility for the integrity of the article as a whole and was involved with conception, design, data management and analysis, and primary author of the manuscript. All authors were part of a critical review of the manuscript and gave final approval with FE providing additional help in drafting the manuscript. All authors, except WW, were involved with the conception and design of the original A900-1140 study. BW and MPLG were principle investigators of the study. DH provided meniscus results, EV provided information from radiographs, and WW and FE played a lead role in providing cartilage thickness measurements.

Conflict of interest

Robert Buck was employed by Pfizer and is a consultant with Chondrometrics. Brad Wyman and Marie-Pierre Hullo Le Graverand are employed by Pfizer. David Hunter and Eric Vignon were consultants to Pfizer. Wolfgang Wirth is a shareholder and
Acknowledgements

We would like to thank the dedicated group of A9001140 investigators: Cecil Charles, Julia Crim, Gary Hutchins, Chris Jackson, Virginia Byers Kraus, Thomas M. Link, Sharmila Majumdar, Prasad Pottumarthi, Thomas Schnitzer, Mihra Taljanovic, Berchman Vaz. Also to the dedicated study coordinators whose skills were essential in assuring the successful conduct of this research: Emily Brown, Sandra Chapman, Eugene Dunkle, Kristen Fredley, Donna Gilmore, Joyce Goggins, Mohsen Haddad-Kaveh, Norine Hall, Thelma Munoz, and Kim Tally. We would also like to express our thanks to the dedicated MRI technologists, the Duke Image Analysis Laboratory staff: Cecil Charles, Maureen Ainslie, April Davis, Allison Fowlkes, Mark Ward and Scott White, the Pfizer A9001140 Team: Lydia Brunstetter, Peggy Coyle, Yevgenia Davidoff, Charles Packard, Ann Remmers, Mark Tengowski, Jeff Evelhoch and John Kotyk, and the Chondrometrics GmbH readers: Gudrun Goldmann, Linda Jakobi, Manuela Kunz, Dr Susanne Maschek, Sabine Mühlheimer, Annette Thebis, and Dr Barbara Wehr for dedicated data segmentation. David Hunter is supported by an ARC Future Fellowship.

Grant supporter(s): Pfizer Inc.

References


