APPENDICES
Updates to May 2011 Submission
Working Group: Prevention or Risk Reduction
Working Group: Assessment of Structural Change

Clinical Development Programs for Human Drugs, Biological Products, and Medical Devices for the Treatment and Prevention of Osteoarthritis

DOCKET NO. 1998D-0077 (FORMERLY 98D-0077)
Federal Register: August 14, 2007 (Volume 72, Number 156)

Working Group: Prevention or Risk Reduction Recommendations

What are potential outcome measures?

Recommendation: For these purposes, primary and secondary prevention and risk reduction of structural and symptomatic indicators of OA would require outcomes relevant to these domains. As definitions of "at risk" populations change and measurements of the disease process and outcomes advance, it is expected that design features and relevant outcomes of prevention trials would necessarily evolve as well.[1] Additionally, the working group focused on knee OA; outcomes, study design issues, populations at risk, duration of trials may vary depending upon the joint site under evaluation.

For example, if the prevention trial hypothesis is that an intervention among obese adults with no or doubtful evidence of radiographic knee OA (Kellgren Lawrence [KL] radiographic score = 0, 1) will be associated with a delayed onset of knee OA compared to the placebo group, this delay could be reflected in two co-primary outcomes: less symptom report and minimal structural change in relation to the untreated group. Candidate measures to detect these areas include changes in: (1) KL score and (2) questionnaire-based pain assessment (PRO's). Other potentially relevant outcome measures could include newer technologies once validated, such as MRI or T2 mapping to assess morphological changes in joint structures or articular cartilage degradation and/or bone marrow lesions. As imaging and molecular techniques advance to the stage where they could be surrogates of downstream clinical outcomes, it may be that an intervention might be able to show a primary effect on structure of the OA process, regardless of its immediate effect on symptoms. Examples in other conditions abound, that is, interventions directed toward lowering serum cholesterol or altering lipid profiles to prevent future cardiovascular events,[2] or altering bone mineral density to prevent osteoporotic fractures.[3]

Secondary outcomes could include some or all of the following largely predicated on the nature of the proposed intervention: (1) clinical measures of function, pain, and mobility; (2) mechanistic measures of the OA disease pathways such as knee alignment, knee external adductor moment, knee joint compressive, and shear forces; (3) biomarker measures of pro-inflammatory molecules (e.g., IL-6, Tumor Necrosis Factor-α, C-RP) and joint metabolism (e.g., CTX-II, COMP); (4) lower extremity strength and power; (5) limb proprioception; and (6) abdominal and thigh fat depots measured by CT.

In addition to OA outcome measures, investigators need to select or develop appropriate measures of intervention-related processes and adherence to the intervention.

What is the desirable duration of a trial for prevention?

Recommendation: A primary prevention trial is likely to require a 10-year follow-up with further follow-ups at 1- or 2-year intervals with the interval distance based on time required to detect meaningful differences in the measures of interest and motivate subjects to maintain optimal participation in the trial. Shorter duration trials could be envisioned with improvement in the sensitivity and responsiveness of outcome measures and the OA process.

What is the desirable population of a trial for duration?

Recommendation: In a prevention trial, the optimal study population should be at high risk for future OA but free of full evidence satisfying the disease definition. Therefore, the study population selection is dependent upon the definition of disease that is employed.

A prevention trial study population can be selected to represent the three major domains of disease definition related to OA: 1) structural compromise, 2) pain and other symptoms, and 3) impaired function. Additionally, physiological/immunological locally or systemically measured biomarkers, such as synovial fluid aggrecan, serum C-reactive protein (CRP) or cartilage oligomeric matrix protein (COMP), urinary type II collagen telopeptides (uCTX-II), or combinations of biomarkers, might be incorporated to either define an at-risk population or to exclude individuals from selection into a prevention trial.[4] Further, population selection can be predicated on addressing each of these domains singularly or in combination.[5-7]

If the eligible population for a prevention trial is to be free of structurally-defined OA, one option for defining a “disease-free” population includes enrollment of persons with K-L radiographic grades 0 or 1. Decision-making based on the selection of a population with a K-L score of K-L = 0 vs K-L = 1, which is designated as "doubtful OA," must acknowledge that there is an embedded probability that individuals with a K-L = 1 have early OA, or the underlying conditions leading to OA, but which has not yet been identified definitively on the radiograph. This probability should be factored into estimating the sample size and in the development of data analytic strategies. Efforts are underway to define OA by media other than the standing knee radiograph. For instance, static MRI to define OA based on morphologic changes in cartilage, bone, or other soft tissues or functional magnetic resonance imaging (fMRI) or other types of MRI (such as DGEMRIC, T2- mapping, T1rho, sodium imaging, etc.) to define OA based on compositional changes in cartilage, bone or other soft tissues,[8-10] may become modalities of choice. Currently, there is no agreed upon definition of OA based on these technologies. However, the field is rapidly evolving, and these developments must be anticipated in developing a trial.

If the eligible population lacks characteristics defining symptoms, especially pain or stiffness, the limits of allowable symptoms must be carefully defined, including how pain is to be assessed, its severity and duration, and the allowable frequency for transient pain, and potentially whether or not pain in joints apart from the target joints are considered informative. The use of usual and rescue medications, such as analgesics or NSAIDs also needs to be factored into the methodologic strategy to assess symptoms of OA.[5-7]

If the eligible study population is to be free of functional performance impediments, investigators will need to determine whether inclusion criteria are based on self-report instruments or performance-based assessments. There are numerous questionnaire-based instruments to characterize functional status. For the selection of a study population, it is particularly important to choose an instrument or combination of instruments that have a known specificity (the known probability of truly being free of functional compromise), and that specificity should be relevant to the population from which the prevention trial

population will be recruited. The use of performance-based assessment in prevention trial recruitment is limited by the relative absence of normative data in persons younger than age 65, thereby precluding the ability to estimate the probability of any specific assessment value's actually representing the disease-free state for a prevention trial. Further, there are many determinants of function, which may or may not be directly relevant to OA. Alternatively, these measures may be considered to be estimates of an "at-risk" state and therefore eligible for study in a prevention trial; it is important that the predictive capacity of these performance measures over a period of time for increased compromise be known. [5-7]

If the eligible study population will be selected based on physiological or immunological biomarker measures, there are at least two expectations. First, there must be adequate information to discern when a specific value of the biomarker(s) truly represents a "disease-free" state and, second, information about the rapidity of the biomarker change (if treated as a continuous variable) or conversion (if treated as a discrete variable) in relation to the development of disease, must be known and available. Additionally, the biomarker must have been previously validated against a clinically relevant endpoint for its use as a surrogate measure.[11] Even if the biomarker is used only as a criterion for inclusion or exclusion for participation in a prevention trial, it must have sufficient evidence of predictive relevance to warrant its application.

What is an appropriate safety database for prevention? Is any risk acceptable in a therapy designed to be given to someone with no signs or symptoms of disease?

Recommendation: Because a prevention trial for OA could involve an intervention with active agents administered to otherwise healthy individuals or to individuals with comorbid conditions for extended periods of time, the safety database must be extensive and involve information from multiple organ systems. The extent of this safety database may depend upon the intervention. For example, some interventions may have pleiotropic effects (e.g., statins or bisphosphonates[12-14]), reinforcing the need to monitor multiple organ systems for toxicity. A more localized intervention, such as an unloading brace, might not require the same degree of vigilance for safety in remote organ systems. Observations must also be long in duration, particularly for agents that might impact the immune system and be associated with infections or subsequent development of cancer.

Safety Working Group Comments

The Safety Working Group determined that since "a prevention trial for OA could involve an intervention with active agents administered to otherwise healthy individuals or to individuals with co-morbid conditions for extended periods of time" and the subjects of the study would not yet suffer the disease of interest; for such a preventive therapy used BEFORE a disease state is established there is a different level of acceptance of the potential for AEs than that tolerated in an observed treatment for an established disease state. Thus, consensus was that such a therapeutic would need to be very safe. In this context it is important to keep in mind the concept of benefit to risk and number needed to treat for benefit and number needed to harm. Thus, a preventative therapeutic would need to have a very low number of patients to be treated for clinical benefit with a very large number needed for exposure to lead to harm. At this time, there are no drugs for treating OA once it has been established that fulfill these criteria and certainly no therapeutic to prevent the disease has been developed.

What does prevention or risk reduction mean in terms of a clinical study and therapeutic intervention?

Recommendation: For these purposes, prevention refers to those agents or actions that curtail or delay the onset or new occurrence of clinically diagnosed OA at the joint site of interest in someone initially without evidence satisfying the clinical definition of the condition; components of this definition may include structural evidence (e.g., on radiographs) and characteristic signs and symptoms (e.g., bony enlargement, crepitus, pain).

Risk reduction refers to decreasing specific and modifiable risk factors associated with the development of OA in an attempt to decrease the likelihood of developing OA or to delay its onset. For example, since obesity and overweight are strong risk factors for knee OA, a weight loss intervention could be evaluated to determine its ability to reduce the risk of developing knee OA in the obese. Similarly, since joint trauma, with its frequently resultant altered biomechanics, is a strong risk factor for the development of OA, an intervention to alter abnormal biomechanics in those at risk for or with joint injury could also be considered in a preventive context for OA. [15,16] As our ability to identify high-risk groups earlier and earlier with more sophisticated imaging or molecular biomarkers, it might be possible to prevent the development of abnormal levels of such markers, which themselves are surrogates for the future development of OA.

Because OA is frequently generalized (i.e., affects more than one joint in more than one joint group), an intervention could be applied in someone with OA in one joint site, in order to prevent the development of OA in another joint site unaffected at the start of the trial. For example, those with hand OA could be the subject of a prevention trial to prevent the development of OA in the knees or hips. This situation blurs the distinction between incidence of new disease and progression of established disease and may need to be considered on a case-by-case basis, with statistical methodology applied to allow for the nonindependence of multiple joints within the same person. This also suggests that information about joints beyond the target joint should be collected at the beginning and throughout the trial, both for the purpose of recognizing important secondary effects of the intervention and for identifying potential safety signals of the intervention.

What is the research agenda required to inform each of the above questions?

Recommendation: Observational studies with both short and long-term follow-up can be particularly helpful to define molecular, structural, and symptomatic correlates of disease and to identify risk factors predictive of the development of disease and its clinical impact. Attention to gender and minority inclusion, with the requisite consideration of distinct issues regarding their participation in prevention trials, should be part of this research agenda.

- Evaluation of existing datasets with particularly long follow-up times (10, 20, or more years) in order to identify risk factors that may be exposed long in advance of disease onset
- Evaluation of existing datasets with detailed genetic, biomarker, and imaging data to link to various OA phenotypes along the continuum from molecular to preradiographic OA to radiographic to symptomatic OA
- Addition of short follow-up times (i.e., months) to existing cohorts to obtain sensitive, dynamic imaging and other biomarker data to aid prediction of the development of structural and clinical disease

- Evaluation of distinct ethnic/racial subpopulations to ascertain accurate assessment of the burden of disease in these groups, differences in risk factor profiles, and genetic, imaging, and biomarker subtypes in order to tailor trials to relevant groups, (i.e., differences in BMI that might be used to screen Asians or African Americans into prevention trials for the overweight/obese)
- Methodological study of distinct threats to validity of prevention trials and their execution, related to cultural differences in attitudes toward trial participation and risk factor reduction; techniques to maximize adherence and retention; ways to measure and overcome biases such as preventive misconception and behavioral disinhibition. Studies in other diseases have shown that study participants may have misconceptions about the potential effectiveness of a preventive intervention and/or may have inflated estimates of the likelihood that they will be randomized to get the active agent, and may have exaggerated impressions of the likelihood that the intervention will be personally effective for them. Simon and colleagues have called this the "preventive misconception," defined as "the overestimate in probability or level of personal protection that is afforded by being enrolled in a trial of a preventive intervention."[17] This can be particularly problematic when accompanied by "behavioral disinhibition" or the adoption of behaviors that may pose a risk to the participant or others.[17]

References (updated)

1. Roos EM and Arden NK. Strategies for the prevention of knee osteoarthritis. *NR Rheum* 2016; 12:92-101.
2. Cholesterol Treatment Trialists' (CTT) Collaboration et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomized trials. *Lancet* 2010; 13:376(9753): 1670-1681.
3. West SL, Lok CE, Langsetmo L, et al. Bone mineral density predicts fractures in chronic kidney disease. *J Bone Min Res* 2015; 30(5): 913-919.
4. Mobasheri A and Henrotin Y. Biomarkers of (osteo)arthritis. *Biomarkers*, 2015; 20(8): 513-518.
5. McAlindon TE, Driban JB, Henrotin Y, et al. OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis. *Osteoarthritis Cartilage* 2015; 23: 747-760.
6. Lane NE, Hochberg MC, Nevitt MC, et al. OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for hip osteoarthritis. *Osteoarthritis Cartilage* 2015; 23: 761-771.
7. Kloppenburg M, Maheu E, Kraus VB, et al. OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for hand osteoarthritis. *Osteoarthritis Cartilage* 2015; 23: 772-786.
8. Hunter DJ, Altman RD, Cicuttini F, et al. OARSI Clinical Trials Recommendations: Knee imaging in clinical trials in osteoarthritis. *Osteoarthritis Cartilage* 2015; 23: 698-715.
9. Gold GE, Cicuttini F, Crema MD, et al. OARSI Clinical Trials Recommendations: Hip imaging in clinical trials in osteoarthritis. *Osteoarthritis Cartilage* 2015; 23: 716-731.
10. Hunter DJ, Arden N, Cicuttini F, et al. OARSI Clinical Trials Recommendations: Hand imaging in clinical trials in osteoarthritis. *Osteoarthritis Cartilage* 2015; 23: 732-746.
11. Kraus VB, Blanco FJ, Englund M, et al. OARSI Clinical Trials Recommendations: Soluble biomarker assessments in clinical trials in osteoarthritis. *Osteoarthritis Cartilage* 2015; 686-697.
12. Egom EE, Rose RA, Neyses L, et al. Activation of sphingosine-1-phosphate signaling as a potential underlying mechanism of the pleiotropic effects of statin therapy. *Critical Rev Clin Lab Sci* 2013; 50: 79-89.
13. Giannopoulos S, Katsanos AH, Kosmidou M, et al. Statins and vascular dementia: a review. J Alzheimer's Dis 2014; 42(3): S315-S320.
14. Xing RL, Zhao LR, Wang PM. Bisphosphonates therapy for osteoarthritis: a meta-analysis of randomized controlled trials. *Springerplus* 2016; 5(1):1704.
15. Emery CA, Roos EM, Verhagen E, et al. OARSI Clinical Trials Recommendations: Design and conduct of clinical trials for primary prevention of osteoarthritis by joint injury in sport and recreation. *Osteoarthritis Cartilage* 2015; 23: 815-825.
16. Fitzgerald GK, Hinman RS, Zeni J, et al. OARSI Clinical Trials Recommendations: Design and conduct of clinical trials rehabilitation interventions for osteoarthritis. *Osteoarthritis Cartilage* 2015; 23: 803-814.

17. Simon AE, Wu AW, Lavori PW, Sugarman J. Preventive misconception: its nature, presence, and ethical implications for research Am J Prev Med 2007; 32(5):370-4.

Working Group: Assessment of Structural Change

Questions posed to the Working Group assigned to consider Assessment of Structural Change:

1. For the current tools for assessing structure modification (e.g., x-ray, MRI, biomarkers): What are the performance metrics for each individual feature that they detect? How can they be used optimally in clinical trials? What are the relative strengths and weaknesses of these assessment tools?
2. What do these putative tools measure? How to determine change over time?
3. How can rapid structural progression patients be identified? Is that necessary?
4. What is the relationship between symptoms and structural progression? What is the relationship between disability and measured structural change?
5. Could the need for a joint replacement be a clinical outcome, which might supplant imaging as a measurement?
6. What is the research agenda required to inform each of the above questions?

Summary and Recommendations of the Assessment of Structural Change Working Group

The underlying assumption of these recommendations is that the manifestations of joint pain and disability currently associated with OA are strongly related to the pathophysiology of OA seen in joint structures. This postulate is strongly supported by epidemiological evidence of the association between radiographic OA, joint pain, and disability in the general population. Further, the systematic reviews suggest there is a direct relationship between structural severity of the disease and severity of symptoms (pain and disability), that between-patient variability in symptom severity can be explained by variations in structural severity of OA, and that worsening symptoms of OA can be accounted for by progressive changes of OA in structures of the joint.

Much of the published evidence in this area relates to OA of the knee, with much less evidence (especially for modern imaging modalities) relating to OA of the hip and very limited information available for hand OA. This report must therefore be seen as largely related to trials for OA of the knee and to a lesser extent, the hip. Importantly most of the therapeutic studies on OA have included symptomatic and structural moderate-to-severe OA, but there is an absence of literature and definitions for “early” OA, especially studies entering people before the currently recognized clinical syndrome is apparent and when structural pathology is presumably minimal. Therefore, the literature on the performance of existing imaging modalities in this important area is sparse. When mentioned, the term *therapies* refers to drugs, devices, and biological products entered into the treatment of OA. The summary and recommendations outlined below should be read in conjunction with the subsequent section on Research Recommendations.

Summary and Recommendations

Conventional radiography (CR)

• Conventional radiographs have traditionally been the method of choice in clinical trials because of their relative feasibility.

• CR presents an image of the joint space of a diarthrodial joint, the width of which represents the thickness of articular cartilage. In some joints, notably the knee, JSW also reflects the presence, location, and condition of other structures (e.g., meniscus), and JSW is a composite measure of the combined thickness of those structures.

• Much is now known of the performance metrics of CR JSW in the knee and to a lesser extent in the hip. In the knee, the use of fluoroscopic positioning and semi-flexed views improve responsiveness, although it is acknowledged that access to 50 fluoroscopic facilities is restricted. Studies will generally need to be at least 12 and more likely, 24 months duration.

• It is possible to 'enrich' a study population to increase the rate of joint space loss, for example, by including higher KL grade.

• Automated methods for assessing parameters of JSW offer promise of improved precision, and therefore, improved responsiveness.

• In terms of correlations with concurrent symptoms, there is a weak association between progression in JSN and progression of symptoms. There is little information on how progression in JSN during the course of a study reflects poststudy change in symptoms. JSN progression is associated with increased rate of subsequent total joint replacement, but these may not be truly independent events as JSN is one of the features used to select people for joint replacement surgery.

• The natural history of hip OA appears different to that of knee OA, and although the literature concerning the hip is much less extensive, there is some evidence for better performance metrics for JSW at the hip. Hip JSW as a construct does not include a meniscus. There is little evidence on enriching cohorts for purposes of increasing rate of JSN progression.

Recommendation

For assessing CR JSW, there is some evidence for construct and predictive validity, with good evidence for reliability and responsiveness. At the knee, JSW represents a composite construct, and a semi-flexed acquisition is recommended for knee trials. We support continued use of CR JSW.

We support use of a validated, quantitative JSW method, and care should be taken about adequately powering studies. (**Update)**

MRI

• There has been a growing awareness that symptomatic OA represents a process involving all the tissues in the OA joint. Structure modification should therefore be considered in a broader context than that of cartilage alone. MRI has evolved substantially over the last decade, and its strengths include its ability to visualize individual tissue pathologies, as well as the interrelationship between tissue pathologies.

• Using MRI, it is possible to accurately and feasibly measure change in cartilage morphology over 12 months for knee OA.

• It is possible to 'enrich' a knee OA study population to increase the rate of cartilage loss, for example, by including higher KL grade or using centrally measured metric JSW between 2-4mm (a way to find KL3 by taking variability of reader out). Modern studies have not demonstrated new ways to enrich. **(Update)**

• In terms of correlations with concurrent symptoms, there is a weak association between progression of cartilage loss and increasing symptoms. There is little information on how change in cartilage parameters during the course of a study reflects poststudy change in symptoms. There is some predictive validity with progression of cartilage loss predicting subsequent total joint replacement.

• More information is required on the performance metrics of MRI semi-quantitative and compositional measures of cartilage morphology. There may be a role for semi-quantitative assessments for assessing focal cartilage defects.

• Since MRI alone has the capacity to image the other tissues, further work is needed on the quantification and predictive validity of noncartilage MRI pathologies. The performance metrics of noncartilage MRI features have not been extensively studied but there is a rapidly emerging literature in this field.

Recommendation

For assessing MRI cartilage morphometry in knee OA, there is some evidence for construct and predictive validity, with good evidence for reliability and responsiveness. We recommend inclusion of MRI cartilage morphometry in the next guidance document.

Bone morphology has now demonstrated construct validity, predictive validity for joint replacement, very high reliability, and responsiveness to change (including in an RCT). We recommend inclusion of bone morphometry in the next guidance document. **(Update; Reference: bone morphology update)**

Other imaging modalities

• Ultrasound is currently the other imaging modality with most information available, and at this stage it appears it is most promising as a tool for evaluating OA synovitis. Ultrasound detected pathologies have been associated with current OA symptoms. Further work is required to better understand the performance methods of ultrasonographic quantification of pathology.

Recommendation

The potential for non-CR or MRI modalities to assess relevant noncartilage tissues should be considered.

Reference Document – Bone Morphology Update

Measurement of Bone Shape and Progression of OA

Because knee OA is thought to be a largely mechanically driven process [1], a promising target for an OA imaging biomarker may be to exploit the ability of bone to adapt to mechanical influences [2]. In particular, bone can readily change its shape in response to stresses acting upon it (Wolff's law), suggesting that such alterations may be feasibly assessed in a practical timeframe, making it attractive as a potential imaging end point for trials. Additionally, subtle differences in bone shape or geometry itself could lead to abnormal joint loading and a predisposition to OA.

This method involves the automatic segmentation of MR images using 3D active appearance models, which are a type of shape model designed for searching inside images [3]. Using a statistical shape model in this way allows for the full parameterization of the shape of each subject knee in terms of the population mean and shape variation learnt during the model training phase. This parameterization can then be used to construct a vector of OA versus non-OA shape. This vector, trained on OA and non-OA shapes can identify knees without OA that are at risk of developing OA 12 months later and beyond, and that the position along this vector is associated with OA incidence [4].

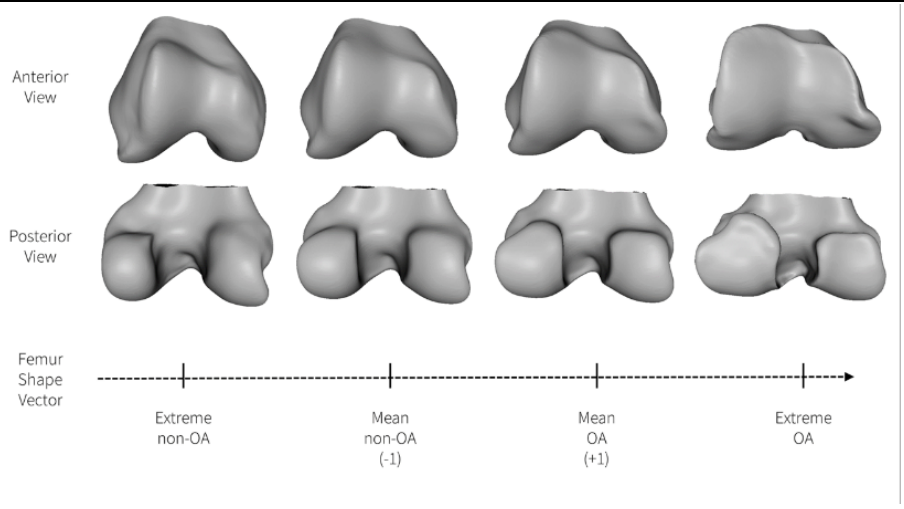

Figure 3: 3-dimensional change in the femur
This figure demonstrates the systematic change in femur shape with OA, represented as a position along a vector of OA shape scaled such that (-1) and (+1) represent the mean non-OA and mean OA shapes respectively.

In a study on the association between bone shape and total knee replacement (TKR), a nested case-control study within the Osteoarthritis Initiative cohort identified case knees with confirmed TKR for OA and controls that were matched using propensity scores. Active appearance modelling quantification of the bone shape of all knee bones identified vectors between knees having or not having OA. Compared to controls (n = 310), TKR cases (n = 310) had a more positive mean baseline 3D bone shape vector, indicating more advanced structural OA, for the femur tibia and patella. Odds ratios (95% CI) for TKR per normalized unit of 3D bone shape vector for the femur, tibia and patella were: 1.85 (1.59, 2.16), 1.64 (1.42, 1.89) and 1.36 (1.22, 1.50), respectively, all P<0.001. After including Kellgren & Lawrence grade in a multivariable analysis, only the femur 3D shape vector remained significantly associated with TKR [5].

Measurement of Changes in Bone Area

It is recognized that OA is a whole joint disease that may involve multiple tissues which confer different phenotypes [6], and subchondral bone is integral to the pathogenesis and progression of OA [6], [7]. The area of subchondral bone at the femorotibial articulation is larger in OA knees than healthy controls, and correlates with knee joint space narrowing, osteophytes and Kellgren Lawrence (KL) grade after adjusting for appropriate confounders in cross-sectional studies. Radiographic measures, derived from a single radiographic projection, are only weakly associated with OA-attributable bone area measured in

3D. This may reflect the additional 3D MRI structural information, unaccounted for by these 2D radiographic measures [8], [9].
The use of 3D statistical shape models means that each bone surface is fitted with a dense set of landmarks during auto-segmentation. The landmarks correspond between subjects and time-points such that they have the same anatomical meaning and position on each bone, and the high density means that the femur model includes over 100 000 points. The density and correspondence of landmarks allows for the accurate analysis of differences in anatomical areas across the bony surfaces. Using this method, we have confirmed that height explains the majority of variance in bone area, confirming an allometric relationship between body and joint size. Radiographic measures of OA, derived from a single radiographic projection, appear to account for only a small amount of variation in 3D knee OA total bone area [10]. Further studies have shown that changes in bone area discriminate people with OA from controls, and are more responsive than the current and impending standards for assessing OA progression [11], see Figures 4 and 5.

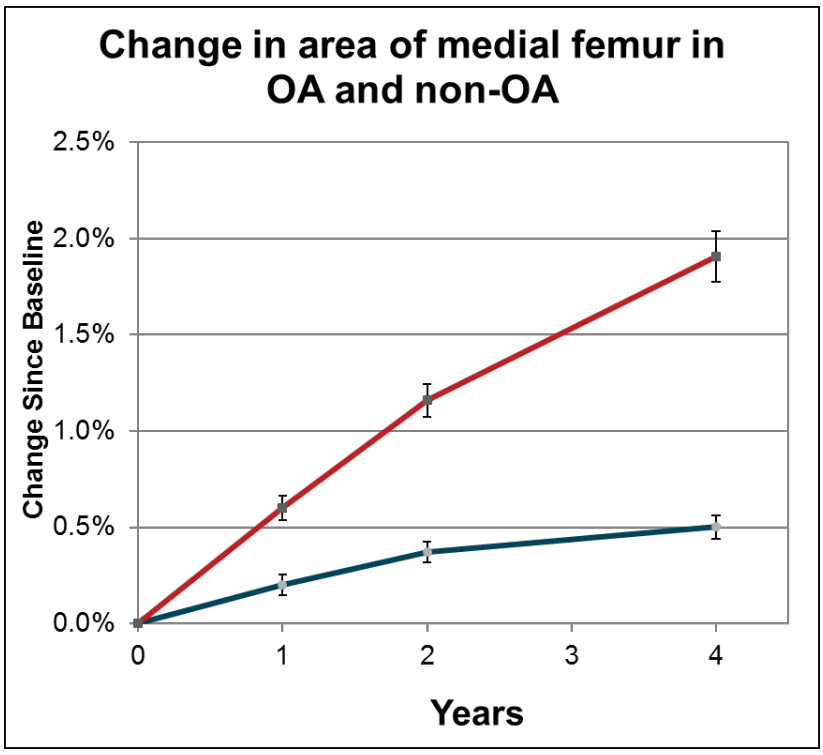

Figure 4: Percentage change in bone area (tAB) for medial regions in OA and non-OA groups. The graph shows percentage change from baseline for OA and non-OA groups, error bars are 95% confidence intervalI. All changes were highly significant ($p<0.0001$).
In a further analysis of data in a subset of 352 participants from the Osteoarthritis Initiative, responsiveness of bone area change was compared with change in radiographic joint space width (JSW) and MRI cartilage thickness over a 2-year period. Responsiveness measured by the standardized response mean (SRM) at 12 months for bone area was 0.66, for JSW it was 0.19 and for cartilage thickness, 0.28. The increased sensitivity of the method is, in part, due to the improved repeatability. AAM segmentations were highly repeatable, with CoVs of less than 1%, compared with around 5% for JSW. Cartilage thickness CoV has been reported from the group performing the measures at around 2–3%. In a clinical trial expected to produce a 50% effect with a double-sided, 80% power, L=0.05 design,

then the cohort size assuming a 1-year trial would be as follows, JSW: 1298 patients per group, cartilage thickness: 459 patients, bone area: 149 patients, see Table 1 for a summary of these comparisons.

Method	Typical CoV	Typical cohort size for 80% power, 50% effect size, L=0.05 study
Radiographic JSW	5%	1298
MR cartilage thickness	2-3%	459
MR bone area	1%	149

Table 1: Responsiveness and typical cohort size for various methods with a 50% effect size

Another recent large-scale study using MRI data from 600 subjects enrolled in the Osteoarthritis Initiative [12] looked at longitudinal validation of bone area measurement and 3D shape as biomarkers for knee OA progression. This study indeed did show that greater increases in bone area and shape markers over 24 months in knees with mild-to-moderate radiographic OA are associated with increased likelihood of clinically relevant progression consisting of a combination of radiographic and symptomatic progression over 48 months. The change in bone found in these studies provides an exciting new window on pathogenesis of the disease, and suggests that bone can now provide a new focus for clinical trials.

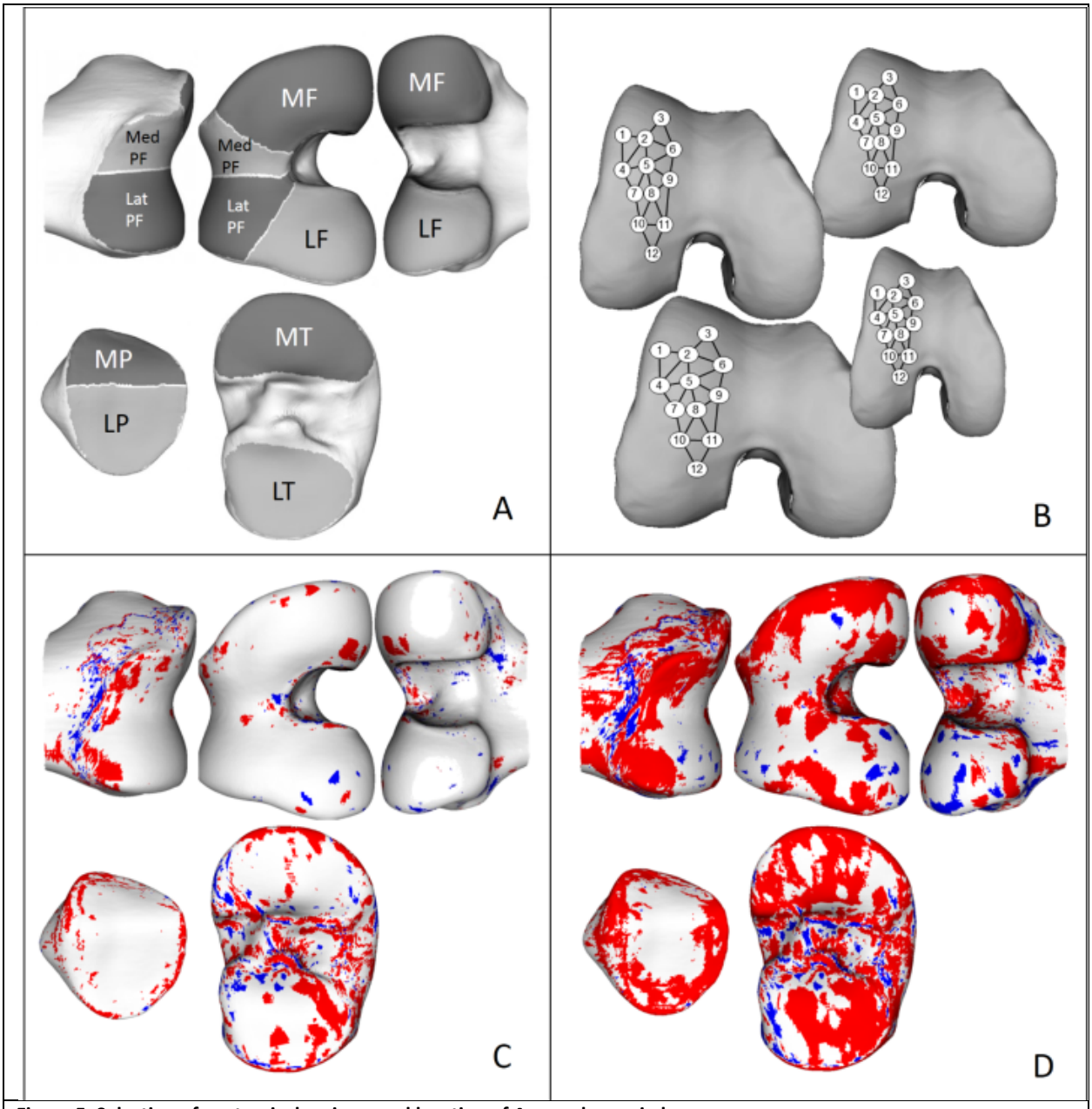

Figure 5. Selection of anatomical regions, and location of 4-year change in bone area
(A) shows the regions used in this study, displayed on the mean shape for each bone. MF, medial femur; LF, lateral femur; MT, medial tibia; LT, lateral tibia; MedPF, medial trochlear femur; LatPF, lateral trochlear femur; MP, medial patella; LP, lateral patella. The MF/MedPF and the LF/LatPF boundaries were defined as a line on the bone corresponding to the anterior edge of the medial or lateral meniscus in the mean model. The MedPF/LatPF boundary was defined as the centre of the trochlear groove in the mean model. (B) shows schematic results for an active appearance model fit to 4 different femurs; each bone surface is fitted with a dense set of landmarks during auto-segmentation, which corrects for individual shape differences. It is impractical to display the actual density of the model, for example, the femur model includes over 100 000 points. Location of 4-year change for the non-OA group is displayed in (C) and for the OA group in (D). Areas which increase in size more than measurement error are coloured red; those with a similar decrease are coloured blue.

References

[1] D. T. Felson *et al.*, “Osteoarthritis: New insights - Part 1: The disease and its risk factors,” in *Annals of Internal Medicine*, 2000, vol. 133, no. 8, pp. 635–646.

[2] S. R. Goldring and M. B. Goldring, “Clinical aspects, pathology and pathophysiology of osteoarthritis,” *J. Musculoskelet. Neuronal Interact.*, vol. 6, no. 4, pp. 376–378, 2006.

[3] T. G. Williams *et al.*, “Anatomically Corresponded Regional Analysis of Cartilage in Asymptomatic and Osteoarthritic Knees by Statistical Shape Modelling of the Bone,” *IEEE Trans. Med. Imaging*, vol. 29, no. 8, pp. 1541–1559, Aug. 2010.

[4] T. Neogi *et al.*, “Magnetic Resonance Imaging-Based Three-Dimensional Bone Shape of the Knee Predicts Onset of Knee Osteoarthritis: Data From the Osteoarthritis Initiative,” *Arthritis Rheum.*, vol. 65, no. 8, pp. 2048–2058, Aug. 2013.

[5] A. J. Barr *et al.*, “The relationship between three-dimensional knee MRI bone shape and total knee replacement—a case control study: data from the Osteoarthritis Initiative,” *Rheumatology*, vol. 55, no. 9, pp. 1585–1593, Sep. 2016.

[6] M. A. Karsdal *et al.*, “The coupling of bone and cartilage turnover in osteoarthritis: opportunities for bone antiresorptives and anabolics as potential treatments?,” *Ann. Rheum. Dis.*, vol. 73, no. 2, pp. 336–348, Feb. 2014.

[7] C. Ding, F. Cicuttini, and G. Jones, “Tibial subchondral bone size and knee cartilage defects: relevance to knee osteoarthritis,” *Osteoarthr. Cartil.*, vol. 15, no. 5, pp. 479–486, May 2007.

[8] G. Jones, C. Ding, F. Scott, M. Glisson, and F. Cicuttini, “Early radiographic osteoarthritis is associated with substantial changes in cartilage volume and tibial bone surface area in both males and females11Sources of support: National Health and Medical Research Council of Australia, Masonic Centenary Medical R,” *Osteoarthr. Cartil.*, vol. 12, no. 2, pp. 169–174, Feb. 2004.

[9] A. E. Wluka, “Tibial plateau size is related to grade of joint space narrowing and osteophytes in healthy women and in women with osteoarthritis,” *Ann. Rheum. Dis.*, vol. 64, no. 7, pp. 1033–1037, Jul. 2005.

[10] A. J. Barr *et al.*, “The relationship between clinical characteristics, radiographic osteoarthritis and 3D bone area: data from the Osteoarthritis Initiative,” *Osteoarthr. Cartil.*, vol. 22, no. 10, pp. 1703–1709, Oct. 2014.

[11] M. A. Bowes, G. R. Vincent, C. B. Wolstenholme, and P. G. Conaghan, “A novel method for bone area measurement provides new insights into osteoarthritis and its progression,” *Ann. Rheum. Dis.*, vol. 74, no. 3, pp. 519–525, Mar. 2015.

[12] D. Hunter *et al.*, “Longitudinal validation of periarticular bone area and 3D shape as biomarkers for knee OA progression? Data from the FNIH OA Biomarkers Consortium,” *Ann. Rheum. Dis.*, vol. 75, no. 9, pp. 1607–1614, Sep. 2016.