FNIH
Foundation for the
National Institutes of Health

November 20, 2018

Beth Walton MBA, PMP, RAC
Regulatory Project Manager
Office of New Drugs (ONO)
Immediate Office (10)
Biomarker Development and Regulatory Science
U.S. Food and Drug Administration

RE: Legacy Biomarker Qualification Project: Status Update for DDT # (DDTBMQ000038)

Dear Ms. Walton,

Attached to this email is a status update on the Letter of Intent submitted by the Foundation for the National Institutes of Health (application DDTBMQ000038) to the FDA in 2015. The team has since made significant progress based on the feedback received from the FDA. We are happy to share our project update per the guidelines requested by you for the legacy biomarker project status. Attached also is the original LOI and specific responses to FDA's comments, wherever possible.

Upon your review of the updated documents, the team would greatly appreciate advice on the project update. Our goal is to address these concerns to FDA's satisfaction before we submit a qualification plan (QP) and full qualification package (FQP) per the new section 507 process

Please do not hesitate to contact me or Dr. Stephanie Cush at scush@fnih.org, should you need additional documents or wish to set up a meeting.

Sincerely,

Joe Menetski, PhD
Associate Vice President, Research Partnerships
Director, Biomarkers Consortium
Foundation for the National Institutes of Health

cc: David Hunter, MBBS, PhD
Virginia Byers-Kraus, MD, PhD
Stephanie Cush, PhD
Steve Hoffmann

11400 Rockville Pike, Suite 600, North Bethesda, MD 20852
phone 301.402.5311 fax 301.480.2752
fnih.org

Legacy Biomarker Qualification Project Status Update[1]

Administrative Information

Requesting Organization

Name: Foundation for the National Institutes of Health (FNIH) Biomarkers Consortium

Address: 11400 Rockville Pike, Suite 600, North Bethesda, MD 20852

Phone: (301) 402-5311

Email: foundation@fnih.org

Website: https://fnih.org/what-we-do/biomarkers-consortium

Primary Contacts

Name: Stephanie Cush, PhD, Scientific Project Manager; Steven Hoffmann, MS., Director, Inflammation and Immunity; and Joseph Menetski, PhD, Associate Vice President of Research Partnerships

Address: 11400 Rockville Pike, Suite 600, North Bethesda, MD 20852

Phone: (301) 594-6649; (301) 443-2102; (301) 594-6596

Emails: scush@fnih.org ; shoffmann@fnih.org ; jmenetski@fnih.org

Alternate Contacts

Name: David Hunter, MBBS, MSc, PhD

Address: Institute of Bone and Joint Research; University of Sydney, Australia

Phone: +61 2 9463 1887

Email: david.hunter@sydney.edu.au

Name: Virginia Byers Kraus, MD, PhD

Address: 300 North Duke St, Durham, NC; Duke University

Phone: (919) 681-6652

Email: kraus004@duke.edu

Submission Date (MM/DD/YYYY): **11/20/2018**

[1] The content you provide in this completed Status Update will be publicly posted as part of the section 507 transparency provisions.

I. Context of Use

A. Biomarker Category

We propose that 11 knee joint magnetic resonance imaging (MRI) markers under five major MRI feature groups be approved by FDA as **prognostic biomarkers** for the enrichment/identification of subjects with knee osteoarthritis who are likely to experience long-term (up to 36 months) disease progression in the absence of treatment.

B. Intended Use in Drug Development

Intended for use in the enrichment/identification of subjects in clinical trials that are likely to experience long-term (up to 36 months) disease progression in the absence of treatment.

C. Context of Use Statement

Primary COU:

Prognostic baseline MRI markers to enrich enrollment/identification of osteoarthritis patients that are likely to experience long-term disease progression in the absence of treatment in order to test disease-modifying drugs for knee osteoarthritis in phase 2 and phase 3 clinical trials.

Secondary or allied COUs:

In future parallel or follow-on biomarker qualification submissions, we anticipate providing evidence packages for FDA review to demonstrate:

1. Prognostic short-term change (baseline to 12 months) in MRI markers to enrich enrollment/identification of osteoarthritis patients who are likely to experience long-term disease progression in the absence of treatment.
2. Prognostic change (baseline to 24 months) in MRI markers to enrich enrollment/identification of osteoarthritis patients who are likely to experience long-term disease progression in the absence of treatment.

II. Drug Development Need

Osteoarthritis (OA) is a highly prevalent, disabling disease, with a commensurate tremendous individual and societal burden [1]. Recent estimates suggest that 250 million people worldwide are affected by knee OA [2]. The risk of mobility disability (defined as needing help walking or climbing stairs) attributable to knee OA alone is greater than that attributable to any other medical condition in people aged 65 years and older [3,4].

Unfortunately, at present, there is no OA equivalent to measuring high lipid levels, hypertension, or high glucose and glucose tolerance, as we have for cardiovascular disease and diabetes, where one can detect and treat the disease precursors preemptively before the associated processes lead to end-organ failure [5,6]. In addition, in OA, even if we had such a biomarker, there are currently no therapies proven to reduce the risk of OA progression [7]; this is due, at least in part, to the lack of qualified biomarkers to intelligently guide OA drug development and OA trial design and conduct. Biomarkers enhance the

success of every phase in the drug development process; they increase the frequency of successful phase transitions (chances of a drug candidate advancing to the next phase of development) [8]. It is estimated that two in four drugs fail in phase 3 trials without biomarkers whereas fewer (i.e., one in four) drug development programs fail with selection biomarkers [8].

Further refinement and improvement of measures of joint structural change are needed to overcome the limited responsiveness of existing imaging biomarkers, such as the poor relationship in individual patients between joint structural pathology (e.g., joint space narrowing on radiographs) and symptomatic disease [5]. To overcome these obstacles, the FNIH OA Biomarkers Consortium undertook an extensive phase 1 biomarker qualification study from 2012–2015 using a nested case-control sample of progressive knee OA within the Osteoarthritis Initiative (OAI) [9]. The overarching project objective was to establish the predictive validity of disease progression biomarkers and assess the responsiveness of several imaging and biochemical markers pertinent to knee OA. The results of this study are now complete, and we are now proposing to pursue phase 2 qualification of the biomarkers in extant clinical trials.

III. Biomarker Information

A. Biomarker Name, Source, Type and Description

Qualification of a set of 11 MRI markers under five major MRI feature groups at baseline to predict knee osteoarthritis progression. DDT # (DDTBMQ000038)

Type of Biomarker (Check relevant type(s))			
	Molecular	√	Radiologic/Imaging
	Histologic		Physiologic Characteristic
	Other (please describe):		

B. For molecular biomarkers, please provide a unique ID.

Scheme/ID: Not applicable.

MRI markers

The MRI markers (Table 1A and 1B) include 11 markers under **five** major MRI feature groups that will be assessed for all levels of proposed qualification on the basis that they performed well in phase 1 multivariable models (manuscript in preparation [12]) to predict longer-term clinical outcome of clinically relevant (pain and radiographic worsening) knee OA progression. These markers are derived by semiquantitative analysis (cartilage morphology, meniscus morphology, synovitis, and osteophytes) and quantitative cartilage morphometry analysis (medial tibio-femoral compartment and central medial femur) of MR images.

Table 1A. Semi-quantitative Imaging Biomarkers Proposed for FDA Approval and Their Performance in Phase 1 of the FNIH OA Biomarkers Consortium Project

Imaging Modality	Biomarkers (units)	Comparators Baseline Mean (SD) Median (n=406)	Cases Baseline Mean (SD) Median (n=194)	Intra/Inter-rater Reliability (weighted kappa)	Biological Process Indicated	Image Analysis Method or Provider	Reference
Cartilage Morphology (MRI)	Maximum thickness score (0–3) 0 1 2 3	 118 (29.1) 83 (20.4) 184 (45.3) 21 (5.2)	 32 (16.5) 44 (22.7) 106 (54.6) 12 (6.2)	0.78 / 0.86	Cartilage degradation	BICL	[13,14]
	N of subregions with thickness score >0 0 1–2 >3	 118 (29.1) 194 (47.8) 94 (23.2)	 32 (16.5) 98 (50.5) 64 (33.0)	—			
	Maximum surface area score (0–3) 0–1 2 3	 34 (8.4) 297 (73.2) 75 (18.5)	 6 (3.1) 133 (68.6) 55 (28.4)	0.87 / 0.89			
	N of subregions with surface area score >0 0–1 2–4 5–7 >8	 40 (9.9) 174 (42.9) 153 (37.7) 39 (9.6)	 5 (2.6) 56 (28.9) 96 (49.5) 37 (19.1)	—			

Imaging Modality	Biomarkers (units)	Comparators Baseline Mean (SD) Median (n=406)	Cases Baseline Mean (SD) Median (n=194)	Intra/Inter-rater Reliability (weighted kappa)	Biological Process Indicated	Image Analysis Method or Provider	Reference
Synovitis (MRI)	Hoffa-synovitis (0–3) 0 1 2–3	 186 (45.8) 190 (46.8) 30 (7.4)	 60 (30.9) 112 (57.7) 22 (11.3)	0.68 / 0.68	Synovitis	BICL	[13,14]
	Effusion-synovitis (0–3) 0–1 2–3	 332 (81.8) 74 (18.2)	 151 (77.8) 43 (22.2)	0.95 / 0.91	Synovitis	BICL	[13,14]
Meniscus Scoring (MRI)	Morphology (maximum score 0–8) 0–1 2–5 6-8	 178 (43.8) 114 (28.1) 114 (28.1)	 75 (38.7) 65 (33.5) 54 (27.8)	0.97 / 0.99	Meniscal damage	BICL	[13,14]
	Extrusion: Medial (maximum score 0–3) 0 1 2 3	 151 (37.4) 119 (29.5) 106 (26.2) 28 (6.9)	 51 (26.3) 57 (29.4) 58 (29.9) 28 (14.4)	0.83 / 0.58 (medial) 0.88 / 0.76 (lateral)	—	BICL	[13,14]
Osteophytes (MRI)	Number of locations affected by osteophyte 0–2 3–6 >6	 81 (20.0) 120 (29.6) 205 (50.5)	 14 (7.2) 39 (20.1) 141 (72.7)	Femoral 0.64 / 0.80 Tibial 0.70 / 0.49 Patella 0.84 / 0.64	Subchondral bone remodeling	BICL	[15]

Abbreviations: BICL = Boston Imaging Core Lab; FNIH = Foundation for the National Institutes of Health; MRI = magnetic resonance imaging; OA = osteoarthritis; SD = standard deviation.

— No findings.

Table 1B. Quantitative Imaging Biomarkers Proposed for FDA Approval and Their Performance in Phase 1 of the FNIH OA Biomarkers Consortium Project

Imaging Modality	Biomarkers (units)	Comparators Baseline Mean (SD) Median (n=406)	Cases Baseline Mean (SD) Median (n=194)	Intra/Inter-rater Reliability (coefficient of variation)	Biological Process Indicated	Image Analysis Method or Provider	Reference
Quantitative Cartilage Morphometry (MRI)	Mean cartilage thickness—central medial femur (ccMF.ThCtAB) [mm] z-score	0.11 (0.89) 0.05	−0.17 (1.14)	3.3%	Cartilage degradation	Chondrometrics	[12,16]
	Mean cartilage thickness—medial tib-fem compartment (MFTC.ThCtAB) [mm] z-score	0.06 (0.94) −0.04	−0.08 (1.09) −0.26	1.3%			

Abbreviations: FNIH = Foundation for the National Institutes of Health; MRI = magnetic resonance imaging; OA = osteoarthritis; SD = standard deviation.

C. Rationale for Biomarker

Mechanistic rationale or biologic plausibility for the biomarker.

Osteoarthritis (OA) is characterized by an active and complex process involving mechanical, inflammatory, and metabolic alterations that may affect multiple joint structures including the hyaline articular cartilage, subchondral bone, synovium, and soft-tissue structures such as the menisci [20]. Change in all these structures has been shown to be associated with clinically relevant progression of the disease. The biomarkers proposed herein, evaluated on MR images, reflect the multitude of simultaneous biological processes involved in disease progression including cartilage degradation (semi-quantitative and quantitative cartilage analysis), synovial inflammation (semi-quantitative analysis), subchondral bone remodeling (semi-quantitative analysis of osteophytes), and meniscal damage.

Natural history of the disease and associated risk factors.

OA is a highly prevalent disease characterized by several steps, including a progressive loss of articular cartilage accompanied by new bone formation and synovial proliferation that may culminate in pain, loss of joint function, and finally, disability [21]. A variety of genetic and environmental risk factors and pathophysiologic processes contribute to the progressive advance of the disease over a period of years, resulting in the typical features of OA: degradation of articular cartilage, osteophyte formation, subchondral sclerosis, meniscal degeneration, bone marrow lesions, and synovial proliferation.

The lifetime risk of knee OA is estimated to range from 14–44% [22,23]. It substantially impacts quality of life and is responsible for elevated healthcare utilization and cost [24] and excess mortality (https://www.oarsi.org/research/oa-serious-disease [25]). Although risk factors have been identified [26,27], disease progression is slow, with periods of structural and symptomatic stasis interposed with periods of worsening [28]. The disease is often characterized by a prolonged pre-symptomatic phase of molecular pathology, a pre-radiographic phase, and a recalcitrant later radiographic phase with evident structural joint changes, frequent pain, and loss of function.

Despite the substantial individual and societal burden, there are few pharmacologic agents to control symptoms beyond analgesic treatments, and no disease-modifying therapies proven to reduce the risk of progression of OA. Instead, the “watchful waiting” of steady decline to end-stage joint disease may typically be followed by knee joint replacement surgery with variable levels of successful outcome [29]. This lack of available therapies is starkly evident in the fact that the vast majority of all knee and hip joint replacements (83% and 79%, respectively) are for osteoarthritis, as well as in the rising rate of knee and hip joint replacements, totaling 860,080 in the US for the 5-year period 2012–2016 [30].

The magnitude and duration of change in the biomarker required to demonstrate a clinically meaningful effect/impact or outcome.

Not applicable. We are proposing prognostic baseline biomarkers to enrich enrollment/identification of osteoarthritis patients that are likely to experience long-term disease progression in the absence of treatment in this submission.

If biomarker is an index/scoring system or a model, please provide information on how the index/model is derived (e.g., algorithm).

In phase 1 of the project, the association and predictive validity of baseline imaging biomarkers and disease progression over 48 months was assessed both individually and in combination in a multivariable model aimed to determine the biomarkers that best describe the risk of future OA progression. The multivariable model was derived using logistic regression to evaluate the association between cases status and biomarkers. The models were evaluated unadjusted, adjusted for covariates (sex, race, baseline minimum joint space width, baseline WOMAC pain score, age, body mass index, KLG, and use of pain medications), and adjusted with 10-fold cross-validation. Three different stepwise selection methods were used to determine the best subset of predictors, selection based on Akaike Information Criterion (AIC), Schwarz Bayesian Criterion (SBC), and p-value (p=0.2 for entry/0.1 for retention). The AIC and SBC differ with respect to model fitting: the AIC tends to favor more complex models that risk overfitting, while the SBC tends to favor less complex models that risk underfitting. Area under the curve (AUC) of the receiver operating characteristic (ROC) curve (C-statistic), the integrated discrimination improvement (IDI), and the category-less net reclassification (NRI) were assessed for each model.

In this phase 2 of the project, we aim to assess the markers with the best performance in phase 1 in a similar manner.

IV. Biomarker Measurement Information

A. General Description of Biomarker Measurement

Semi-quantitative scoring of cartilage morphology, synovitis, meniscus, and osteophytes on MRI will be performed at Boston Imaging Core Lab (BICL) using the MRI Osteoarthritis Knee Score (MOAKS) [15]. Quantitative cartilage segmentation and measurements on MRI will be performed by Chondrometrics GmbH, a company based in Ainring, Bavaria, Germany, which is a leading provider of quantitative medical image analysis services to researchers in academia and the pharmaceutical industry.

B. Test/Assay Information

Indicate whether the biomarker test/assay is one or more of the following:

i. Laboratory Developed Test (LDT) ○ Yes ○ No

ii. Research Use Only (RUO) ○ Yes ○ No

iii. FDA Cleared/Approved ○ Yes √ No
If yes, provide 510(k)/PMA #:
Neither BICL or Chondrometrics have been FDA cleared/approved.

iv. If the biomarker is qualified, will the test/assay be performed in a Clinical Laboratory Improvement Amendments (CLIA)–certified laboratory? n/a
○Yes ○ No

v. Is the biomarker test currently under review by the Center for Devices and Radiological Health or the Center for Biologics Evaluation and Research?

√ Yes √ No √ Don't Know (*See above*)

vi. Is there a standard operating procedure (SOP) for sample collection and storage?

√ Yes ○ No

vii. Is there a laboratory SOP for the test/assay methodology?

√ Yes ○ No

C. Biomarker Measurement

i. Quality Control

Information about quality control material or procedures

The imaging vendors for the phase 2 of this project have previously assessed MR images of several trials that will provide data for this project. These images have appropriate quality standards for reliable reading of data. The imaging vendors are experts in the field and will provide quality control of the MRI readings as also described in the "Image acquisition, analysis, and interpretation" section below. The reliability of the MRI readings is provided in Table 1.

D. Additional Considerations for Imaging Biomarkers

Image acquisition, analysis, and interpretation

Quantitative MRI

MRI acquisition was performed in phase 1 using a 3 Tesla MRI system (Trio, Siemens Healthcare, Erlangen, Germany) at the OAI clinical sites. Additional parameters of the full pulse sequence protocol and sequence parameters have been published in detail [32]. Cartilage thickness measurements were based on manual segmentations by highly trained, experienced, and specialized readers as described previously [28,33]. After quality control of each MR data set by one expert, segmentation of the weight-bearing femorotibial cartilages in paired images was performed, with blinding to acquisition order (baseline vs follow-up) and radiographic status. All segmentations were quality controlled by one expert and were subsequently corrected by the readers, if necessary. Segmentation of the total subchondral bone area (tAB) and the articular cartilage surface area (AC) was performed in the medial and the lateral tibia (MT/LT), and in the central, weight-bearing medial and the lateral femoral condyle (cMF/cLF) using the Chondrometrics Works 3.0 Software.

Semi-quantitative MRI

The semi-quantitative MRI readings in this phase 2 of the project will be done at BICL using the MOAKS method [15]. This method uses an observer-dependent semi-quantitative approach to score a variety of features that are currently believed to be relevant to the functional integrity of the knee and/or potentially involved in the pathophysiology of OA. These instruments for scoring MRI on OA have shown adequate reliability, specificity, and sensitivity, as well as the ability to detect lesion progression over

1–2 years. The reading will be performed using eFilm™ logger and eFilm™ viewer 4.2. Effusion size and synovitis in the infrapatellar and intercondylar regions of Hoffa's fat pad will be scored semi-quantitatively from 0–3. The MOAKS scoring system differentiates meniscal tears from extrusion and provides their localization. The classification of meniscal tears is similar to criteria commonly used when assessing the meniscus during arthroscopy and provides detail on the type of tear. In addition, it separates out signal from tear from maceration. Cartilage thickness and surface area will be scored from 0–3, and the number of subregions with damage will be computed as the number of subregions with a score >0 (possible range 0–14). Across the entire knee and within each compartment (medial, lateral, and patellofemoral), we computed the number of locations affected by any osteophyte (grade >0, range 0–3) and the maximum osteophyte score across all locations.

Assessment of uncertainty including repeatability, reproducibility (e.g., within site, across sites, equipment model/manufacturer) and reader variability.
Systematic literature review reported that the pooled coefficient of variation (CV) for quantitative cartilage volume and thickness was 3% for both inter- and intra-reader reliability. The inter-reader and intra-reader intraclass correlation coefficients (ICCs) for quantitative cartilage measures were both excellent 0.90 (95% CI 0.86, 0.95) and 0.92 (95% CI 0.88, 0.96), respectively [35]. The intra- and inter-rater reliability for semi-quantitative measurements (scoring of cartilage morphology, synovitis, meniscus, and osteophytes on MRI) is summarized in Table 1.

Data to support proposed cutpoint(s) if imaging results are not reported as a continuous variable.
We will initiate analyses under the phase 2 project to examine both the discrete cutpoint approach and continuous variables. We are not at a point at this time where we could advocate for specific cutpoints.

Performance characteristics including sensitivity, specificity, accuracy and agreement.
Responsiveness (sensitivity to change) of quantitative cartilage thickness was analyzed using the standardized response mean (SRM). This measure can be interpreted as the number of standard deviations of change. The pooled SRM [mean change divided by the standard deviation of the change over one year] for quantitative measures of cartilage thickness for the medial tibiofemoral joint was −0.86 (95% CI −1.26 to −0.46) and for lateral tibiofemoral joint was −1.01 (95% CI −2.04 to 0.02) [35]. Recent data have shown that change in mean cartilage thickness is slightly more reliable and responsive to change in knee OA than a change in cartilage volume [36]. Thus the reliability of cartilage thickness on MRI compares favorably to radiographic medial tibiofemoral JSW which has an SRM of 0.33 (95% CI: 0.26 to 0.41) [37].

Device imaging performance characteristics such as resolution, field of view, distortion, contrast, depth of penetration, signal to noise ratio and other imaging parameters as necessary.
The final OAI knee MRI protocol is shown in Tables 5 and 6 (see Appendix I). Subject positioning and scan setup can be found in detail in the OAI MRI Operator's Manual available on the website (https://oai.epi-ucsf.org/datarelease/operationsManuals/MRI_ManualRev.pdf). In brief, the knee MRI acquisition begins with a three-plane localizer, followed by a coronal intermediate-weighted (IW) 2D

turbo spin-echo (TSE) (COR IW 2D TSE) 7 for evaluating the medial collateral ligament (MCL) and lateral collateral ligament (LCL), marginal femoral and tibial osteophytes, medial and lateral meniscal body segments, and presence/extent of subchondral bone cysts and bone attrition. All 2D and 3D coronal acquisitions are oriented coronal to the joint based on anatomic landmarks using a double oblique prescription in order to improve the reproducibility of cross-sectional anatomy depicted on serially acquired MRI exams [32].

The five acquisitions comprising the final OAI MRI protocol were assembled based on the study goals for the imaging protocol, the image evaluation results, and the need to image both knees within a 75-minute time slot, including positioning. For quantitative cartilage morphometry, fat-suppressed 3D DESS acquisitions appear to provide the best universal cartilage discrimination.

However, other machines, sequences, and coils may also be appropriate for biomarkers analysis by BICL and Chondrometrics. For example, the precision of knee cartilage thickness measurement between different MRI acquisition protocols (sagittal DESS and coronal FLASH) was compared in the OAI study and found to be similar, with high correlation for most knee regions (r=0.90–0.97), except for DESS vs. FLASH medial central femur mean cartilage thickness (r=0.81–0.83) [38]. In addition, precision of the FLASH acquisitions at 1.5T and 3T for quantitative cartilage measurements have been shown to be equivalent [38]. Another study using OAI data compared two types of coils (phased-array and quadrature knee coils) and found no significant differences in log (CV%) precision error values and high correlation coefficients between the two coils for cartilage thickness measurements using FLASHwe (r≥0.94) and DESSwe images (r≥0.90) [16]. Analysis of the semi-quantitative MRI markers (MOAKS) can also be done in images acquired by different machines (e.g., Philips, GE, Siemens, Hitachi), field strengths (1.5T or 3T), and coils (quadrature coil, flex coil, birdcage coil, extremity coil).

Algorithms used to interpret the image or data contained in the image. Please provide a full description of these algorithms and validation data or validation plan to confirm the algorithms function as intended.

The semi-quantitative MRI markers reading at BICL will be done according to MOAKS by trained observers using a viewing platform (eFilm™ logger and eFilm™ viewer 4.2). A detailed description of the scoring system (i.e., MOAKS) has been previously published [13–15]. Similarly, femorotibial cartilages are segmented manually by readers for quantitative cartilage assessment by Chondrometrics as previously described [39,40].

Provide the name(s) and version(s) of the software package(s) to be used for image acquisition and analysis.

Chondrometrics: Chondrometrics Works 3.0 Software
BICL: eFilm™ logger and eFilm™ viewer 4.2

V. Assessment of Benefits and Risk

A. Anticipated Benefits

The potential public health benefits of this project are substantial. This project will address several of the most fundamental obstacles to the development of new treatments for osteoarthritis, a disease that presents a large and growing global health burden. The results of these analyses will improve both rheumatologic drug development and testing as well as the clinical use of skeletal and joint repair treatments. For patients with knee joint OA who are considering treatment, this will result in improved treatments, a wider choice of treatments, and better guidance for their clinicians regarding how to use these treatments. Overall, this will improve public health by facilitating the development of new disease-modifying alternatives to knee joint replacement surgeries and by allowing the more effective use of existing osteoarthritis therapies.

Given the potential for more efficient clinical trials and streamlined regulatory approval based on a prognostic enrichment for subjects at increased risk for knee joint disease progression, the benefits for the target group (MRI marker-positive subjects at high risk of progression) could include earlier access to alternative novel treatments that are able to reduce or modify their risk of further knee joint disease morbidity. Perhaps more effective treatments could be developed, compared with the limited number of approved drug or cell therapy interventions now available. Also, since the development costs would not need such large and lengthy trials as are currently needed to prove joint space narrowing reduction, the development costs would be less, ultimately resulting in lower costs of treatment for patients.

B. Anticipated Risks

1. Given the lack of available disease-modifying drugs, the current risk is the failure to develop a more robust clinical trial pathway for OA drug development. This project only mitigates this risk.
2. If a treatment that is shown to preserve joint space cartilage and bone-imaging features and/or systemic biochemical markers of skeletal metabolism results in regulatory approval but is not, in fact, effective in reducing the risk of progression of arthropathy, then the relative risk is dependent upon the drug adverse effect profile.
3. Given the current paradigm and unmet need for disease modification in OA, the greater risk is in type II error—i.e., effective drugs that cannot be shown to be effective due to insensitive clinical trial outcome measures. This project helps to mitigate this risk.
4. There may be safety issues with the drug that are not apparent in the small trials that would be sufficient for approval based solely on knee OA biomarker cohort enrichment. This risk would be mitigated by performing larger and longer studies, perhaps in a post-marketing context, to ensure adequate safety of the drug. The specific size, design, and duration of these trials (or studies) would need to be negotiated between the sponsor and FDA.

C. Risk Mitigation Strategy

If the baseline threshold imaging OA marker progression score is insufficiently prognostic to reliably stratify or enrich the subjects for a definitive clinical trial, a higher threshold score or a longer duration of follow-up may have to be invoked.

D. Conclusions

On balance, the risk-benefit calculus is favorable in the setting of knee OA when MRI markers are utilized to assess the progression of joint disease when compared with the prognostic method used currently, radiographic JSN. When employed for the enrichment of cohorts likely to progress to disability without treatment or to require total knee replacement, these improved biomarkers should reduce the size and duration of clinical trials evaluating disease-modifying candidate therapies.

VIII. Evaluation of Biomarker in Data Collection

Describe the available data used in support of the pre-clinical and/or clinical application of the biomarker for the proposed COU, as well as any ongoing or planned data collection in the relevant sections below.

A. Completed Pre-Clinical Information

Not applicable.

B. Ongoing Data Collection (pre-clinical and clinical)

Please provide an overall timeline for completion of these studies. Include discussion of what gaps these studies are intended to fill in the support of the biomarker for the proposed COU.

The phase 2 OA project will pursue qualification of the previously explored biomarkers in a number of independent validation cohorts and completed clinical trials. The core hypothesis for this study is that single and combinatorial MRI biomarkers will be more prognostic for clinically relevant anatomic and symptomatic worsening of disease than current methods of identifying individuals at high risk of progression based on gender, age, and body mass index. We also hypothesize that select imaging biomarkers will be more sensitive indicators of OA progression than the existing reference standard biomarker, radiographic joint space width. MRI data are available from several of these trials (see Table 4 below). These trials tested a range of therapeutic interventions and the data sets provide a rich resource for further investigations. The number of pain progressors, defined as an increase of ≥9 points on the WOMAC pain subscale (0–100 scale), and radiographic progressors (defined as JSN ≥0.7 mm over the follow-up period) are available for each trial. The method used by each study to measure JSW has been recorded; however, we plan to generate a uniform JSW measurement across all trials with KneeTool (Optasia Medical) radiographic image analyses at baseline and study end point for all included participants. Successful validation of any of these biomarkers in this phase 2 OA FNIH study will therefore likely provide sufficient data for FDA

qualification review and approval. The expectation is that these analyses will be completed over the next 3 years in keeping with the FNIH phase 2 OA Project Plan timeline.

Table 4 Relevant Data Available from Retrospective Knee OA Disease-Modifying Therapy Trials to Be Used in Biomarker Validation Analysis

Trial Name	Agent	Sponsor	Trial duration	MRI images suitable	X-ray at follow up	MRI already assessed	Biospecimens available for assay	Sample size / Total N of MRI / biospecimen data	Type of Biomarker Qualification Possible
Cindunistat (NCT00565812)	Cindunistat	Pfizer	24 months	Yes	Yes	Yes	No	MRI: 27	MRI: bl-PROG, change-PROG
Calcitonin (NCT00486434, NCT00704847)	Calcitonin	Novartis	24 months	No	Yes	No	Yes	Serum: 809 Urine: 660	Serum and urine: bl-PROG; change-PROG
VIDEO (Arden) (ISRCTN94818153)	VIDEO-Vitamin D	Oxford	36 months	Yes	Yes	No	Yes	MRI: 69	MRI: bl-PROG; change-PROG
								Plasma, Serum, Urine: 237	Serum and urine: bl-PROG; change-PROG
Sprifermin (NCT01033994)	Sprifermin	Merck SERONO	12 months	Yes	Yes	Yes	No	MRI: 48	MRI: Bl-PROG; change-PROG
Sprifermin II (NCT01919164)	Sprifermin	Merck SERONO	36 months	Yes	Yes	Yes	No	MRI: 108	MRI: Bl-PROG; change-PROG
TissueGene-C (NCT02072070)	TissueGene-C	Invossa/ Kolon Life Science	12 months	Yes	Yes	No	No	MRI: 81	MRI: bl-PROG; change-PROG;
Strontium Ranelate (SEKOIA) (NCT02072070)	Strontium ranelate	Servier	36 months	Yes	Yes	Yes	No	MRI: 120	MRI: bl-PROG; pharma-EFF
Total								**MRI = 453** **Serum = 1046** **Urine = 660**	

Abbreviations: MRI = magnetic resonance imaging; OA = osteoarthritis.

X. Knowledge Gaps in Biomarker Development

A. List and describe any knowledge gaps, including any assumptions that exist in the application of the biomarker for the proposed COU.

There are a few knowledge gaps in the field of OA biomarkers research related to the proposed COU [9]. First, the positive and negative predictive values for specific cutoff values of the various biomarkers are not yet known. Second, there is a lack of consensus regarding the optimal surrogate measures and the definition of a meaningful clinical end point in OA. Third, OA is extraordinarily complex with marked heterogeneity in onset, clinical presentation, rate of disease progression, pattern of joint involvement, and synovial tissue structure affected.

B. List and describe the approach/tools you propose to use to fill in the above-named gaps when evidence is unknown or uncertain (i.e., statistical measures and models, meta-analysis from other clinical trials).

We aim to investigate different cutoffs of the biomarkers proposed and to perform analyses using different definitions of progression (i.e., clinical and radiographic, clinical only, and radiographic

only). We will investigate the biomarkers individually and also in combination, as it is possible that a combination of markers will be more useful for prediction of OA progression than a single marker.

C. Describe the status of other work currently underway and planned for the future toward qualification of this biomarker for the proposed context of use.

In the future, we aim to analyze the 12- and 24-month time points of each study as well as the treatment arms of these trials to test (1) whether these more sensitive biomarkers identify treatment benefits not recognized with the less sensitive radiographic end points, and (2) whether a reanalysis of subjects, selected on the basis of a biomarker(s) cutoffs at baseline, yield a sample set showing drug benefit on the basis of the radiographic outcomes.

Additional parallel analyses of phase 1 data to inform phase 2 analyses when the data are available, such as mimicking a clinical trial screening process and modeling the cost savings, would add helpful information for the real-world application of the proposed biomarkers. Secondly, as biochemical alterations also occur in the joint and can be measured at the systemic level, 9 serum and urinary markers with the greatest predictive ability among the 18 biochemical markers tested in phase 1 will be assessed in phase 2 of this project. Analysis of XR images will also occur in parallel to the current proposed plan and are planned to be submitted to the FDA in a separate biomarker qualification submission.

Appendix I

Table 5 Final OAI Knee MRI Protocol Acquisition Time (min)

	Scan	Right Knee	Left Knee	Total
1	Localizer (3-plane)	0.5	0.5	1.0
2	COR IW 2D TSE	3.4	3.4	6.8
3	SAG 3D DESS WE	10.6	10.6	21.2
4	COR MPR SAG 3D DESS WE	0.0	0.0	0.0
5	AXIAL MPR SAG 3D DESS WE	0.0	0.0	0.0
6	COR T1W 3D FLASH WE*	8.6	–	8.6
7	SAG IW 2D TSE FS	4.7	4.7	9.4
8	SAG 2D MESE*	10.6	–	10.6
Total		38.4	19.2	57.6

*Acquired on only right knee, unless right knee contains metal in which case, acquired on only left knee.

Table 6 Final OAI Knee MRI Protocol Acquisition Parameters

Scan	Localizer	COR IW 2D TSE	SAG 3D DESS WE	COR T1W 3D FLASH WE	SAG 2D MESE	SAG IW 2D TSE FS
Plane	3-plane	Coronal	Sagittal	Coronal	Sagittal	Sagittal
FS	No	No	WE	WE	No	FS
Matrix (phase)	128	307	307	512	269	313
Matrix (frequency)	256	384	384	512	384	448
No. of slices	21	35	160	80	21	37
FOV (mm)	200	140	140	160	120	160
Slice thickness/gap (mm/mm)	5/1	3/0	0.7/0	1.5/0	3/0.5	3/0
Flip angle (°)	40	180	25	12	n/a	180
TE/TR (ms/ms)	5/10	29/3700	4.7/16.3	7.57/20	10, 20, 30, 40, 50, 60, 70/2700	30/3200
Bandwidth (Hz/pixel)	250	352	185	130	250	248
Chemical shift (pixels)	1.8	1.3	0	0	1.8	0

Scan	**Localizer**	**COR IW 2D TSE**	**SAG 3D DESS WE**	**COR T1W 3D FLASH WE**	**SAG 2D MESE**	**SAG IW 2D TSE FS**
No. excitations averaged	1	1	1	1	1	1
ETL	1	7	1	1	1	5
Phase encode axis	A/P, R/L	R/L	A/P	R/L	A/P	A/P
Distance factor (%)	50	0	0	0	16	0
Phase oversampling	0	20	0	0	0	40
Slice oversampling	0	0	10	0	0	0
Phase resolution	50	80	80	100	70	70
Phase partial Fourier (8/8 = 1)	1	1	1	1	0.875	1
Readout partial Fourier (8/8 = 1)	1	1	1	1	1	1
Slice partial Fourier (8/8 = 1)	1	1	0.75	0.75	0.75	1
X-resolution (mm)	0.391	0.365	0.365	0.313	0.313	0.357
Y-resolution (mm)	0.781	0.456	0.456	0.313	0.446	0.511

Appendix II: Publications/References

1. Hunter DJ, Schofield D, Callander E. The individual and socioeconomic impact of osteoarthritis. Nat Rev Rheumatol 2014;10:437–41.
2. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, Abraham J, Ackerman I, Aggarwal R, Ahn SY, Ali MK, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Bahalim AN, Barker-Collo S, Barrero LH, Bartels DH, Basanez MG, Baxter A, Bell ML, Benjamin EJ, Bennett D, Bernabe E, Bhalla K, Bhandari B, Bikbov B, Bin Abdulhak A, Birbeck G, Black JA, Blencowe H, Blore JD, Blyth F, Bolliger I, Bonaventure A, Boufous S, Bourne R, Boussinesq M, Braithwaite T, Brayne C, Bridgett L, Brooker S, Brooks P, Brugha TS, Bryan-Hancock C, Bucello C, Buchbinder R, Buckle G, Budke CM, Burch M, Burney P, Burstein R, Calabria B, Campbell B, Canter CE, Carabin H, Carapetis J, Carmona L, Cella C, Charlson F, Chen H, Cheng AT, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahiya M, Dahodwala N, Damsere-Derry J, Danaei G, Davis A, De Leo D, Degenhardt L, Dellavalle R, Delossantos A, Denenberg J, Derrett S, Des Jarlais DC, Dharmaratne SD, Dherani M, Diaz-Torne C, Dolk H, Dorsey ER, Driscoll T, Duber H, Ebel B, Edmond K, Elbaz A, Ali SE, Erskine H, Erwin PJ, Espindola P, Ewoigbokhan SE, Farzadfar F, Feigin V, Felson DT, Ferrari A, Ferri CP, Fevre EM, Finucane MM, Flaxman S, Flood L, Foreman K, Forouzanfar MH, Fowkes FG, Franklin R, Fransen M, Freeman MK, Gabbe BJ, Gabriel SE, Gakidou E, Ganatra HA, Garcia B, Gaspari F, Gillum RF, Gmel G, Gosselin R, Grainger R, Groeger J, Guillemin F, Gunnell D, Gupta R, Haagsma J, Hagan H, Halasa YA, Hall W, Haring D, Haro JM, Harrison JE, Havmoeller R, Hay RJ, Higashi H, Hill C, Hoen B, Hoffman H, Hotez PJ, Hoy D, Huang JJ, Ibeanusi SE, Jacobsen KH, James SL, Jarvis D, Jasrasaria R, Jayaraman S, Johns N, Jonas JB, Karthikeyan G, Kassebaum N, Kawakami N, Keren A, Khoo JP, King CH, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lalloo R, Laslett LL, Lathlean T, Leasher JL, Lee YY, Leigh J, Lim SS, Limb E, Lin JK, Lipnick M, Lipshultz SE, Liu W, Loane M, Ohno SL, Lyons R, Ma J, Mabweijano J, MacIntyre MF, Malekzadeh R, Mallinger L, Manivannan S, Marcenes W, March L, Margolis DJ, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGill N, McGrath J, Medina-Mora ME, Meltzer M, Mensah GA, Merriman TR, Meyer AC, Miglioli V, Miller M, Miller TR, Mitchell PB, Mocumbi AO, Moffitt TE, Mokdad AA, Monasta L, Montico M, Moradi-Lakeh M, Moran A, Morawska L, Mori R, Murdoch ME, Mwaniki MK, Naidoo K, Nair MN, Naldi L, Narayan KM, Nelson PK, Nelson RG, Nevitt MC, Newton CR, Nolte S, Norman P, Norman R, O'Donnell M, O'Hanlon S, Olives C, Omer SB, Ortblad K, Osborne R, Ozgediz D, Page A, Pahari B, Pandian JD, Rivero AP, Patten SB, Pearce N, Padilla RP, Perez-Ruiz F, Perico N, Pesudovs K, Phillips D, Phillips MR, Pierce K, Pion S, Polanczyk GV, Polinder S, Pope CA, 3rd, Popova S, Porrini E, Pourmalek F, Prince M, Pullan RL, Ramaiah KD, Ranganathan D, Razavi H, Regan M, Rehm JT, Rein DB, Remuzzi G, Richardson K, Rivara FP, Roberts T, Robinson C, De Leon FR, Ronfani L, Room R, Rosenfeld LC, Rushton L, Sacco RL, Saha S, Sampson U, Sanchez-Riera L, Sanman E, Schwebel DC, Scott JG, Segui-Gomez M, Shahraz S, Shepard DS, Shin H, Shivakoti R, Singh D, Singh GM, Singh JA, Singleton J, Sleet DA, Sliwa K, Smith E, Smith JL, Stapelberg NJ, Steer A, Steiner T, Stolk WA, Stovner LJ, Sudfeld C, Syed S, Tamburlini G, Tavakkoli M, Taylor HR, Taylor JA, Taylor WJ, Thomas B, Thomson WM, Thurston GD, Tleyjeh IM, Tonelli M, Towbin JA, Truelsen T, Tsilimbaris MK, Ubeda C, Undurraga EA, van der Werf MJ, van Os J, Vavilala MS, Venketasubramanian N, Wang M, Wang W, Watt K, Weatherall DJ, Weinstock MA, Weintraub R, Weisskopf MG, Weissman MM, White RA, Whiteford H, Wiersma ST, Wilkinson JD, Williams HC, Williams SR, Witt E, Wolfe F, Woolf AD, Wulf S, Yeh PH, Zaidi AK, Zheng ZJ, Zonies D, Lopez AD, Murray CJ, AlMazroa MA, Memish ZA. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380:2163–96.
3. Prevalence of disabilities and associated health conditions among adults--United States, 1999. MMWR Morb Mortal Wkly Rep 2001;50:120–5.

4. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, Kelly-Hayes M, Wolf PA, Kreger BE, Kannel WB. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. Am J Public Health 1994;84:351–8.
5. Hunter DJ, Losina E, Guermazi A, Burstein D, Lassere MN, Kraus V. A pathway and approach to biomarker validation and qualification for osteoarthritis clinical trials. Curr Drug Targets 2010;11:536–45.
6. Kraus VB, Nevitt M, Sandell LJ. Summary of the OA biomarkers workshop 2009--biochemical biomarkers: biology, validation, and clinical studies. Osteoarthritis Cartilage 2010;18:742–5.
7. Matthews GL, Hunter DJ. Emerging drugs for osteoarthritis. Expert Opin Emerg Drugs 2011;16:479–91.
8. Thomas DW, Burns J, Audette J, Carroll A, Dow-Hygelund C, Hay M. Clinical Development Success Rates 2006-2015. BIO, Biomedtracker, and Amplion, 2016. https://www.bio.org/sites/default/files/Clinical%20Development%20Success%20Rates%202006-2015%20-%20BIO,%20Biomedtracker,%20Amplion%202016.pdf.
9. Hunter DJ, Nevitt M, Losina E, Kraus V. Biomarkers for osteoarthritis: current position and steps towards further validation. Best Pract Res Clin Rheumatol 2014;28:61–71.
10. Kraus VB, Collins JE, Hargrove D, Losina E, Nevitt M, Katz JN, Wang SX, Sandell LJ, Hoffmann SC, Hunter DJ. Predictive validity of biochemical biomarkers in knee osteoarthritis: data from the FNIH OA Biomarkers Consortium. Ann Rheum Dis 2017;76:186–95.
11. Kraus VB, Hargrove DE, Hunter DJ, Renner JB, Jordan JM. Establishment of reference intervals for osteoarthritis-related soluble biomarkers: the FNIH/OARSI OA Biomarkers Consortium. Ann Rheum Dis 2017;76:179–85.
12. Hunter DJ, Collins JE, Losina E, Nevitt MC, Roemer FW, Guermazi A, Bowes M, Dam E, Eckstein F, Hoffmann S, Deveza L, Lynch JA, Katz JN, Kwoh CK, Kraus VB. Multivariable modelling of biomarker data from the phase 1 foundation of NIH osteoarthritis biomarkers consortium. Manuscript in process 2018.
13. Collins JE, Losina E, Nevitt MC, Roemer FW, Guermazi A, Lynch JA, Katz JN, Kent Kwoh C, Kraus VB, Hunter DJ. Semiquantitative imaging biomarkers of knee osteoarthritis progression: Data from the Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Rheumatol 2016;68:2422–31.
14. Roemer FW, Guermazi A, Collins JE, Losina E, Nevitt MC, Lynch JA, Katz JN, Kwoh CK, Kraus VB, Hunter DJ. Semi-quantitative MRI biomarkers of knee osteoarthritis progression in the FNIH biomarkers consortium cohort - Methodologic aspects and definition of change. BMC Musculoskelet Disord 2016;17:466.
15. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, Roemer FW. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage 2011;19:990–1002.
16. Eckstein F, Kunz M, Hudelmaier M, Jackson R, Yu J, Eaton CB, Schneider E. Impact of coil design on the contrast-to-noise ratio, precision, and consistency of quantitative cartilage morphometry at 3 Tesla: a pilot study for the osteoarthritis initiative. Magn Reson Med 2007;57:448-54.
17. Hunter D, Nevitt M, Lynch J, Kraus VB, Katz JN, Collins JE, Bowes M, Guermazi A, Roemer FW, Losina E. 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 2016;75:1607–14.
18. Kraus VB, Collins JE, Charles HC, Pieper CF, Whitley L, Losina E, Nevitt M, Hoffmann S, Roemer F, Guermazi A, Hunter DJ. Predictive validity of radiographic trabecular bone texture in knee osteoarthritis: The Osteoarthritis Research Society International/Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Rheumatol 2018;70:80–87.
19. Kraus VB, Feng S, Wang S, White S, Ainslie M, Brett A, Holmes A, Charles HC. Trabecular morphometry by fractal signature analysis is a novel marker of osteoarthritis progression. Arthritis Rheum 2009;60:3711–22.
20. Loeser RF, Goldring SR, Scanzello CR, Goldring MB. Osteoarthritis: a disease of the joint as an organ. Arthritis Rheum 2012;64:1697-707.
21. Migliore A, Massafra U. Towards the identification of early stage osteoarthritis. Clin Cases Miner Bone Metab 2014;11:114–6.

22. Murphy L, Schwartz TA, Helmick CG, Renner JB, Tudor G, Koch G, Dragomir A, Kalsbeek WD, Luta G, Jordan JM. Lifetime risk of symptomatic knee osteoarthritis. Arthritis Rheum 2008;59:1207–13.
23. Losina E, Weinstein AM, Reichmann WM, Burbine SA, Solomon DH, Daigle ME, Rome BN, Chen SP, Hunter DJ, Suter LG, Jordan JM, Katz JN. Lifetime risk and age at diagnosis of symptomatic knee osteoarthritis in the US. Arthritis Care Res (Hoboken) 2013;65:703–11.
24. Wright EA, Katz JN, Cisternas MG, Kessler CL, Wagenseller A, Losina E. Impact of knee osteoarthritis on health care resource utilization in a US population-based national sample. Med Care 2010;48:785–91.
25. OARSI.OA as a Serious Disease. *In* ", Vol. 2018. OARSI, https://www.oarsi.org/research/oa-serious-disease, 2013.
26. Belo JN, Berger MY, Reijman M, Koes BW, Bierma-Zeinstra SM. Prognostic factors of progression of osteoarthritis of the knee: a systematic review of observational studies. Arthritis Rheum 2007;57:13–26.
27. Collins JE, Katz JN, Dervan EE, Losina E. Trajectories and risk profiles of pain in persons with radiographic, symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage 2014;22:622–30.
28. Eckstein F, Nevitt M, Gimona A, Picha K, Lee JH, Davies RY, Dreher D, Benichou O, Le Graverand MP, Hudelmaier M, Maschek S, Wirth W. Rates of change and sensitivity to change in cartilage morphology in healthy knees and in knees with mild, moderate, and end-stage radiographic osteoarthritis: results from 831 participants from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2011;63:311–9.
29. Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. BMJ Open 2012;2:e000435.
30. American Joint Replacement Registry (AJRR). Annual Report 2017. http://www.ajrr.net/images/annual_reports/AJRR-2017-Annual-Report---Final.pdf.
31. Kraus VB, Burnett B, Coindreau J, Cottrell S, Eyre D, Gendreau M, Gardiner J, Garnero P, Hardin J, Henrotin Y, Heinegard D, Ko A, Lohmander LS, Matthews G, Menetski J, Moskowitz R, Persiani S, Poole AR, Rousseau JC, Todman M. Application of biomarkers in the development of drugs intended for the treatment of osteoarthritis. Osteoarthritis Cartilage 2011;19:515–42.
32. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis Cartilage 2008;16:1433–41.
33. Wirth W, Nevitt M, Hellio Le Graverand MP, Benichou O, Dreher D, Davies RY, Lee J, Picha K, Gimona A, Maschek S, Hudelmaier M, Eckstein F. Sensitivity to change of cartilage morphometry using coronal FLASH, sagittal DESS, and coronal MPR DESS protocols--comparative data from the Osteoarthritis Initiative (OAI). Osteoarthritis Cartilage 2010;18:547–54.
34. Bowes MA, Vincent GR, Wolstenholme CB, Conaghan PG. A novel method for bone area measurement provides new insights into osteoarthritis and its progression. Ann Rheum Dis 2015;74:519-25.
35. Hunter DJ, Zhang W, Conaghan PG, Hirko K, Menashe L, Reichmann WM, Losina E. Responsiveness and reliability of MRI in knee osteoarthritis: a meta-analysis of published evidence. Osteoarthritis Cartilage 2011;19:589–605.
36. Hudelmaier M, Wirth W, Wehr B, Kraus V, Wyman BT, Hellio Le Graverand MP, Eckstein F. Femorotibial cartilage morphology: reproducibility of different metrics and femoral regions, and sensitivity to change in disease. Cells Tissues Organs 2010;192:340–50.
37. Reichmann WM, Maillefert JF, Hunter DJ, Katz JN, Conaghan PG, Losina E. Responsiveness to change and reliability of measurement of radiographic joint space width in osteoarthritis of the knee: a systematic review. Osteoarthritis Cartilage 2011;19:550–6.
38. Schneider E, Nevitt M, McCulloch C, Cicuttini FM, Duryea J, Eckstein F, Tamez-Pena J. Equivalence and precision of knee cartilage morphometry between different segmentation teams, cartilage regions, and MR acquisitions. Osteoarthritis Cartilage 2012;20:869-79.
39. Eckstein F, Collins JE, Nevitt MC, Lynch JA, Kraus VB, Katz JN, Losina E, Wirth W, Guermazi A, Roemer FW, Hunter DJ. Brief report: Cartilage thickness change as an imaging biomarker of knee osteoarthritis progression: Data from the Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Rheumatol 2015;67:3184–9.

40. Wirth W, Hunter DJ, Nevitt MC, Sharma L, Kwoh CK, Ladel C, Eckstein F. Predictive and concurrent validity of cartilage thickness change as a marker of knee osteoarthritis progression: data from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2017;25:2063–71.

Attachments

Optional*: If you have other supporting information you would like to provide, please submit as attachment(s).

Attachment 1: Original LOI
Attachment 2: Combined BQRT comments
Attachment 3: Response letter to the FDA comments on the LOI
Attachment 4: Multivariable manuscript DRAFT- Phase I Results
*Optional information will not be posted publicly.

Letter of Intent to Propose Biomarker Qualification

1. **Administrative structure**

Description of the Submitter including, but not limited to Principal Investigator(s), Working Group Member(s), institutions, and contact information not contained within the cover letter

Principal Investigator

David Hunter MBBS, MSc (Clin Epi), M SpMed, PhD, FRACP (Rheum)
Florance and Cope Chair of Rheumatology, Professor of Medicine
Institute of Bone and Joint Research and Kolling Institute, University of Sydney

Working Group

Felix Eckstein, Chondrometrics

Colin Miller, BioClinica

Michael Nevitt, UCSF

2. **Biomarker Qualification Overview**

a. **Introduction**

Osteoarthritis (OA), the most common of all arthritides, is a heterogeneous (in etiology and spatial pattern) disease characterized by failure of the synovial joint organ. The risk of mobility disability (defined as needing help walking or climbing stairs) attributable to knee OA alone is greater than that due to any other medical condition in people aged 65 and over (1;2). Recent estimates suggest the global burden of knee osteoarthritis affects approximately 250 million people (3). Although ageing is a significant risk factor, the majority of those affected with OA (64%) are of working age (15-64 years) accounting for 11% of the workforce (4;5). There are presently no therapies approved by regulatory authorities that modify the onset or progression of OA structural damage, and available symptom-modifying (analgesic) treatments have only moderate long-term effect sizes with the majority of patients dissatisfied with their efficacy (6;7). As a result of the failure of pharmacological approaches to manage the condition, the number of joint replacement surgeries, over 95% of which are done for OA, is increasing by ~10% annually. In the US alone, the financial burden has been estimated to be $81 billion in medical costs and $128 billion in total cost, given approximately 21 million people with OA associated limitations, 36 million outpatient visits and 750,000 hospitalizations per year (8). This formidable individual and socioeconomic impact of OA will continue to grow as the population ages and obesity

rates continue to grow, with the number of persons affected predicted to double by 2020 (4;9).

It is clear that finding effective disease- and symptom-modifying therapies for OA is a global unmet need whose amelioration should be an international medical priority. There have been major research advances that have significantly increased our understanding of the molecular pathophysiology of joint destruction and pain in OA. Despite this pre-clinical progress however, no new structure-modifying therapies have translated into treatments for patients. Indeed, the recent failure of a number of phase II and III clinical trials for OA structure-modifying drugs has resulted in a considerable decline in the number and size of pharmaceutical company research programs in this area (6). The reasons for the translational failure of anti-OA drugs are likely multifold, but include the poor relationship in individual patients between joint structural pathology (especially joint space narrowing on radiographs) and symptomatic disease, and limited responsiveness of existing biomarkers (10).

The draft regulatory (FDA) guidance and current gold standard for measuring clinical efficacy in disease modifying therapy development in OA is radiographic joint space narrowing (JSN) (11). From JSN outcomes the health, integrity and thickness of hyaline articular cartilage are inferred (12;13). This FDA guidance describes a process for drug approval for specific indications in OA, including treatment of symptoms, delays in structural progression and even discusses prevention of OA. The JSN measure is currently recommended by both the FDA and European Agency for the Evaluation of Medicinal Products (EMEA) guidance documents as the imaging endpoint for clinical trials of disease-modifying OA drugs (DMOADs).

The current regulatory standard of demonstration of both clinical (symptoms, function) efficacy and structural efficacy (JSN) for DMOADs is unlikely to change given the structure / symptoms uncoupling characteristic of OA. Hence we concede that clinical and structure effects need to be assessed with different endpoints. However if we choose the current recommended endpoint, namely JSN, due to limited responsiveness we would require many hundreds of subjects, followed for at least 2-3 years, to demonstrate a significant incremental benefit of a novel therapy. The direct costs of conducting such trials and the costs resulting from the overall duration of the therapeutic development and regulatory review process has dampened enthusiasm for development of therapeutic agents in this area and, in some instances, has rendered advancement of novel treatments prohibitively expensive. On the other hand, if other, more efficient means of establishing the benefit of new drugs exist, the promise of timely access to new therapies remains. There is, therefore, potentially tremendous value to public health in accelerating the discovery and development processes for OA therapeutics through shorter studies, using validated endpoints other than radiographic JSN.

b. Proposed context of use

The purpose of this letter of intent is to propose a measure of cartilage thickness from an MRI scan for qualification as a BioMarker.

To be used in disease-modifying regimens in knee OA and more specifically to assist in planning and design of clinical trials focused on facilitating scientific discovery and establishing efficacy of future disease modifying regimens for knee OA. More specifically we propose MRI cartilage thickness as a:

- Structure endpoint for evaluation of DMOAD effect on health and integrity of articular cartilage in the knee.
- To be used instead of JSW as a primary endpoint
- To be used in POC and phase II and potentially also phase III trials

c. High-level data description (1 to 2 pages in length). This description should provide a data overview that not only supports the use of the biomarker for the proposed context of use, but also encourages FDA engagement because of drug development applicability.

The OARSI-FDA OA Assessment of Structural Change (ASC) Working Group reviewed and synthesized published data on the performance metrics of the most common imaging tools used to assess structural change in OA, focusing predominantly on conventional radiographs and magnetic resonance imaging (MRI). A search of plain radiography and MRI literature in OA was conducted using articles published up to the time of the search, April 2009. These systematic reviews depict the responsiveness of quantitative JSW on plain radiographs, and the responsiveness, reliability and validity of MRI cartilage volume and thickness measurements (14-16).

MRI measures of cartilage thickness were recommended by the Working Group for clinical trials of knee OA treatments with structural outcomes on the basis of their preferable validity and responsiveness (17). The basis for this recommendation is on superior validity, reliability and responsiveness to existing plain radiographic standard (15;16).

Reliability (15)

The pooled CV for quantitative cartilage volume and thickness was 3% for both inter and intra-reader reliability. The inter-reader and intra-reader ICCs for quantitative cartilage measures were both excellent 0.90 (95%CI 0.86, 0.95) and 0.92 (95%CI 0.88, 0.96) respectively. This compares favorably with radiographic JSW with pooled inter-reader ICC was estimated at 0.93 (95% CI: 0.86, 0.99) and the inter-reader CV estimated at 3.4% (95% CI: 1.3%, 5.5%) (14).

Responsiveness

The pooled SRM [mean change divided by the standard deviation of the change over one year] for quantitative measures of cartilage thickness for the medial tibiofemoral joint was -0.86 (95%CI --1.26 to -0.46) and for lateral tibofemoral joint was -1.01 (95%CI -2.04 to 0.02). Recent data have shown that change in cartilage thickness is slightly more reliable and responsive to change in knee OA than change in cartilage volume (18). Thus the reliability of cartilage thickness on MRI compares favorably to radiographic medial tibiofemoral JSW which has an SRM of 0.33 [95% confidence interval (CI): 0.26, 0.41] (14).

Validity (16)

An important obstacle to biomarker validation and qualification is the adequate delineation of a gold standard. Unlike other diseases where surrogate endpoints exist, OA does not have a clear gold standard clinical endpoint and further is a remarkably heterogeneous disease. Therefore, the 'clinical endpoint' is more difficult to establish. A number of experts in the field have advocated that joint replacement be the clinical outcome of interest but due to constraints over comorbidities, insurance status and a number of other factors that influence determining if a person receives a joint replacement, alternate suggestions have been recommended including the use of virtual TKR (vTKR) (19). This is a composite endpoint that includes domains of pain, physical function and joint structure on X-rays (20). At this point, virtual TKR remains to be validated and as a consequence the constituent literature in this review does not include this endpoint to establish the predictive validity of MRI. In general there is a strong correlation of cartilage thickness measures to histologic findings (21). Cartilage thickness, measured with MRI in vitro (22-26); and in vivo (21;24), has been shown to be highly consistent with gold standard methods, such as physical measurement on anatomical sections (i.e. histology (22-24), CT- arthrography (25), stereophotogrammetry (26), and water displacement of surgically removed tissue (21;24). In vivo, the pairwise differences between results obtained with MRI and by direct morphology were for cartilage thickness were ±8.9%, systematic differences were not statistically significant (+2.2%), the correlation coefficient was r= 0.92 and the standard error (x/y) was 9.6% (24). In contrast, only one study addressed the accuracy of radiography (27) and found JSW to be reasonably accurate in the medial, but not in the lateral femorotibial compartment.

Predictive validity:

Joint replacement: One study investigated the relation of change in quantitative cartilage volume to risk of knee replacement. For every 1% increase in the rate of tibial cartilage loss there was a 20% increase risk of undergoing a knee replacement at four years (95% confidence interval, 10% to 30%). Those in the highest tertile of tibial cartilage loss had 7.1 (1.4 to 36.5) higher odds of undergoing a knee replacement than those in the lowest tertile (28).

A more recent study provides further support for the predictive value of quantitative, magnetic resonance imaging-based measures of cartilage for the clinically relevant endpoint of knee replacement, providing support for their utility in clinical trials to evaluate the effectiveness of structure modifying intervention (29). Cartilage thickness loss in the central and total medial femorotibial compartment in the year prior to KR was significantly greater in KR case than matched non-KR control knees (AUC=0.59/0.58), which were in the same radiographic disease stratum [Kellgren Lawrence grade 0 KLG]. Differences in cartilage loss were greater at earlier (KLG 1-2) than later (KLG 3-4) radiographic disease stages ($p<0.01$ for interaction with KLG).

Although cartilage thickness loss in the central subregion of the medial tibia was the most predictive longitudinal measure in context of KR (AUC=0.64), other medial compartment measures showed similar separation, with the central subregion of the medial femur

generally exhibiting greater sensitivity to change than the tibia. Therefore, for the time being, cartilage thickness change in the total medial femorotibial compartment (MFTC) appears to be the measure of choice [at least as long as the cohort studied has predominantly medial compartment disease], because it covers the entire compartment and does not require specific subregional measurement technology (30).

d. Additional resources that support the context of use as well as data the submitter plans to obtain from ongoing or future studies.

AS part of the recent OARSI FDA recommendations we reviewed the literature as it pertains to the measurement of structure in OA (14-16) and made a number of recommendations with regards conduct an design of DMOAD clinical trials (17). The measures proposed are also currently being further evaluated within the Foundation of NIH OA Biomarkers Consortium.

e. Indicate if there are plans to submit the biomarker for qualification by other international regulatory agencies

We are not aware of plans to submit the biomarker for qualification by other international regulatory agencies.

3. Process-related questions for FDA

What additional information is required to support this application?

Reference List

(1) Prevalence of disabilities and associated health conditions among adults--United States, 1999. [erratum appears in MMWR Morb Mortal Wkly Rep 2001 Mar 2;50(8):149.]. MMWR - Morbidity & Mortality Weekly Report 2001 Feb 23;50(7):120-5.

(2) Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. American Journal of Public Health 1994 Mar;84(3):351-8.

(3) Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2013 Dec 15;380(9859):2197-223.

(4) Painful Realities: The economic impact of Arthritis in Australia in 2007. Access Economics: Arthritis Australia; 2007.

(5) Losina E, Weinstein AM, Reichmann WM, Burbine SA, Solomon DH, Daigle ME, et al. Lifetime risk and age of diagnosis of symptomatic knee osteoarthritis in the US. Arthritis Care Res (Hoboken) 2012 Nov 30.

(6) Matthews GL, Hunter DJ. Emerging drugs for osteoarthritis. Expert Opin Emerg Drugs 2011 May 27.

(7) Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage 2010 Apr;18(4):476-99.

(8) Arthritis: The Nation's Most Common Cause of Disability. 2012.
Ref Type: Internet Communication

(9) Arthritis prevalence and activity limitations--United States, 1990. MMWR - Morbidity & Mortality Weekly Report 1994 Jun 24;43(24):433-8.

(10) Hunter DJ, Losina E, Guermazi A, Burstein D, Lassere MN, Kraus V. A pathway and approach to biomarker validation and qualification for osteoarthritis clinical trials. Curr Drug Targets 2010 May;11(5):536-45.

(11) Food and Drug Administration. Guidance for Industry. Clinical Development Programs for Drugs, Devices, and Biological Products Intended for the Treatment of Osteoarthritis (OA). http://www fda gov/Cber/gdlns/osteo htm 1999 [cited 2009 Mar 17];Available from: URL: http://www.fda.gov/Cber/gdlns/osteo.htm

(12) Mazzuca SA, Brandt KD. Is knee radiography useful for studying the efficacy of a disease-modifying osteoarthritis drug in humans?. [Review] [22 refs]. Rheumatic Diseases Clinics of North America 2003 Nov;29(4):819-30.

(13) Mazzuca SA, Brandt KD, Buckwalter KA, Lequesne M. Pitfalls in the accurate measurement of joint space narrowing in semiflexed, anteroposterior radiographic imaging of the knee. Arthritis & Rheumatism 2004 Aug;50(8):2508-15.

(14) Reichmann WM, Maillefert JF, Hunter DJ, Katz JN, Conaghan PG, Losina E. Responsiveness to change and reliability of measurement of radiographic joint space width in osteoarthritis of the knee: a systematic review. Osteoarthritis Cartilage 2011 May;19(5):550-6.

(15) Hunter DJ, Zhang W, Conaghan PG, Hirko K, Menashe L, Reichmann WM, et al. Responsiveness and reliability of MRI in knee osteoarthritis: a meta-analysis of published evidence. Osteoarthritis Cartilage 2011 May;19(5):589-605.

(16) Hunter DJ, Zhang W, Conaghan PG, Hirko K, Menashe L, Li L, et al. Systematic review of the concurrent and predictive validity of MRI biomarkers in OA. Osteoarthritis Cartilage 2011 May;19(5):557-88.

(17) Conaghan PG, Hunter DJ, Maillefert JF, Reichmann WM, Losina E. Summary and recommendations of the OARSI FDA osteoarthritis Assessment of Structural Change Working Group. Osteoarthritis Cartilage 2011 May;19(5):606-10.

(18) Hudelmaier M, Wirth W, Wehr B, Kraus V, Wyman BT, Hellio Le Graverand MP, et al. Femorotibial cartilage morphology: reproducibility of different metrics and femoral regions, and sensitivity to change in disease. Cells Tissues Organs 2010;192(5):340-50.

(19) Gossec L, Hawker G, Davis AM, Maillefert JF, Lohmander LS, Altman R, et al. OMERACT/OARSI initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis. [12 refs]. Journal of Rheumatology 2007 Jun;34(6):1432-5.

(20) Ornetti P, Brandt K, Hellio Le Graverand M, Hochberg M, Hunter D, Kloppenburg M, et al. OARSI-OMERACT definition of relevant radiological progression in hip/knee osteoarthritis. Osteoarthritis and Cartilage. In press 2009.

(21) Burgkart R, Glaser C, Hyhlik-Durr A, Englmeier KH, Reiser M, Eckstein F. Magnetic resonance imaging-based assessment of cartilage loss in severe osteoarthritis: accuracy, precision, and diagnostic value. Arthritis & Rheumatism 2001 Sep;44(9):2072-7.

(22) Eckstein F, Gavazzeni A, Sittek H, Haubner M, Losch A, Milz S, et al. Determination of knee joint cartilage thickness using three-dimensional magnetic resonance chondro-crassometry (3D MR-CCM). Magnetic Resonance in Medicine 1996 Aug;36(2):256-65.

(23) Sittek H, Eckstein F, Gavazzeni A, Milz S, Kiefer B, Schulte E, et al. Assessment of normal patellar cartilage volume and thickness using MRI: an analysis of currently available pulse sequences. Skeletal Radiology 1996 Jan;25(1):55-62.

(24) Graichen H, Eisenhart-Rothe R, Vogl T, Englmeier KH, Eckstein F. Quantitative assessment of cartilage status in osteoarthritis by quantitative magnetic resonance imaging: technical validation for use in analysis of cartilage volume and further morphologic parameters. Arthritis & Rheumatism 2004 Mar;50(3):811-6.

(25) Eckstein F, Schnier M, Haubner M, Priebsch J, Glaser C, Englmeier KH, et al. Accuracy of cartilage volume and thickness measurements with magnetic resonance imaging. Clinical Orthopaedics & Related Research 1998 Jul;(352):137-48.

(26) Cohen ZA, McCarthy DM, Kwak SD, Legrand P, Fogarasi F, Ciaccio EJ, et al. Knee cartilage topography, thickness, and contact areas from MRI: in-vitro calibration and in-vivo measurements. Osteoarthritis & Cartilage 1999 Jan;7(1):95-109.

(27) Buckland-Wright JC, MacFarlane DG, Lynch JA, Jasani MK, Bradshaw CR. Joint space width measures cartilage thickness in osteoarthritis of the knee: high resolution plain film and double contrast macroradiographic investigation. Annals of the Rheumatic Diseases 1995 Apr;54(4):263-8.

(28) Cicuttini FM, Jones G, Forbes A, Wluka AE. Rate of cartilage loss at two years predicts subsequent total knee arthroplasty: a prospective study. Annals of the Rheumatic Diseases 2004 Sep;63(9):1124-7.

(29) Eckstein F, Kwoh CK, Boudreau RM, Wang Z, Hannon MJ, Cotofana S, et al. Quantitative MRI measures of cartilage predict knee replacement: a case-control study from the Osteoarthritis Initiative. Ann Rheum Dis 2012 Jun 23.

(30) Wirth W, Eckstein F. A technique for regional analysis of femorotibial cartilage thickness based on quantitative magnetic resonance imaging. IEEE Trans Med Imaging 2008 Jun;27(6):737-44.

Comments for Dr. David Hunter regarding his Letter of Intent (LOI) submission for evaluating cartilage thickness on knee MRI as a biomarker of osteoarthritis

The comments and questions contained in this document represent initial thinking by the Biomarker Qualification Review Team (BQRT) about topics that may be useful to consider during the course of developing this biomarker for your proposed context of use. Please consider these comments in preparing your briefing document. We recognize that you will not be able to completely address all the comments at this juncture. While some of this information was included in your LOI, a discussion as complete as feasible should be included in your briefing document. Please include current knowledge, your perspective on the topic, and a description of the additional work you propose to conduct, in the briefing document. This will be very useful in holding a productive meeting.

For additional information on the structure of the initial briefing package, please see http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm284667.htm

Please see below for comments and recommendations from the Biomarker Qualification Review Team.

A. Biomarker Considerations:

1. Since osteoarthritis manifests as degeneration not only of the cartilage but also structural and biochemical changes to the entire joint, cartilage thickness alone may not serve as a useful predictor for effectiveness of osteoarthritis drugs. MRI is capable of visualizing many different joint features that have been associated with osteoarthritis including effusions, meniscal damage, and cartilage and bone lesions, among many others. Specialized MRI protocols may provide a means to assess cartilage biochemistry, which is strongly altered in osteoarthritis. Therefore, please discuss your rationale for the selection of cartilage thickness for the proposed biomarker over other MR structural parameters (for example, cartilage volume, regional analysis or a composite score of MR measurements) as potential biomarkers. In addition, please include the rationale for the proposed use of cartilage thickness as compared to MRI-derived features that may more directly reflect cartilage biochemistry (for example, dGEMRIC, T2-mapping, diffusion).
2. Please discuss the strengths and limitations of the proposed biomarker.

B. Context of Use (COU) Considerations:

The COU has an important influence on the type of data that is needed to achieve Qualification. While it is important to state the targeted COUs at this time, the COU statement that is ultimately qualified evolves with the development of the biomarker(s) based on the study results and the interpretation of those results. In order to be useful at this stage, your target COU should be sufficiently precise and specific to understand how biomarker results would be interpreted and the intended drug development decisions that would be made relying on the biomarker.

3. The type and amount of clinical data that may be expected to support the proposed COU for your biomarker would depend on the context of use. Examples to consider include:

- Early clinical development (e.g., Proof of Concept (POC), Dose-Ranging studies)
  - Sufficient clinical data to justify a scientific rationale that the biomarker is relevant to the treatment's clinical effect of interest
    - E.g., Correlations of MRI cartilage thickness and OA severity in a cross-sectional study. Sufficient longitudinal information to understand what type of changes might be expected naturally over the time-frame of a clinical trial.
    - Other necessary covariates such as baseline severity, gender, age, genetic status that may influence the correlation in the evaluation shouls be included in these evaluations.
    - Caveat: POC and Dose-ranging studies should also include the clinical outcomes that are being considered for pivotal studies, as the approval of a treatment for OA would likely require demonstration of benefit for a clinical outcome. The role of a pharmacodynamic measure using MRI in these situations would be to provide supportive information to help with development program decision-making.
- Phase 3 clinical trials
  - Adequate clinical data to understand the clinical meaningfulness of changes in MRI cartilage thickness
    - E.g., Longitudinal data correlating MRI cartilage thickness with long-term clinical outcomes. The long-term clinical outcomes have to be identified and agreed upon, but could include measures of functional status as well as a measure of "joint failure". By long-term, we mean long-enough to capture the natural history of OA progression in a cohort of OA patients, such as the OA Initiative.
    - Ideally, this would also include information addressing questions such as, the variability of MRI cartilage thickness over time (i.e. do changes wax and wane?); what level of improvement would be clinically meaningful and over what time would it need to be sustained to make a difference in long-term clinical outcomes?
    - Caveat: correlation of MRI cartilage thickness with short-term clinical outcomes (such as short-term pain or physical functioning) would not be sufficiently informative with respect to "DMOAD" type claims. Additionally, direct measures of patient benefit are generally preferable, so if MRI cartilage thickness is only demonstrated to correlate with short-term clinical outcomes, this would not be helpful, as short-term clinical outcomes can be measured directly. Correlation with short-term clinical outcomes also does not provide information on whether the MRI outcome would be predictive of long-term clinical outcomes related to progression of structural damage in OA.

C. Methodology and Analysis Considerations:

4. Deriving a one-dimensional measurement (for example, thickness) from a volumetric dataset allows for many possible analysis procedures. Please summarize your proposed image analysis procedure(s) for measuring cartilage

thickness from an MR image. Please include more detail about how thickness information will be used as an endpoint. For example, will average thickness or maximum thickness in each knee compartment (patella, femur, tibia) be used as the endpoint? Please specify steps in the analysis that are manual (require human input), automatic (require no user interaction), or semi-automatic. Please describe any advanced image processing procedure(s).

5. Please describe the relationship between MRI-derived values for cartilage thickness and the actual cartilage thickness present in the joint. Please summarize existing literature supporting the validity of MRI as a measurement of cartilage thickness. Please include in this summary any comparisons to anatomic sectioning, stereophotogrammetric, or alternative physical measurements for both healthy and diseased patients. Please discuss measurement limitations, bias, and uncertainty for both the MRI and comparison method. Please consider the clinically meaningful difference of cartilage thickness in osteoarthritis in relation to the accuracy of the measurement technique.

6. The choice of image acquisition protocols may influence the results of cartilage thickness measurements. Please provide a summary of the image acquisition protocols you plan to use for ongoing or future studies, as well as acceptance criteria for past studies. In particular, items you may wish to consider include:
    a. Pulse sequence(s)
    b. Field strength
    c. Manufacturers (e.g., Philips, GE)
    d. Coils used (body T/R only, extremity coils, etc.)
    e. Use of contrast agents
    f. Spatial resolution and slice thickness
    g. Patient positioning (weight bearing, non-weight bearing, or semi-weight bearing)

7. Please describe your planned strategies to account for variability introduced by different image acquisition protocols in existing (and possibly future) datasets.

8. Please provide a brief synopsis of your plans to assess bias, statistical linearity, uncertainty, repeatability, reproducibility, and sensitivity of your measurement system for assessing cartilage thickness.

D. Clinical Considerations:

9. As you have noted in your Letter of Intent, the relationship between structural changes and clinical outcomes in osteoarthritis (OA) is not well defined or tightly correlated. At the present timea proposed treatment would need to demonstrate a benefit on a clinical outcome as a basis for drug approval. Although the 1999 OA FDA draft guidance mentions “Normalizing the x-ray” and “Improving the x-ray,” using joint space narrowing (JSN) as an endpoint, as claims requiring no formal parallel evidence of improvement in clinical

outcomes, this appears to be predicated on the assumption that it would take a large treatment effect to be able to demonstrate improvement on JSN in standard x-rays. Thus, it is not likely that one would be able to extrapolate the potential for these claims supported by the stated basis to apply to improvements in a much more sensitive structural outcome such as MRI, and an approach intending to link MRI to radiographic JSN (or joint space width [JSW]) does not address fundamental concerns about the clinical meaningfulness of changes in a given MRI biomarker.

10. Please discuss the patient population and disease state of the cartilage you plan to use in your investigations. Please assess any potential differences in the imaging or analysis protocol for thickness measurements of cartilage in patients with healthy compared to diseased cartilage.

11. Please clarify whether the utility of the proposed biomarker applies to all stages of the disease.

12. Please discuss the potential utility as a biomarker of the change in cartilage thickness over multiple (two or more) MRI sessions. Please also discuss the potential utility of single session MRI measurement of cartilage thickness as a biomarker. Please discuss the relationships of cartilage thickness and change in cartilage thickness to osteoarthritis disease state, osteoarthritis progression, and clinical outcomes such as pain levels in osteoarthritis patients. Please discuss the relationship between cartilage thickness as measured on MR images and radiographs to symptomatic osteoarthritis. Please discuss any plans for demonstrating the superiority as a drug development tool of cartilage thickness measured by MRI as compared to radiography.

13. For background information, please provide list of relevant publications or internal reports, to support the assay in the desired Context of Use (COU). Please include copies, where feasible, of a limited number of the most important references that will be helpful to read in evaluating your project.

14. We recommend you to provide data and analysis which can demonstrate that the biomarker selected is suitable for describing longitudinal changes in disease progression of OA. We recommend you to evaluate if the proposed biomarker and JSN have the same trend of changes with disease progress.

Response letter to the FDA's comments on the LOI

Comments for Dr. David Hunter regarding his Letter of Intent (LOI) submission for evaluating cartilage thickness on knee MRI as a biomarker of osteoarthritis

The comments and questions contained in this document represent initial thinking by the Biomarker Qualification Review Team (BQRT) about topics that may be useful to consider during the course of developing this biomarker for your proposed context of use. Please consider these comments in preparing your briefing document. We recognize that you will not be able to completely address all the comments at this juncture. While some of this information was included in your LOI, a discussion as complete as feasible should be included in your briefing document. Please include current knowledge, your perspective on the topic, and a description of the additional work you propose to conduct, in the briefing document. This will be very useful in holding a productive meeting.

For additional information on the structure of the initial briefing package, please see http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm284667.htm

Please see below for comments and recommendations from the Biomarker Qualification Review Team.

A. Biomarker Considerations:

1. Since osteoarthritis manifests as degeneration not only of the cartilage but also structural and biochemical changes to the entire joint, cartilage thickness alone may not serve as a useful predictor for effectiveness of osteoarthritis drugs. MRI is capable of visualizing many different joint features that have been associated with osteoarthritis including effusions, meniscal damage, and cartilage and bone lesions, among many others. Specialized MRI protocols may provide a means to assess cartilage biochemistry, which is strongly altered in osteoarthritis. Therefore, please discuss your rationale for the selection of cartilage thickness for the proposed biomarker over other MR structural parameters (for example, cartilage volume, regional analysis or a composite score of MR measurements) as potential biomarkers. In addition, please include the rationale for the proposed use of cartilage thickness as compared to MRI-derived features that may more directly reflect cartilage biochemistry (for example, dGEMRIC, T2-mapping, diffusion).

Response: We strongly agree with the reviewer's comment. We will focus on 5 major MRI parameters for proposed qualification on the basis that they performed well in the recently completed phase 1 FNIH project. Specifically, these markers had the greatest ability to predict longer-term clinical and radiographic progression in the multivariable models and showed promise as potential biomarkers for the proposed context of use. Several aspects of the OA pathogenesis are encompassed in this selection including cartilage degradation (semi-quantitative analysis of cartilage morphology and quantitative cartilage morphometry), meniscal damage (semi-quantitative analysis of meniscal morphology), inflammation (semi-quantitative analysis of synovitis), and subchondral bone changes (osteophytes). The metrics for assessment of cartilage degeneration selected for this project have been more extensively studied than other markers of cartilage morphometry or composition. Quantitative measures of cartilage composition require specialized MRI sequences that are not widely available. In addition, dGEMRIC requires intravenous contrast and a long time for examination which is a disadvantage for use in clinical trials and large epidemiological studies (1).

2. Please discuss the strengths and limitations of the proposed biomarker.

Response: The proposed biomarkers are responsive, valid and reliable measures for use in knee OA trials which reflect important components of OA pathogenesis (2, 3). There is robust literature evidence supporting their usefulness as prognostic biomarkers in knee OA (4, 5). These markers are obtained on non-contrast enhanced MRI, which is a widely available, non-invasive procedure that does not involve ionizing radiation and has very few contra-indications. Potential disadvantages are the current relative high costs associated with MRIs and need for trained personnel for MRI analysis.

B. Context of Use (COU) Considerations:
The COU has an important influence on the type of data that is needed to achieve Qualification. While it is important to state the targeted COUs at this time, the COU statement that is ultimately qualified evolves with the development of the biomarker(s) based on the study results and the interpretation of those results. In order to be useful at this stage, your target COU should be sufficiently precise and specific to understand how biomarker results would be interpreted and the intended drug development decisions that would be made relying on the biomarker.

Response: We have refined the COU initially proposed in the LOI form. We propose qualification of MRI markers for the identification of subjects likely to experience long-term disease progression in the absence of treatment (prognostic biomarkers). Prognostic biomarkers would allow enrichment of clinical trials with progressors and decrease the sample size required thereby increasing trials cost-effectiveness and decreasing their duration.

3. The type and amount of clinical data that may be expected to support the proposed COU for your biomarker would depend on the context of use. Examples to consider include:

- Early clinical development (e.g., Proof of Concept (POC), Dose-Ranging studies)
    - Sufficient clinical data to justify a scientific rationale that the biomarker is relevant to the treatment's clinical effect of interest
        - E.g., Correlations of MRI cartilage thickness and OA severity in a cross-sectional study. Sufficient longitudinal information to understand what type of changes might be expected naturally over the time-frame of a clinical trial.
        - Other necessary covariates such as baseline severity, gender, age, genetic status that may influence the correlation in the evaluation should be included in these evaluations.
        - Caveat: POC and Dose-ranging studies should also include the clinical outcomes that are being considered for pivotal studies, as the approval of a treatment for OA would likely require demonstration of benefit for a clinical outcome. The role of a pharmacodynamic measure using MRI in these situations would be to provide supportive information to help with development program decision-making.

- Phase 3 clinical trials
    - Adequate clinical data to understand the clinical meaningfulness of changes in MRI cartilage thickness
        - E.g., Longitudinal data correlating MRI cartilage thickness with long-term clinical outcomes. The long-term clinical outcomes have to be identified and agreed upon, but could include measures of functional status as well as a measure of "joint failure". By long-term, we mean long-enough to capture the natural history of OA progression in a cohort of OA patients, such as the OA Initiative.
        - Ideally, this would also include information addressing questions such as, the variability of MRI cartilage thickness over time (i.e. do changes wax and wane?); what level of improvement

would be clinically meaningful and over what time would it need to be sustained to make a difference in long-term clinical outcomes?

- Caveat: correlation of MRI cartilage thickness with short-term clinical outcomes (such as short-term pain or physical functioning) would not be sufficiently informative with respect to "DMOAD" type claims. Additionally, direct measures of patient benefit are generally preferable, so if MRI cartilage thickness is only demonstrated to correlate with short-term clinical outcomes, this would not be helpful, as short-term clinical outcomes can be measured directly. Correlation with short-term clinical outcomes also does not provide information on whether the MRI outcome would be predictive of long-term clinical outcomes related to progression of structural damage in OA.

C. Methodology and Analysis Considerations:

4. Deriving a one-dimensional measurement (for example, thickness) from a volumetric dataset allows for many possible analysis procedures. Please summarize your proposed image analysis procedure(s) for measuring cartilage thickness from an MR image. Please include more detail about how thickness information will be used as an endpoint. For example, will average thickness or maximum thickness in each knee compartment (patella, femur, tibia) be used as the endpoint? Please specify steps in the analysis that are manual (require human input), automatic (require no user interaction), or semi-automatic. Please describe any advanced image processing procedure(s).

Response: The tibiofemoral cartilage surface is segmented in 16 subregions (8 in the medial and 8 in the lateral tibiofemoral compartment) and the mean cartilage thickness is measured at each of these locations. This segmentation is performed manually and the reliability of these measurements and their feasibility in clinical trials have been reported previously (1, 6, 7). Cartilage thickness at the medial tibiofemoral compartment and central medial femur had the greatest predictive ability for progression in multivariable models and will be our focus. Cartilage thickness change will be computed as the absolute difference between follow up (e.g. 12 months) and the baseline value (mm).
Adequate images for measurement of cartilage thickness can be obtained with widely available MRI scanners (e.g. 1.5 T field strength) and sequences (SPGR/FLASH). In addition, there is a high agreement between cartilage thickness measures as determined from different MRI sequences (i.e. FLASH and DESS) cross-sectionally and over time (8).

5. Please describe the relationship between MRI-derived values for cartilage thickness and the actual cartilage thickness present in the joint. Please summarize existing literature supporting the validity of MRI as a measurement of cartilage thickness. Please include in this summary any comparisons to anatomic sectioning, stereophotogrammetric, or alternative physical measurements for both healthy and diseased patients. Please discuss measurement limitations, bias, and uncertainty for both the MRI and comparison method. Please consider the clinically meaningful difference of cartilage thickness in osteoarthritis in relation to the accuracy of the measurement technique.

Response: Our team has previously conducted a systematic review of the literature to examine the validity of MRI biomarkers in OA (2). There is a very good correlation between cartilage thickness measurement on MRI and histological sections in the middle of each articular region examined in knees from cadavers with varying degrees of cartilage damage ($r = 0.88$), which was independent of the grade of cartilage lesions (9). In addition, there were statistically significant correlations between radiography and MR cartilage loss in the medial (r 0.7142, $p=0.0001$) and lateral compartments ($r = 0.4004$, $p=0.0136$) (10). Cartilage volume and thickness were also less in patients with OA compared to normal

controls (11). The annual changes in cartilage volume /thickness exceeded the precision errors and appear to be associated with clinical symptoms as well as with time to knee arthroplasty (12, 13).

6. The choice of image acquisition protocols may influence the results of cartilage thickness measurements. Please provide a summary of the image acquisition protocols you plan to use for ongoing or future studies, as well as acceptance criteria for past studies. In particular, items you may wish to consider include:
a. Pulse sequence(s): we would suggest that the primary sequence would be PD FSE/T2 TSE with FS with secondary consideration given to 3D T1-GRE with FS or WE.
b. Field strength: the majority of the studies that have been conducted in the space have been conducted at 1.5 T and 3T, with preference given to the latter.
c. Manufacturers (e.g., Philips, GE): the proposed sequences or variants of these are generically available on multiple different manufacturer platforms including Siemens, Philips, GE.
d. Coils used (body T/R only, extremity coils, etc.) : to facilitate optimal acquisition we would advocate for the use of a dedicated extremity coil.
e. Use of contrast agents : no contrast is required for the measures proposed.
f. Spatial resolution and slice thickness: section thickness 0.7 to 1.0mm and in plan 0.3-0.5mm.
g. Patient positioning (weight bearing, non-weight bearing, or semi-weight bearing) : nonweight bearing, supine.
Further details supporting this information can be found in our recent publication (14).

7. Please describe your planned strategies to account for variability introduced by different image acquisition protocols in existing (and possibly future) datasets.

Response: The MRI acquisition protocol of the Osteoarthritis Initiative study which was used for the phase 1 of this project has been published previously (6). The cartilage thickness analysis used sagittal double-echo steady-state imaging acquired by standardized 3 Tesla MRI (3T-MRI) scanners. However, other protocols have also been shown to be similarly suitable for cartilage thickness analysis. The precision of knee cartilage thickness measurement between different MRI acquisition protocols (sagittal DESS and coronal FLASH) was compared in the OAI study and found to be similar, with high correlation for most knee regions (r=0.90–0.97), except for DESS versus FLASH medial central femur mean cartilage thickness (r=0.81–0.83) (7). In addition, precision of the FLASH acquisitions at 1.5T and 3T for quantitative cartilage measurements have been shown to be equivalent (7). Another study using OAI data compared two types of coils (phased-array and quadrature knee coils) and found no significant differences in log (CV%) precision error values and high correlation coefficients between the two coils for cartilage thickness measurements using FLASHwe ($r \geq 0.94$) and DESSwe images ($r \geq 0.90$) (15).

8. Please provide a brief synopsis of your plans to assess bias, statistical linearity, uncertainty, repeatability, reproducibility, and sensitivity of your measurement system for assessing cartilage thickness.

Response: Chondrometrics GmbH, a Germany-based company located in Ainring, Bavaria, is a leading provider of quantitative medical image analysis services to researchers in academia and the pharmaceutical industry. Their primary focus is the quantitative analysis of articular cartilage from MRIs and the research of OA. The company has developed a highly efficient software platform, and has formed a team of specialized and highly experienced readers to provide quantitative imaging surrogates of tissue adaptation and disease progression for large-scale studies. After initial quality control and conversion to a proprietary format at the image analysis center (Chondrometrics GmbH, Ainring,

Germany), manual segmentation of the femorotibial cartilages will be performed by trained technicians with at least 3 years of experience in cartilage segmentation. The image data will be processed in parallel (baseline and follow up images will be read side by side blinded to timepoint), the readers being blinded to the order of the image acquisition. Chondrometrics will perform quality control readings of the imaging data before segmentation, experienced Chondrometrics readers will perform the segmentations, and one of three available expert readers will perform quality control readings of all segmented slices of each knee throughout the study. Validated Chondrometrics software will be used to compute subregional cartilage morphology endpoints and ordered values, consistent with previous OA Biomarker study and OAI analyses and with analyses performed in other large cohorts. Chondrometrics has been involved in numerous large-scale studies. For example, they have analyzed cartilage morphology in three large NIH-funded longitudinal cohort studies, including 1,100 knee datasets from the Framingham cohort, 800 knee data sets (i.e., bilateral evaluation at baseline and follow-up on 200 participants) in the Mechanical Factors in Arthritis of the Knee (MAK) study, and 600 knees in the Multicenter Knee Osteoarthritis (MOST) study. They have also analyzed 180 knee data sets in the first 3T MRI cross-validation study (Pfizer A 9001191), and about 3,000 knee data sets (coronal and sagittal, 7 time points) in a longitudinal method trial using 3T MR imaging (Pfizer A9001140). The Project Team has selected Chondrometrics to execute this part of the project based upon extensive work validating their approach and their performance in phase 1 of this OA Biomarkers Consortium Project.

D. Clinical Considerations:

9. As you have noted in your Letter of Intent, the relationship between structural changes and clinical outcomes in osteoarthritis (OA) is not well defined or tightly correlated. At the present time a proposed treatment would need to demonstrate a benefit on a clinical outcome as a basis for drug approval. Although the 1999 OA FDA draft guidance mentions “Normalizing the x-ray” and “Improving the x-ray,” using joint space narrowing (JSN) as an endpoint, as claims requiring no formal parallel evidence of improvement in clinical outcomes, this appears to be predicated on the assumption that it would take a large treatment effect to be able to demonstrate improvement on JSN in standard x-rays. Thus, it is not likely that one would be able to extrapolate the potential for these claims supported by the stated basis to apply to improvements in a much more sensitive structural outcome such as MRI, and an approach intending to link MRI to radiographic JSN (or joint space width [JSW]) does not address fundamental concerns about the clinical meaningfulness of changes in a given MRI biomarker.

10. Please discuss the patient population and disease state of the cartilage you plan to use in your investigations. Please assess any potential differences in the imaging or analysis protocol for thickness measurements of cartilage in patients with healthy compared to diseased cartilage.

Response: This project will utilize a number of extant clinical trial resources with existing biospecimens and imaging repositories. These trials test a range of therapeutic interventions targeted to different synovial joint tissues. The target population is individuals with mild to moderate radiographic OA defined by Kellgren-Lawrence grade or joint space width on radiographs. We are not aware of differences in MRI protocols for assessing healthy and OA cartilage. The trials that will be utilized are listed below:

1) Cindunistat study (Pfizer) (NCT00565812)

The A6171016 or iTIC study (iNOS Trial to Investigate Chondroprotection) was sponsored by Pfizer and targeted persons with medial tibiofemoral OA. The efficacy of SD-6010 was evaluated by radiography using joint space narrowing in the medial tibiofemoral compartment of the study knee as the primary endpoint. A total of 1400 persons were enrolled in the main cohort (Xray + Outcome Measures) and 100

persons were enrolled in an MRI sub-cohort (patients who underwent MRI of the knee). The duration of the trial for individual participants was 22 months.

2) Vitamin D (VIDEO-Arden) Study (ISRCTN 94818153)

The VIDEO study was design as a double-blind, randomized, placebo-controlled trial to assess the effect of vitamin D supplementation in the rate of knee OA progression. Four hundred and seventy four patients aged > 50 years, with knee pain and radiographically confirmed knee OA were randomized to receive either placebo or 800 IU cholecalciferol daily and outcomes were assessed at 12,24 and 36 months. The study's primary outcome was difference in rate of medial joint space narrowing (JSN) between the groups, and secondary outcomes included changes in lateral JSN, KLG, WOMAC pain, function, stiffness and the Get up and Go test. MRI with gadolinium enhancement was further performed in a subset of patients (n=150).

3) Intraarticular Sprifermin (Merck Serono) (NCT01033994 - Phase 1)

This multicenter, randomized, double-blind, placebo-controlled trial tested the hypothesis that sprifermin (Recombinant Human Fibroblast Growth Factor 18) could reduce the loss of joint cartilage, using femorotibial cartilage thickness on MRI. It was conducted at 30 sites in Europe, South Africa, and North America and included patients aged ≥40 years with femorotibial knee OA according to ACR clinical and radiologic criteria and KLG 2 or 3. Between October 2008 and December 2010, 192 patients were randomized 3:1 to receive sprifermin or placebo. Multiple-dose regimens (3 doses of either 10 μg, 30 μg, or 100 μg) of sprifermin were evaluated in 6 cohorts (24 patients to the single-dose cohorts and 168 to the multiple-dose cohorts). The primary efficacy end point was the longitudinal change from baseline in central medial femorotibial compartment cartilage thickness at 6 months and 12 months, as assessed using quantitative MRI (qMRI). Secondary imaging end points included total and compartment (both medial and lateral) femorotibial cartilage thickness and volume as assessed by qMRI at 3, 6, and 12 months after the first injection, quantitative measurement of joint space width (JSW) by fixed-flexion weight-bearing radiography at 12 months, and assessment of bone marrow lesions, cartilage, menisci, effusion, and synovitis by semi-quantitative MRI at 3, 6, and 12 months. Symptom efficacy was evaluated as change at 3, 6, and 12 months from baseline in the WOMAC.

4) TissueGene-C (Kolon Life Science) (NCT02072070 – phase 3)

This double-blind, randomized, parallel-group, multi-center study aimed to determine the efficacy and safety of tissuegene-C (allogeneic human chondrocytes expressing Transforming Growth Factor (TGF)-b1). The trial included patients with a diagnosis of knee OA according to ACR clinical and radiographic criteria, radiographic KLG stage 3 and International Cartilage Repair Society (ICRS) Grade III or IV cartilage damage in the major lesions, as confirmed through an MRI scan. One hundred and fifty six outpatients were randomized to TissueGene-C or placebo in 1:1 ratio and outcomes were assessed at 26, 39 and 52 weeks. Primary outcomes were change in IKDC scores and VAS pain from 0 to 52 weeks and secondary outcomes included changes in WOMAC and KOOS scores, changes in MRI scan, JSW, levels of serum and urinary markers and use of rescue medication.

5) Strontium ranelate (Servier) (ISRCTN41323372 – phase 3)

The aim of this 3-year multicenter, double-blind, randomized, placebo-controlled trial—Strontium ranelate Efficacy in Knee Osteoarthritis trial (SEKOIA) — was to evaluate the effect of strontium ranelate on radiological and clinical progression of knee OA. The study included patients with knee OA according to ACR criteria, KLG 2or 3 on radiograph and joint space width (JSW) of 2.5 to 5 mm with predominant knee OA of the medial tibiofemoral compartment. The trial randomly allocated 1683 patients to three treatment groups (strontium ranelate 1g [n=558] or 2 g/day [n=566] or placebo [n=559]). The primary endpoint was radiographical change in JSW (medial tibiofemoral compartment) over 3 years versus placebo. Secondary endpoints included radiological progression, WOMAC score, knee pain, and urinary CTX-II levels.

11. Please clarify whether the utility of the proposed biomarker applies to all stages of the disease.

Response: We propose cartilage thickness as a prognostic biomarker for individuals with mild to moderate knee osteoarthritis (e.g, Kellgren Lawrence grade 2 and 3) which is the focus of our investigations.

12. Please discuss the potential utility as a biomarker of the change in cartilage thickness over multiple (two or more) MRI sessions. Please also discuss the potential utility of single session MRI measurement of cartilage thickness as a biomarker. Please discuss the relationships of cartilage thickness and change in cartilage thickness to osteoarthritis disease state, osteoarthritis progression, and clinical outcomes such as pain levels in osteoarthritis patients. Please discuss the relationship between cartilage thickness as measured on MR images and radiographs to symptomatic osteoarthritis. Please discuss any plans for demonstrating the superiority as a drug development tool of cartilage thickness measured by MRI as compared to radiography.

Response: Baseline and short-term change of biomarkers that predict likelihood of disease progression in the absence of treatment may facilitate early identification of subjects likely to progress without treatment who can be targeted for therapy. Biomarkers that are responsive over time may also be useful to reflect a response to therapy in clinical trials. This is particularly important since radiographic JSW, which is the current standard technique for assessment of structural progression in trials, has limited responsiveness to change over time (3, 16, 17), limited responsiveness to structural treatment effects (18, 19) and is not able to distinguish cartilage loss from meniscal lesions (1). Previous direct comparisons of MRI measures of cartilage thickness and volume and radiographic JSW have shown that MRI has greater responsiveness in knee OA than radiography (16, 17). Longitudinal changes in cartilage thickness have been associated with progression of pain, structural damage and higher incidences of total joint replacement in individuals with knee OA (4, 12). In another study, cartilage thinning over 12 months was found to be more frequent in knees with frequent pain compared with asymptomatic knees (20).

13. For background information, please provide list of relevant publications or internal reports, to support the assay in the desired Context of Use (COU). Please include copies, where feasible, of a limited number of the most important references that will be helpful to read in evaluating your project.

Response: A list of relevant references is provided.

14. We recommend you to provide data and analysis which can demonstrate that the biomarker selected is suitable for describing longitudinal changes in disease progression of OA. We recommend you to evaluate if the proposed biomarker and JSN have the same trend of changes with disease progress.

Response: We have demonstrated in phase 1 of this project that cartilage thickness loss over 24 months was significantly associated with long-term (48 months) clinical and structural disease progression based on WOMAC knee pain and JSW loss, respectively (4). A stronger association was observed for radiographic progression vs. no progression with OR 4.0 (2.9–5.3). Radiographic JSW and cartilage thickness on MRI have been shown to have the same trend of progression over time with greater responsiveness of MRI (16, 17).

References

1. Eckstein F, Guermazi A, Gold G, Duryea J, Hellio Le Graverand MP, Wirth W, et al. Imaging of cartilage and bone: promises and pitfalls in clinical trials of osteoarthritis. Osteoarthritis and cartilage. 2014;22(10):1516-32.
2. Hunter DJ, Zhang W, Conaghan PG, Hirko K, Menashe L, Li L, et al. Systematic review of the concurrent and predictive validity of MRI biomarkers in OA. Osteoarthritis and cartilage. 2011;19(5):557-88.
3. Hunter DJ, Zhang W, Conaghan PG, Hirko K, Menashe L, Reichmann WM, et al. Responsiveness and reliability of MRI in knee osteoarthritis: a meta-analysis of published evidence. Osteoarthritis and cartilage. 2011;19(5):589-605.
4. Eckstein F, Collins JE, Nevitt MC, Lynch JA, Kraus VB, Katz JN, et al. Brief Report: Cartilage Thickness Change as an Imaging Biomarker of Knee Osteoarthritis Progression: Data From the Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis & rheumatology. 2015;67(12):3184-9.
5. Collins JE, Losina E, Nevitt MC, Roemer FW, Guermazi A, Lynch JA, et al. Semiquantitative Imaging Biomarkers of Knee Osteoarthritis Progression: Data From the Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Rheumatol. 2016;68(10):2422-31.
6. Eckstein F, Wirth W, Nevitt MC. Recent advances in osteoarthritis imaging--the osteoarthritis initiative. Nature reviews Rheumatology. 2012;8(10):622-30.
7. Schneider E, Nevitt M, McCulloch C, Cicuttini FM, Duryea J, Eckstein F, et al. Equivalence and precision of knee cartilage morphometry between different segmentation teams, cartilage regions, and MR acquisitions. Osteoarthritis and cartilage. 2012;20(8):869-79.
8. Altman R, Scazziota A. Why aspirin cannot prevent arterial thrombosis [letter; comment]. Circulation. 1996;94(11):3002-3.
9. Kladny B, Martus P, Schiwy-Bochat KH, Weseloh G, Swoboda B. Measurement of cartilage thickness in the human knee-joint by magnetic resonance imaging using a three-dimensional gradient-echo sequence. International orthopaedics. 1999;23(5):264-7.
10. Chan WP, Lang P, Stevens MP, Sack K, Majumdar S, Stoller DW, et al. Osteoarthritis of the knee: comparison of radiography, CT, and MR imaging to assess extent and severity. AJR American journal of roentgenology. 1991;157(4):799-806.
11. Lindsey CT, Narasimhan A, Adolfo JM, Jin H, Steinbach LS, Link T, et al. Magnetic resonance evaluation of the interrelationship between articular cartilage and trabecular bone of the osteoarthritic knee. Osteoarthritis and cartilage. 2004;12(2):86-96.
12. Eckstein F, Kwoh CK, Boudreau RM, Wang Z, Hannon MJ, Cotofana S, et al. Quantitative MRI measures of cartilage predict knee replacement: a case-control study from the Osteoarthritis Initiative. Ann Rheum Dis. 2013;72(5):707-14.
13. Wluka AE, Wolfe R, Stuckey S, Cicuttini FM. How does tibial cartilage volume relate to symptoms in subjects with knee osteoarthritis? Annals of the rheumatic diseases. 2004;63(3):264-8.
14. Hunter DJ, Altman RD, Cicuttini F, Crema MD, Duryea J, Eckstein F, et al. OARSI Clinical Trials Recommendations: Knee imaging in clinical trials in osteoarthritis. Osteoarthritis Cartilage. 2015;23(5):698-715.
15. Eckstein F, Kunz M, Hudelmaier M, Jackson R, Yu J, Eaton CB, et al. Impact of coil design on the contrast-to-noise ratio, precision, and consistency of quantitative cartilage morphometry at 3 Tesla: a pilot study for the osteoarthritis initiative. Magnetic resonance in medicine. 2007;57(2):448-54.
16. Wirth W, Duryea J, Hellio Le Graverand MP, John MR, Nevitt M, Buck RJ, et al. Direct comparison of fixed flexion, radiography and MRI in knee osteoarthritis: responsiveness data from the Osteoarthritis Initiative. Osteoarthritis and cartilage. 2013;21(1):117-25.

17. Cromer MS, Bourne RM, Fransen M, Fulton R, Wang SC. Responsiveness of quantitative cartilage measures over one year in knee osteoarthritis: comparison of radiography and MRI assessments. Journal of magnetic resonance imaging : JMRI. 2014;39(1):103-9.
18. Roemer FW, Aydemir A, Lohmander S, Crema MD, Marra MD, Muurahainen N, et al. Structural effects of sprifermin in knee osteoarthritis: a post-hoc analysis on cartilage and non-cartilaginous tissue alterations in a randomized controlled trial. BMC musculoskeletal disorders. 2016;17:267.
19. Eckstein F, Wirth W, Guermazi A, Maschek S, Aydemir A. Brief report: intraarticular sprifermin not only increases cartilage thickness, but also reduces cartilage loss: location-independent post hoc analysis using magnetic resonance imaging. Arthritis & rheumatology. 2015;67(11):2916-22.
20. Buck RJ, Wirth W, Dreher D, Nevitt M, Eckstein F. Frequency and spatial distribution of cartilage thickness change in knee osteoarthritis and its relation to clinical and radiographic covariates - data from the osteoarthritis initiative. Osteoarthritis and cartilage. 2013;21(1):102-9.

Multivariable modelling of biomarker data from the phase 1 foundation of NIH osteoarthritis biomarkers consortium

David J. Hunter[1]

Jamie E. Collins[2,3]

Elena Losina[2,3]

Michael C. Nevitt[4]

Frank W. Roemer[4,5]

Ali Guermazi[4]

Mike Bowes

Erik Dam

Felix Eckstein

Steve Hoffman

Leticia Deveza

John A. Lynch[6]

Jeffrey N. Katz[2,3]

C. Kent Kwoh[7]

Virginia B. Kraus[8]

[1]Department of Rheumatology, Royal North Shore Hospital and Institute of Bone and Joint Research, Kolling Institute, University of Sydney, Sydney, Australia.

[1]Orthopaedic and Arthritis Center for Outcomes Research, Department of Orthopedic Surgery, Brigham and Women's Hospital Boston, MA USA

[3]Harvard Medical School, Boston, MA USA

[4]Quantitative Imaging Center, Department of Radiology, Boston University School of Medicine, Boston, MA USA

[5]Department of Radiology, University of Erlangen-Nuremberg, Erlangen, Germany

[6]Department of Epidemiology and Biostatistics, University of California at San Francisco, San Francisco, CA USA

[7]University of Arizona Arthritis Center & Division of Rheumatology, University of Arizona College of Medicine, Tucson, AZ USA

[8]Duke Molecular Physiology Institute and Division of Rheumatology, Department of Medicine, Duke University School of Medicine, Durham, NC 27701 USA

Abstract word count: XXX

Body word count: X,XXX

Tables and Figures: X

Correspondence:

David J. Hunter, Rheumatology Department, Royal North Shore Hospital - Reserve Road, St. Leonards, NSW 2065, Australia.

Email: David.Hunter@sydney.edu.au

Phone: 61 2 9463 1887

Fax: 61 2 9463 1077

Running Title: FNIH Biomarkers of Knee Osteoarthritis Progression

Key words: knee osteoarthritis, biochemical markers, predictive validity, MRI

ABSTRACT

Objective: To determine the association between baseline and change in several imaging and biochemical biomarkers over 24 months and radiographic and pain progression over 48 months in knees with mild to moderate osteoarthritis.

Methods: We undertook a nested case-control study as part of the Osteoarthritis Biomarkers Consortium Project. We built multivariable logistic regression models to examine the association between baseline and change over 24 months in several imaging and biochemical biomarkers and knee OA radiographic and pain progression.

Results: Five hundred and fifty two (92%) of subjects had complete baseline and 24 month data on all biomarkers and thus were included in the analysis. The average age of the cohort was 62 years, 59% were females and average BMI was 31kg/m^2. The results of multivariable modelling for baseline biomarkers reveal consistent inclusion of central media femur cartilage thickness, medial meniscal volume and number of subregions affected by osteophyte, and inconsistent inclusion of some other parameters. The biomarker parameters that are consistently included in the multivariable modelling for 24 month change in biomarkers are effusion-synovitis, meniscal morphology and medial femorotibial cartilage thickness.

Conclusion: The study highlights the potential biomarkers that could provide prognostic utility in the context of OA disease modifying clinical trials.

Introduction

Osteoarthritis (OA) is a highly prevalent, disabling disease, with a commensurate tremendous individual and societal burden (1). Recent estimates suggest that 250 million people worldwide are affected by knee OA (2). The risk of mobility disability (defined as needing help walking or climbing stairs) attributable to knee OA alone is greater than that attributable to any other medical condition in people aged 65 years and older (3, 4).

Despite the substantial individual and societal burden, there are no therapies proven to reduce the risk of progression to OA (5). Instead, the "watchful waiting" of steady decline to end-stage joint disease is a major cause of disablement and loss of quality of life (6).

There are currently few pharmacologic agents beyond analgesic therapy for OA management. A major shift in the focus of OA research is critically needed to overcome barriers to the development of pharmacological treatments if an impact is to be made for the millions living with the chronic pain and disability of OA. Biomarkers enhance the success of every phase of the drug development process; they increase the frequency of successful phase transitions (chances of a drug candidate advancing to the next phase of development) (7, 8).

Further refinement and improvement of measures of joint structural change based on imaging and/or biochemical markers are needed to overcome the limited responsiveness of existing imaging biomarkers such as the poor relation in individual patients between joint structural pathology (e.g. joint space narrowing on radiographs) and symptomatic disease (9). To overcome these obstacles, the Foundation of National Institutes of Health (FNIH) OA Biomarkers Consortium undertook an extensive phase I biomarker qualification study from 2012 to 2015 using a nested case-control sample of progressive knee OA within the Osteoarthritis Initiative (OAI) (10). The overarching project objective was to establish the predictive validity of disease progression biomarkers and assess the responsiveness of several imaging and biochemical markers pertinent to knee OA. The results of this study have now been published in a series of papers focused on individual biomarkers (11-16).

The purpose of this paper is to take all of the biomarkers that were assessed in the phase 1 FNIH biomarkers study and conduct multivariable analyses. More specifically we will determine the association between biomarkers and knee osteoarthritis progression over 48

months. We investigated whether baseline biomarkers and changes in biomarkers from baseline to 24 months predicted radiographic and pain progression from baseline to 48 months in knees with mild to moderate OA.

Methods

Study Design

We undertook a nested case-control study, entitled The FNIH Osteoarthritis Biomarkers Consortium, using data from the Osteoarthritis Initiative (OAI) (17, 18). We selected subjects from the OAI with at least one knee with a Kellgren Lawrence grade (KLG) of 1, 2 or 3 at baseline based on the central reading of a standardized fixed-flexion radiograph and availability at baseline and 24 months of medial tibiofemoral (TF) joint space width (JSW) from knee radiographs, knee MRI, stored serum and urine specimens and clinical data. Minimum joint space width (minJSW) in the medial femorotibial compartment (MFTC) was measured using automated software (19).

We selected a predetermined number of index knees based on assessment of outcome from 24 to 48 months (one knee per subject) in four mutually exclusive groups: 1) knees with both radiographic and pain progression; 2) knees with radiographic but not pain progression; 3) knees with pain but not radiographic progression; and 4) knees with neither radiographic nor pain progression. For the purposes of this analysis we used the single contrast, comparing knees with both radiographic and pain progression (cases) vs. all other knees (controls). Radiographic progression, based on medial TF minJSW loss, and pain progression, based on an increase in WOMAC knee pain score above a minimum clinically important difference (MCID) were determined as previously described (18). Briefly, radiographic progression was defined as minJSW loss of ≥0.7mm; pain progression was defined as a persistent (sustained at ≥2 time points) increase of ≥9 points on the WOMAC pain subscale (0-100 scale)(17, 20, 21). If both of a subject's knees fell into any one group, one was randomly selected as the index knee. Knees with radiographic and pain progression by 12 months of follow-up, with radiographic lateral joint space narrowing (JSN) grade 2 or 3 at baseline, or with total knee replacement by 24 months, were excluded.

For better covariate balance among the groups, and to the extent feasible, knees selected for the four groups were frequency matched for KLG and individuals were matched for body

mass index (BMI) (kg/m^2) categories (<25; 25 to <27.5; 27.5 to <30; 30 to <35; ≥35). MRIs of the selected index knees were reviewed for artefacts that would interfere with image analysis. If artifacts were present, the knee and subject were excluded and a replacement selected. The sample size for the four groups was 194 (radiographic and pain progression), 103 (radiographic only), 103 (pain only) and 200 (neither radiographic nor pain progression) knees, respectively. In this analysis, we compared cases (radiographic and pain progression) vs. the remaining subjects, whom we considered controls. In order to compare models with different subsets of predictors, for this analysis we additionally required that subjects have complete data on all parameters in order to be included. As a sensitivity analysis we re-ran the models including all subjects, excluding the bone trabecular integrity biomarkers, as this class of biomarkers contributed the vast majority of the missing data.

Knee MRI Acquisition

MRI acquisition was performed using a 3 Tesla MRI system (Trio, Siemens Healthcare, Erlangen, Germany) at the four OAI clinical sites. Additional parameters of the full OAI pulse sequence protocol and sequence parameters have been published in detail elsewhere (22).

Knee MRI Measurement

Semi-Quantitative Analyses

Semi-quantitative scoring of MRIs is a method for performing a multi-feature assessment of the knee using conventional MRI acquisitions (23-26). This method uses an observer dependent semi-quantitative approach to score a variety of features that are currently believed to be relevant to the functional integrity of the knee and/or potentially involved in the pathophysiology of OA. Semi-quantitative scoring of MRI for OA has shown adequate reliability, specificity and sensitivity, and ability to detect lesion progression over 1-2 years (14, 15). The semi-quantitative scorings for the FNIH study were done by two musculoskeletal radiologists with 15 and 13 years of experience in semi-quantitative MRI assessment of OA (AG, FWR), respectively. MRIs were read according to the MRI Osteoarthritis Knee Score (MOAKS) system in sequential order and unblinded to the time point of acquisition to maximize sensitivity to change (27, 28). The readers were blinded to all clinical characteristics and case/control status. Based upon the results of the previous analyses in the FNIH cohort (14, 15), the measures used in the multivariable analyses were limited to cartilage morphology (number of subregions with surface area score >0), bone

marrow lesions (BML) (number of subregions with BML score >0), osteophytes (number of locations affected by any osteophyte) and effusion-synovitis (score >0).

Quantitative Cartilage Morphometry

The three dimensional (3D) coverage of an entire cartilaginous region by MRI allows for the direct quantification of cartilage volume, surface areas and thickness. Measurements of cartilage volume via MRI have been previously shown to correlate well with the ex vivo assessments of cartilage volume (stripped away from the bone) (29, 30). The annual changes in cartilage volume/thickness exceeded the precision errors and appear to be associated with clinical symptoms as well as with time to knee arthroplasty (31, 32). These methods are described in more detail elsewhere (33-36). Cartilage thickness analysis for this study relied on sagittal double-echo steady-state (DESS) imaging (11). Segmentation of the femorotibial cartilage surfaces, i.e. medial and lateral tibia and weight-bearing femur, was performed by 7 readers who had received continuous training for ≥5 years. All time-points of one knee (baseline, 12M and 24M) were processed as triplets by the same reader. The analysis center was blinded to case/control status and image acquisition order so that an unbiased rate of change could be determined in each group. All segmentations were quality control checked by one of two experts. Based upon the results of the previous analyses in the FNIH cohort (11), the cartilage thickness measures used in the multivariable analyses were limited to medial femorotibial (MFTC) compartment, external medial tibia (eMT), external central medial femoral (ecMF), central weight-bearing femoral (ccMF), and sum of the central medial tibia (cMT) and central weight-bearing femoral (ccMF) cartilage thickness (cMFTC).

Bone shape and area

Femur, tibia and patella bone surfaces were automatically segmented from 3T DESS-we images using active appearance models (AAM) provided by Imorphics (Manchester, UK) in a multistage process (12). Two measures of bone shape were used as primary independent variables: (1) the baseline measure and change in subchondral bone area (tAB) (mm^2)on the medial and lateral femur, tibia and patella; and, (2) the baseline and change in position on 3D shape vectors for the femur, tibia and patella bones. The OA vector is the line passing through the mean shape for the 2 populations; one with OA and one without. The shape vector for each bone is calculated by taking the principal components of the mean non-OA

shape and the mean OA shape, and drawing a straight line through them. The individual bone shape of study participants is represented by the same principal components as a consequence of using the AAM to search the images. Individual bone shapes are projected orthogonally onto the vector. Distances along the vector are normalized to a z-scale with the mean OA shape represented as +1 and the mean non-OA shape as -1. Based upon the results of the previous analyses in the FNIH cohort (11, 12), the bone area and shape measures used in the multivariable analyses were limited to medial patella area, and patella and femur shape.

Medial and lateral meniscal volumes

Medial and lateral meniscus volumes were automatically quantified using a computer-based framework (KneeIQ) by Biomediq A/S, Copenhagen, Denmark. The framework combines multi-atlas registration and supervised classification to segment the knee tissues (37). To correct for scanner drift, the geometric transformation by the multi-atlas registration was augmented with an affine intensity transformation that normalized the scan intensities across sites and over time. We used the measure of medial meniscus volume in the multivariable analyses.

Bone Trabecular Integrity (BTI)

Bone trabecular integrity (BTI) is a way of representing the state of the vertical and horizontal trabeculae of a standardized region of interest of bone. BTI as a biomarker measure has been extensively validated and is an excellent predictor of structural progression (38-40). Although only a modest predictor of clinically relevant progression (structural AND pain worsening) as demonstrated in phase I of the OA FNIH study, it is a strong predictor of structural radiographic progression in the OAI cohort on the whole (40).

Quantification of BTI is a two-step process(13, 41). In the initial step, the fractal signature analysis (FSA) is performed on the tibial subchondral bone of the medial compartment of a knee radiograph using a semi-automated software originally designed by Optasia Medical (Manchester, UK). In the second data reduction step, the fractal data are reduced to 6 parameters suitable for multivariable regression for evaluation of association with progression status or evaluation by cut-off scores for use as a screening tool to enrich OA

trials for progressors. Automated software for performing this step has been developed at Duke and was deployed on the extracted fractal data.

Based upon the results of the previous analyses in the FNIH cohort (13), the BTI measure used in the multivariable analyses was baseline quadratic term (vertical) z-score.

Biochemical markers

A total of 18 biomarkers were quantified at baseline, 12 and 24 months, in both serum and/or urine of the 600 selected subjects for this study as described in depth previously (16). These biomarkers were selected for this analysis based on their ability to predict OA progression in the primary analysis, including the following: serum hyaluronan (HA), N-terminal propeptide of collagen IIA (PIIANP), and the C-terminal crosslinked telopeptide of type I collagen (CTXI); serum and urine cross linked N-telopeptide of type I collagen (NTXI); and urine C-terminal crosslinked telopeptide type II collagen (*CTXII*), Col2-3/4 C terminal cleavage product of type II collagen (C2C HUSA), alpha and beta isomerized versions of the C-terminal crosslinked telopeptide of type I collagen (CTXIα and CTXIβ) and creatine (Cr) with which to normalize urinary biomarkers. All analyses were conducted by a CLIA and CAP certified division within LabCorp. Inter-assay coefficients of variation (CVs) ranged from 3.0% to 12.3% as previously described. The technical performance metrics of these biomarkers have been extensively evaluated for sensitivity (lower limit of quantitation and limit of detection), accuracy, spike and recovery, precision (intra- and inter-assay), specificity, and freeze thaw stability (reports available at OAI website). Based upon the results of the previous analyses in the FNIH cohort (16), the biochemical marker measures used in the multivariable analyses were serum-HA and serum- PIIANP.

Statistical Analysis

We used logistic regression to evaluate the association between cases status and biomarkers. We used three different stepwise selection methods to determine the best subset of predictors, selection based on AIC (Akaike Information Criterion), SBC (Schwarz Bayesian Criterion), and p-value (p=0.2 for entry/0.1 for remain). The AIC and SBC differ with respect to model fitting: the AIC tends to favor more complex models that risk overfitting, while the SBC tends to favor less complex models that risk underfitting. For both sets of

models, baseline and change in biomarkers, we first considered models with imaging parameters only. Then, we added the biochemical markers to the models.

We present the Area Under (AUC) the Receiver Operating Characteristic (ROC) curve (C-statistic), the integrated discrimination improvement (IDI) and the category-less net reclassification (NRI) for each model. The AUCs are presented for the unadjusted, adjusted for covariates (sex, race, baseline minJSW, baseline WOMAC pain score, age, BMI, KLG, use of pain medications), and adjusted with 10 fold cross validation. The IDI and NRI are calculated as improvement vs. the model with covariates only and are calculated under 10 fold cross-validation. We subset on patients with complete biomarker data.

For the BTI and biochemical markers change over 24 month was quantified both as absolute change (24 month value minus baseline value) and time-integrated concentration (TICs). TICs are equivalent to the area under the curve defined by the individual values for the specific time interval. We ran two sets of analyses for the 24 month change models, one including absolute change parameters for BTI and biochemical markers, and another including TICs for these markers.

Results

Demographic characteristics

Five hundred and fifty two (92%) of subjects had complete baseline and 24 month data on all biomarkers and thus were included in the analysis, 173 cases and 379 controls. The demographic characteristics of the analytic cohort and allocated groups are presented in Table 1. The average age of the cohort was 62 years, 59% were females and average BMI was 31kg/m^2. The cases and controls were well balanced on all demographic and clinical covariates, except baseline KLG. The sample with complete biomarker data is broadly comparable to the total sample (n=600), presented in Supplementary Table 1.

Baseline prediction

The results of multivariable modelling for baseline biomarkers are presented in Table 2. For the imaging biomarkers, the results are broadly consistent for models one through three with consistent inclusion of central media femur cartilage thickness (ccMF), medial meniscal volume and number of subregions affected by osteophyte. When biochemical markers are added to the imaging biomarkers we see that the patellar vector of three-dimensional

shape, cartilage thickness of the central medial femorotibial compartment (cMFTC), cartilage thickness of the external medial tibia (eMT) and serum PIIANP are also included in the stepwise AIC model (M4). There are a couple of things that are important to note: the overall C-statistic is not much improved from the earlier models with imaging biomarkers only. This model (M4) includes multiple cartilage thickness measures, despite the fact that the markers are highly correlated. This results in some unusual parameter estimates, for example, the odds ratios for ccMF and eMT are less than one, indicating that more cartilage is associated with a decreased odds of being a case while for cMFTC the odds ratio is greater than one, indicating that more cartilage is associated with increased odds of being a case which is in contrast with univariate results (the baseline univariate associations are presented in Supplementary Table 2). This is a consequence of forcing several highly correlated markers into the same model.

Change over 24 months prediction

The results of multivariable modelling for 24 month change in biomarkers are presented in Table 2. The change in biomarker parameters that are consistently included in each model are effusion-synovitis, meniscal morphology and medial femorotibial cartilage thickness (MFTC). Some other biomarker parameter changes are inconsistently included in these models, including: number of areas of cartilage morphology with worsening in thickness, number of areas of cartilage morphology with worsening in surface area, quadratic BTI term (vertical), lateral femur total area of bone, serum HA and PIIANP. In stepping from the imaging biomarker models to those that include biochemical biomarkers, again there is little change in the overall C-statistic. The 24 month change univariate associations are presented in Supplementary Table 3.

Discussion

The baseline biomarkers that consistently predicted subsequent radiographic and pain progression were cartilage thickness of the central media femur, medial meniscal volume and number of subregions affected by osteophyte. Some other markers were inconsistently included in these models including patellar vector of three-dimensional shape, cartilage thickness of the central medial femorotibial compartment, cartilage thickness of the medial tibia and serum PIIANP. The change in biomarkers that consistently predicted radiographic

and pain progression were effusion-synovitis, meniscal morphology and medial femorotibial cartilage thickness. Some other biomarker parameter changes are inconsistently included in these models including number of areas of cartilage morphology with worsening in thickness, number of areas of cartilage morphology with worsening in surface area, quadratic BTI term (vertical), lateral femur total area of bone, serum HA and PIIANP.

The FNIH cohort has a number of important strengths. Firstly the careful selection of the case and control samples represents a rich substudy within the OAI. The panel of biomarkers, both imaging and biochemical, represents the most comprehensive analysis of these measures in a direct head-to-head comparison to date.

This study also demonstrates that the biomarkers that predict change are not entirely consistent when identifying baseline and comparing that to changes in biomarkers and their ability to predict radiographic and pain progression. This is important to note as they likely serve distinct purposes. Firstly, the baseline biomarkers that predict subsequent clinical outcomes may be helpful in enriching study samples by stratifying those most likely to have adverse clinical outcomes that potentially would be able to demonstrate a response in the context of a clinical trial. This is in contrast to the change in biomarkers which may provide utility for alternate study endpoints to the less responsive standards that we currently use. This is particularly important in enhancing the efficiency and shortening the duration of phase II clinical trials and by virtue of this reducing cost and improving time to market (42).

It is important to acknowledge the univariate analyses that are presented in this manuscript and independently presented in a number of other publications (11-16). The multivariable analyses facilitate decondensing that complexity through direct head-to-head comparison for clear delineation of the biomarkers that perform better in the roles of baseline and change prediction of longer term structural and clinical outcome.

There are a number of important limitations that are important to be cognisant of when reading this manuscript. Firstly, the analyses are based on a subsample of the FNIH cohort. This occurred largely due to missing data in FSA. It is important to note that the total sample is broadly consistent with the subsample with regards demographic characteristics. The selection of biomarkers in the study reflects an expert consensus on valid and commercially available biomarkers which were selected now approximately five years ago. There are potentially other markers that could be analysed and potentially have better performance.

We would like to highlight the opportunity that the FNIH and OAI cohorts provide with regards providing this validation subsample to allow direct comparison with the results that we have provided.

The study highlights the potential biomarkers that could provide utility in the context of OA disease modifying clinical trials. It is important to differentiate and recognise markers that may provide baseline prognostic information as distinct from those that may be efficacy of intervention biomarkers. Further work is required to qualify these biomarkers so that they can be used for registered clinical trial purposes.

Author contributions

- Study conception and design: Hunter, Collins, Nevitt, Kraus, Katz, Losina
- Acquisition of data: Bowes, Collins, Nevitt, Lynch
- Analysis & interpretation of data: Hunter, Collins, Nevitt, Kraus, Katz, Losina, Roemer, Guermazi, Bowes
- Writing of first manuscript draft: Hunter
- Critical manuscript revision and approval of final manuscript: All authors
- David Hunter had full access to all of the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis.

Acknowledgements and financial support

Scientific and financial support for the FNIH OA Biomarkers Consortium and the study are made possible through grants, direct and in-kind contributions provided by: AbbVie; Amgen Inc.; Arthritis Foundation; Bioiberica S.A.; DePuy Mitek, Inc.; Flexion Therapeutics, Inc.; GlaxoSmithKline; Merck Serono; Rottapharm | Madaus; Sanofi; Stryker; The Pivotal OAI MRI Analyses (POMA) Study, NIH HHSN2682010000. We thank the Osteoarthritis Research Society International (OARSI) for their leadership and expertise on the FNIH OA Biomarker Consortium project. The OAI is a public-private partnership comprised of five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National Institutes of Health. Funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc. Private sector funding for the Consortium and OAI is managed by the FNIH.

Dr Hunter is funded by an NHMRC Practitioner Fellowship and support from the FNIH OA Biomarkers Consortium. Dr. Kraus is funded by grants from the FNIH, the NIH/NIA Claude D Pepper 5P30 AG028716 and NIH/NIAMS 5P01-AR050245. Dr. Losina is funded by grants from NIAMS K24 AR057827, P60 AR47782.

Disclosure

- David Hunter is funded by an NHMRC Practitioner Fellowship and is supported from the FNIH OA Biomarkers Consortium.
- Virginia Kraus, Michael Nevitt, Jamie Collins and Elena Losina are funded by grants from the FNIH.

- Elena Losina is funded by grants from NIAMS R01 AR064320, K24 AR057827, P60 AR47782
- Mike Bowes is a part time employee and co-owner of Imorphics.
- Ali Guermazi is President of Boston Imaging Core Lab, LLC (BICL), a company providing MRI reading services to academic researchers and to industry. He has provided consulting services to Merck Serono, Genzyme, OrthoTrophix and TissueGene.
- Frank Roemer is CMO and co-owner of BICL.
- Jamie Collins, Michael C. Nevitt, John A. Lynch, Joanne Jordan and Jeff N. Katz have no conflicts of interest to report

References

Table 1. Cohort Characteristics for Analytic Sample (n=552). (mean (SD) presented for continuous variable. n (%) presented for categorical variables.

Label	Level	Cases (radiographic and pain progression) n=173	Controls			
			All three groups n=379	Radiographic progression only n=93	Pain progression only n=97	Neither radiographic or pain progression n=189
Age		61.9 (8.8)	61.6 (8.9)	63.1 (8.3)	59.2 (8.7)	61.5 (9.1)
Sex	Male	73 (42%)	151 (40%)	50 (54%)	34 (35%)	67 (35%)
	Female	100 (58%)	228 (60%)	43 (46%)	63 (65%)	122 (65%)
BMI		30.8 (4.9)	30.7 (4.8)	30.7 (4.7)	31.1 (5.0)	30.5 (4.8)
Baseline KLG	1	22 (13%)	50 (13%)	14 (15%)	13 (13%)	23 (12%)
	2	76 (44%)	208 (55%)	44 (47%)	58 (60%)	106 (56%)
	3	75 (43%)	121 (32%)	35 (38%)	26 (27%)	60 (32%)
White Race	No	36 (21%)	75 (20%)	10 (11%)	27 (28%)	38 (20%)
	Yes	137 (79%)	304 (80%)	83 (89%)	70 (72%)	151 (80%)
Baseline pain med	No	115 (66%)	275 (73%)	74 (80%)	62 (64%)	139 (74%)
	Yes	58 (34%)	104 (27%)	19 (20%)	35 (36%)	50 (26%)
Baseline WOMAC Pain		10.2 (12.7)	12.6 (16.3)	16.5 (19.8)	9.6 (13.3)	13.0 (16.1)
Baseline minJSW		3.8 (1.4)	3.9 (1.0)	3.8 (1.2)	3.9 (1.0)	3.9 (1.0)

Table 2. Results of Multivariable Modeling – Baseline Biomarker to predict JSL + Pain cases (n=552)

	Imaging Biomarkers Only			Imaging + Biochemical Biomarkers		
MODEL **Selection Method**	**M1** **Stepwise, AIC**	**M2** **Stepwise, SBC**	**M3** **Stepwise, P-value**	**M4** **Stepwise, AIC**	**M5** **Stepwise, SBC**	**Stepwise, P-value**
Model characteristics					*Same as imaging only, M2*	*Same as imaging only, M3*
AUC (unadjusted)	0.686	0.646	0.683	0.694		
AUC (adjusted)	0.723	0.685	0.716	0.724		
AUC (adjusted, 10 fold cross-validation)	0.677	0.641	0.671	0.676		
IDI	0.0957	0.0562	0.0876	0.0961		
NRI	0.5142	0.4039	0.4792	0.5504		
%cases correctly reclassified[1]	27%	29%	24%	28%		
%controls correctly reclassified[2]	25%	11%	24%	27%		
Biomarkers Included						
BICL: BL OST - Number of subregions affected by any Osteophyte Category (0-2/3-5/6+)	0.0181	<0.0001	0.0160	0.0172		
0-2	REF	REF	REF	REF		
3-5	1.7 (0.8, 3.7)	1.9 (0.9, 3.9)	1.7 (0.8, 3.6)	1.7 (0.8, 3.7)		
6+	2.8 (1.3, 5.9)	4.0 (2.0, 7.8)	2.7 (1.3, 5.8)	2.8 (1.3, 5.9)		
znPatellaOAVector0: IMORPH: BL Patella Vector of 3D Shape z-Score	0.0210 0.8 (0.6, 1.0)	0.0015 0.7 (0.6, 0.9)		0.0215 0.8 (0.6, 1.0)		

zV00EBMFMTH: CHON: BL ecMF.ThCtAB mean cart thickness - central medial femur (external) [mm] z-score	0.0015 0.7 (0.5, 0.9)		0.0009 0.7 (0.5, 0.8)	0.0012 0.7 (0.5, 0.8)		
zV00MedMeniscus: BIOMEDIQ: BL Medial Meniscus volume	0.0833 1.3 (1.0, 1.7)		0.0448 1.3 (1.0, 1.8)	0.0872 1.3 (1.0, 1.7)		
BICL: BL CART - Surface Area - Number of subregions with score>0 across entire knee category (0-1/2-4/5-7/8+	0.0054		0.0020	0.0063		
0-1	REF		REF	REF		
2-4	1.8 (0.6, 5.3)		1.8 (0.6, 5.3)	1.8 (0.6, 5.2)		
5-7	3.7 (1.2, 11.4)		3.9 (1.3, 12.1)	3.6 (1.2, 11.1)		
8+	5.2 (1.4, 18.6)		5.8 (1.6, 20.8)	5.0 (1.4, 18.0)		
zSer_PIIANP: Biochem: Serum PIIANP (interpolated research value if below lower limit) z-score				0.2414 0.9 (0.7, 1.1)		

IDI = integrated discrimination improvement

NRI = category-less net reclassification

[1]%cases correctly reclassified = % of cases with a higher probability of being a case in new model vs. old;
[2] %controls correctly reclassified = % of controls with a lower probability of being a case in new model vs. old;

Table 3. Results of Multivariable Modeling – 24 Month change in Biomarker to predict JSL + Pain cases (n=552)

	Imaging Biomarkers Only			Imaging + Biochemical Biomarkers		
MODEL	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**
Selection Method	**Stepwise, AIC**	**Stepwise, SBC**	**Stepwise, p-value**	**Stepwise, AIC**	**Stepwise, SBC**	**Stepwise, p-value**
Model characteristics						
AUC (unadjusted)	0.747	0.709	0.726	0.754	0.709	0.737
AUC (adjusted)	0.758	0.723	0.742	0.767	0.723	0.747
AUC (adjusted, 10 fold cross-val)	0.708	0.680	0.697	0.715	0.680	0.698
IDI (vs. covariates only model)	0.1392	0.1096	0.1296	0.1451	0.1096	0.1330
NRI (vs. covariates only model)	0.6860	0.5377	0.6504	0.7070	0.5377	0.6338
%cases correctly reclassified[1]	20%	12%	16%	22%	12%	18%
%controls correctly reclassified[2]	48%	42%	49%	49%	42%	45%
Biomarkers Included						
BICL: 24M Chg Change in MOAKS Whole Knee Effusion Category	0.0218	0.0007	0.0088	0.0293	0.0007	0.0097
Worsen vs No Change	1.8 (1.1, 2.9)	2.2 (1.4, 3.4)	1.9 (1.2, 3.1)	1.8 (1.1, 2.9)	2.2 (1.4, 3.4)	1.9 (1.1, 3.0)
Worsen vs Improve	2.5 (1.2, 5.1)	3.2 (1.6, 6.4)	2.6 (1.3, 5.4)	2.3 (1.1, 4.8)	3.2 (1.6, 6.4)	2.7 (1.3, 5.5)
BICL: 24M Chg Cart Morphology - # of areas with worsening in thickness	0.1125		0.0581	0.0568		0.0874
1 vs. 0	1.4 (0.8, 2.3)		1.5 (0.9, 2.4)	1.4 (0.9, 2.4)		1.5 (0.9, 2.4)
2 vs. 0	1.4 (0.7, 2.7)		1.5 (0.8, 2.9)	1.5 (0.8, 3.0)		1.4 (0.7, 2.7)
3+ vs. 0	2.6 (1.2, 5.8)		2.8 (1.3, 6.2)	3.0 (1.3, 6.8)		2.6 (1.2, 5.9)
BICL: 24M Chg Meniscal Morph: Any regions with worsening	0.0633	0.0017	0.0094	0.0387	0.0017	0.0608
Yes vs. No	1.7 (1.0, 3.1)	2.4 (1.4, 4.1)	2.1 (1.2, 3.7)	1.9 (1.0, 3.4)	2.4 (1.4, 4.1)	1.7 (1.0, 3.1)

BICL: 24M Chg Worsening in MOAKS Inter-Condylar Synovitis	0.0587		0.0258	0.0759		
Yes vs. No	1.9 (1.0, 3.7)		2.1 (1.1, 4.1)	1.8 (0.9, 3.6)		
BICL: 24M Chg MOAKS Cartilage Morphology - entire knee number of areas with worsening in surface area (include within-grade change)	0.0821			0.0575		0.0559
1 vs. 0	1.4 (0.8, 2.3)			1.4 (0.8, 2.3)		1.4 (0.9, 2.3)
2 vs. 0	1.7 (0.9, 3.1)			1.7 (0.9, 3.2)		1.7 (0.9, 3.2)
3+ vs. 0	2.3 (1.2, 4.5)			2.5 (1.3, 4.9)		2.4 (1.3, 4.6)
CHON: 24M Chg Change in mean cartilage thickness - medial tib-fem compartment (MFTC.ThCtAB) [mm]	0.0052	<.0001	0.0012	0.0045	<.0001	0.0023
OR for each 1 unit increase in SD[3]	1.4 (1.1, 1.8)	1.6 (1.3, 2.0)	1.5 (1.2, 1.9)	1.4 (1.1, 1.8))	1.6 (1.3, 2.0)	1.5 (1.1, 1.9)
BTI: Quadratic Term (Vertical) 24M Change Z-Score	0.1520					
OR for each 1 unit increase in SD[3]	1.2 (0.9, 1.4)					
Serum PIIANP: 24m Chg (numeric, interpolated research value if below lower limit)				0.1344		
OR for each 1 unit increase in SD[3]				1.2 (0.9, 1.5)		
Serum HA: 24M Chg (numeric, interpolated research value if below lower limit) 24 Month Change				0.0855		
OR for each 1 unit increase in SD[3]				1.2 (1.0, 1.5)		
IMORPH: 24M Chg Lateral Femur (tAB), mm				0.1790		
OR for each 1 unit increase in SD[3]				0.9 (0.7, 1.1)		

IDI = integrated discrimination improvement; NRI = category-less net reclassification

[1]%cases correctly reclassified = % of cases with a higher probability of being a case in new model vs. old;
[2] %controls correctly reclassified = % of controls with a lower probability of being a case in new model vs. old;
[3]coded such that increasing OR = increasing change

Table 4. Results of Multivariable Modeling – 24 Month change in Biomarker to predict JSL + Pain cases; TICs used for biochemical and BTI biomarkers (n=552)

	Imaging Biomarkers Only			Imaging + Biochemical Biomarkers		
MODEL	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**
Selection Method	**Stepwise, AIC**	**Stepwise, SBC**	**Stepwise, p-value**	**Stepwise, AIC**	**Stepwise, SBC**	**Stepwise, p-value**
Model characteristics						
AUC (unadjusted)	0.749	0.708	0.735	0.751	0.717	0.740
AUC (adjusted)	0.767	0.722	0.757	0.774	0.733	0.764
AUC (adjusted, 10 fold cross-val)	0.720	0.679	0.712	0.721	0.690	0.714
IDI (vs. covariates only model)	0.1495	0.1104	0.1407	0.149	0.117	0.142
NRI (vs. covariates only model)	0.7358	0.5378	0.6607	0.710	0.547	0.650
%cases correctly reclassified[1]	25%	12%	21%	27%	17%	23%
%controls correctly reclassified[2]	48%	42%	45%	44%	38%	42%
Biomarkers Included						
BICL: 24M Chg Change in MOAKS Whole Knee Effusion Category	0.0399	0.0007	0.0172	0.0328	0.0007	0.0144
No Change vs. Improve	1.4 (0.7, 2.7)	1.4 (0.8, 2.8)	1.4 (0.7, 2.6)	1.4 (0.7, 2.6)	1.4 (0.7, 2.7)	1.3 (0.7, 2.6)
Worsen vs Improve	2.3 (1.1, 4.8)	3.2 (1.6, 6.4)	2.5 (1.2, 5.1)	2.3 (1.1, 4.8)	3.1 (1.5, 6.3)	2.5 (1.2, 5.1)
BICL: 24M Chg Cart Morphology - # of areas with worsening in thickness	0.1020		0.0507	0.1385		0.0767
1 vs. 0	1.4 (0.8, 2.3)		1.5 (0.9, 2.4)	1.3 (0.8, 2.2)		1.4 (0.8, 2.3)
2 vs. 0	1.4 (0.7, 2.7)		1.5 (0.8, 3.0)	1.3 (0.7, 2.6)		1.5 (0.8, 2.8)
3+ vs. 0	2.7 (1.2, 6.0)		2.9 (1.3, 6.4)	2.6 (1.1, 5.8)		2.7 (1.2, 6.1)
BICL: 24M Chg Meniscal Morph: Any regions with worsening	0.0615	0.0016	0.0088	0.0553	0.0017	0.0082
Yes vs. No	1.8 (1.0, 3.2)	2.4 (1.4, 4.2)	2.2 (1.2, 3.7)	1.8 (1.0, 3.2)	2.4 (1.4, 4.2)	2.1 (1.2, 3.8)

BICL: 24M Chg Worsening in MOAKS Inter-Condylar Synovitis	0.0457		0.0207	0.0557		0.0286
Yes vs. No	2.0 (1.0, 3.9)		2.4 (1.4, 4.2)	1.9 (1.0, 3.8)		2.1 (1.1, 4.1)
BICL: 24M Chg MOAKS Cartilage Morphology - entire knee number of areas with worsening in surface area (include within-grade change)	0.0861			0.1079		
1 vs. 0	1.4 (0.8, 2.3)			1.4 (0.9, 2.4)		
2 vs. 0	1.7 (0.9, 3.2)			1.6 (0.8, 3.0)		
3+ vs. 0	2.3 (1.2, 4.5)			2.2 (1.2, 4.4)		
CHON: 24M Chg Change in mean cartilage thickness - medial tib-fem compartment (MFTC.ThCtAB) [mm]	0.0039	<0.0001	0.0014	0.0032	<.0001	0.0013
OR for each 1 unit increase in SD[3]	1.4 (1.1, 1.8)	1.6 (1.3, 2.0)	1.5 (1.2, 1.9)	1.4 (1.1, 1.9)	1.6 (1.3, 2.0)	1.5 (1.2, 1.9)
BIOCHEM: 24M TIC Urine-CTXIalpha (Creatinine adjusted)				0.0526	0.0145	0.0503
OR for each 1 unit increase in SD				1.2 (1.0, 1.5)	1.3 (1.1, 1.6)	1.2 (1.0, 1.5)
BTI: 24M TIC Composite Score (horiz int, vert slope, vert quad term (vert params reverse coded)	0.0103 1.3 (1.1, 1.7)		0.0094	0.0235		0.0204
OR for each 1 unit increase in SD			1.3 (1.1, 1.6)	1.3 (1.0, 1.6)		1.3 (1.0, 1.6)

IDI = integrated discrimination improvement; NRI = category-less net reclassification

[1]%cases correctly reclassified = % of cases with a higher probability of being a case in new model vs. old;
[2] %controls correctly reclassified = % of controls with a lower probability of being a case in new model vs. old;
[3]coded such that increasing OR = increasing change

Appendix

File 1. Baseline variables Selected for modelling based upon univariate analyses (11-16)

a. **Cartilage (Chondrometrics)**
    i. eMT.ThCtAB
    ii. ccMF.ThCtAB
    iii. ecMF.ThCtAB
    iv. MFTC.ThCtAB
    v. cMFTC.ThCtAB
b. **Bone Area and Shape (Imorphics)**
    i. Medial Patella Area
    ii. Femur shape (vector)
    iii. Patella Shape (vector)
c. **Biomediq**
    i. zV00MedMeniscus
d. **Bone Trabecular Integrity (BTI)**
    i. BL Quadratic Term (Vertical) Z-Score
e. **Semi-quantitative measures (BICL)**
    i. Cartilage morphology: number of subregions with surface area score >0
    ii. BMLs: number of subregions with bml score >0
    iii. Osteophytes: number of locations affected by any osteophyte
    iv. Effusion-synovitis score
f. **Biochem**
    i. Serum-HA
    ii. Serum-PIIANP

File 2. Change variables Selected for modelling based upon univariate analyses

- **a. Biomediq**
  - i. zPatCart24_0: Change in Patellar Cartilage volume
- **b. Cartilage (Chondrometrics)**
  - i. ΔcMFTC
  - ii. ΔcMT
  - iii. ΔccMF
  - iv. ΔMFTC
- **c. Bone area and shape (Imorphics)**
  - i. Medial Area
    1. Femur
    2. Tibia
    3. Patella
    4. Trochlea
  - ii. Lateral Area
    1. Femur
    2. Tibia
    3. Patella
    4. Trochlea
  - iii. Shape (Vector)
    1. Femur
    2. Tibia
    3. Patella
- **d. Semi-quantitative measures (BICL) (taken from Multivariate models (Model 7))**
  - i. Cartilage: number of areas with worsening in thickness
  - ii. Cartilage: number of areas with worsening in surface area
  - iii. Meniscal morphology: any regions with worsening
  - iv. Synovitis-Effusion : change in effusion category
  - v. Hoffa-Synovitis: change in synovitis category

e. Bone Trabecular Integrity (BTI)

i. Change

1. Quadratic Term (Vertical) 12M and 24M Change Z-Score
2. Composite Score (sum of horizontal intercept, vertical slope, and vertical quadratic term (vertical parameters reverse coded))

ii. TIC

1. Quadratic Term (Vertical) 12M and 24M Change Z-Score
2. Composite Score (sum of horizontal intercept, vertical slope, and vertical quadratic term (vertical parameters reverse coded))

f. Biochemical Biomarkers

i. Change

1. Serum-CTXI
2. Serum-HA
3. Urine CTX-1beta
4. Urine CTX-1alpha
5. Urine C2C
6. Urine CTXII
7. Urine NTXI

Supplementary Table 1. Cohort Characteristics for Complete Sample (n=600). (mean (SD) presented for continuous variable. n (%) presented for categorical variables.

Label	Level	Cases (radiographic and pain progression) n=194	Controls			
			All three groups n=406	Radiographic progression only n=103	Pain progression only n=103	Neither radiographic or pain progression n=200
Age		62.0 (8.8)	61.3 (8.9)	63.1 (8.3)	59.2 (8.7)	61.5 (9.1)
Sex	Male	84 (43%)	163 (40%)	57 (55%)	36 (35%)	70 (35%)
	Female	110 (57%)	243 (60%)	46 (45%)	67 (65%)	130 (65%)
BMI		30.7 (4.8)	30.7 (4.8)	30.7 (4.7)	31.1 (5.0)	30.5 (4.8)
Baseline KLG	1	24 (12%)	51 (13%)	14 (14%)	13 (13%)	24 (12%)
	2	84 (43%)	222 (55%)	47 (46%)	61 (59%)	114 (57%)
	3	86 (44%)	133 (33%)	42 (41%)	29 (28%)	62 (31%)
White Race	No	39 (20%)	86 (21%)	12 (12%)	29 (28%)	45 (23%)
	Yes	155 (80%)	320 (79%)	91 (88%)	74 (72%)	155 (78%)
Baseline pain med	No	131 (68%)	292 (72%)	81 (79%)	66 (64%)	145 (73%)
	Yes	63 (32%)	114 (28%)	22 (21%)	37 (36%)	55 (28%)
Baseline WOMAC Pain		10.2 (13.0)	13.0 (16.7)	16.5 (19.8)	9.6 (13.3)	13.0 (16.1)
Baseline minJSW		3.8 (1.4)	3.9 (1.1)	3.8 (1.2)	3.9 (1.0)	3.9 (1.0)

Supplementary Table 2. Baseline univariate associations (one biomarker at a time, adjusted for covariates, n=552)

Biomarker	n	Case* (n=173)	Control* (n=379)	OR**	P-value**
CHON: BL eMT.ThCtAB mean cart thickness - medial tibia (external) [mm] z-score	552	-0.13 (1.08) -0.07	0.11 (0.93) 0.15	0.655	0.0013
CHON: BL ccMF.ThCtAB mean cart thickness - central medial femur (center) [mm] z-score	552	-0.17 (1.14) -0.15	0.11 (0.89) 0.05	0.591	0.0002
CHON: BL ecMF.ThCtAB mean cart thickness - central medial femur (external) [mm] z-score	552	-0.15 (1.11) -0.18	0.07 (0.89) 0.03	0.706	0.0046
CHON: BL MFTC.ThCtAB mean cart thickness - medial tib-fem compartment [mm] z-score	552	-0.08 (1.09) -0.26	0.06 (0.94) -0.04	0.733	0.0290
CHON: BL cMFTC.ThCtAB - mean cart thickness central medial tib-fem compartment (weight bearing) [mm] z-score	552	-0.07 (1.11) -0.16	0.07 (0.93) -0.03	0.754	0.0584
IMORPH: BL Medial Patella (tAB), mm2 Baseline z-Score (normalized 0,1)	552	0.15 (1.03) 0.11	-0.07 (0.98) -0.16	1.380	0.0132
IMORPH: BL Femoral Vector of 3D Shape z-Score	552	-0.17 (1.03) -0.08	0.11 (0.97) 0.14	0.750	0.0129
IMORPH: BL Patella Vector of 3D Shape z-Score	552	-0.22 (1.07) -0.21	0.12 (0.96) 0.15	0.691	0.0004
BIOMEDIQ: BL Medial Meniscus volume	552	0.11 (1.04) -0.10	-0.07 (0.96) -0.25	1.393	0.0181
FSA: BL Quadratic Term (Vertical) z-score	552	-0.07 (0.92) -0.09	0.03 (1.03) -0.11	0.854	0.1138
BL Composite Score: sum of horz int, vert slope, vert quad term (vert parameters reverse coded)	552	0.13 (0.95) 0.17	-0.07 (1.01) -0.03	1.253	0.0220
Serum HA (numeric, interpolated research value if below lower limit) STANDARDIZED by visit	550	0.04 (0.91) -0.22	-0.03 (1.05) -0.32	1.031	0.7659
Serum PIIANP (numeric, interpolated research value if below lower limit) STANDARDIZED by visit	552	-0.09 (1.04) -0.16	0.02 (0.95) -0.03	0.892	0.2434
BICL: BL BML - Number of subregions affected by any BML (0/1/2/3/4/5+)	552				0.0015
0		11 (6%)	53 (14%)	REF	

Biomarker	n	Case* (n=173)	Control* (n=379)	OR**	P-value**
1		22 (13%)	71 (19%)	1.464	
2		35 (20%)	85 (22%)	2.063	
3		38 (22%)	83 (22%)	2.302	
4		26 (15%)	46 (12%)	2.749	
5+		41 (24%)	41 (11%)	5.271	
BICL: BL CART - Surface Area - Number of subregions with score>0 across entire knee category (0-1/2-4/5-7/8+	552				<.0001
0-1		5 (3%)	39 (10%)	REF	
2-4		48 (28%)	162 (43%)	2.563	
5-7		88 (51%)	144 (38%)	6.184	
8+		32 (18%)	34 (9%)	10.373	
BICL: BL OST - Number of subregions affected by any Osteophyte Category (0-2/3-5/6+)	552				<.0001
0-2		13 (8%)	77 (20%)	REF	
3-5		37 (21%)	113 (30%)	1.822	
6+		123 (71%)	189 (50%)	4.050	
BICL: BL SYN: Inter-Condylar Synovitis Category (0,1,2-3)	552				0.0037
0		53 (31%)	175 (46%)	REF	
1		102 (59%)	179 (47%)	1.912	
2-3		18 (10%)	25 (7%)	2.228	

*mean(SD); median presented for continuous variables, n(%) presented for categorical variables.
** adjusted for covariates: sex, race, baseline age, BMI, JSW, PAIN, Pain medications, KLG

Supplementary Table 3. 24 month change univariate associations (one biomarker at a time, adjusted for covariates, n=552)

Label	Case* (n=173)	Control* (n=379)	OR**	P-value**
CHON: 24M Chg Change in mean cartilage thickness - central medial tib-fem compartment (weight bearing) (cMFTC.ThCtAB) [mm] Z-score	-0.34 (1.14) -0.14	0.20 (0.82) 0.32	1.854	<.0001
CHON: 24M Chg Change in mean cartilage thickness - central medial femur (center) (ccMF.ThCtAB) [mm] Z-score	-0.34 (1.11) -0.09	0.19 (0.82) 0.33	1.837	<.0001
CHON: 24M Chg Change in mean cartilage thickness - medial tibia (center) (cMT.ThCtAB) [mm] Z-score	-0.23 (1.18) -0.09	0.14 (0.83) 0.23	1.471	0.0001
CHON: 24M Chg Change in mean cartilage thickness - medial tib-fem compartment (MFTC.ThCtAB) [mm] Z-score	-0.34 (1.13) -0.15	0.20 (0.82) 0.33	1.866	<.0001
IMORPH: 24M Chg Lateral Femur (tAB), mm2 (24M minus BL) Z-Score	0.07 (1.08) 0.04	-0.04 (0.94) -0.01	1.160	0.1414
IMORPH: 24M Chg Lateral Patella (tAB), mm2 (24M minus BL) Z-Score	0.19 (0.93) 0.08	-0.12 (1.01) -0.19	1.405	0.0009
IMORPH: 24M Chg Lateral Tibia (tAB), mm2 (24M minus BL) Z-Score	0.20 (0.99) 0.18	-0.14 (0.92) -0.14	1.471	0.0002
IMORPH: 24M Chg Medial Femur area of bone (tAB), mm2 (24M minus BL) Z-Score	0.24 (1.12) 0.04	-0.16 (0.86) -0.28	1.583	<.0001
IMORPH: 24M Chg Medial Patella (tAB), mm2 (24M minus BL) Z-Score	0.17 (0.96) 0.09	-0.11 (0.99) -0.13	1.371	0.0020
IMORPH: 24M Chg Medial Tibia (tAB), mm2 (24M minus BL) Z-Score	0.20 (1.03) 0.13	-0.13 (0.92) -0.25	1.447	0.0003
IMORPH: 24M Chg Lateral PF region on Femur (tAB), mm2 (24M minus BL) Z-Score	0.25 (1.00) 0.15	-0.12 (0.98) -0.15	1.493	<.0001
IMORPH: 24M Chg Medial PF region on Femur (tAB), mm2 (24M minus BL) Z-Score	0.25 (1.07) 0.15	-0.18 (0.88) -0.31	1.566	<.0001
IMORPH: 24M Chg Femoral Vector of 3D Shape, Normalised units +1 = mean OA shape, -1 = mean non-OA shape (BL minus 24M) Z-Score	0.23 (1.13) 0.03	-0.16 (0.89) -0.34	1.549	<.0001
IMORPH: 24M Chg Patella Vector of 3D Shape, Normalised units +1 = mean OA shape, -1 = mean non-OA shape (BL minus 24M) Z-Score	0.08 (0.92) -0.00	-0.08 (1.02) -0.12	1.194	0.0691
IMORPH: 24M Chg Tibial Vector of 3D Shape, Normalised units +1 = mean OA shape, -1 = mean non-OA shape (BL minus 24M) Z-Score	0.27 (1.03) 0.10	-0.15 (0.96) -0.20	1.568	<.0001
BIOMEDIQ: 24M Chg Patellar Cartilage volume Z-Score	0.15 (0.96) 0.05	-0.06 (0.92) -0.08	1.278	0.0196

Label	Case* (n=173)	Control* (n=379)	OR**	P-value**
BTI: Quadratic Term (Vertical) 24M Change Z-Score	0.08 (1.03) 0.05	-0.04 (0.98) -0.06	1.148	0.1470
BTI: Quadratic Term (Vertical) 24M TIC Z-Score	-0.07 (1.02) -0.22	0.03 (0.98) -0.02	0.842	0.0769
BTI 24m TIC Composite Score: sum of horz int, vert slope, vert quad term (vert parameters reverse coded) Z-Score	0.16 (0.98) 0.24	-0.08 (1.00) -0.04	1.340	0.0035
BTI 24m CHG Composite Score: sum of horz int, vert slope, vert quad term (vert parameters reverse coded) Z-Score	-0.06 (1.07) -0.06	0.03 (0.97) 0.05	0.920	0.3799
Serum HA (numeric, interpolated research value if below lower limit) 24 Month Change Z-SCORE	0.07 (1.06) 0.01	-0.05 (0.96) -0.05	1.160	0.1183
Serum PIIANP (numeric, interpolated research value if below lower limit) 24 Month Change Z-SCORE	0.13 (0.96) 0.12	-0.03 (0.98) 0.06	1.208	0.0595
Serum CTXI (numeric, interpolated research value if below lower limit) 24 Month Indiv Longitudinal Burden Z-SCORE	0.14 (1.03) 0.02	-0.07 (0.97) -0.25	1.271	0.0117
Serum HA (numeric, interpolated research value if below lower limit) 24 Month Indiv Longitudinal Burden Z-SCORE	0.08 (0.98) -0.13	-0.06 (0.99) -0.36	1.144	0.1950
Urine C2C creatinine adj (numeric, interpolated research value if below lower limit) 24 Month Indiv Longitudinal Burden Z-SCORE	0.10 (0.94) -0.09	-0.08 (0.93) -0.25	1.212	0.0731
Urinary CTXII creatinine adj (numeric, interpolated research value if below lower limit) 24 Month Indiv Longitudinal Burden Z-SCORE	0.15 (0.92) 0.03	-0.09 (1.00) -0.31	1.296	0.0104
Urine NTXI creatinine adj (numeric, interpolated research value if below lower limit) 24 Month Indiv Longitudinal Burden Z-SCORE	0.12 (0.96) -0.03	-0.08 (0.98) -0.22	1.269	0.0156
Urine CTX-1a (Ur_alpha) creatinine adj (numeric, interpolated research value if below lower limit) 24 Month Indiv Longitudinal Burden Z-SCORE	0.14 (1.01) 0.04	-0.09 (0.96) -0.31	1.285	0.0108
Urine CTX-1ß (Ur_beta) creatinine adj (numeric, interpolated research value if below lower limit) 24 Month Indiv Longitudinal Burden Z-SCORE	0.10 (1.05) -0.09	-0.07 (0.94) -0.28	1.239	0.0278
BICL: 24M Chg Meniscal Morphology: Any regions with worsening (0=No 1=Yes)				<.0001
No	127 (73%)	340 (90%)	REF	
Yes	46 (27%)	38 (10%)	3.621	
BICL: 24M Chg MOAKS Cartilage Morphology - entire knee number of areas with worsening in thickness (0/1/2/3+)				<.0001
No Change	74 (43%)	250 (66%)	REF	

Label	Case* (n=173)	Control* (n=379)	OR**	P-value**
Worsen in 1 subregion	44 (25%)	78 (21%)	1.891	
Worsen in 2 subregions	33 (19%)	36 (9%)	3.132	
Worsen in 3+ subregions	22 (13%)	15 (4%)	5.181	
BICL: 24M Chg MOAKS Cartilage Morphology - entire knee number of areas with worsening in surface area (include within-grade change) (0/1/2/3+)				<.0001
No Change	47 (27%)	182 (48%)	REF	
Worsen in 1 subregion	49 (28%)	116 (31%)	1.673	
Worsen in 2 subregions	34 (20%)	47 (12%)	2.806	
Worsen in 3+ subregions	43 (25%)	34 (9%)	5.268	
BICL: 24M Chg Change in MOAKS Whole Knee Effusion Category				<.0001
Improvement	16 (9%)	60 (16%)	0.717	
No Change	91 (53%)	249 (66%)	REF	
Worsen	66 (38%)	70 (18%)	2.752	
BICL: 24M Chg Worsening in MOAKS Inter-Condylar Synovitis				<.0001
No Change	143 (83%)	354 (93%)	REF	
Worsen	30 (17%)	25 (7%)	3.297	

*mean(SD); median presented for continuous variables, n(%) presented for categorical variables.
** adjusted for covariates: sex, race, baseline age, BMI, JSW, PAIN, Pain medications, KLG

1. Hunter DJ, Schofield D, Callander E. The individual and socioeconomic impact of osteoarthritis. NatRevRheumatol. 2014;10(7):437-41.
2. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2163-96.
3. Prevalence of disabilities and associated health conditions among adults--United States, 1999. [erratum appears in MMWR Morb Mortal Wkly Rep 2001 Mar 2;50(8):149.]. MMWR - Morbidity & Mortality Weekly Report. 2001;50(7):120-5.
4. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. American Journal of Public Health. 1994;84(3):351-8.
5. Matthews GL, Hunter DJ. Emerging drugs for osteoarthritis. ExpertOpinEmergDrugs. 2011.

6. Hunter DJ. Osteoarthritis. BestPractResClinRheumatol. 2011;25(6):801-14.
7. Thomas D BJ, Audette J, Carroll A, Dow-Hygelund C, Hay M. Clinical development success rates 2006-2015: Biotechnology Innovation Organization (BIO), Biomedtracker and Amplion. 2016.
8. Lotz M, Martel-Pelletier J, Christiansen C, Brandi ML, Bruyere O, Chapurlat R, et al. Value of biomarkers in osteoarthritis: current status and perspectives. Ann Rheum Dis. 2013;72(11):1756-63.
9. Hunter DJ, Losina E, Guermazi A, Burstein D, Lassere MN, Kraus V. A pathway and approach to biomarker validation and qualification for osteoarthritis clinical trials. CurrDrug Targets. 2010;11(5):536-45.
10. Hunter DJ, Nevitt M, Losina E, Kraus V. Biomarkers for osteoarthritis: current position and steps towards further validation. BestPractResClinRheumatol. 2014;28(1):61-71.
11. Eckstein F, Collins JE, Nevitt MC, Lynch JA, Kraus VB, Katz JN, et al. Brief Report: Cartilage Thickness Change as an Imaging Biomarker of Knee Osteoarthritis Progression: Data From the Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Rheumatol. 2015;67(12):3184-9.
12. Hunter D, Nevitt M, Lynch J, Kraus VB, Katz JN, Collins JE, et al. Longitudinal validation of periarticular bone area and 3D shape as biomarkers for knee OA progression? Data from the FNIH OA Biomarkers Consortium. AnnRheumDis. 2015.
13. Kraus VB, Collins JE, Charles HC, Pieper CF, Whitley L, Losina E, et al. Predictive Validity of Radiographic Trabecular Bone Texture in Knee Osteoarthritis: The Osteoarthritis Research Society International/Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Rheumatol. 2018;70(1):80-7.
14. Roemer FW, Guermazi A, Collins JE, Losina E, Nevitt MC, Lynch JA, et al. Semi-quantitative MRI biomarkers of knee osteoarthritis progression in the FNIH biomarkers consortium cohort - Methodologic aspects and definition of change. BMC Musculoskelet Disord. 2016;17(1):466.
15. Collins JE, Losina E, Nevitt MC, Roemer FW, Guermazi A, Lynch JA, et al. Semiquantitative Imaging Biomarkers of Knee Osteoarthritis Progression: Data From the Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Rheumatol. 2016;68(10):2422-31.
16. Kraus VB, Collins JE, Hargrove D, Losina E, Nevitt M, Katz JN, et al. Predictive validity of biochemical biomarkers in knee osteoarthritis: data from the FNIH OA Biomarkers Consortium. Ann Rheum Dis. 2017;76(1):186-95.
17. Hunter DJ, Nevitt M, Losina E, Kraus V. Biomarkers for osteoarthritis: current position and steps towards further validation. Best practice & research Clinical rheumatology. 2014;28(1):61-71.
18. Eckstein F, Collins J, Nevitt M, Lynch J, Kraus V, Katz J, et al. Cartilage thickness change as an imaging biomarker of knee osteoarthritis progression–data from the fnih OA biomarkers consortium. Arthritis & Rheumatology. 2015.
19. Neumann G, Hunter D, Nevitt M, Chibnik LB, Kwoh K, Chen H, et al. Location specific radiographic joint space width for osteoarthritis progression. Osteoarthritis Cartilage. 2009;17(6):761-5.
20. Ornetti P, Brandt K, Hellio-Le Graverand MP, Hochberg M, Hunter DJ, Kloppenburg M, et al. OARSI-OMERACT definition of relevant radiological progression in hip/knee osteoarthritis. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society. 2009;17(7):856-63.
21. Angst F, Aeschlimann A, Michel BA, Stucki G. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. The Journal of rheumatology. 2002;29(1):131-8.
22. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society. 2008;16(12):1433-41.
23. Peterfy CG, Guermazi A, Zaim S, Tirman PF, Miaux Y, White D, et al. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. Osteoarthritis & Cartilage. 2004;12(3):177-90.
24. Biswal S, Hastie T, Andriacchi TP, Bergman GA, Dillingham MF, Lang P. Risk factors for progressive cartilage loss in the knee: a longitudinal magnetic resonance imaging study in forty-three patients. Arthritis & Rheumatism. 2002;46(11):2884-92.
25. Hunter D, Gale D, Grainger G, Lo G, Conaghan P. The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment: BLOKS (Boston Leeds Osteoarthritis Knee Score). Annals of the Rheumatic Diseases. 2008;67(2):206-11.

26. Kornaat PR, Ceulemans RY, Kroon HM, Riyazi N, Kloppenburg M, Carter WO, et al. MRI assessment of knee osteoarthritis: Knee Osteoarthritis Scoring System (KOSS)--inter-observer and intra-observer reproducibility of a compartment-based scoring system. Skeletal Radiology. 2005;34(2):95-102.
27. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society. 2011;19(8):990-1002.
28. Roemer FW, Nevitt MC, Felson DT, Niu J, Lynch JA, Crema MD, et al. Predictive validity of within-grade scoring of longitudinal changes of MRI-based cartilage morphology and bone marrow lesion assessment in the tibio-femoral joint--the MOST study. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society. 2012;20(11):1391-8.
29. Burgkart R, Glaser C, Hyhlik-Durr A, Englmeier KH, Reiser M, Eckstein F. Magnetic resonance imaging-based assessment of cartilage loss in severe osteoarthritis: accuracy, precision, and diagnostic value. Arthritis & Rheumatism. 2001;44(9):2072-7.
30. Graichen H, Eisenhart-Rothe R, Vogl T, Englmeier KH, Eckstein F. Quantitative assessment of cartilage status in osteoarthritis by quantitative magnetic resonance imaging: technical validation for use in analysis of cartilage volume and further morphologic parameters. Arthritis & Rheumatism. 2004;50(3):811-6.
31. Wluka AE, Wolfe R, Stuckey S, Cicuttini FM. How does tibial cartilage volume relate to symptoms in subjects with knee osteoarthritis? Annals of the Rheumatic Diseases. 2004;63(3):264-8.
32. Cicuttini FM, Jones G, Forbes A, Wluka AE. Rate of cartilage loss at two years predicts subsequent total knee arthroplasty: a prospective study. Annals of the Rheumatic Diseases. 2004;63(9):1124-7.
33. Eckstein F, Burstein D, Link TM. Quantitative MRI of cartilage and bone: degenerative changes in osteoarthritis. [Review] [238 refs]. NMR in Biomedicine. 2006;19(7):822-54.
34. Cicuttini F, Forbes A, Morris K, Darling S, Bailey M, Stuckey S. Gender differences in knee cartilage volume as measured by magnetic resonance imaging. Osteoarthritis & Cartilage. 1999;7(3):265-71.
35. Peterfy C, van Dijke C, Janzen D, Gluer C, Namba R, Majumdar S, et al. Quantification of articular cartilage in the knee with pulsed saturation transfer subtraction and fat-suppressed MR imaging: optimization and validation. Radiology. 1994;192(2):485-91.
36. Kshirsagar AA, Watson PJ, Tyler JA, Hall LD. Measurement of localized cartilage volume and thickness of human knee joints by computer analysis of three-dimensional magnetic resonance images. Investigative Radiology. 1998;33(5):289-99.
37. Dam EB, Lillholm M, Marques J, Nielsen M. Automatic segmentation of high- and low-field knee MRIs using knee image quantification with data from the osteoarthritis initiative. J Med Imaging (Bellingham). 2015;2(2):024001.
38. Kraus VB, Feng S, Wang S, White S, Ainslie M, Brett A, et al. Trabecular morphometry by fractal signature analysis is a novel marker of osteoarthritis progression. Arthritis and rheumatism. 2009;60(12):3711-22.
39. Kraus VB, Feng S, Wang S, White S, Ainslie M, Graverand MP, et al. Subchondral bone trabecular integrity predicts and changes concurrently with radiographic and magnetic resonance imaging-determined knee osteoarthritis progression. Arthritis and rheumatism. 2013;65(7):1812-21.
40. Janvier T, Jennane R, Valery A, Harrar K, Delplanque M, Lelong C, et al. Subchondral tibial bone texture analysis predicts knee osteoarthritis progression: data from the Osteoarthritis Initiative: Tibial bone texture & knee OA progression. Osteoarthritis and cartilage. 2016.
41. Kraus VB, Feng S, Wang S, White S, Ainslie M, Brett A, et al. Trabecular morphometry by fractal signature analysis is a novel marker of osteoarthritis progression. Arthritis Rheum. 2009;60(12):3711-22.
42. Berndt E, Gottschalk A, Strobeck M. Opportunities for improving the drug development process: results from a survey of industry and FDA. MIT-FDA-Industry White Paper. http://webmitedu/cbi/docs/berndt-et-al6-3-05pdf [Internet]. 2006 3/17/2009. Available from: http://web.mit.edu/cbi/docs/berndt-et-al6-3-05.pdf.