



UNIVERSITY OF LEEDS

This is a repository copy of *FDA/Arthritis Foundation osteoarthritis drug development workshop recap: Assessment of long-term benefit*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/189843/>

Version: Accepted Version

---

**Article:**

Kim, JS, Borges, S, Clauw, DJ et al. (21 more authors) (2022) FDA/Arthritis Foundation osteoarthritis drug development workshop recap: Assessment of long-term benefit. *Seminars in Arthritis and Rheumatism*, 56. 152070. ISSN 0049-0172

<https://doi.org/10.1016/j.semarthrit.2022.152070>

---

© 2022, Elsevier. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# FDA-AF OA Drug Development Workshop Recap: Assessment of Long-term Benefit

## A. AUTHORS AND AFFILIATIONS

Jason S. Kim\*  
Silvana Borges  
Daniel Clauw  
Philip Conaghan  
David Felson  
Thomas Fleming  
Rachel Glaser  
Elizabeth Hart  
Marc Hochberg  
Yura Kim  
Virginia Kraus  
Larissa Lapteva  
Xiaojuan Li  
Sharmila Majumdar  
Tim McAlindon  
Ali Mobasher  
Tuhina Neogi  
Frank Roemer  
Rebecca Rothwell  
Rob Shibuya  
Jeff Siegel  
Lee Simon  
Kurt Spindler  
Nikolay Nikolov\*

\*indicates corresponding author

## B. KEYWORDS

Osteoarthritis, Structure-modifying therapy, Regulatory approval, Drug Development, Disease-modifying therapy, Clinical benefit, Long-term benefit, biomarkers,

## C. SUMMARY: OBJECTIVE, DESIGN, RESULTS, CONCLUSIONS

*Objective: To summarize proceedings of a workshop convened to discuss the current state of science in the disease of osteoarthritis (OA), identify the knowledge gaps, and examine the developmental and regulatory challenges in bringing these products to market.*

*Design: Summary of the one-day workshop held virtually on June 22<sup>nd</sup>, 2021.*

*Results: Speakers selected by the Planning Committee presented data on the current approach to assessment of OA therapies, biomarkers in OA drug development, and the assessment of disease progression and long-term benefit.*

*Conclusions: Demonstrated by numerous failed clinical trials, OA is a challenging disease for which to develop therapeutics. The challenge is magnified by the slow progression of disease and the need for clinical trials of long duration and/or large sample size to demonstrate the effect of an intervention. The OA science community, including academia, pharmaceutical companies, regulatory agencies, and patient communities, must continue to develop and test better clinical endpoints that meaningfully reflect disease modification related to long-term patient benefit.*

*Biomarkers may be considered as surrogate endpoints and used for regulatory approval in OA trials if there is clinical data providing strong evidence that an effect on the biomarker predicts a specific clinical benefit and is supported by a clear mechanistic rationale. There may be other opportunities for biomarker use in OA treatment development including their use to better understand treatment effects, as a prognostic factor to enrich enrollment in trials, inform dose selection and other uses.*

## BODY

### OPENING

The serious disease of osteoarthritis (OA) has no available drugs or biologics proven to arrest or reverse progression.<sup>1</sup> OA is a heterogeneous disease and a challenge with respect to developing therapeutics, demonstrated by numerous failed clinical trials. Under the 21<sup>st</sup> Century Cures Act, the United States Food and Drug Administration (FDA) is required to report on patient experiences, perspectives, needs, and priorities in regulatory decision-making.<sup>2</sup> On March 8, 2017, the Arthritis Foundation, FDA, osteoarthritis (OA) patients, drug developers, health care providers and academic researchers came together to discuss the serious burden of OA disease, particularly the most significant OA-related symptoms, the impact of those symptoms, currently available treatments and ideal treatments, at the 2017 OA Patient-focused Drug Development (PFDD) meeting (REF). To further this collaboration and to include the patient voice in the ongoing scientific discussions on the topic, the FDA partnered with the Arthritis Foundation to co-sponsor a drug development workshop in June of 2021.<sup>3</sup> In preparation for the workshop, a series of planning committee teleconferences and patient perspective sessions were facilitated by the FDA and Arthritis Foundation, with the participation of important organizations including National Institutes of Health (NIH), Osteoarthritis Research Society International (OARSI), and Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT). The virtual workshop brought various stakeholders together to discuss the current state of drug development in OA, insights from promising research, considerations on the assessment of clinical benefit, and the regulatory challenges. This paper summarizes the workshop presentations and discussions.

Overviewed are (a) the need for treatments and the current regulatory approach to evaluating drugs and biologics, (b) the role and current state of biomarkers in drug development, and (c) the challenges and recent developments in assessing disease progression and long-term clinical benefit of interventions. We close with suggestions for important next steps for the scientific community to address uncertainties about assessment of OA disease progression and potential discussion topics for future meetings.

## CURRENT APPROACH TO THE DISEASE OF OA

### *Overview of OA and Burden of Disease, Tuhina Neogi, MD/PhD*

OA is a large public health problem affecting over 500 million adults or 15% of adults globally. In the US, OA affects over 32 million adults.<sup>4</sup> The cost burden to the US healthcare system is high, representing hundreds of billions annually with over 900,000 hospitalizations.<sup>5-7</sup> Joint pain is the primary reason for doctor's office visits in the US, and musculoskeletal pain is the second leading reason for years lived with disability worldwide. People are living longer with the burden of OA with reduced quality-of-life and a higher risk of mortality.<sup>8</sup>

Initial symptoms of the disease occur intermittently, often with use of the joint such as weight-bearing activities for the knee, and over time may become more severe and persistent.<sup>9</sup> There is good evidence that weight loss and exercise manage OA symptoms.<sup>10</sup> Joint replacement is considered "definitive treatment" for knee and hip osteoarthritis, but such surgery is not available for all joints affected by OA, and surgical options may be limited for patients with medical comorbidities. The treatment options for patients, some of which are only minimally or moderately efficacious for pain, include topical, intraarticular, and oral pharmacologic agents.<sup>10,11</sup> Non-steroidal anti-inflammatory drugs (NSAIDs) have precautions and contraindications that limit the use of these medication in OA patients and may have partially contributed to the increase of opioid use.<sup>12</sup> There is a large unmet need for more treatments to alleviate symptoms, as well as treatments that modify/prevent joint structural damage in early-to-moderate OA. The development of therapies to address structural damage is complicated by the multiple tissues involved within the structure of the joint. The question of whether accretion of new cartilage or alteration of subchondral bone alone will have benefit in terms of both signs and symptoms as well as structural improvement remains unanswered.

## A PATIENT'S PERSPECTIVE

Patient-focused drug development is intended to better incorporate the patient's voice in drug development and evaluation. Listening to the story of the patient is an important aspect of enhancing understanding of patient preferences, potential acceptability of benefits, risks, and burden. The FDA and Arthritis Foundation convened multiple patient perspective sessions for patient participants to prepare for this workshop. One such participant provided remarks, excerpted here:

### *Testimony Excerpt from Bill Agee*

*"I have been living with OA since 1995. I have had 23 knee surgeries, including a left knee replacement. I have had one peroneal nerve surgery on my left leg from scar tissue due to multiple knee surgeries. In 2020 I had a lumbar spinal fusion from L3-L5 due to severe osteoarthritis. I have also had two surgeries on my right ankle from a spiral fracture in the fibula and now I have post traumatic osteoarthritis following the removal of the hardware from the surgical repair to my ankle. Due to the swelling from the arthritis, it took 7 months for the incision to heal fully. I was extremely lucky to not develop an infection which would have been tragic for my ankle, my knee replacement, and potentially my life. At this very moment I am facing a choice between ankle fusion or ankle replacement and I have no idea how to make this decision."*

*"I have arthritis pain daily; some are good pain days, and some are bad pain days. On the good days, the pain is there but it does not prevent me from being able to work full time in my office job. On the bad days, the pain is debilitating. However, I do believe all days are good days, some are just better than others. Because the severity of the pain varies throughout the day, I*

*must be very careful with my day-to-day decisions on what I choose to do. I must plan my activities and errands because I have learned the hard way that I cannot do everything in one day. This includes errands that require me to leave the house.*

*My osteoarthritis journey has been, and continues to be, difficult. I tried all of the NSAIDS, so my prescription pain relief options are limited. I still wake up on the average of 14 times a night to turn and get comfortable. I cannot remember sleeping an entire night without waking up. Nothing works consistently so it is trial and error.*

*But I've also come to know my fellow patient representatives [...] They are at a different stage of their battle with osteoarthritis. They are at an earlier stage, but we experience similar symptoms and effects to our lives. My experience gives me an ability to see a possible storyline for their future. I hope in their future, they will have a treatment that will improve the health of their joints and a menu of treatment options for pain AND structure. So, I hope this community will come together to bring meaningful drugs to market that can positively affect their storylines."*

#### **REGULATORY PERSPECTIVE - DIVISION OF ANESTHESIOLOGY, ADDICTION MEDICINE, AND PAIN MEDICINE (DAAP), Robert Shibuya, MD**

Drug products for the treatment for OA have historically fallen under the regulatory jurisdiction of DAAP, the FDA Division responsible for regulating drugs for pain. In studies for pain of OA, eligibility criteria typically select for (a) adults diagnosed with OA of the hip or knee based on ACR criteria and confirmed with knee radiograph (Kellgren Lawrence grade  $\geq 2$ ) and (b) the Western Ontario and McMaster Universities Arthritis Index (WOMAC) pain or Numerical Pain Rating Scale (NPRS)  $\geq 5$ , with exclusions based on other causes of index joint pain (such as inflammatory joint disease, crystalline disease, etc.), patients with a planned surgical procedure, and patients who have received intra-articular therapies within a certain period of time.

To support an indication of "pain of OA," an acceptable primary measurement instrument is a 0-10-point NPRS as well as the WOMAC pain subscale. The secondary outcome measures may include (a) WOMAC function, (b) Patient Global Assessment, (c) Joint-specific assessments of pain and function, (d) Brief Pain Inventory, (e) OARSI-OMERACT Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP), and (g) Quality of life instruments. The primary efficacy endpoint is the change in pain intensity (PI) from baseline (average PI in the week immediately prior to randomization) to end-of-treatment (average PI in last week of treatment). Registrational trials must include at least 12-weeks of double-blind treatment. DAAP is considering adding other aspects of the experience of OA to the primary efficacy assessment for drugs for symptomatic treatment of OA.

However, the development of treatments intended to alter the natural history of the disease, includes other considerations, in addition to the effects on pain, that require additional expertise. These aspects were further discussed at this workshop.

#### **REGULATORY PERSPECTIVE - CENTER FOR BIOLOGICS EVALUATION AND RESEARCH (CBER), Elizabeth Hart, MD**

Office of Tissues and Advanced Therapies (OTAT) in CBER regulates different types of biological products developed for the treatment of OA. Specifically, CBER-regulated products include autologous or allogeneic tissue-based and cellular therapies obtained from different sources, including stem cells and somatic tissue-specific differentiated cell lines, and gene

therapy products, including genetically modified cells, viral vector-based therapies, and others. As for any new drug or biological product aiming to treat OA, regulatory approval for these products is based on demonstration of substantial evidence of effectiveness and safety obtained from adequate and well-controlled trials. For cellular, gene and tissue-based therapies, important considerations also include durability of product effect, short-term and long-term risks, and patient preferences.

Consistent with FDA's broader approach, CBER considers improvement in pain and function measured by validated scales as clinically meaningful outcomes and acceptable primary endpoints in OA trials. Structural biomarkers from imaging studies have not yet been validated as surrogate endpoints and therefore are not currently used as primary endpoints in clinical development programs for novel OA treatments. Validation of such endpoints should be performed to enable their use and to demonstrate their ability to reasonably predict clinically meaningful outcomes. Evidence to support biomarker and surrogate endpoint validation may be developed in individual product programs or via the FDA's Drug Development Tool Qualification Program.

An adequate safety database is essential for understanding of risks associated with cellular, gene, or tissue-based therapies. Duration of the safety monitoring depends on product characteristics, and the duration of that monitoring in clinical trials should be sufficient to identify important safety signals to inform benefit-risk decisions. Common short-term risks include inflammatory and immunogenic reactions and procedure-related complications. Potential long-term risks may include ectopic tissue formation, acceleration of cartilage degradation, and tumorigenicity.

## **BIOMARKERS IN OA DRUG DEVELOPMENT**

### *Regulatory Considerations on Biomarkers and Assessment of Long-term Benefit in OA*

*Nikolay Nikolov, MD*

The benefit-risk assessment is the basis for the FDA's regulatory decision-making process for approval of drugs. In OA, benefit refers to clinical benefit and an improvement in how a patient (a) feels or experiences pain or other symptoms, (b) functions or physically performs with their affected joint(s), and (c) survives or maintains a healthy joint(s). Endpoints in trials of OA treatments need to demonstrate the clinical benefit directly or at least be interpretable with respect to the clinical benefit to be expected. The safety assessment should be guided by drug class, prior experience, events of interest, etc. Currently, drugs approved for OA have been approved based on patient-reported outcomes (PROs) assessing two key OA domains: pain and function, as discussed by the two previous speakers.

However, clinical benefits related to inhibition of structural damage remain elusive to capture in OA and represent an unmet need and a scientific knowledge gap. The reliance on structural outcomes for the assessment of clinical efficacy is challenging, with questions remaining on what clinical outcomes can be reliably predicted by changes in structural measures. The pathogenesis of OA is complex and a treatment may affect one of multiple pathways, so the required magnitude and duration of effect on structural outcome must be considered. Additionally, the benefit-risk of on-target effects compared to off-target effects must also be carefully considered.<sup>13</sup> There are complex relationships between the disease, potential biomarkers, and clinical outcomes. Importantly, a correlation between a biomarker and a clinical endpoint is not sufficient to demonstrate that an effect on the proposed surrogate endpoint will reliably predict an effect on the clinical outcomes of interest; this point was further discussed at the workshop by Dr. Fleming, as detailed below. Ideally, this demonstration would

be based on empirical evidence from randomized, controlled comparisons from clinical trials and/or on a comprehensive understanding of the disease process and drug mechanism of action.

Ultimately, the goal of OA treatments is to provide clinical benefit to the patient and thus the goal of clinical trials is to demonstrate this benefit. However, in OA, there is a complex relationship between pathophysiology, structural damage, and clinical outcomes. Knowledge gaps in the relationship between the structural and pathophysiological elements of OA is a major challenge to use of biomarkers as a reliable assessment of a product's ability to alter disease progression. To use structural outcomes, via imaging or other biomarkers, in the benefit-risk assessment, we need to be able to describe the clinical benefit expected from the structural change. Structural outcomes could be used in addition to clinical outcomes in OA trials. Approaches to use of structural or other biomarkers in OA trials will depend on level of information available to characterize a clinical benefit. With less information, structural outcomes may still be useful as adjunctive or secondary endpoints. To be used as the primary endpoint to support approval, a high level of characterization would be needed about the relationship of the endpoint to the anticipated clinical benefit, which is yet to be defined.

For products for which early symptomatic benefit is not expected or would require quantitative changes in the structural outcome to occur first, a different study design may be needed to capture the direct clinical benefit. The OA research community is challenged to design feasible studies to assess direct clinical benefit of such therapies. However, there are potential strategies to help address some of these challenges. For example, composite endpoints that utilize joint replacement or "end-stage" joint disease (i.e., the severe, irreversible, intolerable pain or functional impairment) is one strategy that could improve the feasibility of such studies. {Kim, 2020 #10} Enrichment strategies, such as models of accelerated OA and trials in subjects prior to knee replacement, are other potential opportunities to demonstrate clinical benefit in a reasonable timeline of a controlled clinical trial. Innovative clinical trials, i.e., platform, pragmatic trials, should also be considered.

### *Role of Biomarkers in Drug Development, Jeffrey Siegel, MD*

The FDA recognizes that biomarkers are an important drug development tool to measure biological processes, pathogenic processes, or responses to therapeutic interventions. The NIH-FDA Biomarker Working Group developed the Biomarkers, Endpoints, and other Tools (BEST) resource, and categorized biomarker types into those that (a) measure disease presence and status (susceptibility / risk biomarker, diagnostic biomarker, prognostic biomarker) and those that (b) measure aspects of response to treatment (monitoring biomarker, predictive biomarker, pharmacodynamic / response biomarker – including surrogate endpoints, safety biomarker).<sup>14</sup> Context of Use (COU) is a critical consideration in establishing biomarker utility to impact a clinical trial or drug development program. Biomarkers may be used to address (a) inclusion/exclusion criteria for prognostic or predictive enrichment, (b) alter treatment allocation based on biomarker status, (c) result in cessation of a patient's participation in a clinical trial because of safety concern, (d) result in adaptation of the clinical trial design, (e) establish proof of concept for the patient population of interest, (f) support clinical dose selection for first in human or Phase III studies, (g) evaluate treatment response (e.g. pharmacodynamic effect), and (h) support regulatory acceptability of a surrogate endpoint for accelerated or traditional approval.

There are several different ways to establish acceptance of biomarkers by the FDA, such as from data generated by corporate sponsors during the drug approval process, scientific

community consensus, and FDA's biomarker qualification program. These pathways do not exist in isolation and many times parallel efforts are underway within or between pathways. All pathways share common core concepts, are data-driven, and involve regulatory assessment and outcomes based on the available data. The 21<sup>st</sup> Century Cures Drug Development Tool (DDT) Legislation established the Biomarker Qualification Program (BQP).<sup>15</sup> Qualified biomarkers can advance public health by encouraging efficiencies and innovation in drug development.

Of high interest to drug developers are surrogate endpoints, which measure pharmacodynamic response. These biomarkers can be used to support approval of a drug, both traditional and accelerated. The FDA requires substantial evidence of effectiveness--showing that a drug improves meaningful clinical outcomes (how a patient feels, functions, or survives). For validated surrogate endpoints (often used in traditional approval of a drug product), the FDA accepts that the effect on the biomarker predicts a specific clinical outcome. Validated endpoints have strong and diverse evidence supporting the relationship of the biomarker and the outcome. For a "reasonably likely" surrogate endpoint (which may provide the basis of accelerated approval of a drug product), the FDA expects the biomarker be supported by strong mechanistic and/or epidemiologic rationale such that an effect on the surrogate endpoint is expected to be correlated with a clinical benefit, but not yet reaching the standard for validation.

Limitations exist, as the surrogate endpoint is not a direct measure of how a patient feels, functions, or survives. Instead, the surrogate endpoint is intended to reflect and predict clinical benefit and not measure the outcome. The benefit-risk assessment of a drug is therefore based upon assumptions / predictions of benefit. Further, the extent of clinical benefit is translated from an indirect measure, and with use of a limited dataset on risk to assess harms. Surrogates have also been shown to fail to predict clinical benefit itself. These are rooted in the relation of the surrogate endpoint to the causal pathway modulated by the drug to achieve the clinical outcomes. Challenges may occur if the biomarker is not on the causal pathway, if multiple pathways lead to the clinical outcome, or if the drug induces an adverse effect on the desired clinical outcome via a pathway not reflected by the surrogate biomarker.

When efficacy is established via the effect on a surrogate endpoint with unquantifiable clinical benefit, the benefit-risk assessment must balance an unmeasured clinical benefit against unmeasured risks. In the FDA's accelerated approval pathway, used for serious and/or life-threatening conditions, the endpoint is often a "reasonably likely" surrogate endpoint. This approval pathway requires post-marketing studies to confirm clinical benefit. While the accelerated approval pathway may allow faster access to promising treatments, patients may be exposed to risks of a drug that does not show benefit and may have less safety data. For potential surrogate endpoints in OA trials, the expected clinical benefit, the strength of the scientific evidence tying the biomarker to clinical outcomes, and the magnitude of biomarker change that would indicate a clinically meaningful benefit to patients should all be considered.

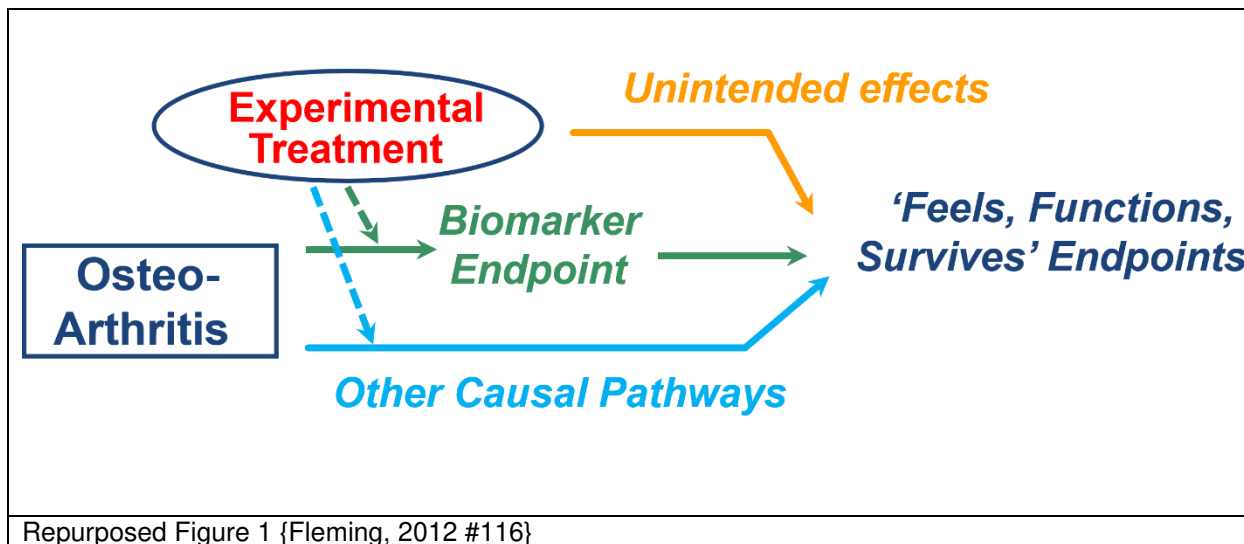
#### *Statistical Considerations on the Use of Surrogate Endpoints, Thomas Fleming, PhD*

"A correlate does not a surrogate make."<sup>16</sup> Utilization of a biomarker as a surrogate endpoint for a clinical trial intended to provide reliable evidence about efficacy and safety of an OA drug is enticing due to the long time course of the disease. Clinical researchers frequently assume that a strong correlation between their favored biomarker and clinical efficacy measure justifies the acceptance of the biomarker as a surrogate.<sup>17</sup> For example, longer survival in cancer patients who experience substantial tumor shrinkage following therapy may lead researchers to believe



that an increased response rate should predict improvement in overall survival. However, such correlations do not allow a true understanding of the causal pathway.<sup>13</sup>

Biomarkers can appropriately be used to measure biological activity of experimental treatments in proof-of-mechanism or proof-of-concept trials. In registrational evaluations, their use as a surrogate or replacement endpoint for efficacy may be justifiable in trials designed to refine dosing/schedules to address safety risks, or to generalize results to broader categories of patients. However, justifying their use as replacement endpoints is challenging in trials designed to evaluate new interventions in a class of established effective treatments, and even more challenging when evaluating treatments that are in new classes.



To establish validity of a biomarker as a surrogate endpoint, the net effect of the treatment on the biomarker should reliably predict its net effect on how a patient 'feels, functions, and survives'. FIGURE 1 provides insights into why individual-level correlation between a biomarker and a direct measure of clinical benefit does not provide adequate justification that a treatment effect on the biomarker would reliably predict its clinical benefit.{Fleming, 2012 #116} The treatment's effect on the biomarker endpoint could underestimate its true clinical efficacy if its intended effects are not only on the surrogate endpoint but also on other causal pathways, and could overestimate that true efficacy if the treatment has unintended effects, not captured by the surrogate, that are harmful to how a patient 'feels, functions or survives'. Furthermore, the breadth, magnitude, and duration of the effect of the treatment on the biomarker needed to achieve meaningful clinical benefit often are not known.

Validation of a biomarker as a surrogate endpoint, therefore, requires an in-depth clinical understanding of (a) the causal pathways of the disease process, and (b) the treatment's intended and unintended mechanisms of action. Meta-analyses of clinical trials also are needed that directly show the relationship between (a) the net effect of the treatment on the biomarker, and (b) its net effect on direct measures of clinical benefit.

The OA field should continue conducting properly controlled studies that are randomized, have standard-of-care controls, and evaluate effects on both biomarkers and clinical endpoints. Doing so not only enables reliable evaluations of promising interventions, but also allows meta-analyses to be conducted to assess trial-level correlations between effects on potential

surrogate endpoints and on direct measures for how a patient ‘feels, functions and survives’, in turn increasing the potential for reliable validation of biomarkers as surrogate endpoints.

### *Biochemical Markers in OA, Virginia Byers Kraus, MD/PhD*

In the weeks prior to this workshop, the FDA approved aducanumab for Alzheimer’s Disease (AD) based on the product’s ability to clear plaques of beta-amyloid as documented by imaging.<sup>18,19</sup> Like AD, OA has no available treatments that modify the disease course and is a disease of large unmet need. In fact, OA has a significantly higher prevalence in the US with 5.8 million affected by AD compared to 32.5 million affected by OA. Both AD and OA are diseases that are problematic with respect to conducting clinical trials due to their long disease time course. In addition, both AD and OA have biomarkers that are associated with progression of disease that may predict important benefits to patients. For the approval of aducanumab, the beta-amyloid biomarker was used as a surrogate endpoint based on being “reasonably likely to predict important benefits to patients” in the context of the FDA’s Accelerated Approval pathway for serious diseases; this pathway necessitates a Phase 4 post-marketing trial to show drug impact on patient relevant outcomes. While the approval of aducanumab remains controversial at the time of writing, historically the first drug in a new category invigorates a field.

In effort to clinically evaluate and qualify biochemical markers for OA, the Progress OA program was formed through the FNIH Biomarkers Consortium.<sup>20,21</sup> The group includes industry stakeholders as well as the Arthritis Foundation as sponsors, academic investigators and OARSI. Based on results from phase I analyses of eighteen biomarkers in the Osteoarthritis Initiative (OAI) cohort,<sup>22</sup> nine biochemical markers were chosen for full qualification in phase 2 using samples from placebo arms of extant clinical trials (n=871). The chosen markers include the following: (Marker-1) urinary (u)CTXII, (2) uC2C-HUSA, (3) serum (s)PIIBNP, (4) uNTXI, (5) sNTXI, (6) sCTXI, (7) uCTXI-alpha, (8) uCTXI-beta, and (9) sHA. Baseline and 12-month samples will be evaluated for their ability to predict radiographic and clinically relevant (combination of radiographic and symptomatic) knee OA progression at 36 months. It is anticipated that a combinatorial biomarker will yield the most robust predictor of knee OA progression, composed of one biomarker from each of four groups: (A) Type II Collagen Degradation [1-2], (B) Type II Collagen Synthesis [3], (C) Type I Collagen Degradation [4-8], and (D) Inflammation [9]. To date, the most promising of these biomarkers is uCTXII, which has been shown to predict clinically relevant knee OA progression based on the combination of patient reported outcomes (pain worsening) and radiographic progression.<sup>22</sup> Additionally, uCTXII has also been shown to predict total joint replacement of the knee and hip in the Calcitonin Trials<sup>23,24</sup> and OFELY Cohort<sup>25</sup>. In an effort to establish uCTXII and these other high level OA biomarkers as OA drug development tools, a qualification package to the FDA Biomarker Qualification Program is in development by the FNIH Biomarkers Consortium Progress OA investigators.

An emerging biomarker, cartilage acidic protein 1 precursor (CRTAC1), should be highlighted as a potentially superior prognostic compared to existing biomarkers. CRTAC1 was discovered recently in the plasma of OA patients and was found to be associated with clinically relevant knee OA progression (pain and radiographic worsening);<sup>26</sup> moreover, circulating concentrations have been observed to decline following total joint replacement strongly suggesting a joint tissue origin of this biomarker.<sup>27</sup> In the Chingford study,<sup>28</sup> CRTAC1 was found to be elevated in patients as much as 8 years prior to incident radiographic knee OA.<sup>29</sup>

Biochemical markers have been linked to clinically relevant outcomes in OA.<sup>27</sup> Notably, incident and progressive knee OA share biomarkers and therefore molecular pathophysiology—

suggesting the false dichotomy of using the radiograph to define “OA” versus “no OA”. Demonstration of biomarker relationships to a causal pathway in OA awaits existence of a disease modifying drug; thus, more work is needed before biochemical biomarkers are used by drug developers and accepted by regulators as surrogate endpoints.

#### *Imaging Biomarkers in OA: MRI vs. X-Ray – Frank Roemer, MD*

In current OA clinical trials, radiographs are used to help determine eligibility, to measure outcome as a structural endpoint, and/or safety monitoring.<sup>31</sup> Radiographs are used to define structural severity of the joint at baseline and to exclude patients with ongoing disease that may not be amenable by the intervention under evaluation. Currently, patients are included in clinical trials for OA products based on radiographic severity using Kellgren-Lawrence (KL) grading. The most commonly used grades, KL2 and KL3, are considered to represent early to moderate OA, but with current findings demonstrating the grades to be highly heterogenous.<sup>32</sup> Furthermore, radiographs do not have the sensitivity of MRI to screen for structural findings that are unlikely to be modified by a drug or considered exclusionary for inclusion, such as bone marrow infiltration, meniscal root tears, or occult fractures.<sup>33</sup> Radiography, as a projectional imaging technique, is also highly dependent on standardized image acquisition to be able to monitor progression over time. Additionally, joint space width by radiograph provides an indirect measure of cartilage, meniscus damage, and meniscal extrusion.

MRI has promise as an enrichment tool due to superior structural disease characterization and potential to identifying the subset of the patient population for whom an intervention would have a clinically meaningful benefit-risk profile. MRI is able to stratify patient populations into different structural endo- or phenotypes such as inflammatory, subchondral bone, atrophic or cartilage /meniscus considering the mode of action of a given pharmacologic compound.<sup>33</sup> Like biochemical markers, the mode of action of the product must be considered when using MRI markers. For instance, with an anabolic compound, knees without cartilage loss should be excluded and knees with cartilage damage need to be included to show re-growth of cartilage. With an anticatabolic compound, widespread full-thickness cartilage loss should be excluded to be able to monitor slowing of cartilage loss or preservation of cartilage tissue. Technical advances in the field of MRI research allow image acquisition in a very short time of 3-5 minutes which supports the use of MRI to be used as a screening instrument due to feasibility that was not available until recently.

OA is a disease of multiple tissues including bone as well as soft tissues such as cartilage, meniscus, ligaments, bone marrow, the joint capsule, and muscle. Different MRI modalities to OA joint assessment are available, including quantitative analysis for cartilage, meniscus, or muscle (discussed in further detail below), semiquantitative analysis for all joint tissues and determination of eligibility, and Dynamic Contrast Enhanced (DCE) MRI for synovitis/inflammation. Also available are exploratory approaches such as compositional analysis for cartilage, meniscus, muscle, bone shape, and metabolic imaging by Positron Emission Tomography-Computed Tomography (PET-CT) or PET-MRI. Additionally, the role of the target tissue and the mode of action of the intervention on the target tissue should be considered.

#### *Imaging Biomarkers in OA: Quantitative MRI – Sharmila Majumdar, PhD*

Quantitative MRI (qMRI) can show biochemical and microstructural tissue changes at early stages of disease before changes become visible with conventional MRI or arthroscopy.<sup>34,35</sup> qMRI can also assess changes in cartilage morphology, muscle composition, and bone shape. One qMRI modality of interest is T2 mapping, which reflects water and collagen content as well

as collagen orientation.<sup>36,37</sup> Another complementary modality of interest is Spin-lattice relaxation time constant in rotating frame (T1rho) imaging, which reflects the proteoglycan content across the articular cartilage.<sup>38,39</sup> The Quantitative Imaging Biomarkers Alliance (QIBA) sponsored by the Radiological Society of North America (RSNA) has published a profile to provide recommendations to improve reproducibility and to standardize the two qMRI modalities.<sup>40</sup>

When using anterior cruciate ligament (ACL) injury as a model for early OA, qMRI is able to detect cartilage matrix changes.<sup>38</sup> T2 and T1rho predict cartilage loss over 2-4 years in subjects who start with no or mild OA.<sup>39,40</sup> The biochemical changes to the cartilage matrix can be detected in the various cartilage layers. T2 and T1rho values in various compartments and layers of cartilage have low correlations with patient reported outcomes; but, localizing beyond the averaged values of compartments and layers have more meaning, with greater degeneration found to correlate to higher pain symptoms.

An area of learning has been uncovered with artificial intelligence and machine learning, particularly when applied to rich datasets such as in the OAI.<sup>44-48</sup> These methods have assisted in phenotyping large data sets<sup>49</sup> and predicting total knee replacements<sup>50</sup> While the methods could not be correlated to pain using a single qMRI biomarker or region of interest, opportunity exists for combining qMR biomarkers for multiple tissues and morphology within the joint.<sup>51</sup>

## **ASSESSMENT OF OA DISEASE PROGRESSION AND LONG-TERM BENEFIT**

### *Challenges with Assessment of Disease Progression – Clinical and Structural, Timothy McAlindon, MD/MPH*

OA continues to suffer from historical biases where the disease was characterized to be “wear and tear”. OA was also thought to be a disease of cartilage degeneration, popularized by the developments of the KL grading system and JSW measurement by radiograph. This outdated thinking has led to the consideration of reduction of rate of structural progression as the treatment goal in disease modification of OA. Yet, current structural assessment tools show accumulated cartilage damage and do not measure clinical severity. While structural progression does not accurately capture disease severity, patient reported outcome measures are also inadequate. Measurement of pain is subjective and contextualized due to the role of the nervous system in signaling joint damage and tissue responses. Nociception, sensitization, placebo effect, exaggerated pain effect, flares and other factors result in numerous measurement issues in randomized clinical trials.

Clinical severity of disease should be measured by connecting the many pathophysiologic pathways and clinical manifestations to treatment target. An “End of the Road” endpoint, such as total joint replacement (TJR), a proxy for joint survival, is being thoughtfully considered, as it encompasses pain, function, and structure. Yet, such an assessment is problematic in OA drug development due the long time course of the disease. While subjects should be studied for 10 years or more, typical clinical trials can only study subjects for ~2 years. A highly sensitive and selective predictive biomarker to detect OA prior to clinical disease would be helpful by shifting the focus away from end-stage disease to early-disease in disease modifying drug development.

The OA scientific community has come to define the disease as a complex, multifactorial whole joint disorder involving numerous tissues. Yet, outdated heuristics of OA, specifically cartilage degeneration, continue to trouble the community—leading to an overemphasis on structural progression to achieve disease modification. Pain also does not adequately represent joint

health. Instead, a core measurement of disease severity, unifying the concepts of “feel”, “function”, and joint “survival” is necessary. All the challenges overviewed highlight the inadequate understanding of the relationship between disease process, joint structure, and patient reported outcomes.

#### *Lessons Learned – Sprifermin Case Study, Marc Hochberg, MD, MPH*

Sprifermin is a recombinant human fibroblast growth factor-18 (rhFGF-18) being developed as a potential disease-modifying OA drug (DMOAD) due to the ability to induce hyaline cartilage formation by increasing chondrocyte proliferation<sup>46</sup>. A 1-year Phase Ib study and 5-year Phase II study (FORWARD) reported statistically significant, dose-dependent effects on total cartilage thickness in patients treated with intra-articular sprifermin after 2 years.<sup>47,48</sup> Eligible participants in FORWARD were aged 40-85 with symptomatic knee OA (KL grade 2 or 3, medial minimal joint space width (mJSW)  $\geq$  2.5 mm and Score of 4-9 on WOMAC A1). The primary endpoint was change in total femorotibial joint cartilage thickness in the index knee from baseline to 2 years measured by quantitative MRI and the secondary endpoint was change in WOMAC total and subscale scores.

The FORWARD study was designed to have a treatment period of two years with four treatment groups (with two doses and two injection schedules per dose) and a placebo group. After two years, there was a three-year extended follow-up period. At baseline, cartilage thickness was similar in all treatment arms (~1.8 mm average) at total femorotibial joint (TFTJ). The 0.05 mm mean increase in TFTJ cartilage thickness with the treatment (at the highest dose and most frequent injection schedule) compared to placebo after two years was statistically significant and was sustained from Year 2 to Year 5. The 50% improvement in WOMAC pain at Year 2 was maintained to Year 5 in all cohorts in the intent-to-treat (ITT) population, and there was no significant treatment effect comparing any dose of sprifermin with placebo. Results for secondary imaging endpoints were consistent for both medial and lateral femorotibial joint cartilage thickness as well as both the central medial and central lateral femorotibial subregions.<sup>52,53</sup> There were no clear differences in the nature, severity or type of reported adverse or severe adverse events between treatment and placebo groups. While there was a statistically significant increase in cartilage thickness, uncertainty exists surrounding the clinical significance of the modest increase in cartilage thickness without a significant reduction in signs and symptoms as measured by the WOMAC.

The FORWARD study is the longest phase II trial reported and had an intent-to-treat population of 549 subjects.<sup>51,54</sup> Post-hoc analyses has suggested translation of the structural benefit to clinical benefit;<sup>55</sup> based on these analyses, a target dose and patient population have been identified for future Phase III studies of this promising treatment.

#### *Lessons Learned – Canakinumab Case Study, Philip Conaghan, MBBS, PhD*

Interleukin-1 (IL-1) has been a target in OA clinical trials due to previous in vitro studies demonstrating its key roles in OA inflammation and catabolic processes.<sup>56</sup> Previously explored anti-IL1 compounds in knee OA include anakinra,<sup>57</sup> AMG108,<sup>58</sup> lutikizumab,<sup>59-61</sup> and canakinumab.<sup>62</sup> The duration of these trials were often short, with the longest anti-IL-1 trial running for one year; however they provided little in terms of positive results.

In 2017, Ridker et al. published the results of the Canakinumab Anti-inflammatory Thrombosis Outcomes Study (CANTOS), where canakinumab was used to treat the inflammation underpinning atherothrombosis.<sup>63</sup> In this study of 1,091 clinical sites, 10,061 participants with previous myocardial infarction and high sensitivity C-reactive protein (hsCRP) levels  $\geq$  2 mg/L,

subjects were randomised to canakinumab dose of 50, 150, 300 mg or placebo. Subjects were followed for five years and monitored every 3 months for the primary outcomes of nonfatal myocardial infarction, nonfatal stroke, and cardiovascular death--also monitored were adverse events including the detection of osteoarthritis. A post-hoc analysis of the CANTOS trial examined time-to-incident total knee replacement (TKR) and total hip replacement (THR) and time-to-first OA-reported adverse events. In this analysis, the median follow-up 3.7 years and the median BMI was approximately 30 and 40% of participants were diabetic.<sup>64</sup> The analyses found “40% to 50% reductions in the hazard for incident arthroplasty at all three active canakinumab doses” as well as reductions in OA adverse events.

The CANTOS data resulted in reconsidering the role of IL-1 and anti-inflammatory pathway inhibition. There were important differences with previous anti-IL-1 trials, such as inclusion criteria, massive size of study, long duration of treatment, and utilization of total joint replacement (TJR, an adverse event) as an endpoint. As many of these factors are not feasible to replicate in future OA clinical trials, perhaps the community may consider endpoints such as “virtual TJR”<sup>62,63</sup>{Gossec, 2011 #93} or composite endpoints that may reduce duration of follow-up and required sample size.<sup>64</sup>

### *Considerations on Assessment of Long-term Pain and Function, Daniel Clauw, MD*

In considering the challenges of measuring long-term outcomes, one solution may be to use an objective assessment of functional performance in conjunction with a measurement of pain. To date, there is not a good relationship between self-reported and objective measures of physical function, such as actigraphy, in healthy individuals or in individuals with disease. When using actigraphy to measure real-life activity or function, there is a consistently nonexistent or weak relationship ( $r = 0 - 0.40$ ) between average activity levels and measures of functional status or activity.<sup>65-68</sup> Fibromyalgia patients are amongst the patient populations with the lowest self-reported functional status of any chronic illness, often scoring two standard deviations below the population’s mean activity levels.<sup>69</sup> Yet, peak activity was significantly lower in the fibromyalgia group relative to the control ( $p=0.008$ ), and the variability of peak activity was also significantly different between groups ( $p=0.001$ ).{Kop, 2005 #115} Also, a better correlation was found between self-reported functional status with peak activity in fibromyalgia patients. Previous work has considered the intriguing measurement of peak activity through Moderate to Vigorous Physical Activity (MVPA).<sup>70,71</sup>

Composite measures of pain and activity may be more responsive than pain-alone measures in patients with OA.<sup>72</sup> Pain should be considered a “state”, while function should be considered a “behavior”. Pain can change rapidly, for example, in analgesic procedures (drug, surgical procedures, or other), many patients experience pain relief but do not experience corresponding functional benefit. Function does not change as rapidly as pain, and relates to what the patient can do or think they can do. In a randomized clinical trial of 47 OA patients, the WOMAC pain subscale was the most responsive of all five pain measures. Pain-activity composites resulted in a statistically significant difference between an analgesic (celecoxib) and placebo, but were not more responsive than pain measures alone. However, a composite responder defined as having 20% improvement in pain or 10% improvement in activity yielded much larger differences. The most responsive actigraphy measure was peak activity, with a between-group difference of 91.9 counts/min ( $p = 0.090$ ) with mean activity and total activity not approaching statistical significance. Actigraphy was more responsive than the WOMAC functional scale, possibly due to lower placebo responsiveness.

A strong relationship between self-reported and objective functional performance should not be expected. In sleep and memory/cognition, there is generally a poor relationship between the subjective measures and objective performance.<sup>73-75</sup> Self-reported and objective measures are separate domains and provide different information about a patient's functional performance. An objective assessment of performance based on activity (such as tests for walk, stair, climb, or chair-stand) with a measure of pain during the activity may be a promising composite assessment of long-term pain and function.

### *Proposed Concept Clinical Endpoints of Long-term Benefit - Yura Kim, PhD*

To address the challenge of defining appropriate measures for OA clinical trials, Kim et al. presented concept clinical endpoints that capture clinical benefit in a long-term clinical setting and evaluate the feasibility of their use.<sup>64</sup> The analysis used the multi-center, longitudinal, observational Osteoarthritis Initiative (OAI) database, which consisted of 4,796 participants primarily with knee OA.<sup>76</sup> The proposed endpoints were based on TKR composite endpoints defined by TKR and thresholds of PROs of pain and function. When *Time-to-TKR* is used as an endpoint in a study with average follow-up time of three years, 3,000 to 18,000 participants are required, depending on effect size. In comparison, using a composite endpoint such as *Time-to-TKR or Severe Pain or Severely-Impaired Function*, the required sample sizes range from approximately 2,000 to 11,000 for a 3-year study. The presented approach demonstrates that composite measures could reduce sample size within the 3-5 year study follow-up periods compared to the use of *Time-to-TKR* alone. This approach directly incorporates pain and function so the patients who did not get a TKR surgery due to factors beyond pain and function (e.g., race, gender, socioeconomic status, access to care, surgeon preference, and health care systems) but have severe pain or disability also can be captured.

## **CONCLUSIONS**

The FDA and Arthritis Foundation recognize the tremendous unmet need for OA treatment options and thus convened the stakeholders of OA science, including academia, pharmaceutical companies, regulatory agencies, and patient communities, to encourage the continued open public dialogue and progress in the field. The expert participants at the virtual meeting discussed the current state of science in OA, identified knowledge gaps, and examined the challenges in bringing products to market.

Osteoarthritis is a complex disease with variable course and slow evolution, in most cases making drug development in that space challenging. It is widely recognized by patient communities and other stakeholders that lack of effective therapies contributes to disability, poor quality of life, and significant economic burden associated with osteoarthritis.

Currently, there are no accepted definitions of disease progression, and clinical outcomes that meaningfully reflect long-term patient benefit related to disease modification are needed. Structural outcomes, defined by morphological tissue changes observed on imaging studies including radiography or MRI, have been proposed as endpoints to capture the long-term patient benefit. Radiographs are widely used; however, their limitations include insufficient technical performance (accuracy, precision, reproducibility) for monitoring the effect of investigational products on all the various joint tissues in interventional clinical trials. MRI is superior as an imaging tool but is more expensive and less accessible. Machine learning approaches which combine qMRI biomarkers for multiple tissues and morphology within the

joint and non-imaging factors have been applied in observational data to develop a predictive model for end-stage joint replacement. Other biomarkers are also actively researched to help advance the science and the understanding of the disease of OA. Biochemical markers in urine and blood continue to be identified, with the most promising markers being CTXII and CRTAC1. However, none of the current biochemical or imaging biomarkers are validated as surrogate endpoints predictive of change in the disease course to enable their use in clinical trials.

The research community needs to continue collecting biomarker data in the context of clinical trials to invest in the future efficiencies that surrogate endpoints can provide. Importantly, the scientific community needs to define the clinical benefit of which these biomarkers are surrogates and to generate the supportive data. While qualifying one or more of these biomarkers as surrogate endpoints should be the goal, there may be other opportunities for biomarker use in OA treatment development, including understanding treatment effects, allowing for enrichment of trials, and informing dose selection. Continued innovation in clinical imaging tools may improve visualization of the cartilage and bone and may identify biomarkers prognostic of disease progression to support future development of surrogate endpoints to accelerate the availability of potentially promising products to treat OA.

The model of short-term clinical trials that may have worked well for analgesics and NSAIDs, is not the right model for developing products intended to change the disease course of OA. A unified concept to encompass the currently disconnected measures of pain, function, and structure is challenging due to the complexity of the disease. Similarly, defining endpoints that assess long-term patient benefit, besides the utilization of joint replacement data that represents end-stage OA, is a major knowledge gap and a significant unmet need in OA drug development. The workshop participants noted that in addition to the assessment of various degrees of durable improvement in pain and function of the affected joints, clinical outcomes sensitive to OA change, including performance-based assessment of pain, pain and movement perception evaluated by brain MRI, and use of mobile devices collecting clinically meaningful data, need to be further developed and validated. Based on analyses of the OAI data, the FDA proposed a conceptual approach with a composite endpoint using total knee replacement rates and unacceptable pain and disability, i.e., critical thresholds to define severe disease, as demonstrating delay in time to the development of severe disease would be a clinically meaningful long-term benefit of a product intended to change the disease course in OA. Once this conceptual approach is tested in prospective study designs, the scientific community will better understand for what product types it may be suitable in order to further optimize its use. Clinical studies using this approach may be larger and longer than the traditional studies for other rheumatological conditions or for analgesic products but could be feasible and will address many of the challenges discussed during the workshop.

#### D. AUTHOR CONTRIBUTIONS

- Conception and Design: All authors
- Analysis and interpretation of data: All authors
- Drafting of the article: Jason Kim
- Critical revision of the article for important intellectual content: All authors
- Final approval of the article: All authors
- Administrative, technical, or logistical support: Jason Kim
- Collection and assembly of data: All authors



## E. ROLE OF FUNDING SOURCE

The workshop was co-sponsored by the United States Food and Drug Administration and the Arthritis Foundation.

## F. CONFLICTS OF INTEREST

One or more of the authors has received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this article.

## G. ACKNOWLEDGEMENTS

The authors, planning committee, and speakers thank Bill Agee, Emily Hunt, and Stephanie Rosado for serving as patient panelists as well as Sadaf Nabavian of the FDA, Katie Bitner of the Arthritis Foundation for logistic and technical support.

## H. REFERENCES

- 1 Hawker, G. A. Osteoarthritis is a serious disease. *Clin Exp Rheumatol* **37 Suppl 120**, 3-6 (2019).
- 2 *21st Century Cures Act, FDA Website*, <<https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act>> (2020).
- 3 *FDA-Arthritis Foundation Osteoarthritis Drug Development Workshop: Assessment of Long-term Benefit*, <<https://www.fda.gov/drugs/news-events-human-drugs/fda-arthritis-foundation-osteoarthritis-drug-development-workshop-assessment-long-term-benefit>> (2021).
- 4 Cisternas, M. G. *et al.* Alternative Methods for Defining Osteoarthritis and the Impact on Estimating Prevalence in a US Population-Based Survey. *Arthritis Care Res (Hoboken)* **68**, 574-580, doi:10.1002/acr.22721 (2016).
- 5 Torio, C. M. & Moore, B. J. in *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs* (2006).
- 6 Kotlarz, H., Gunnarsson, C. L., Fang, H. & Rizzo, J. A. Insurer and out-of-pocket costs of osteoarthritis in the US: evidence from national survey data. *Arthritis Rheum* **60**, 3546-3553, doi:10.1002/art.24984 (2009).
- 7 Wang, S. X. *et al.* Healthcare resource utilization and costs by age and joint location among osteoarthritis patients in a privately insured population. *J Med Econ* **20**, 1299-1306, doi:10.1080/13696998.2017.1377717 (2017).
- 8 Safiri, S. *et al.* Global, regional and national burden of osteoarthritis 1990-2017: a systematic analysis of the Global Burden of Disease Study 2017. *Ann Rheum Dis* **79**, 819-828, doi:10.1136/annrheumdis-2019-216515 (2020).
- 9 Neogi, T. & Zhang, Y. Epidemiology of osteoarthritis. *Rheum Dis Clin North Am* **39**, 1-19, doi:10.1016/j.rdc.2012.10.004 (2013).
- 10 Kolasinski, S. L. *et al.* 2019 American College of Rheumatology/Arthritis Foundation Guideline for the Management of Osteoarthritis of the Hand, Hip, and Knee. *Arthritis Care Res (Hoboken)* **72**, 149-162, doi:10.1002/acr.24131 (2020).
- 11 Seed, S. M., Dunican, K. C. & Lynch, A. M. Osteoarthritis: a review of treatment options. *Geriatrics* **64**, 20-29 (2009).
- 12 Hochberg, M. C. & Dougados, M. Pharmacological therapy of osteoarthritis. *Best Pract Res Clin Rheumatol* **15**, 583-593, doi:10.1053/berh.2001.0175 (2001).
- 13 Fleming, T. R. & Powers, J. H. Biomarkers and surrogate endpoints in clinical trials. *Stat Med* **31**, 2973-2984, doi:10.1002/sim.5403 (2012).
- 14 Group, F.-N. B. W. (2016).

- 15 *Biomarker Qualification Program*, <<https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/biomarker-qualification-program>> (
- 16 DeMets, D. L., Psaty, B. M. & Fleming, T. R. When Can Intermediate Outcomes Be Used as Surrogate Outcomes? *JAMA* **323**, 1184-1185, doi:10.1001/jama.2020.1176 (2020).
- 17 Fleming, T. R. & DeMets, D. L. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* **125**, 605-613, doi:10.7326/0003-4819-125-7-199610010-00011 (1996).
- 18 *FDA's Decision to Approve New Treatment for Alzheimer's Disease*, <<https://www.fda.gov/drugs/news-events-human-drugs/fdas-decision-approve-new-treatment-alzheimers-disease>> (
- 19 *Alzheimer's drug approved despite doubts about effectiveness*, <<https://www.science.org/news/2021/06/alzheimer-s-drug-approved-despite-doubts-about-effectiveness>> (
- 20 Hunter, D. J., Nevitt, M., Losina, E. & Kraus, V. Biomarkers for osteoarthritis: current position and steps towards further validation. *Best Pract Res Clin Rheumatol* **28**, 61-71, doi:10.1016/j.berh.2014.01.007 (2014).
- 21 *Biomarkers Consortium - PROGRESS OA: Clinical Evaluation and Qualification of Osteoarthritis Biomarkers*, <<https://fnih.org/our-programs/biomarkers-consortium/programs/progress-oa>> (
- 22 Kraus, V. B. *et al.* Predictive validity of biochemical biomarkers in knee osteoarthritis: data from the FNIH OA Biomarkers Consortium. *Ann Rheum Dis* **76**, 186-195, doi:10.1136/annrheumdis-2016-209252 (2017).
- 23 Karsdal, M. A. *et al.* Treatment of symptomatic knee osteoarthritis with oral salmon calcitonin: results from two phase 3 trials. *Osteoarthritis Cartilage* **23**, 532-543, doi:10.1016/j.joca.2014.12.019 (2015).
- 24 Bjerre-Bastos J, B.-J. A.-C., Karsdal M, Byrjalsen I, Andersen J, Riis B, Christiansen C, Bihlet A. Biomarkers of bone and cartilage turnover CTX-I and CTXII predict total joint replacements in osteoarthritis. *Osteoarthritis Cartilage* **27**, S31-32 (2019).
- 25 Garnero, P., Sornay-Rendu, E. & Chapurlat, R. The cartilage degradation marker, urinary CTX-II, is associated with the risk of incident total joint replacement in postmenopausal women. A 18 year evaluation of the OFELY prospective cohort. *Osteoarthritis Cartilage* **28**, 468-474, doi:10.1016/j.joca.2019.12.012 (2020).
- 26 Zhou K, L. Y.-J., Soderblom E, Reed A, Sun S, Moseley M, Kraus V. . Qualification of proteomic biomarkers for knee osteoarthritis progression. *Osteoarthritis & Cartilage* **29**, S7-8 (2021).
- 27 Styrkarsdottir, U. *et al.* The CRTAC1 Protein in Plasma Is Associated With Osteoarthritis and Predicts Progression to Joint Replacement: A Large-Scale Proteomics Scan in Iceland. *Arthritis Rheumatol*, doi:10.1002/art.41793 (2021).
- 28 Driban, J. B. *et al.* The incidence and characteristics of accelerated knee osteoarthritis among women: the Chingford cohort. *BMC Musculoskelet Disord* **21**, 60, doi:10.1186/s12891-020-3073-3 (2020).
- 29 Sun S, L. Y.-J., Soderblom E, Moseley M, Zhou K, Reed A, Arden N, Kraus V. Serum prognostic biomarkers for incident radiographic knee osteoarthritis. *Osteoarthritis & Cartilage* **29**, S8-9 (2021).
- 30 Kraus, V. B. *et al.* OARSI Clinical Trials Recommendations: Soluble biomarker assessments in clinical trials in osteoarthritis. *Osteoarthritis Cartilage* **23**, 686-697, doi:10.1016/j.joca.2015.03.002 (2015).
- 31 Hunter, D. J. *et al.* OARSI Clinical Trials Recommendations: Knee imaging in clinical trials in osteoarthritis. *Osteoarthritis Cartilage* **23**, 698-715, doi:10.1016/j.joca.2015.03.012 (2015).

- 32 Guerhazi, A. *et al.* Severe radiographic knee osteoarthritis--does Kellgren and Lawrence grade 4 represent end stage disease?--the MOST study. *Osteoarthritis Cartilage* **23**, 1499-1505, doi:10.1016/j.joca.2015.04.018 (2015).
- 33 Roemer, F. W. *et al.* MRI-based screening for structural definition of eligibility in clinical DMOAD trials: Rapid OsteoArthritis MRI Eligibility Score (ROAMES). *Osteoarthritis Cartilage* **28**, 71-81, doi:10.1016/j.joca.2019.08.005 (2020).
- 34 Roemer, F. W. *et al.* State of the Art: Imaging of Osteoarthritis-Revisited 2020. *Radiology* **296**, 5-21, doi:10.1148/radiol.2020192498 (2020).
- 35 Li, X. & Majumdar, S. Quantitative MRI of articular cartilage and its clinical applications. *J Magn Reson Imaging* **38**, 991-1008, doi:10.1002/jmri.24313 (2013).
- 36 Xia, Y., Farquhar, T., Burton-Wurster, N. & Lust, G. Origin of cartilage laminae in MRI. *J Magn Reson Imaging* **7**, 887-894, doi:10.1002/jmri.1880070518 (1997).
- 37 Nieminen, M. T. *et al.* T2 relaxation reveals spatial collagen architecture in articular cartilage: a comparative quantitative MRI and polarized light microscopic study. *Magn Reson Med* **46**, 487-493, doi:10.1002/mrm.1218 (2001).
- 38 Duvvuri, U. *et al.* T1rho-relaxation in articular cartilage: effects of enzymatic degradation. *Magn Reson Med* **38**, 863-867, doi:10.1002/mrm.1910380602 (1997).
- 39 Li, X. *et al.* Quantitative MRI using T1rho and T2 in human osteoarthritic cartilage specimens: correlation with biochemical measurements and histology. *Magn Reson Imaging* **29**, 324-334, doi:10.1016/j.mri.2010.09.004 (2011).
- 40 Chalian, M. *et al.* The QIBA Profile for MRI-based Compositional Imaging of Knee Cartilage. *Radiology*, 204587, doi:10.1148/radiol.2021204587 (2021).
- 41 Theologis, A. A. *et al.* Comparison of T1rho relaxation times between ACL-reconstructed knees and contralateral uninjured knees. *Knee Surg Sports Traumatol Arthrosc* **22**, 298-307, doi:10.1007/s00167-013-2397-z (2014).
- 42 Prasad, A. P., Nardo, L., Schooler, J., Joseph, G. B. & Link, T. M. T(1)rho and T(2) relaxation times predict progression of knee osteoarthritis. *Osteoarthritis Cartilage* **21**, 69-76, doi:10.1016/j.joca.2012.09.011 (2013).
- 43 Kretzschmar, M. *et al.* Spatial distribution and temporal progression of T2 relaxation time values in knee cartilage prior to the onset of cartilage lesions - data from the Osteoarthritis Initiative (OAI). *Osteoarthritis Cartilage* **27**, 737-745, doi:10.1016/j.joca.2018.10.016 (2019).
- 44 Bowes, M. A. *et al.* Machine-learning, MRI bone shape and important clinical outcomes in osteoarthritis: data from the Osteoarthritis Initiative. *Ann Rheum Dis*, doi:10.1136/annrheumdis-2020-217160 (2020).
- 45 Brahim, A. *et al.* A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the OsteoArthritis Initiative. *Comput Med Imaging Graph* **73**, 11-18, doi:10.1016/j.compmedimag.2019.01.007 (2019).
- 46 Ashinsky, B. G. *et al.* Predicting early symptomatic osteoarthritis in the human knee using machine learning classification of magnetic resonance images from the osteoarthritis initiative. *J Orthop Res* **35**, 2243-2250, doi:10.1002/jor.23519 (2017).
- 47 Namiri, N. K. *et al.* Deep learning for large scale MRI-based morphological phenotyping of osteoarthritis. *Sci Rep* **11**, 10915, doi:10.1038/s41598-021-90292-6 (2021).
- 48 Razmjoo, A. *et al.* T2 analysis of the entire osteoarthritis initiative dataset. *J Orthop Res* **39**, 74-85, doi:10.1002/jor.24811 (2021).
- 49 Namiri, N. K. *et al.* Deep learning for large scale MRI-based morphological phenotyping of osteoarthritis. *Scientific Reports* **11**, 10915, doi:10.1038/s41598-021-90292-6 (2021).
- 50 Tolpadi, A. A., Lee, J. J., Pedoia, V. & Majumdar, S. Deep Learning Predicts Total Knee Replacement from Magnetic Resonance Images. *Scientific Reports* **10**, 6371, doi:10.1038/s41598-020-63395-9 (2020).

- 51 Morales, A. G. *et al.* Uncovering associations between data-driven learned qMRI biomarkers and chronic pain. *Scientific Reports* **11**, 21989, doi:10.1038/s41598-021-01111-x (2021).
- 52 Gigout, A., Guehring, H., Froemel, D., Meurer, A., Ladel, C., Reker, D., Bay-Jensen, A. C., Karsdal, M. A., & Lindemann, S. Sprifermin (rhFGF18) enables proliferation of chondrocytes producing a hyaline cartilage matrix. *Osteoarthritis and cartilage*, 1858-1867 (2017).
- 53 Lohmander, L. S. *et al.* Intraarticular sprifermin (recombinant human fibroblast growth factor 18) in knee osteoarthritis: a randomized, double-blind, placebo-controlled trial. *Arthritis Rheumatol* **66**, 1820-1831, doi:10.1002/art.38614 (2014).
- 54 Hochberg, M. C. *et al.* Effect of Intra-Articular Sprifermin vs Placebo on Femorotibial Joint Cartilage Thickness in Patients With Osteoarthritis: The FORWARD Randomized Clinical Trial. *JAMA* **322**, 1360-1370, doi:10.1001/jama.2019.14735 (2019).
- 55 Eckstein, F. *et al.* Intra-articular sprifermin reduces cartilage loss in addition to increasing cartilage gain independent of location in the femorotibial joint: post-hoc analysis of a randomised, placebo-controlled phase II clinical trial. *Ann Rheum Dis* **79**, 525-528, doi:10.1136/annrheumdis-2019-216453 (2020).
- 56 Roemer, F. W. *et al.* Evaluating the structural effects of intra-articular sprifermin on cartilage and non-cartilaginous tissue alterations, based on sqMRI assessment over 2 years. *Osteoarthritis Cartilage* **28**, 1229-1234, doi:10.1016/j.joca.2020.05.015 (2020).
- 57 Eckstein, F. *et al.* Long-term structural and symptomatic effects of intra-articular sprifermin in patients with knee osteoarthritis: 5-year results from the FORWARD study. *Ann Rheum Dis*, doi:10.1136/annrheumdis-2020-219181 (2021).
- 58 Guehring, H. *et al.* The effects of sprifermin on symptoms and structure in a subgroup at risk of progression in the FORWARD knee osteoarthritis trial. *Semin Arthritis Rheum* **51**, 450-456, doi:10.1016/j.semarthrit.2021.03.005 (2021).
- 59 Chow, Y. Y. & Chin, K. Y. The Role of Inflammation in the Pathogenesis of Osteoarthritis. *Mediators Inflamm* **2020**, 8293921, doi:10.1155/2020/8293921 (2020).
- 60 Chevalier, X. *et al.* Intraarticular injection of anakinra in osteoarthritis of the knee: a multicenter, randomized, double-blind, placebo-controlled study. *Arthritis Rheum* **61**, 344-352, doi:10.1002/art.24096 (2009).
- 61 Cohen, S. B. *et al.* A randomized, double-blind study of AMG 108 (a fully human monoclonal antibody to IL-1R1) in patients with osteoarthritis of the knee. *Arthritis Res Ther* **13**, R125, doi:10.1186/ar3430 (2011).
- 62 Wang, S. X. *et al.* Safety, tolerability, and pharmacodynamics of an anti-interleukin-1alpha/beta dual variable domain immunoglobulin in patients with osteoarthritis of the knee: a randomized phase 1 study. *Osteoarthritis Cartilage* **25**, 1952-1961, doi:10.1016/j.joca.2017.09.007 (2017).
- 63 Kloppenburg, M. *et al.* Phase IIa, placebo-controlled, randomised study of lutikizumab, an anti-interleukin-1alpha and anti-interleukin-1beta dual variable domain immunoglobulin, in patients with erosive hand osteoarthritis. *Ann Rheum Dis* **78**, 413-420, doi:10.1136/annrheumdis-2018-213336 (2019).
- 64 Fleischmann, R. M. *et al.* A Phase II Trial of Lutikizumab, an Anti-Interleukin-1alpha/beta Dual Variable Domain Immunoglobulin, in Knee Osteoarthritis Patients With Synovitis. *Arthritis Rheumatol* **71**, 1056-1069, doi:10.1002/art.40840 (2019).
- 65 *To Determine the Safety, Tolerability, Pharmacokinetics and Effect on Pain of a Single Intra-articular Administration of Canakinumab in Patients With Osteoarthritis in the Knee*, <<https://clinicaltrials.gov/ct2/show/study/NCT01160822>> (
- 66 Ridker, P. M. *et al.* Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N Engl J Med* **377**, 1119-1131, doi:10.1056/NEJMoa1707914 (2017).

- 67 Schieker, M. *et al.* Effects of Interleukin-1beta Inhibition on Incident Hip and Knee Replacement : Exploratory Analyses From a Randomized, Double-Blind, Placebo-Controlled Trial. *Ann Intern Med* **173**, 509-515, doi:10.7326/M20-0527 (2020).
- 68 Gossec, L. *et al.* OMERACT/OARSI initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis. *J Rheumatol* **34**, 1432-1435 (2007).
- 69 Gossec, L. *et al.* The role of pain and functional impairment in the decision to recommend total joint replacement in hip and knee osteoarthritis: an international cross-sectional study of 1909 patients. Report of the OARSI-OMERACT Task Force on total joint replacement. *Osteoarthritis Cartilage* **19**, 147-154, doi:10.1016/j.joca.2010.10.025 (2011).
- 70 Kim, Y., Levin, G., Nikolov, N. P., Abugov, R. & Rothwell, R. Concept Endpoints Informing Design Considerations for Confirmatory Clinical Trials in Osteoarthritis. *Arthritis Care Res (Hoboken)*, doi:10.1002/acr.24549 (2020).
- 71 Kashikar-Zuck, S., Flowers, S. R., Strotman, D., Sil, S., Ting, T. V., & Schikler, K. N. Physical activity monitoring in adolescents with juvenile fibromyalgia: findings from a clinical trial of cognitive-behavioral therapy. *Arthritis care & research*, 398-405 (2013).
- 72 Chandonnet, N., Saey, D., Alm eras, N., & Marc, I. French Pregnancy Physical Activity Questionnaire compared with an accelerometer cut point to classify physical activity among pregnant obese women. *PLoS one*, e38818 (2012).
- 73 Ferrioli, E., Skipworth, R. J., Hendry, P., Scott, A., Stensteth, J., Dahele, M., Wall, L., Greig, C., Fallon, M., Strasser, F., Preston, T., & Fearon, K. C. Physical activity monitoring: a responsive and meaningful patient-centered outcome for surgery, chemotherapy, or radiotherapy? *Journal of pain and symptom management*, 1025-1035 (2012).
- 74 Evenson, K. R., Herring, A. H., & Wen, F. Self-Reported and objectively measured physical activity among a cohort of postpartum women: the PIN Postpartum Study. *Journal of physical activity & health*, 5-20 (2012).
- 75 Kop, W. J., Lyden, A., Berlin, A. A., Ambrose, K., Olsen, C., Gracely, R. H., Williams, D. A., & Clauw, D. J. Ambulatory monitoring of physical activity and symptoms in fibromyalgia and chronic fatigue syndrome. *Arthritis and rheumatism*, 296-303 (2005).
- 76 Lee, J. *et al.* Sedentary behavior and physical function: objective evidence from the Osteoarthritis Initiative. *Arthritis Care Res (Hoboken)* **67**, 366-373, doi:10.1002/acr.22432 (2015).
- 77 Master, H. *et al.* Joint Association of Moderate-to-vigorous Intensity Physical Activity and Sedentary Behavior With Incident Functional Limitation: Data From the Osteoarthritis Initiative. *J Rheumatol* **48**, 1458-1464, doi:10.3899/jrheum.201250 (2021).
- 78 Trudeau, J. *et al.* Assessment of pain and activity using an electronic pain diary and actigraphy device in a randomized, placebo-controlled crossover trial of celecoxib in osteoarthritis of the knee. *Pain Pract* **15**, 247-255, doi:10.1111/papr.12167 (2015).
- 79 Weaver, E. M., Kapur, V., & Yueh, B. Polysomnography vs self-reported measures in patients with sleep apnea. *Archives of otolaryngology--head & neck surgery*, 453-458 (2004).
- 80 Bastien, C. H., Valli eres, A., & Morin, C. M. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep medicine*, 297-307 (2001).
- 81 Spencer, R. J., Drag, L. L., Walker, S. J., & Bieliauskas, L. A. Self-reported cognitive symptoms following mild traumatic brain injury are poorly associated with neuropsychological performance in OIF/OEF veterans. *Journal of rehabilitation research and development*, 521-530 (2010).
- 82 *The Osteoarthritis Initiative*, <<https://oai.epi-ucsf.org/datarelease/>> (