LUND UNIVERSITY

Application of biomarkers in the development of drugs intended for the treatment of osteoarthritis

Kraus, V. B.; Burnett, B.; Coindreau, J.; Cottrell, S.; Eyre, D.; Gendreau, M.; Gardiner, J.; Garnero, P.; Hardin, J.; Henrotin, Y.; Heinegård, Dick; Ko, A.; Lohmander, Stefan; Matthews, G.; Menetski, J.; Moskowitz, R.; Persiani, S.; Poole, A. R.; Rousseau, J. -C.; Todman, M.

Published in:
Osteoarthritis and Cartilage

DOI:
10.1016/j.joca.2010.08.019

2011

Link to publication

Citation for published version (APA):
Kraus, V. B., Burnett, B., Coindreau, J., Cottrell, S., Eyre, D., Gendreau, M., Gardiner, J., Garnero, P., Hardin, J., Henrotin, Y., Heinegård, D., Ko, A., Lohmander, S., Matthews, G., Menetski, J., Moskowitz, R., Persiani, S., Poole, A. R., Rousseau, J. .-C., & Todman, M. (2011). Application of biomarkers in the development of drugs intended for the treatment of osteoarthritis. *Osteoarthritis and Cartilage*, *19*(5), 515-542. https://doi.org/10.1016/j.joca.2010.08.019

Total number of authors:
20

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 12. Oct. 2024

Application of Biomarkers in the Development of Drugs Intended for the Treatment of Osteoarthritis

OARSI FDA Osteoarthritis Biomarkers Working Group

Virginia Byers Kraus, Bruce Burnett, Javier Coindreau, Susan Cottrell, David Eyre, Michael Gendreau, Jennifer Gardiner, Patrick Garnero, John Hardin, Yves Henrotin, Dick Heinegård, Amy Ko, Stefan Lohmander, Gloria Matthews, Joseph Menetski, Roland Moskowitz, Stefano Persiani, Robin Poole, Jean Charles Rousseau, Martin Todman

Author Affiliations

Virginia Byers Kraus: Department of Medicine, Duke University Medical Center, Durham, NC, UCA

Bruce Burnett: Duke Translational Medicine Institute, Duke University School of Medicine, Durham, NC, USA

Javier Coindreau: Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA

Sue Cottrell: Amgen, Seattle, WA, USA

David Eyre: Department of Orthopaedics and Sports Medicine, University of Washington, Seattle, WA, USA

Michael Gendreau: Cypress Biosciences, Inc, San Diego, CA, USA

Jennifer Gardiner: previously, Discovery Translational Medicine, Wyeth Research, Collegeville, PA, USA

Patrick Garnero: INSERM Research Unit 831, Université Lyon, Lyon, France

John Hardin: Albert Einstein College of Medicine, Bronx, NY, USA; Montefiore Medical Center, Bronx, NY, USA; Arthritis Foundation National Office, Atlanta, GA, USA

Yves Henrotin: Bone and Cartilage Research Unit, Institute of Pathology, level 5, CHU Sart-Tilman, 4000 Liège, Belgium

Dick Heinegard: Department of Experimental Medical Science, Lund University, Lund, Sweden

Amy Ko: Merck & Company, Inc, Rahway, NJ, USA

Stefan Lohmander: Department of Orthopaedics, Clinical Sciences Lund, Lund University, Lund, Sweden

Gloria Matthews: Genzyme Corporation, Cambridge, MA, USA

Joseph Menetski: Merck & Company, Inc, Rahway, NJ, USA

Roland W. Moskowitz: University Hospitals, Case Western Reserve University, Cleveland, OH, USA

Stefano Persiani: Division of Research & Development, Rottapharm SpA, Monza, Italy

A Robin Poole: Department of Surgery, McGill University, Montreal, Quebec, Canada; Canadian Arthritis Network, Canada

Jean Charles Rosseau: INSERM 831 Research Unit, Université de Lyon, Lyon, France

Martin Todman: Smith & Nephew Research Centre, Heslington, York, UK

Correspondence can be sent to:
Virginia Byers Kraus, MD, PhD
Box 3416
Duke University Medical Center
Durham, NC 27710
Phone: 919-681-6652
Fax: 919-684-8907
Email: vbk@duke.edu

Abstract

Objective. Osteoarthritis (OA) is a chronic and slowly progressive disease for which biomarkers may be able to provide a more rapid indication of therapeutic responses to therapy than is currently available; this could accelerate and facilitate OA drug discovery and development programs. The goal of this document is to provide a summary and guide to the application of *in vitro* (biochemical and other soluble) biomarkers in the development of drugs for OA and to outline and stimulate a research agenda that will further this goal.

Methods. The Biomarkers Working Group representing experts in the field of OA biomarker research from both academia and industry developed this consensus document between 2007-2009 at the behest of the Osteoarthritis Research Society International (OARSI FDA initiative).

Results. This document summarizes definitions and classification systems for biomarkers, the current outcome measures used in OA clinical trials, applications and potential utility of biomarkers for development of OA therapeutics, the current state of qualification of OA-related biomarkers, pathways for biomarker qualification, critical needs to advance the use of biomarkers for drug development, recommendations regarding practices and clinical trials, and a research agenda to advance the science of OA-related biomarkers.

Conclusions. Although many OA-related biomarkers are currently available they exist in various states of qualification and validation. The biomarkers that are likely to have the earliest beneficial impact on clinical trials fall into two general categories, those that will allow targeting of subjects most likely to either respond and/or progress (prognostic value) within a reasonable and manageable time frame for a clinical study (for instance within one to two years for an OA trial), and those that provide early feedback for preclinical decision-making and for trial organizers that a drug is having the desired biochemical effect. As *in vitro* biomarkers are increasingly investigated in the context of specific drug treatments, advances in the field can be expected that will lead to rapid expansion of the list of available biomarkers with increasing understanding of the molecular processes that they represent.

1. INTRODUCTION

It is said that a disease starts when detected by the best marker available to define it. To date, this usually requires the presence of a clinical symptom, which often occurs well into the progression of an illness or disease. However, there is significant evidence that there are often early, pre-symptomatic biomarkers of illness and disease, which if detected, may allow for earlier treatment. Therein lies the power and importance of applying biomarkers to osteoarthritis (OA), a disease often characterized by a prolonged asymptomatic molecular phase, a preradiographic phase, and a recalcitrant later radiographic phase with evident structural joint changes, frequent pain, and loss of function (Figure 1). Biomarkers have the potential to provide an early warning of the initiation of matrix breakdown that could prompt earlier treatment to prevent the cartilage and bone destruction that leads to disability. Thus, there currently exists a great need and opportunity for biomarkers to provide a method for earlier diagnosis of OA, and to inform the prognosis, monitoring and therapeutic strategies for OA. Wagner has predicted that the next few years will see a rapid increase in the number of drugs approved with biomarker data in their labels, and older drugs that will have biomarker data added to their labels [1]. OA may be chief among them due to the current lack of a gold standard that comprehensively captures the disease in all of its manifestations. In addition, OA is a chronic and slowly progressive disease for which biomarkers may be able to provide a more rapid indication of therapeutic response to disease structure modifiers than is available through currently established means; this could streamline and optimize the discovery and development programs of new therapeutic agents. The mandate of the OARSI FDA Biomarkers Working Group was twofold. First to create a critical appraisal of fundamentals of the science related to biomarkers of OA, particularly as they relate to the development of drugs intended for the treatment of OA. Second, to address specific queries posed by the FDA related to OA biomarkers, namely: What biomarkers now exist? What is their utility? What evidence is available to support surrogacy for clinical outcomes? What is the face validity? What is the practicality? What is the research agenda required to inform each of the above questions? Thus this document is intended to address this twofold purpose in the hopes of helping to advance the development of drugs for OA.

1.1. Scope of the Document

A previous broad ranging biomarker white paper was commissioned and prepared for the launch of the National Institutes of Health Osteoarthritis Public/Private Research Initiative and was published on line in 2000 (and now found at the OARSI website http://www.oarsi.org/index2.cfm?section=OARSI_Initiatives&content=Biomarkers) The present document has a much more specific focus. It also covers the great increase in biomarker research activity in the present decade and utilizes definitions and nomenclature that are harmonized with and expand upon those proposed to date in Federal Drug Administration (FDA) draft guidance documents. This current paper covers biochemical/molecular and genomic (RNA-gene expression, DNA-genetic polymorphisms) biomarkers of OA but excludes imaging biomarkers and clinical risk factors such as obesity, malalignment, and gender because other working groups are covering these topics in companion documents. We include a brief summary of issues related to the current methods of OA diagnosis, treatment and response criteria for therapeutic trials, and the challenges posed by the current ‘gold standard’ radiographic trial criteria, in order to provide a

framework in which to conceptualize the role to be played by biomarkers in the development of drugs for OA. The concept of OA as a continuum that includes early stages that may be amenable to treatment if appropriate biomarkers are defined, which in turn could complement current treatment paradigms for established radiographic OA, traditionally referred to as primary and secondary prevention, respectively.

Potential uses and challenges for each type of biomarker based on the BIPEDS classification scheme (described below) in the drug development process are discussed. Summary tables illustrating study power for treatment effects based on varying effect sizes are provided utilizing a theoretical biomarker as well as known soluble biochemical OA biomarkers, and their current level of qualification based on published clinical trials.

A summary of the pathways required for biomarker qualification is included that lists the regulatory agencies involved with biomarker development, as well as recommendations for biomarker endpoints in trials. Clinical and scientific issues are also raised that would benefit from more research. Appendices are provided containing recommendations for sample collection, processing and storage, as well as a glossary of biomarker terms.

1.2. Definition of Biomarkers

A biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention [2]. This is in contrast to a clinical endpoint that is a marker or variable that measures how a patient feels, functions or survives. A biomarker becomes a surrogate endpoint when it is appropriately qualified to substitute for a clinical endpoint. The technical revolution in molecular biology has led to the expansion of the notion of what constitutes a potential biomarker to include, not only proteins and protein fragments, but also metabolites, carbohydrate biomarkers, genomic biomarkers (RNA and DNA) [3], cellular biomarkers (captured as the cell pellet from body fluids), and imaging biomarkers. Based on their characteristics, we can divide biomarkers into two major groups: the so called soluble or "wet biomarkers", usually measured in a selected body fluid such as blood, serum, plasma, urine, or synovial fluid and usually representing modulation of an endogenous substance in these fluids; and the so called "dry biomarkers" usually consisting of visual analog scales, questionnaires, performed tasks, or imaging. These two types of biomarkers can also be referred to as *in vitro* biomarkers (derived from *in vitro* diagnostics) versus *in vivo* biomarkers respectively. Although many of the concepts presented here are applicable to all of these types of biomarkers, imaging biomarkers are dealt with more specifically in a companion document so we focus herein on the non-imaging, *in vitro*, soluble biomarkers.

1.3. Processes of Biomarker Qualification and Validation

<u>Qualification</u> is a process applied to a particular biomarker to support its use as a surrogate endpoint in drug discovery, development or post-approval and, where appropriate, in regulatory decision-making [2]. In contrast, <u>validation</u> of a biomarker is much broader and can relate to verification of analytical performance characteristics (such as precision, accuracy, stability, etc) as well as clinical correlation of a biomarker with a biological process or clinical outcome. Current practice however is to supplant the term validation with qualification when the focus is on the portent (meaning) as opposed to the performance (analytical aspects) of the biomarker. A major difference between validation and qualification resides in the fact that the latter only has meaning in a context. For example,

qualification of a biomarker may take into consideration the particular level of progression of the disease and its severity, thereby leading to the qualification for some states of the disease, but not for others. A systematic process has been in development for accurate and comprehensive qualification of biomarkers for use in drug development [4]. To date, draft guidelines exist on qualification of genomic biomarkers [2], produced by the International Conference on Harmonisation (ICH), whose goal has been to create a harmonized structure for qualifying the biomarkers that will lead to consistent applications and discussions among regulatory authorities and sponsors. Qualification endpoints in OA could include structural outcomes (identified with MRI, or x-ray etc), and/or clinical outcomes (pain, function etc); biochemical and/or genomic biomarkers are linked to modifications in these outcomes through the process of biomarker qualification.

1.4. Classification Systems for Biomarkers

1.4.1. BIPEDS

In this document we refer to and use two main classification systems for biomarkers with modifications as described here. The first, a system called BIPED, classifies the major types of biomarkers[5] into 5 categories corresponding to Burden of disease, Investigational, Prognostic, Efficacy of Intervention, and Diagnostic biomarkers. We have added a Safety category to the BIPED system, and hereafter, throughout this document, refer to the BIPEDS classification system. This change facilitates the goal of this document to provide a guide to the comprehensive application of biomarkers to the study and treatment of osteoarthritis. Biomarkers of safety can be considered biomarkers able to reflect tissue and or organ toxicity of an agent or intervention and are analogous to biomarkers of toxicity in the process of evaluation and validation by the Critical Path Initiative for diverse organ systems (see home page http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/default.htm).

1.4.2. Qualification levels for biomarkers

The second useful classification system referred to here divides biomarkers into four categories according to their current level of qualification described further in section 5.2.1 [1]:

Exploration level biomarkers are research and development tools accompanied by *in vitro* and/or preclinical evidence for which there is no consistent information linking the biomarker to clinical outcomes in humans (these are used for hypothesis generation);

Demonstration level biomarkers are associated with clinical outcomes but have not been reproducibly demonstrated in clinical studies (this category corresponds to ‘‘probable valid biomarkers’’ in nomenclature suggested in draft guidance from the FDA [6] and are useful for decision-making by providing evidence to support the primary clinical evidence);

Characterization level biomarkers are reproducibly linked to clinical outcomes in more than one prospective clinical study in humans (this category corresponds to ‘‘known valid biomarkers’’ in nomenclature suggested in guidance by the FDA[6] and are useful for decision-making, dose finding, and secondary and tertiary claims); and

Surrogacy level biomarkers can substitute for a clinical endpoint (this category corresponds to ‘‘surrogate end point’’ and requires agreement with regulatory authorities as an FDA registrable endpoint).

1.5. Summary

As noted in a recent FDA guidance document [7], the use of biomarkers in drug discovery,

development and post-approval has the potential to facilitate development of safer and more effective medicines; in fact, one of the main objectives of a biomarker for drug development is to allow the construction of the dose-exposure-response curve in patients for both the therapeutic and toxicity effects. This will facilitate dose selection in order to reach the best benefit-risk ratio of an approved medicine. In the OA field, the potential also exists for biomarkers to enhance the probability of obtaining early indications of success during clinical drug development for OA. The selection of a new biomarker test depends critically upon the ability of the test to link the mechanism of action of a new agent with a therapeutic response. The therapeutic response usually addresses an unmet medical need, and in the case of OA, there are currently no qualified biomarkers that can be considered as surrogate clinical endpoints. Thus it is a two-edged sword: the ultimate degree of biomarker uptake and use is intimately tied to the ability to act on the biomarker information provided, which in turn is dependent on the ability of biomarkers to enhance the success of clinical trials to achieve the actionable result needed for biomarkers to be adopted for clinical use.

It is worth noting here that the field of drug development for OA is currently analogous to osteoporosis 30 years ago [8], namely a disease in search of a robust gold standard outcome measure to inform clinical trials. The 1979 FDA Osteoporosis Guidelines acknowledged that evaluating the clinical effectiveness of osteoporosis drugs posed special challenges because of the "difficulties in assessing the state of skeletal bone quantitatively *in vivo*, the relatively small changes that are usually encountered and the duration of studies necessary to show significant effects" [8, 9]. By 1984, the FDA Osteoporosis Guidelines upgraded dual-energy photon absorptiometry from investigational to a valid and reliable method for measuring trabecular bone mass of the spine and this was critical to the subsequent approach to the development and regulation of osteoporosis drugs [8, 9]. OA is at a similar crossroads to which biomarkers may contribute substantively at this time. Given the urgent need for OA therapies, it is hoped that the concepts advanced in this document will facilitate and stimulate the inclusion of biomarkers as secondary endpoints in all future OA trials, and lay the groundwork for the evolution to the use of biomarkers, in some cases, as primary endpoints.

2. OSTEOARTHRITIS DIAGNOSIS, TREATMENT AND TRIALS

2.1. Diagnosis

The American College of Rheumatology (ACR) has developed a set of clinical and radiological criteria for the diagnosis of hip, knee and hand OA [10-12]. The ACR diagnostic criteria are based on the association of many clinical, or clinical and radiological criteria, and are commonly used for patient inclusion in clinical trials. These ACR criteria are very specific and thus are useful for differentiating patients with OA from those with inflammatory joint diseases. Their sensitivity is less impressive, illustrating their limited ability to discriminate patients with early OA from healthy controls. The most commonly used radiographic grading system is that of Kellgren and Lawrence (KL) [13], based on the presence of osteophytes, joint space narrowing (JSN), subchondral bone sclerosis and cyst formation. This scoring system divides OA into five grades (0-4) mainly based on the presence and number of osteophytes. A score of 2 or more has traditionally been considered to be a definitive radiographic diagnosis of OA and has been widely used in clinical trials as an inclusion criterion. However, evidence suggests that KL grade 1 is *bona fide* OA and distinct from KL grade 0 based on subsequent risk of progression [14]. Based on the concept

of the disease continuum that includes a molecular stage and a pre-radiographic stage of OA as presented in Figure 1 and supported by the literature [15, 16], even with inclusion of KL grade 1 as bona fide OA, radiographic criteria will identify only late stage OA.. Because the KL scoring system relies predominantly on osteophytes to determine OA severity, the atrophic form of OA, which consists mainly of JSN, is underestimated. The KL grading system is also known for its poor correspondence of radiographic severity with hip or knee pain. MRI, ultrasound or biochemical markers are not yet included in any set of diagnostic criteria for OA.

2.2. Treatment

A cure for OA remains elusive and the management of OA is largely palliative, focusing on the alleviation of symptoms. Current recommendations by the European League Against Rheumatism (EULAR), the ACR, and the Osteoarthritis Research Society International (OARSI) for the management of OA include a combination of non-pharmacological interventions and pharmacological treatments. One of the main obstructions to efficient development of new structure modifying therapies for OA is the low sensitivity to change of the plain radiographic endpoints that necessitates long-term trials involving a large number of patients to show a significant difference between placebo and active-drug treated groups. Biomarkers are promising sensitive tools, but they have to demonstrate specificity for OA pathology and ideally, provide earlier information than JSN measurement by X-ray. The current paucity of 1) biomarker data from human OA clinical trials (summarized in Table 2), and 2) data on the role of biochemical markers for monitoring the treatment of OA, can chiefly be ascribed to the absence of therapies with structure modifying activity. Without a structure modifying agent and a practically useful gold standard for monitoring structural change, it is challenging to qualify a biomarker to be “fit for purpose” for monitoring structural modification. Nevertheless, preclinical studies of DMOADs using biomarkers offer significant promise in terms of early indications of responses to treatment that may translate into the clinic. Experiences with biomarkers in the context of biologic therapies in rheumatoid arthritis offer promise for OA in that short-term changes in serum levels of biomarkers following initiation of therapy may predict long-term clinical and radiographic outcomes [17]. These kinds of data need to be generated in OA trials [18, 19].

Another issue regarding treatment monitoring using biochemical markers is the heterogeneity of OA subsets. Results may differ considerably between subsets with differences in pathobiology. OA may be localized in one joint or generalized, hypertrophic with osteophytes and subchondral bone sclerosis or atrophic, slowly or rapidly progressing or showing no progression. Finally, a therapy may act on OA through a variety of mechanisms and pathways. This suggests that a biomarker may need to be specific for the particular molecular target of the therapy in question. For instance, neoepitopes generated by collagenase activity could be sensitive to collagenase inhibitors but not to drugs acting on proteoglycan turnover. Even if a biomarker reflects the effects of a particular therapy, it may not reflect all the mechanisms of action of the drug, thus underestimating the therapeutic efficacy or missing the toxicity of the particular therapy. This means that the sensitivity to change of a biomarker in a clinical trial may be dependent on the characteristics of the population and the mechanisms of action of the therapy. For these reasons, it would be advantageous to develop a panel of biomarkers and use a wide variety of biomarkers during the preclinical and clinical drug development processes.

2.3. Therapeutic Trials

OA clinical trials are commonly focused on the investigation of symptoms or structure modification. In general, trial participants fulfill the validated OA criteria of the American College of Rheumatology (ACR). In addition, trials of symptom-modifying agents include patients whose disease is likely to respond to treatment, for example those with at least moderate intensity of symptoms (VAS $\geq$50 mm), and those with a flare of symptoms upon withdrawal of their standard therapy (flare trials). These trials are generally limited to 3 or 6-months follow-up.

Trials of structure-modifying agents include patients without end-stage disease and often those with a perceived high risk for structural progression, for example, middle age, overweight women, although these traditional selection criteria are generally poor for identifying risk of knee OA progression [20, 21]. Structure-modifying trials generally span one to three years. A series of disappointing late-stage terminations of clinical trials investigating new potential disease modifying OA drugs (DMOADs) has led to the call for a new development paradigm for DMOADs, with a stronger focus on the biology of the joint and the redesign of clinical trials to include new and more sensitive biomarkers [22].

One very important issue that is usually ignored in recruiting patients for clinical trials is the phasic nature of OA in some patients resulting in much variability in rates of disease progression. Some patients with knee OA observed over prolonged periods (5 years) may experience periods of progressive structural damage and then relative inactivity [23]. Often non-progressors have been found to predominate in OA clinical trials for disease modification making the detection of therapeutic efficacy very difficult if not impossible. Importantly, there are a few studies that demonstrate the potential ability to identify progressors using biomarkers thereby enabling enrichment of trial populations with disease progressors as opposed to non-progressors, and providing a significant advantage over existing practice [18, 19, 24, 25]. In future, recruitment for clinical trials should take advantage of such biomarker-directed opportunities to enrich for progressors.

The OMERACT-OARSI consensus has recommended a core set of clinical outcome measures that should be included in clinical trials in OA. No OMERACT-OARSI guidelines have yet been developed for the use of non-imaging *in vitro* biomarkers in clinical trials. The core set of clinical items includes pain, physical function, patient global assessment, and for studies of at least 1 year duration, joint imaging [26]. It was subsequently found that successful trial designs must include both absolute and relative change, as well as measures of pain and function as primary domains [27]. Each of these types of clinical outcome measures (pain, physical function, patient or physician global responses), as well as imaging outcomes, can serve as clinical trial endpoints and endpoints for biomarker qualification.

The success of biomarker qualification on a structural modifying endpoint depends critically on the performance and specificity of the endpoint. Although the, methodological limitations are well recognized [28], to date, assessment of the inter-bone distance and loss of joint space on a plain radiograph of the hip or knee is the only validated measure of OA progression recommended for use in randomized clinical trials in OA. Unfortunately, the limitations of the traditional clinical trial outcome, joint space narrowing (JSN), are considerable and have hampered the qualification of biomarkers as well as the registration of disease-modifying OA drugs (DMOADS). To date, no therapeutic agent has met this definition, and it remains

unclear how best to identify structural outcomes, whether by radiographs, magnetic resonance imaging, biomarkers, or direct visualization using arthroscopy or a combination of these approaches.

General limitations of joint space narrowing (JSN) that hamper the qualification of biomarkers include the following:

- It is an indirect measure of the alterations in articular cartilage;
- It fails to measure a dynamic process;
- Assessment of knee OA is confounded by the presence of meniscal lesions and meniscal extrusion [29];
- Changes in the knee over time are small, and typically occur in only a subset (progressor) of patients [mean estimated annual JSN rate 0.13 $\pm$ 0.15 mm/year for knee OA)] [30];
- It is poorly reproducible when measured from conventional weight-bearing radiographs of the hip or especially, of the knee in full extension;
- Apparent joint space narrowing occurs in the absence of structural changes due to varying degrees of knee flexion;
- Bone marrow and synovial abnormalities may go undetected;
- X-ray features appear only after deterioration of surrounding hard and soft tissues;
- It is poorly correlated with joint function and pain.

A variety of methodological approaches have been proposed to improve the reproducibility of the assessment of the joint space width in randomized controlled trials (RCTs) including semi-flexed views [31], and fluoroscopically assisted protocols; it remains unclear which approaches are preferable, or whether other imaging techniques are preferable and more promising. Among the new imaging techniques, MRI is the most promising and a more sensitive imaging modality for use in the immediate future. MRI allows assessment of cartilage biochemical and biomechanical integrity. It permits quantification of cartilage volume and changes in cartilage contour and can be tailored to assess pathological changes in associated joint structures, and tissues including bone, synovium (inflammation), ligament, menisci and muscle as well as effusions. Correlations between serum biomarkers and MRI data have already been reported for knee OA [32]. Moreover, a combination of MRI and soluble biomarkers have recently been used to improve the ability to identify patients at highest risk of knee OA progression over either modality used independently [33]. MRI has not yet been recommended as a primary endpoint in structural modifying RCTs in OA. A review of its potential, and recommendations regarding the use of MRI for OA clinical trials, is the subject of a companion OARSI FDA white paper.

Although the consensus reached at OMERACT 3 advocated continued study of biological markers of bone and cartilage degradation and repair, none was recommended for inclusion in clinical trials. Nonetheless, in view of the duration required for phase III structure-modifying trials, identification of a surrogate biomarker for use in earlier phase II trials could considerably improve the safety, cost, and efficiency of clinical development programs. Osteoporosis trials provide a good example in which molecular biomarkers are increasingly used as adjunct measures of effect before initiation of multi-year long phase III trials [34].

3. BIOMARKER APPLICATIONS IN DEVELOPMENT OF THERAPEUTICS FOR OSTEOARTHRITIS

Qualified biomarkers of OA have the potential to greatly expand the knowledge gained from preclinical and clinical trials of disease modifying agents. The BIPEDS system classifies potential OA biomarkers into six categories and encompasses the array of biomarkers that could be used for enhancing clinical trials. The most immediate hurdle facing researchers wishing to test a potential DMOAD in humans is the lack of early information in a clinical trial. In order to test a DMOAD, a trial must presently have a lengthy follow up, enroll many subjects and rely upon an insensitive method of assessment of disease progression. The level of financial investment is daunting, resulting in a negative impact on research and development. In this section, we consider how each category of the BIPEDS classification scheme could be used to improve clinical trial design and outcome. We also address the challenges in developing and qualifying such biomarkers for clinical use.

3.1. Burden of Disease

Burden of disease biomarkers indicate the extent or severity of disease and could be considered tools for the staging of the disease. They reflect the state of the disease at the time of assessment, but do not necessarily predict a likelihood of progression or change in disease burden. A burden of disease biomarker is typically qualified by comparison to a clinically defined gold standard assessment method. A burden of disease biomarker assessed locally, such as from analysis of synovial fluid, would be expected to reflect the disease status in a single joint, while assessment in blood or urine would more likely indicate the extent of the disease in all joints as well as normal physiology. Some molecular biomarkers, such as biomarkers of cartilage turnover, can provide information on the nature and extent of the current active process, but will not indicate the level of tissue damage already accrued or its precise location.

3.1.1. USES

- To provide a global measure of disease burden from all joints and skeletal and soft tissue components thereof;
- Potentially to discriminate between mono- and polyarticular osteoarthritis;
- To identify patients with high burden of active disease for inclusion into clinical trials of DMOADS expected to improve later stage disease;
- To help identify patients with low burden of active disease but with no or limited tissue alterations or structural alterations for inclusion in clinical trials of DMOADS expected to prevent progression of early OA;
- To balance treatment arms in a DMOAD trial for metabolic activity or stage of disease that would not otherwise be obvious from usual randomization criteria;
- To identify where in the body the burden of disease lies and aid in patient stratification, made possible when joint-specific biomarkers or patterns of biomarker expression are discovered.

3.1.2. CHALLENGES

- Requires comparison to gold standard for qualification, but there is no clear gold standard;
- A biomarker may be more sensitive than imaging, picking up a signal of early OA in asymptomatic joints with no obvious imaging changes;
- Uncertainty about what level of burden of disease is the optimal target for a DMOAD as early pathology may differ from more advanced pathology;
- There may be molecular subsets of disease - a biomarker might accurately reflect the burden of disease in one patient but not another;
- The level of a biomarker may change with the disease progression, such that some will be particularly elevated in early phases and others in late phase;
- Due to the complex nature of the joint organ comprised of different tissue types, a true burden of disease measurement might require multiple biomarkers.

3.2. Investigative

Investigative biomarkers are those that may not yet have enough evidence accumulated to be assigned to a particular BIPEDS category but nevertheless show sufficient promise to be incorporated in drug research at early stages to determine utility for subsequent use. In general, investigative biomarkers should be included, along with better-qualified biomarkers, in preclinical studies and clinical trials to advance our understanding of the disease and drug and to provide opportunities for biomarker development and qualification.

3.2.1. USES

- To explore novel biomarkers that could be informative in future preclinical and clinical trials;
- To contribute to biomarker data packages that support qualification of a biomarker or biomarker set for a particular outcome;
- To further understand the pathobiology of osteoarthritis;
- To further understand the mechanism of action of a DMOAD.

3.2.2. CHALLENGES

- Assays for investigative biomarkers might not be well validated and the data produced might not be robust;
- Conversely, investigative assays could produce highly reproducible, robust data that turn out to lack specificity for the molecular or tissue target;
- Clinical trials are not currently designed for testing of investigative biomarkers, making it difficult to achieve statistical power for biomarker evaluation;
- Biomarkers studied in preclinical disease models might not translate to human OA.

3.3. Prognostic

A prognostic biomarker indicates whether a patient’s disease is likely to progress and may also indicate how quickly the progression will occur. A prognostic biomarker may also provide an early response to treatment that is prognostic of subsequent, much later, clinical responses. Similarly, a prognostic biomarker could indicate who is at risk for developing symptomatic OA. There is a need for such markers since current clinical trials designed without the aid of biomarkers, often contain a minority of progressors (mean annual risk 6%,

range 1-20% based on KL grade) [30]. Predictive biomarkers, used to identify a subset of patients likely to respond to a particular drug, constitute a particularly useful subset of prognostic biomarkers. For instance, a threshold PGE2 level in synovial fluid might correlate with the ability of a COX-2 antagonist to be effective in that joint. Prognostic biomarkers include the largest variety of biomarker types, including variant biochemical biomarkers and invariant genetic biomarkers, although the latter may at some point in the future be considered risk factors as opposed to biomarkers.

3.3.1. USES

- To select subjects likely to progress rapidly ('high-risk' patients by biomarker measurement) to reduce the length of time required to see an effect of a DMOAD in a clinical trial thereby shortening the trial and to improve the chances of observing efficacy;
- To select subjects likely to progress rapidly ('high-risk' patients by biomarker measurement) for purposes of stratification;
- To increase the power of a trial to detect a significant drug effect with a limited number of subjects;
- To select subjects likely to progress rapidly ('high-risk' patients by biomarker measurement) who would benefit most from therapy with structure modifying agents;
- To select subjects for primary prevention trials (screen for at risk for developing OA to demonstrate reduction of incidence);
- To select patients likely to respond to a given drug for inclusion in a clinical trial. For instance, patients with high levels of an MMP-13 specific collagen cleavage product could be selected for inclusion in a trial of an MMP-13 inhibitor;
- As a companion diagnostic, to select likely responders for treatment with a marketed product;
- To provide predictive evidence that disease processes have been beneficially impacted by serving as an early indicator of a later trial outcome or response to therapy; this category of markers would therefore form a specific subset of efficacy of intervention markers described below.

3.3.2. CHALLENGES

- The prognostic effect of a biochemical biomarker must be distinguished from prognostic clinical (weight, injury) or genetic variables that may influence biomarker levels;
- Qualification of a prognostic biomarker would require a large, long and financially daunting prospective trial although this challenge may be overcome with the use of legacy samples from the many excellent existing osteoarthritis epidemiology studies.

3.4. Efficacy of Intervention

Biomarkers of efficacy of intervention can range from target engagement and pharmacodynamic assays (which assess whether the compound is hitting the desired target and is having the desired downstream biochemical effects) to strict surrogate endpoints that indicate the drug is having an impact on the clinical manifestations of the disease. Slowly

progressive diseases, such as osteoarthritis, pose a range of drug development challenges, particularly in phase II dose-finding studies [35]. Target engagement and pharmacodynamic biomarkers are likely to have the earliest impact on drug development of all the BIPEDS biomarkers by influencing decisions on dose selection and advancement of drugs to later phase trials. While a surrogate biomarker would be highly desirable, the path to generation and qualification for a 'characterization level' biomarker is likely to be shorter and provide benefit to programs in the near term at decision points in early preclinical studies and clinical trials. In contrast, qualification of a biomarker as a surrogate biomarker will be a painstaking but highly valuable effort (see section 5).

3.4.1. USES

- To demonstrate that a drug is having the desired immediate downstream biochemical effect;
- To understand the pharmacodynamics of a drug intervention and the relationship between pharmacodynamics and pharmacokinetics;
- To provide a basis for the selection of lead candidates for clinical trials;
- To contribute to the understanding of the pharmacology of candidates;
- To characterize subtypes of disease for which a therapeutic intervention is most appropriate;
- To choose a dose and dose schedule via *ex vivo* and *in vivo* studies;
- To support an efficacy endpoint;
- To support go/no go decisions in advance of preclinical and clinical studies and trials;
- To serve as a surrogate biomarker for delay of structural worsening, reduction of pain, or improvement in function.

3.4.2. CHALLENGES

- For drugs administered intra-articularly to treat a single joint, it may be difficult to monitor efficacy of intervention using systemic biomarker assessments (blood or urine), particularly if other joints are involved in OA;
- Qualification as a surrogate biomarker is difficult in the absence of a gold standard;
- In order for a pharmacodynamic or target engagement biomarker to be informative, it must be specific for the mechanism of action of drug being assessed;
- A biomarker might provide an accurate assessment of target engagement, but might not be related to clinical response.

3.5. Diagnostic

A diagnostic biomarker usually indicates whether an individual has the disease or a specific subtype of the disease, but may not reflect disease severity. It also has the potential to identify people at risk for OA based on genetic or other considerations. A biochemical biomarker could be more sensitive than an imaging marker, by detecting the process leading to OA before it is detectable by radiography or other imaging modalities.

3.5.1. USES

- To select subjects with molecular pre-radiographic OA for primary prevention trials;

- To identify patients with different disease subtypes;
- To identify individuals unlikely to have OA as controls in case-control studies.

3.5.2. CHALLENGES

- The processes in OA vary with time and may vary in nature, although common pathobiology is identifiable. A single diagnostic biomarker may therefore not be informative in all patients;
- Qualification of a diagnostic biomarker requires a gold standard. A biochemical assay could potentially be more sensitive than an imaging gold standard. The qualification would then depend on long term cohorts where the diagnosis can be verified in follow up;
- Given the insidious onset and slow progression of OA structural changes, it may take many years, patients, trials, and dollars to achieve correlation between a biochemical biomarker and disease. The NIH/NIAMS/NIA public/private Osteoarthritis Initiative is an example of an effort that could contribute to this end or assessing the correlation between biomarkers and osteoarthritis.

3.6. Safety

There exist important opportunities to use biomarkers to detect pathological changes and cytotoxicity. Safety biomarkers could be used in preclinical and clinical applications to monitor the health of the joint tissues, the whole joint organ, or the skeleton in general. For instance, biomarkers reflecting the synthesis of the main proteins of the joint might provide an index of the "joint-protective" effect of a potential treatment. There are currently no studies exploring specifically this aspect of joint tissue related biomarkers. Potential complications obviously exist with regard to discriminating toxic or pathological effects from beneficial effects in the case of skeletal biomarkers. In the absence of contrary evidence, increased cartilage degradation or decreased synthesis of cartilage based on biomarker data would be considered as potential "red flags" in any treatment regimen. A special circumstance is represented by repair, exemplified by collagen fibrillogenesis, where molecules catalyzing and enhancing this process, may instead prevent fibril formation and hamper repair when produced in relative excess [36, 37].

In contrast, there are emerging examples of toxicity monitoring in OA trials with biomarkers of other organ systems. A notable recent example is provided by the pilot trial of Brune 2009 [38] wherein N-terminal pro-B-type natriuretic peptide concentrations were shown to predict the risk of cardiovascular adverse events from NSAIDs and glucocorticoid rescue medications in a trial of an MMP inhibitor for OA. We anticipate that this will be a growing area that will enhance the goal of personalized medicine and patient safety. Clearly, a broad spectrum of biomarkers will be necessary for a full safety assessment. The safety biomarkers should also be chosen to demonstrate any effects on other similar structural anatomical elements, e.g. tracheal cartilage, intervertebral disc, and rib, to name a few.

3.6.1. USES

- To support other more generalized organ system safety indicators in preclinical and clinical trials;
- To monitor for local and systemic adverse effects both early and advanced;

- To set therapeutic dosages that do not impact on physiology.

3.6.2. CHALLENGES

- Understanding what 'safe' ranges are for joint tissue biomarkers;
- Safety biomarkers will need to be qualified against accepted clinical standards, including pain assessments, functional testing, and imaging;
- The safety threshold for each biomarker might be different across individuals.

3.7. Summary

With the BIPEDS scheme, the biomarkers that are likely to have the earliest beneficial impact on clinical trials fall into two general categories. The first are those that will allow us to target trials to subjects that are likely to either respond and/or progress within a short time frame. For instance, a patient population with high levels of an MMP-13 cleavage product, but without endstage cartilage loss, would be ideal for a trial with an MMP-13 inhibitor. The second category of biomarkers includes those that provide early feedback for preclinical decision-making and for trial organizers that a drug is having the desired biochemical effect. This category of biomarkers is particularly desirable in chronic diseases, such as OA, where clinical outcomes may take years to present [39]. In some cases, the biomarker might be sufficiently qualified that the researchers have confidence in using it to justify advancement to phase 2 trials and to determine a dosing schedule. These two categories reduce the burden and risk of early stage trials by delivering essential early information, making OA a more manageable and therefore a more attractive target for drug developers.

4. QUALIFICATION OF KNOWN OSTEOARTHRITIS BIOMARKERS

4.1. Biomarker Validation versus Qualification

The validation and qualification of a biomarker are two essential processes involved with assessing the level of confidence in a specific biomarker. For scientists who develop new biomarkers, validation means assessing all technical aspects of a specific assay to address the following question: "Under what conditions can we trust this assay and what it tells us?" Conversely, qualification consists of assessing the clinical value of a specific assay and answers the question: "Is this marker useful for learning more about the disease pathobiology or the efficacy of the treatment tested?" Currently there are no biomarkers that have been formally qualified and cleared by the FDA for OA-related outcomes.

4.1.1. Validation

Standard laboratory–based biomarker assays are typically quantitative in nature. Analytical validation of a specific quantitative assay is usually established by five tests: intra-and inter-assay variation, dilution recovery, determination of the detection and quantification limits and spiking recovery, although this latter test is often not performed, especially when standards are synthetic peptides. In addition, the stability of the biomarker (with storage and freeze-thaws) and key reagents should be established to determine the parameters and

stringency of storage necessary to assure reliability of measurements. The exact assay validation process will depend on the intended use of the assay, with assays for surrogate markers undergoing more rigorous validation than assays for exploratory endpoints. Not all biomarker assays are “definitive” quantitative measurements. Some biomarker assays generate “relative” results, due to the nature of the reference materials or sample matrix [39]. One example would be genomic data generated from microarray analysis of RNA. For these sorts of relative quantitative assays it is appropriate to place greater emphasis on relative and temporal changes in biomarker concentrations rather than the absolute concentrations. Another example would be an ELISA that uses a crude extract as standard and for which biomarker results are reported in arbitrary units. For these sorts of assays, the availability and sharing of a common international standard for normalization is highly desirable.

In contrast to quantitative biomarkers, qualitative biomarkers are discrete (discontinuous) and reported in either ordinal or nominal formats. An example of a qualitative assay would be a method to detect the presence of a single nucleotide polymorphism or gene mutation in a sample of DNA [39]. Assay validation for a qualitative assay is more limited than for a quantitative assay since concepts such as precision and dilutional recovery are not relevant [39]. Just as important as pre-study method validation is in-study validation (run acceptance), appropriate control samples and run/sample acceptance criteria should be incorporated into the analytical method for each assay to ensure quality data.

The specificity of the antibody(ies) used in the immunoassay is a very important factor, although this has not been carefully investigated for most biomarkers. Indeed, recognition and cross-reactivity experiments are usually performed using synthetic peptides or *in vitro* generated degradation fragments, which are probably of a different structure than the native immunoreactive forms detected in biological fluids. To date, published results of the structure of the immunoreactive form has only been partly determined for one OA-related biomarker, TIINE, which involves type II collagen cleavage by collagenase [40]. This information can be difficult to generate because the concentrations of the analytes found in serum and/or urine are usually very low and their determination requires complex analysis. This aspect of the biomarker validation process is however of critical importance for correct interpretation of biomarker results [41].

Other critical information is that which concerns the tissue and site(s) of origin of the biomarker. Incorrect assumptions regarding tissues of origin have been led to misinterpretation of biomarker data. Mistakes of this kind may in part account for lack of correlation between clinical and biomarker outcomes.

The STARD initiative [42] has provided a checklist of specific information about biomarker measurement, and the subjects tested, that should be provided in any study validating a biomarker regardless of its intended use. These include the following specific requirements (summarized by Felson et al [43]): to blind those measuring the biomarker as to disease status (in a study of prognosis, this would mean blinding to progression status); to define the rationale for and selection of cutoffs differentiating ‘normal’ from ‘abnormal’ biomarker levels; and importantly, to note the source of subjects in a study, reporting whether they were selected because of their biomarker status or unique clinical findings.

4.1.2. Qualification

Previously, the process of linking a surrogate endpoint to a clinical endpoint has been referred to as *validation* or *evaluation* [2]. However the use of the term validation has now been confined to the assessment of the performance characteristics of a biomarker assay, while linking a biomarker to a clinical endpoint is referred to as qualification [44]. The use of biomarkers as surrogate endpoints in a clinical trial requires the qualification of the biomarker for specific clinical endpoints (such as pain, loss of mobility, or need for a total joint replacement) in a specific population with a particular disease state and/or in the context of a specific class of therapeutic intervention (adapted from [2]). Loss of mobility and total joint replacement occur only after a very long time in most patients (Figure 2), and vary by nation and region due to differences in patient expectations and health-care policies. Consequently, to reduce the time needed to qualify a biomarker, studies use structural endpoints derived from x-ray and more recently from MRI.
For drug development, 'efficacy of intervention' ('E' of BIPEDS) biomarkers are sought. In theory, the optimal efficacy of intervention biomarker would be a perfect clinical outcome surrogate. In the case of the perfect surrogate:

- The effect of the intervention on the surrogate predicts the effect on the clinical outcome;
- The surrogate is in the only causal pathway of the disease process;
- The intervention's entire effect on the true clinical outcome is mediated through its effect on the surrogate;
- The surrogate fully captures the treatment effect.

In reality, it is likely that few if any biomarkers will ultimately achieve surrogate status let alone perfect surrogate status. Several different methods have been proposed for quantifying the strength of the surrogate [45]. This method provides a quantitative score for a biomarker. Wagner et al categorize the strength of a surrogate based on four levels [1]: Exploratory, Demonstration, Characterization and Surrogate biomarkers (summarized in 1.5.2). This mark of the strengths of surrogacy is used in this document.

As the Wagner classification implies, robust linkage of a biomarker with a clinical endpoint is not essential in early clinical development when the goal is confirmation of pharmacologic activity or optimization of dose regimens [2]. As stated by the Biomarkers Definitions Working Group in 2001: "Reliance on a biomarker early in the drug development process, for instance for candidate selection, entails the hazard that failure of a biomarker may lead to the elimination of potentially effective agents. On the other hand, substantial evidence that a biomarker will predict clinical benefit or risk is needed when use of the biomarker as a surrogate endpoint is proposed as the basis for regulatory approval. In this case, erroneous decisions based on invalid surrogate endpoints may have broad public health consequences" [2].

4.2. Qualification Endpoints for OA Biomarkers

As described above, there are many possible qualifying endpoints for an OA-related biomarker including signs (inflammation) and symptoms (pain), structure or functional outcomes in OA. A biomarker could be qualified for different stages of OA such molecular, preradiographic, or radiographic stages of OA. In theory, a biomarker could be qualified for an outcome in a specific joint if the biological findings supported such specificity. We are only beginning to appreciate cartilage matrix biochemistry in this level of detail as exemplified by the differences in matrix biochemistry and response to injury of ankle versus

knee cartilage [46]. In practice, the qualification process is an empiric and gradual one, correlating changes in a biomarker with change in state of a joint(s). To date the process of biomarker qualification has tended to relate a biomarker to a specific tissue component of the whole joint organ such as bone, cartilage or synovial tissue.

4.3. Sources of Biomarker Variability

Biochemical markers in blood and urine provide information on systemic skeletal tissue turnover [47] and are not necessarily specific for the alterations occurring in the signal joint [48]. For example, it has been shown that degenerative disease of the knees, hips, hands and lumbar discs contributed independently and additively to urinary CTX-II levels illustrating the total body contribution to systemic levels [48, 49]. The potential contribution of intervertebral discs is of particular relevance because disc degeneration is common in ageing. Systemic biomarker levels cannot be assumed to reflect total body OA burden based on radiographic damage or cartilage volume estimated by quantitative MRI because these factors alone do not fully account for the differential contribution of soluble biomarkers from different joints [50]. Serum and urinary levels of most markers also vary with gender, age, menopausal status, ethnicity, and OA risk factors such as body mass index. Specific examples include the effects of gender, ethnicity and age on COMP [51, 52] and the effect of BMI on CPII [53].

Biomarker levels can also be influenced by other skeletal alterations, such as osteoporosis or by concomitant medications. It is likely that differential processing by the liver or kidneys occurs before systemic biomarkers reach a steady state in body fluids, and this metabolism may not occur reproducibly in all patients, particularly in the presence of systemic disease [54, 55]. Measurements in urine require correction by creatinine to adjust for variability related to hydration and renal status. One of the main factors affecting pre-analytical variability is diurnal change. The magnitude of diurnal-related changes in the concentration of seven markers (serum HA, COMP, KS-5D4, TGFß1, CPII, and urinary CTX-II and C2C) has been shown to be greater than the analytical inter- and intra-assay related variability, indicating that the diurnal-related variation was predominantly a result of biological variability rather than assay variability [56, 57]. For the biomarkers found to be significantly associated with radiographic severity (serum COMP, KS-5D4, C2C, C1,2C, and urinary CTXII), the biomarker concentrations at the T2 or T3 time points showed the most consistent correlation with radiographic knee OA when the sampling was performed during the afternoon (T2) and the early evening (T3). A study on serum PIIANP and serum HELIX-II concluded that concentrations of these two markers increased significantly from T0 (before arising from bed) to T1 (1 hour after arising) [58]. It was also shown that serum CTX-I and serum HA markers levels are markedly influenced by food intake which also does increase intra-subject variability [59]. These and other data (prior biomarkers white paper http://www.oarsi.org/index2.cfm?section=OARSI_Initiatives&content=Biomarkers) provide a rationale for standardization of sample collection procedures for OA clinical trials.

Limited research has been done to analyze the effects of diet and dietary supplements on biomarker levels. As described above, serum hyaluronan showed significant variation related to food consumption in healthy volunteers [60] and circadian variation of CTX-I was found to be reduced by fasting [61], suggesting that fasting can have a significant effect on the circadian variation of markers of bone resorption. Gordon *et al* 2008 [57] showed that urinary CTX-II was not affected by food consumption or physical activity and may offer an

advantage in the context of clinical trials incorporating morning body fluid sampling. Clearly, pre-analytical factors contribute to intra- and inter-assay variability of biochemical markers levels and consequently need to be investigated and controlled as tightly as possible. Taken together, these studies point to the need for standardization of sample collection within a trial to minimize non-treatment related variation. Recommended methods of sample acquisition, handling and storage are provided in Appendix A.

4.4. Summary of OA Biomarkers

Biochemical markers of bone and cartilage turnover are presently the most advanced with respect to matrix remodeling [35]. Several excellent recent reviews provide a summary of biomarkers in general and several summarize the data to support classification into one or more of the particular BIPEDS categories [35, 62-68]. In this section, we focus on "soluble biomarkers" studied to date in human OA clinical trials, and not genetic/genomic or imaging biomarkers or biomarkers studied in the absence of an intervention. Although a few soluble biomarkers are quantified by mass spectroscopy approaches, most are currently assessed by immunoassay. Tables 1a-b provide a look at the sample sizes required for biomarker studies. Table 2 presents data for all known peer-reviewed publications to date of pharmacologic OA trials with either structural or clinical trial outcomes that included published biomarker analyses. OA-related biomarkers that have been studied to date in OA clinical trials, and an indication of the success or failure of the trial for the primary and biomarker outcomes. The reported assay coefficients of variation (CVs) are provided when they were reported, which may be helpful for assessing needed sample sizes for future studies. In addition, the reported concentrations (and standard deviations when available) before and after treatment are listed to begin to provide a benchmark for comparison across studies, albeit limited at the present time. Table 3 provides a summary of the known tissue sources and current BIPED classification for many of the most common and best-qualified OA-related biomarkers.

4.4.1. Statistical issues and sample size estimates for biomarker studies

Table 1a provides a look at the sample sizes required if the between-subject variability (standard deviation) increases from 1.5 to 2.0 or the power desired changes from 90% to 80% given the same treatment differences. Biomarkers are often not normally distributed due to the potential for a high incidence of values below the limit of quantification. To normalize the distribution the values are usually log-transformed and Table 1b provides some sample size estimates when the biomarker is expressed as ratio or percent differences and analyzed on the log scale. In the papers summarized in Table 2 (below) and others (reviewed by van Spil et al [62]), many biomarkers, such as those measured by radiography, e.g., JSN, were explored for their ability to predict the progression of OA or to change concurrently with OA. However, results were generally not consistent across the studies for multiple reasons: large variability of the assays, unpredictable variability of the biomarkers, under-powering of the study, or slow progression of OA were the most often cited reasons for non-significant or inconsistent findings.

The under-powering of the studies was generally due to the fact that the biomarkers were regarded as exploratory endpoints or the basis for subgroup analyses, hence, were not powered sufficiently at the planning stage. Some studies were designed as pilot studies, which relied on detecting statistical significance instead of meaningful difference as a measure of the importance of the biomarker. These types of studies serve the purpose of

hypothesis generation; however, as experiences with the biomarkers accumulate, an organized effort is necessary to define the following elements so that standards can be established for future studies against which to benchmark:

1. Identify clinically meaningful differences between two active treatments or between an active treatment and placebo with respect to validated clinical endpoints.
2. Define meaningful correlation between the biomarkers and the clinical endpoints, i.e., how large the magnitude of the correlation has to be.
3. Define the meaningful difference between two active treatments or between an active treatment and placebo with respect to the biomarker once it is demonstrated to correlate with the clinical endpoints.

Consideration of these three elements is important to ensure sufficient numbers of subjects in the study, and hence, sufficient power to detect the underlying meaningful difference based on biomarkers. They also prevent statistical significance being reached only because of the large sample size while meaningful difference is not observed. A critical component for the success of these aims will be the establishment of clinical meaningful endpoints related to imaging and symptom-related outcomes which serve as the qualifying endpoints for biomarker studies.

Le Graverand, et al 2006 [69] had also suggested the possibility that no single biomarker is sensitive enough to serve as a surrogate for radiographic outcomes in OA, but the combination of multiple biomarkers, representing different aspects of articular cartilage biochemistry, may significantly improve the detection and prediction of radiographic changes of knee OA. A natural extension of the three elements stated above, therefore, is to identify groups of biomarkers that are correlated with each other and that, in combination, have good predictive value for the progression of OA or change concurrently with radiographic outcomes.

Table 1a. Sample sizes to achieve 80 and 90% power to detect assumed differences between two parallel groups.

Number of Patients/Group Required for 90% Power†	Number of Patients/Group Required for 80% Power†	Underlying Treatment Difference to Detect	Standard Deviation (Between-Subject)	Effect Size (Difference/SD)
15	12	1.84	1.5	1.23
30	23	1.28	1.5	0.85
60	45	0.90	1.5	0.60
100	76	0.69	1.5	0.46
15	12	2.45	2.0	1.23
30	23	1.70	2.0	0.85
60	46	1.19	2.0	0.60
100	76	0.92	2.0	0.46

† Based on 2-sample T-Test (2-sided, alpha=0.05) for difference between groups with null hypothesis that treatment difference = 0.
SD = standard deviation

Table 1b. Sample sizes to achieve 80 and 90% power to detect assumed underlying ratio of treatment effect between two parallel groups.

Number of Patients/Group Required for 90% Power†	Number of Patients/Group Required for 80% Power†	Underlying Mean Ratio Between Groups to Detect	Coefficient of Variation (SD/Mean) in Original Scale
114	85	0.65	1.3
165	124	0.70	1.3
253	189	0.75	1.3
418	313	0.80	1.3
80	60	0.65	1.0
116	87	0.70	1.0
177	133	0.75	1.0
294	220	0.80	1.0
47	35	0.65	0.7
67	51	0.70	0.7
103	77	0.75	0.7
170	127	0.80	0.7

† Based on 2-sample T-Test (2-sided, alpha=0.05) for ratio of treatment effect between groups with null hypothesis that ratio=1, and common coefficient of variation.
SD = standard deviation

4.4.2. Summary of biomarker data generated in OA clinical trials.

Table 2. Summary of biomarker data generated in OA clinical trials to date.

						Treatment		Placebo		COMMENTS	
TRIAL-Intervention (duration)	**Study REF**	**Patient numbers**	**Sample Type**		**CV% (biomarker units)**	**pre**	**post**	**pre**	**post**		**Assay/Cut-points**
Ibuprofen 2400 mg qd for knee pain x 4-6 w (E)	Gineyts 2004 [70]	Human 156/45	NF morning urine	uCTX-II (E)	<10% (ng/mmol Cr)	225 ± 2.16	229 ± 2.06	226 ± 1.88	265 + 2.06	Patients with high levels were responsive to therapy	C- ELISA Cartilaps-Christgau 2001 [71]
				uGlc-Gal-PYD (∅E)	<11% (nmol/mmol Cr)	6.0 ± 1.5	6.2 ± 1.5	5.7 ± 1.4	6.3 ± 1.4		HPLC Gineyts 2001 [72]
Glucosamine sulphate 1500mg/d x 3y (E)	Christgau 2004 [18]	Human 106/106 [n=61 above 1SD cut-off]	NF 2nd morning void urine	uCTX-II/Cr (E)	8.4% (ng/mmol)	All: 216.5 ± 9 at baseline	All: Loss of joint space (0.06 mm) over 3 years	All: 219.5 ± 9 at baseline	Loss of joint space (0.31 mm) over 3 years		C-ELISA Cartilaps with mAb F46 per Christgau 2001 [71]; High turnover group defined as baseline ≥ 261.3 (i.e. ≥1 SD above mean of 169.1 ± 92.3 in reference population)
						High turnover group mean 413 ± 28	High turnover group mean 336 ± 26 [Gain of joint space (0.083 mm; p=0.07) over 3 years; Global WOMAC decreased 24.5%]	High Turnover group mean 375 ± 33	High Turnover group mean 411 ± 252 [Loss of joint space (0.44 mm) over 3 years; Global WOMAC decreased 4.5%]	Promising approach; larger sample size (>61) of high turnover patients likely needed for statistical significance	
						For high turnover group: Change in uCTXII from baseline to 12m correlated with average joint space width loss over 3 years (r=0.43; p<0.05)		For high turnover group: Change in uCTXII from baseline to 12m correlated with average joint space width loss over 3 years (r=0.27; p=0.03)			
Salmon calcitonin (oral) 0.5-1.0mg/d x 48d for knee OA patients with positive knee bone scans (E)	Manicourt 2006 [73]	Human 27/14	F serum & 2nd morning void between 9-11 AM [all median values reported-show baseline and	uCTX-II/Cr (E)	<6% (ng/mM)	395	290	368	370		ELISA Cartilaps-Nordic Bioscience (Herlev, Denmark)
				sHA (E)	<6% (µg/ml)	61	48	60	69		ELISA Method of Manicourt 1999
				sC2C (E)	<6% (ng/ml	30	23	27	30		ELISA-IBEX (Montreal, CA)
				uNTX-I/Cr (∅E)	<4% (BCE mM/mM Cr)	48	43	57	56		ELISA-OSTEX Intl (Seattle, WA)

			day 84 values]	sOC (∅E)	<9% (ng/ml)	12	16	18	16		ELISA BioSource, (Nivelles, Belgium)
				sMMP-1 (∅E)	<8% (ng/ml)	8	9	8	9		ELISA-GE Healthcare (Little Chalfont, UK)
				sMMP-3 (E)	<5% (ng/ml)	20	19	19	24		ELISA-GE Healthcare (Little Chalfont, UK)
				sMMP-8 (∅E)	<5%	5	5	4	4		ELISA-GE Healthcare (Little Chalfont, UK)
				sMMP-13 (E)	<5% (pg/ml)	100	64	52	76		ELISA-GE Healthcare (Little Chalfont, UK)
				sTIMP-1 (∅E)	<5%	173	184	151	149		ELISA-GE Healthcare (Little Chalfont, UK)
				TIMP-2 (∅E)	<5%	11	11	18	14		ELISA-GE Healthcare (Little Chalfont, UK)
BRISK study: Risedronate 5mg/d or 15mg/d x 12m (∅E for JSN, E for WOMAC)	Spector 2005 [74]	Human	F early morning urine and serum	uCTX-II	(ng/mmol Cr) (E)	340.1 (24.0)	-22.8 ± 5.35% (15mg)	312.5 (19.9)	+14.5 ± 5.4% (15mg)		ELISA Cartilaps-Nordic Bioscience (Herlev, Denmark)
				uNTX-1	(nmol/mmol Cr) (E)	38.6 (2.2)	-32.9 ± 4.92% (15mg)	40.3 (2.8	+17.2 ± 4.9% (15mg)		ELISA Osteomark-Orthoclinical Diagnostics (High Wycombe, Bucks, UK)
				sAlk Phos (bone specific)	NR	NR	-29.1 ± 2.6% (15mg)	NR	-2.7 ± 2.5% (15mg)		ELISA Ostase-Beckman-Coulter (San Diego, USA)
KOSTAR study: Risedronate 5mg/d, 15mg/d, or 35-50mg/w x 24m (∅E)	Bingham 2006 [75]	Human 1861/622 (from two cohorts)	F 2nd morning void	uCTX-II/Cr (E)	<10% (ng/nmole Cr)	297.16-360.70 ± 14.87-12.06	(-) 17.9-19.6% (decrease at 24m)	296.47-376.72 ± 17.09-13.72	(+) 10.1-26.3% (increase at 24m)	Treatment effect on biomarkers but not x-ray progression	ELISA Cartilaps-Nordic Bioscience (Herlev, Denmark)
				uNTX-I/Cr (E)	<10% (nmol BCE/ nmole Cr)	38.80-49.91 ± 1.07-2.10	(-) 39.2-41.7% (decrease at 24m)	37.48-49.43 ± 1.96-1.36	(+) 3.0-7.3% (increase at 24m)		Osteomark-OrthoClinical Diagnostics (Rochester, NY)
	Garnero 2008 [19]	Human 1885 (subset of two cohorts)	F early morning urine	uCTX-II/Cr	<10% (ng/mmol Cr)	(-) 39.9 ± 3.0 (treatment effect of biomarker shown for all doses by a mean decrease from baseline to 6m, p<0.05 compared to baseline and placebo [baseline CTX-II and change from baseline to 6m associated with radiographic progression at 24m as absolute change or for progression defined as JSN≥0.6 mm)				Early biomarker endpoint to predict long-term progressor/non-progressor status	ELISA Cartilaps-Nordic Bioscience (Herlev, Denmark); high turnover defined as > 150 ng/mmol Cr
				uNTX-I/Cr	<10% (nmol/mmol	(values reported graphically)					ELISA Osteomark-OrthoClinical Diagnostics

					Cr)						(Rochester, NY)
Chondroitin sulfate 500mg bid x 24w (E)	Mazieres 2007 [76]	Human 139/140	F serum between 7:30-10 AM & 2nd morning void urine	uCTX-II/Cr (∅E)	<15% (ng/mmol Cr)	389 ± 247	406 ± 302	375 ± 238	376 ± 214		ELISA Cartilaps-Nordic Bioscience (Herlev, Denmark)
				sHA (∅E)	<9% (ng/ml)	86 ± 71	100 ± 86	79 ± 61	89 ± 78		ELISA HA-Corgenix, (CA, USA)
				sCTX-I (∅E)	<10% (ng/ml)	0.44 ± 0.27	0.44 ± 0.23	0.39 ± 0.22	0.40 ± 0.22		Automated analyzer Elecsys 2010-Roche (Mannheim, Germany)
Acute activity (∅E)	Andersson 2006 [77]	Human 29/29	NF serum twice with 1 hour apart (after activity and after rest)	sCOMP (E)	NR (U/L)	11.03	(+) 1.3 [median change score after 1h activity]	11.29	(-) 0.6 [median change score after 1h rest]		S-ELISA COMP-Anamar (Uppsala, Sweden)
ADAPT: Exercise and/or diet x 18m (E)	Chua 2008 [78]	Human 138/53 (193 studied)	F between 7-9 AM	sCOMP (∅E) (baseline values reported graphically)	NR (U/L)	10.80 ± 0.49 (at 6m diet & exercise)	11.81 ± 0.46 (at 18m diet & exercise)	11.75 ± 0.45 (at 6m)	11.72 ± 0.42 (at 18m)		S-ELISA COMP-Anamar (Uppsala, Sweden)
				sHA (∅E) (baseline values reported graphically)	NR (ng/ml)	42.28 ± 3.79 (at 6m diet & exercise)	45.33 ± 3.63 (at 18m diet & exercise)	40.46 ± 3.58 (at 6m)	47.67 ± 3.35 (at 18m)		Immunosorbent assay Li 1989 [79]
				sKS (∅E) (baseline values reported graphically)	NR (ng/ml)	310.22 ± 7.62 (at 6m diet & exercise)	310.93 ± 7.32 (at 18m diet & exercise)	308.67 ± 7.17 (at 6m)	286.66 ± 6.71 (at 18m)		ELISA with mAb 5-D-4 per Method of Thonar 1985 [80]
				sTGF-ß1 (∅E) (baseline values reported graphically)	NR (ng/ml)	38.89 ± 1.14 (at 6m diet & exercise)	39.06 ± 1.07 (at 18m diet & exercise)	40.93 ± 1.04 (at 6m)	39.41 ± 0.98 (at 18m)		ELISA-Quantikine R&D (Minn, USA)
Glucosamine sulfate discontinuation x 6m (OE)	Cibere 2005 [81]	Human 63-65/63-65	NF urine or serum	sC2C (∅E)	5.5% (pmol/ml)	Mean change: -3.5 ± 28.5		Mean change: 3.7 ± 23.6			ELISA-IBEX per Method of Poole 2004 [82]
				uC2C/Cr (∅E)	NR (pmol/µmol Cr)	Mean change: -6.9 ± 54.1		Mean change: -0.6 ± 11.8			
				sC1,2C (∅E)	NR (pmol/ml)	Mean change: 8.5 ± 64.2		Mean change: 9.5 ± 80.0			ELISA-IBEX with pAb per Method of Billinghurst 1997 using pAb [83]
				uC1,2C/Cr (∅E)	NR (pmol/µmol Cr	Mean change: -20.2 ± 144.9		Mean change: 0.4 ± 17.1			
MMP inhibitor, variable dosagex3w prior to knee replacement (∅E)	Leff et al, 2003 [84]	Human 22/11	F articular cartilage at arthroplasty	cCPII (∅E)	NR (ng/µg DNA)		2.35 (1.07-8.34) [median and range for max dose]		1.42 (0.33-3.86) [median and range]		Method of Nelson 1998 [85]
				cC1,2C (∅E)	NR (pmol/µg DNA)		29.8 (9.3-134) [median		25.7 (5.1-45.7)		ELISA Billinghurst

							and range for max dose]		[median and range]		1997 [83]
				cCol2-3/4m (∅E)	NR-5% per RP (nmol/µg DNA)		0.21 (0.09-0.64) [median and range for max dose]		0.17 (0.06-1.25) [median and range]		ELISA Hollander 1994 [86]
				cCS-846 (E)	NR (µg/µg DNA)		0.78 (0.18-5.62) [median and range for max dose]		0.35 (0.24-2.86) [median and range]		ELISA Rizkalla 1992 [87]
				cKS (∅E)	NR (µg/µg DNA)		66.0 (28.7-258) [median and range for max dose]		90.7 (24.1-177) [median and range]		ELISA Rizkalla 1992 [87]
Doxycycline x 30m (E increased)	Lohmander 2005 [88]	Human 60/60 subset of main study [21/39 progressors; 30/30 non-progressors]	NF plasma and 2nd morning void urine	pMMP-3 (E)	19.4% ng/ml	Contrary to placebo group – every SD increase in mean MMP-3 was associated with lower rate of JSN (-0.11 mm)		Baseline upper tertile (11.86-41.00) more likely to progress than lower tertile (<6.43); for every SD (4.6 ng/ml) increase in mean MMP-3 - JSN increased 0.18 mm (p=0.001); increase over time in MMP-3 associated with concurrent JSN			ELISA Method of Walakovits 1992 [89]
	Mazzuca 2006 [90]			uCTX-II (∅E)	27.7% (ng/mg Cr)	No association between uCTX-II and JSN progression; Mean values 63.5-66.8; change from baseline at 30m (mean ± SEM): doxycycline group=1.14±1.93; placebo group 0.53±1.75; progressors -0.03±1.88; non-progressors 1.69±1.78				Study designed for 80% power to detect 35% difference between highest and lowest tertiles of baseline uCTX-II in frequency of JSN progression	ELISA with mAb 2B4 and plates coated with matrilysin digested type II collagen (different antibody from Cartilaps assay)
	Mazzuca 2006 [91]			sC2C, sCPII, sCS846, sC1,2C (∅E)	9.7%, 6.4%, 11.5%, 10% respectively (all ng/ml)	1SD change in CS846 associated with concurrent JSN; no biomarker was significant predictor of JSN progression					ELISAs-IBEX (Montreal, CA)
	Le Graverand 2006 [69]			uTIINE/Cr (E)	Up to 12.3% (ng/mM Cr)	1SD (64-68 ng/mM Cr) increase in baseline uTIINE associated with lower rate of JSN (not significant in either group)					Two dimensional LC-MS/MS
	Otterness 2007 [92]	Human 51/69 (subset)		uTIINE/uCr (∅E – increased with treatment)	8% interassay, 30 ± 17% within patient (ng/mmole)	109 ± 68	144 ± 81 ng/mmole	125 ± 62 (overall baseline mean)	115 ± 49		Two dimensional LC-MS/MS; increase due to treatment due

											possibly to decreased fragment metabolism or change in clearance
Chondroitin sulfate x 1y (E)	Uebelhart 1998 [93]	Human 21/19 (21/20 for sOC but 23/23 overall)	NF serum & 2nd morning void urine	sKS (E)	NR (ng/ml)	449 ± 119	420 ± 100	386 ± 133	403 ± 142		C-ELISA with mAb 1/20/5-D-4 Method of Thonar 1984 [80]
				uPYD/Cr (E)	NR (nmol/L/mmol	56 ± 25	53 ± 19	59 ± 40	70 ± 30		RP-HPLC Uebelhart 1990 [94]
				uDPD/Cr (E)	NR (nmol/L/mmol)	7.7 ± 3.0	7.7 ± 2.3	8.5 ± 5.4	11.7 ± 8.1		RP-HPLC Uebelhart 1990 [94]
				sOC (E)	NR (ng/ml)	16 ± 7	16 ± 6	21 ±13	26 ±29		RIA, ELSA-OSTEO CisBiointernational, Gif/Yvette, France
Intra-articular hyaluronan x 5 weekly injections	Hasegawa 2008 [95]	Human 28 (all treated)	SF time not specified	sfKS (E)	NR (nmol/ml)	61.2 ± 35.8	52.8 ± 25.3			ND-no vehicle control	HPLC Method of Yamada 2000 [96]
				sfC6S (E)	NR (nmol/ml)	19.1 ± 6.7	17.8 ± 6.1			ND-no vehicle control	HPLC Method of Yoshida 1989 [97] & Shinmei 1992 [98]
				sfC4S (E)	NR (µg/ml)	6.1 ± 3.7	5.2 ± 2.9			ND-no vehicle control	
				sfTenascin-C (∅E)	NR (ng/ml)	37.4 ± 59.1	39.0 ± 58.1			ND-no vehicle control	ELISA – IBA (Gunma, Japan)
Supplemental soy protein 40g/d x 3m (E)	Arjmandi 2004 [99]	Human 44/44	F serum	YKL-40 (E in men)	6.8% (ng/ml)	All: 89.9 ± 7.6; men:91.0 ± 10.3; women: 93.4 ± 11.4	Change (decrease) in YKL-40 from baseline to 3 months only significant in men (compared to placebo)	All: 67.8 ± 6.3 men: 71.3 ± 10.2; women: 64.6 ± 7.8	Increased in all groups	Clinical and biomarker effects in men, not women	S-ELISA-Metra Biosystems (Mountain View, CA)
				IGF-1 (E in men)	7.6% (ng/ml)	All: 113.3 ± 8.2; men: 125.0 ± 10.7; women: 97.6 ± 12.9	Change (increase) in IGF-1 from baseline to 3 months only significant overall and	All: 135.6 ± 10.6; men: 158.7 ± 14.8; women: 107.9 ± 9.3	Increased in all groups	Clinical and biomarker effects in men, not women	Radioimmunoassay -Diagnostic Systems Labs Inc (Webster, TX)

							in men, not women (compared to placebo)				

F=fasting; NF=non-fasting; h=hour; d=day; bid=twice daily; w=week; m=month; y=year; pAb=polyclonal antibody; LC-MS/MS=liquid chromatography followed by low then high energy mass spectroscopy; RP=reversed-phase; HPLC=high pressure liquid chromatography; sf=synovial fluid; s=serum; p=plasma; u=urine; c=cartilage; NR=not reported.
S-ELISA = sandwich Enzyme Linked Immunosorbent Assay
C-ELISA = competitive (inhibition) Enzyme Linked Immunosorbent Assay

(E) means showed evidence for change with intervention and (∅E) means no evidence of statistical difference of biomarker with intervention (did not meet efficacy of intervention criteria); when the trial produced disease modification an (E) is listed in the first column.

CTX-II=C-telopeptide of type II collagen; COMP=cartilage oligomeric matrix protein; HA=hyaluronan; C2C= collagenase generated neoepitope of type II collagen; C1,2C= collagenase generated neoepitope of types I and II collagens ; TIINE (mAbs 9A4/5109); CPII/PIICP=type II collagen propeptide; Col2-3/4m= type II collagen denaturation epitope; KS-keratan sulfate; CS-846 = aggrecan chondroitin sulfate epitope; NTX-I=N-telopeptide of type I collagen; CTX-I=C-telopeptide of type I collagen; PYD=pyridinoline; DPD=deoxy-pyridinoline; OC=osteocalcin; Glc-Gal-PYD= glucosyl-galactosyl-pyridinoline; C4S and C6S=chondroitin-4 and -6 sulfate; Tenascin-C; YKL-40=human cartilage glycoprotein 39; IGF-1=insulin growth factor-1; MMP=metalloproteinases: -1 (collagenase-1), -3 (stromelysin), -8 (neutrophil collagenase), -13 (collagenase-3); TIMP=tissue inhibitor of metalloproteinase: -1 or -2; TGF-ß1=transforming growth factor-ß1; Alk Phos=alkaline phosphatase

A few details are worth noting regarding the use of biomarkers in published clinical trails. In 137 individuals with knee OA, no significant difference was seen between patients in the placebo or glucosamine sulfate treated groups with respect to the ratio of markers of collagen type II breakdown (lnC1,2C/C2C) in serum or in urine [81]. This study used flare / no flare status as the clinical endpoint. In a study of 201 patients with inflammatory knee OA, a decrease in the levels of urinary Glc-Gal-PYD was observed following treatment with the nonsteroidal anti-inflammatory drug ibuprofen but not with placebo [70]. Finally, a group of 35 patients with OA were randomly selected to receive a potent inhibitor of MMP-3 (BAY 12-9566) or placebo. Levels of the aggrecan 846 epitope were higher in the treated group compared to the control group implying that aggrecan synthesis improved [84]. This study used an original protocol, measuring markers directly in the cartilage samples obtained at the time of surgery 3 weeks after the start of treatment. The advantage of such an approach is direct analysis of cartilage, short duration of treatment, and small numbers of patients.

Most past studies have used structural and/or clinical endpoints to investigate the usefulness of a biomarker. A 30 months study of a subset (60 progressors and 60 non-progressors) of the patients in a clinical trial assessed by radiography (progression limit: JSN$\geq$ 0.33 mm) showed a reduction of JSN in the doxycycline treatment group but a paradoxical increase in uTIINE with treatment [92]. In a study testing the effects of risedronate on 1885 patients suffering from knee OA, CTX-II levels decreased with risedronate in patients with knee OA although there were no differences in the traditional joint space narrowing radiologic outcome or in symptoms in response to the treatment. There was however a dose-related preservation and improvement in tibial subchondral bone architecture [100] with treatment. The utilization of more sensitive imaging methods such as MRI, in future clinical trials, may clarify and resolve such apparent inconsistencies, providing a way forward for biomarkers qualification. In another study, CTX-II levels reached at 6 months were associated with radiological progression at 24 months [19], defined *a priori* as a JSN of $\geq$ 0.6 mm from baseline which corresponds to 3 times the SD of the X-ray measurement method for joint space [19]. Three clinical trials in OA [18, 19, 88] have used baseline levels, or early change in a biomarker of a biomarker (CTX-II, MMP-3), to predict subsequent progression of radiological damage. These studies demonstrate the advantages of selecting a high matrix turnover / progressor patient population for trial inclusion.

Other studies evaluated biomarkers with patient centered (self-reported) clinical endpoints. A study with 53 patients receiving oral calcitonin daily for 84 days showed that CTX-II, C2C, and MMP-13 levels were decreased in the group of patients receiving 1 mg/day of calcitonin. The efficacy of the treatment was evaluated by Lequesne's index [73]. A small Japanese study with 28 patients with knee OA evaluated the effects of repeated injections of hyaluronan and showed a significant reduction of C6S, C4S and KS relative to baseline. However, a vehicle treated control group was not evaluated and it would be important to rule out changes in biomarkers due to synovial fluid aspiration alone. The effect of this treatment was evaluated by change in knee pain assessed by a visual analog scale (VAS) [95].

4.4.3. Level of qualification of OA-related biomarkers.

The following, Tables 3 and 4, summarize OA-related biomarkers used to date in human clinical trials described in Table 2 and/or commercially available. Specifically, Table 3 lists commercially available biomarkers currently recommended as a panel for study in past and future clinical trials (discussed in section 6.3), and Table 4 lists other OA-related biomarkers

qualified for various OA outcomes. The BIPEDS classifications are based on studies in which the biomarker showed a statistically significant difference for a clinical or structural outcome as summarized primarily by per van Spil [62] but also by Cibere 2009 [15]; Conrozier 2008 [101]; and Kraus 2010 [50]. The Surrogate classification is restricted to results based on current published human clinical trials only. These designations could be further refined by a consideration of preclinical results and unpublished results if a repository of this knowledge existed as called for in the recommendations of this document.

Table 3. Recommended panel of informative commercially available OA-related biomarkers qualified for various OA outcomes. +					
Biomarker	**Process* (preliminary)**	**Tissues of Origin (see discussion below Table)**	**BIPEDS Classifications***	**Surrogacy Based on Human Clinical Trials (preliminary)**	**ELISA assay type**
urinary CTX-II	type II collagen degradation, osteophyte burden of disease	mineralized and non-mineralized cartilage, growth plate cartilage, bone	Knee: BPED Hip: BPD	characterization: changed significantly in 3 pharmacologic trials that met primary clinical endpoints [18, 70, 73]	competitive-inhibition for human urinary samples and sandwich for animal serum samples
serum COMP	cartilage degeneration	cartilage > tendon, meniscus, synovium, osteoblasts, arterial wall	Knee: BPD Hip: BPD	exploration: not used to date in published pharmacologic trial	competitive-inhibition & sandwich
serum HA	osteophyte burden of disease and synovitis	cartilage, meniscus, synovium and ubiquitous in body	Knee: BPED Hip: P	demonstration: changed significantly in one pharmacologic trial that met primary clinical endpoints [73]	sandwich protein binding assay
serum and urine C1,2C	Types I and II collagen degradation	cartilage, bone, synovium,etc.	Knee: D(u) Hip: none	exploration: nonsignificant change in one pharmacologic trial that met primary clinical endpoint [73, 91]	competitive-inhibition
serum and urine C2C	type II collagen degradation	cartilage	Knee: E(s), D(u) Hip: B(s)	demonstration: changed significantly in one pharmacologic trial meeting primary clinical endpoints [73]	competitive-inhibition
serum and urine Coll2-1 and Coll2-1NO2	type II collagen degradation	cartilage	Knee: D(s),B(u),P(u) Hip: D(s)	exploration: not used to date in published pharmacologic trial	competitive-inhibition
serum CPII or PIICP	type II collagen synthesis	cartilage	Knee: D(s) Hip: B(s)	exploration: nonsignificant change in one pharmacologic trial that met primary clinical endpoint [91]	competitive-inhibition
PIIANP	Type II collagen synthesis	cartilage	Knee: BPD Hip: none	exploration: not used to date in published pharmacologic trial	competitive-inhibition
urine/serum NTX-1	bone resorption	bone turnover	Knee: P(u),E(u) Hip: P(s)	demonstration: changed significantly in one pharmacologic trial that met primary clinical (WOMAC) endpoint [74]	competitive-inhibition
urine/serum CTX-1	bone resorption	bone turnover	Knee: B(u), D(s/u), P(u) Hip: none	exploration: not used to date in published pharmacologic trial	competitive-inhibition
serum CS846	cartilage aggrecan	cartilage	Knee: P Hip: none	exploration: nonsignificant change in one pharmacologic trial that met primary clinical endpoint [91] but changed associated with concurrent JSN	competitive-inhibition

	synthesis /turnover				
serum MMP-3	Protease stromelysis involved with joint tissue degradation and inflammation	synovium, cartilage	Knee: E Hip: none	characterization: changed significantly in two pharmacologic trials that met primary clinical endpoints [73, 88]	sandwich for total MMP-3 assay

+This list does not include many emerging biomarkers that may prove useful in the future nor cytokines and chemokines that are also worthy of consideration. *These are general recognized processes for which these biomarkers are known. This is very preliminary information at this time and should not be considered definitive but rather in evolution. This information is derived from van Spil [62]; Cibere 2009 [15]; Conrozier 2008 [101]; Kraus 2010 [50]. References in Table as follows: [18, 70, 73, 74, 88, 90]. Table 1 abbreviations: CTX-II=carboxy-telopeptide of type II collagen; COMP=cartilage oligomeric matrix protein; HA=hyaluronan; C1,2C=collagenase-generated neoepitope of types I and II collagen collagenase; C2C= collagenase-generated neoepitope of type II collagen; Col2-3/4m= type II collagen denaturation epitope; CPII/PIICP=type II procollagen carboxy-propeptide; PIIANP=type IIA procollagen amino propeptide; NTX-I=N-telopeptide of type I collagen; CTX-I=carboxy-telopeptide of type I collagen; CS-846=aggrecan chrondroitin sulfate 846 epitope; MMP=metalloproteinases-3 (stromelysin).

Table 4. Other OA-related biomarkers qualified for various OA outcomes.

Biomarker	Process (preliminary)	Tissues of Origin (see discussion below Table)	BIPEDS Classifications	Surrogacy Based on Human Clinical Trials (preliminary)	ELISA assay type
serum KS	Cartilage catabolism, aggrecan	cartilage	Knee: BPED Hip: none	Demonstration: changed significantly in one pharmacologic trial meeting primary clinical endpoints [93]	competitive-inhibition (not commercially available)
serum YKL-40	Catabolic;	macrophages,	Knee: BE	Demonstration: changed significantly in men one	(not commercially

	macrophages, cartilage, synovium, cells of epithelial origin	cartilage, synovium, cells of epithelial origin	Hip: D	pharmacologic trial meeting primary clinical endpoints [99]	available)
urinary TIINE	Cartilage catabolicsm type II collagen	cartilage	Knee: BP Hip: none	Exploration: paradoxical response [102]	(not commercially available)
serum OC	Anabolic bone turnover	bone	Knee: BPED Hip: none	Demonstration: changed significantly in one pharmacologic trial meeting primary clinical endpoints [93]	ELISA
urinary Glc-Gal-PYD	catabolic synovium	synovium	Knee: BD Hip: none	Exploration: insignificant change in one pharmacologic trial meeting primary clinical endpoints [70]	HPLC
urinary PYD	Catabolic bone turnover	bone	Knee: BED Hip: none	Demonstration: changed significantly in one pharmacologic trial meeting primary clinical endpoints [93]	HPLC
urinary DPD	Catabolic bone turnover	bone	Knee: BED Hip: none	Demonstration: changed significantly in one pharmacologic trial meeting primary clinical endpoints [93]	HPLC
MMP-13	protease	synovium, cartilage	Knee: E Hip: none	Demonstration: changed significantly in one pharmacologic trial meeting primary clinical endpoints [73]	sandwich for total MMP-13 assay

Table 2 abbreviations: KS-keratan sulfate; YKL-40=human cartilage glycoprotein 39; uTIINE (mAbs 9A4/5109) urinary type II collagen collagenase-generated neoepitope; OC=osteocalcin; Glc-Gal-PYD= glucosyl-galactosyl-pyridinoline; MMP=matrix metalloproteinases:-13 (collagenase-3); PYD=pyridinoline; DPD=deoxy-pyridinoline.

Although type II collagen is an attractive candidate marker of cartilage degradation, it can be difficult to precisely identify the principle tissue sources of a biomarker and the source within a tissue such as articular cartilage which is composed of both calcified (adjacent to subchondral bone) and non-calcified regions. A case in point is represented by the biomarker CTX-II, the most widely tested OA-related biomarker to date. The CTX-II assay exists in two forms: a sandwich ELISA used for animal serum samples that likely recognizes a dimeric form of the EKGPDP epitope; and a competitive ELISA used for human and animal urine samples that likely recognizes monomeric and dimeric forms of the EKGPDP collagen II telopeptide [67]. Unlike the collagen epitope urinary TIINE [40], the exact nature of the immunoreactive cleavage products in urine has not been reported for CTX-II. EKGPDP is released from denatured human type II collagen upon enzymatic digestion with matrilysin, and MMPs -3, -8, and -13 [103], and in another study from cartilage sections by enzymatic digestions with MMPs-1, -3, -7, -9, and -13 and cathepsin B [104]. CTX-II immunoreactive epitope can also be released *in vitro* from non-mineralized bovine articular cartilage treated with oncostatin M and TNFalpha and its release can be blocked by estrogen [105]. In young animals and skeletally immature humans, a significant amount of this epitope originates from growth plate cartilage [106-108]. In adult human osteoarthritic cartilage CTX-II immunostaining is in uncalcified fibrillated cartilage as well as calcified articular cartilage [109].

Further complicating the interpretation of collagen type II fragment origins, are the many sites where type II collagen is found in skeletally mature adults, including: articular cartilage, fibrocartilage (intervertebral disc, menisci), respiratory tract cartilage, rib cartilage, insertion sites of tendons and ligaments into bone, and to a small extent, in the ear and eye [63]. However, as pointed out by Lohmander and Eyre, type II collagen makes up only ~1% of all collagen in the body but the normal turnover is low suggesting that pathological turnover from a single joint might be expected to raise the systemic level of fragments significantly [63].

Finally, CTX-II urine levels are very low in individuals with pycnodysostosis compared with age-matched controls [63]. Pycnodysostosis is a lysosomal storage disease of the bone caused by mutation of the gene encoding the enzyme cathepsin K, a cysteine protease expressed by osteoclasts and a major protease involved in bone resorption. In pycnodysostosis (OMIM #265800), osteoclasts function normally in demineralizing bone, but do not adequately degrade the organic matrix. This finding has suggested that a major source of CTX-II is the breakdown and remodeling of mineralized cartilage collagen by osteoclasts [63, 105]. In fact, by immunohistochemistry, the EKGPDP epitope is localized in calcified articular cartilage, at the interface between the calcified cartilage and bone, and to some extent at the surface of non-mineralized cartilage lesions, as well as subchondral bone (in a rat model of OA) [109] [105]. Osteophyte formation and remodeling may thus also be a significant source of CTX-II since, like the growth plate, this also involves endochondral ossification and is a fundamental feature of joint degeneration in OA. Urinary CTXII has in fact been shown to correlate with total body burden of osteophyte [50].

In summary, and as illustrated here for the most reported OA biomarker CTX-II, the complexities in structure, the paucity of evidence on tissue origins, and the incompletely understood catabolic, clearance, and regulatory pathways currently make it difficult to be certain of the principal sites of origin of OA-related biomarkers. This serves to illustrate how critically important it is to understand as much as we can about each of these biomarkers from *in vivo* and

in vitro analyses in order to be able to more precisely and correctly interpret biomarker data in preclinical and clinical drug development and assessment.

4.4.4. Summary related to use of biomarkers in clinical trials

There have been few published clinical trials reporting biomarker results. The lack of medications with established chondroprotective activity has limited the availability of clinical trial samples in which to test the utility of biomarkers.

In many cases, especially involving preclinical and clinical trials, biomarker results may not be reported or are not reported in a systematic and standardized manner. So it is difficult to utilize published data from trials to power future trials or to draw conclusions by comparing across studies. Recommendations regarding standardization and access to body fluids can be found at the end of this document.

Of those clinical trials reporting biomarker results, relatively few biomarkers have been tested, often using different methodologies, and very few trials and studies have tested multiple biomarkers in the same samples. Only recently have a variety of biomarkers started to be examined head to head in the same studies [15].

Many promising OA-related biomarkers have never or rarely been tested in clinical trial samples. Existing clinical trials have not used standardized methods of sample collection and assay methods differ among studies for many of the biomarkers tested.

5. PATHWAYS FOR BIOMARKER QUALIFICATION

The increased use of biomarkers is viewed as a critical component in improving the traditional inefficiency of the OA drug development process. Biomarkers can be used in a variety of ways from drug target development in preclinical studies to surrogate endpoints for regulatory approval. How biomarkers are used also defines the level of qualification required.

As described in section 1.2, a biomarker may be defined as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” [2]. The definition has two key components; the measurement of the biomarker and its evaluation as an indicator of some biological process(es). Consequently, any discussion of biomarker qualification must include both characterization of the source of the biomarker, the analytical capabilities of the test used to quantify the biomarker, as well as the evaluation (i.e. qualification) of the association between the biomarker and the pathobiological state and/or clinical outcome.

In general, companies are struggling with defining and developing a process for what *in vitro* (soluble) biomarker data to include in regulatory submissions [39]. Results intended to influence the course of the clinical development process would be considered part of the safety and efficacy evaluation and would need to be part of the regulatory submission [39]. Biomarker data that do not have such a regulatory impact would not need to be part of the regulatory submission. This section describes some of the considerations related to the biomarker qualification process and pathway.

5.1. Assessment of Analytical Capabilities of a Biomarker Test

Analytical validation of a biomarker follows a different pathway from that of validation of a drug. Validation of analytical methods related to the drug itself is a well-defined process. Regulatory agencies require that critical parameters of tests performed to assess the material conform with current Good Manufacturing Practice [110]. The International Conference on Harmonization (ICH) has published detailed guidelines on the validation of analytical procedures included as part of a registration application [111]. In the case of a drug or biologic development process, analytical methods are considered acceptably validated if the assays perform in a manner that demonstrates that the drug substance or drug product has the appropriate identity, strength, quality, and purity. There is an expectation by regulatory authorities that the analytical capabilities of critical test methods will be enhanced during the drug development process and that the methods are fully validated at the time of the market application.

The contrasting process of biomarker validation was described in section 4.1.1 and entails assessment of the accuracy, precision, specificity (what process is it measuring and in which tissue(s), detection limit, quantification limit, linearity and range. In 1988, Congress passed the Clinical Laboratory Improvement Amendments (CLIA) establishing quality standards for all laboratory testing to ensure the accuracy, reliability and timeliness of patient test results regardless of where the test is performed.

5.2. Biomarker Qualification - Association of a Biomarker and a Clinical Outcome

As mentioned earlier, qualification and validation have different meanings; biomarker qualification consists of the process of associating a biomarker with a clinical outcome or biological parameter. Biomarker qualification processes are in a state of evolution but a recent document outlines the current pilot pathway and regulatory agencies involved [44].

5.2.1. Levels of qualification of biomarkers for drug development use

Biomarkers for drug development use can be divided into four categories according to the degree or level that the biomarker can be shown to be associated with the pathobiological state or clinical outcome. An <u>exploration level</u> biomarker has some evidence of an *in vitro* or preclinical association that may predict a clinical outcome. This type of biomarker is often an outcome of scientific research describing a pathway that may impact a clinical state. The consistency of the data or correlation is low and would be used primarily to support hypothesis generation.

A <u>demonstration level</u> biomarker has higher level of evidence of a correlation between the biomarker and the clinical outcome. This correlative data derives from limited clinical studies. This category is equivalent to the “probable valid” biomarker defined in the guidance on the submission of pharmacogenomic data to FDA [6] and the process map proposal for validating genomic biomarkers by Goodsaid and Frueh [112]. Often the association of the biomarker and the clinical outcome is the result of a post-hoc analysis. While the data showing the association between the biomarker and the biological state may be promising, it is limited until further independent verification can be performed.

<u>Characterization level</u> biomarkers are reproducibly linked to clinical outcomes in more than one prospective clinical study in humans and have been independently verified. This category corresponds to the ‘‘known valid” biomarkers in nomenclature suggested in the FDA pharmacogenomic guidance and process map proposal of Goodsaid and Frueh referenced earlier.

These biomarkers have been shown to be associated with clinical outcomes as prospectively defined endpoints and are appropriate for making a range of decisions regarding the therapeutic being studied. The biomarkers can be used to identify responders versus non-responders, individuals that may be at risk for toxicity, or assist in defining the appropriate dose for an individual.

The final category of biomarker qualification is when the biomarker can be used as a surrogate for a clinical outcome, and thus can be used as the basis of a regulatory decision. Surrogate level biomarkers should be considered a subset of characterization level biomarkers. The use of a surrogate endpoint as the basis for approval of a new drug requires prior agreement with the regulatory agency, and is also restricted to drugs that are intended to treat serious and life-threatening illnesses [113].

Biomarker qualification can occur both during development of a therapeutic [114] or independent of a therapeutic. As described in section 3, there are many uses for biomarkers that are independent of a therapeutic; in addition they may assess characteristics related to safety or toxicity, such as biomarkers that are correlated with stress or damage to critical organs.

5.2.2. Biomarker qualification independent of a therapeutic

Biomarkers that correlate with disease progression that are developed independent of a therapeutic may be perfect tools for identification of promising new therapeutics. This type of biomarker would be considered an exploration level biomarker. Clinical studies would be required to develop the necessary data to show if it can be categorized as a demonstration or characterization level biomarker. Agencies involved include Clinical Laboratory Improvement Amendments; (CLIA), the Food and Drug Administration (FDA) and the Center for Devices and Radiological Health (CDRH).

For the qualification of biomarkers independent of a therapeutic, a process map for the validation/qualification of genomic biomarkers of drug safety has been proposed [112]. The proposed process could encompass biomarkers other than safety, such as biomarkers of disease progression. As the authors state, the process can be considered intuitive as it follows well-established processes. Following the identification of a potential safety or disease progression biomarker and the development of an appropriate analytical method, a qualification protocol can be proposed and discussed with the regulatory agency. Once approved, the qualification protocol could be executed and the report submitted for review. If the data support the correlation between the biomarker and the safety signal or disease progression, then the biomarker could be considered qualified. The level of qualification could be dependent upon whether the protocol included independent or cross validation of the biomarker.

5.2.3. Biomarker qualification in conjunction with a therapeutic

Biomarkers have been used in drug development for some time, and this practice is expected to expand with the trend towards personalized medicine. Because qualification of a biomarker in conjunction with development of a therapeutic is usually done within a single company, independent verification is rarely feasible. However, the process for qualification would be comparable to the process described earlier. Lesko and Atkinson describe in detail a strategy for

biomarker qualification, and note that the criteria used in the qualification of any biomarker are dependent upon the regulatory role a biomarker is expected to play [115]. In 2005, the FDA published a concept paper on drug-diagnostic co-development [116]. This draft document addresses issues related to the development of a test that would be mandatory in the therapeutic use of a drug. Due to the critical role the test would assume, the FDA recommends that the co-development pathway should be determined early in development and the sponsor consult with the appropriate drug/biologic/device reviewing centers. The approval of a drug that utilized the analysis of a biomarker as integral in the use of the drug would require the parallel review and approval of the diagnostic. Agencies involved include CLIA, the FDA, and CDRH.

5.3. Examples of Biomarkers Used for Regulatory Approval of a Therapeutic

There are no examples of biomarkers used for OA drug registration. However, examples are emerging in other fields of the successful application of biomarkers in the development of drugs. To date, the primary biomarkers qualified for use with a drug are genomic. In fact, pharmacogenomic information is contained in about 10% of labels for drugs approved by the FDA. The FDA has published a list of valid genomic biomarkers in the context of these FDA-approved drugs [114]. This list, containing approximately 30 drugs, provides the regulatory context in which the biomarker was approved. Currently, only a few drugs recommend or require an assessment of the biomarker in the context of prescribing the drug or arriving at a therapeutic decision.

In summary, the pathway for qualification of a biomarker is defined by how the biomarker will be used, the questions that are addressed, and how closely the biomarker is associated with a clinical outcome. The qualification process can be viewed somewhat as a continuum, with a relatively low bar required of an exploration level biomarker and the highest level required of a surrogate level biomarker.

6. CONCLUSIONS and RECOMMENDATIONS

6.1. General Overview

This guidance document is being prepared at a time of rapid biomarker evolution in this and other fields when studies are revealing many promising and important contributions that could be made by biomarkers to the development of new treatments for OA. The advantages and potential opportunities offered by the use of biomarkers can be traced from preclinical work involving laboratory-based studies, through work with animal models of OA extending into clinical trials and eventually into the treatment of patients. The use and assessment of the value of these biomarkers is seen as very much a work in progress, building on the lessons learned to date and on the ongoing advances in the clinical and imaging biomarker outcomes that form the basis for the qualifying endpoints for non-imaging biomarkers. At the present time biomarker usage will not provide primary outcome measures in OA clinical trials; this in large part stems from the lack of an appropriate gold standard, which allows robust biomarker qualification with regard to symptomatic and structural outcomes. Because OA is a whole organ disease with different tissues and biological processes involved, a combination of a panel of biochemical markers will probably be more powerful for the investigation of joint damage than assessment of a single biomarker [15, 25]. The potential for the effective clinical use of biomarkers may therefore be

more readily realized as biomarkers start to be included in OA clinical trials, used in combination rather than individually, and used in combination with imaging such as MRI [35]. It may, in the not so distant future, become possible to use selected individual markers or combinations thereof to inform decisions in clinical trials and patient diagnosis, treatment and monitoring.

6.2. Summary of Issues Related to the Application of Biomarkers in the Development of Drugs for OA

6.2.1. Preclinical studies

Biomarkers have already proven their relevance in preclinical studies of arthritis onset, progression, treatment and outcomes. This work includes studies of OA development in mice, rats, guinea pigs, rabbits, horses and dogs, both induced and naturally occurring. Studies with surgically induced joint instability can produce significant biomarker changes in peripheral blood and urine within 2-8 weeks of onset that parallel histologically demonstrated cartilage degeneration. Therapeutic interventions during this period are reflected by biomarker changes in these models. Pre-clinical studies can be used to link changes in specific biomarker parameters (i.e. magnitude of change with intervention and time to measure change in biomarker from first dose) to histological benefit and therefore inform regarding the use of these biomarkers in clinical studies. Routine use of biomarkers for dose selection will require establishing a link to structural and clinical outcomes. Preclinical model studies of cartilage collagen biomarkers of degradation and synthesis and COMP have proven to be of special value. Such studies should provide valuable insights in human clinical investigations. If biomarkers reflecting structural and/or symptomatic changes can be identified in preclinical studies these can then be considered for use in a clinical trial.

6.2.2. Clinical studies

- There are currently no recognized and approved “disease modifying” therapeutic agents, therefore there is no valid means by which to test the ability of biomarkers to change with therapy.
- Biomarkers may serve as titration tools, facilitating dose setting in early clinical studies.
- Although systemic biomarkers (serum and urine) potentially reflect generalized OA (analogous to a global outcome measure) and local (intra-articular) biomarkers reflect local OA, in general, therapeutic studies are focused on one joint, often knee or hip; data are not routinely collected on symptoms or structure in other joints that may also be affected as part of generalized OA and that may impact systemic biomarkers.
- Rescue medication and placebo effects confound trial results but biomarkers provide objective outcomes with the potential to overcome some of the inherent limitations of subjective outcomes.
- Therapeutic trials include patients whose disease is likely to respond to treatment based on symptomatic and imaging criteria, but not biological criteria reflecting tissue metabolic activities; whereas biomarkers and biomarker profiles have the potential to identify molecular and/or metabolic subsets of disease activity and progression that may reflect different responses to a particular intervention.
- Biomarkers provide the only current potential means of identifying the early molecular stages of OA as defined in Figure 1. These early changes, having been identified by a biomarker

may be most susceptible to disease modification, and also measurable by that same biomarker, based on experiences with biologic therapy in inflammatory arthritis [17].

- A major reason for failure of OA clinical trials to date has been lack of study power due to insufficient numbers of progressors with regard to imaging outcome.
- Biomarkers should offer both sensitive detection of patients with active disease for inclusion in trials and monitoring of effects on tissues.
- Biomarkers provide potential means of increasing trial power with a specified sample size through enrichment of a predominantly disease progressing patient population.
- Biomarkers provide potential to decrease the length of a trial or facilitate early decision-making regarding the therapeutic value of a treatment if early biomarker changes are predictive of later clinical or structural outcomes; this has been exemplified to date by several biomarkers including CTX-II [18, 19], MMP-3 [88], and considering the combination of collagen degradation and synthesis [24, 25].
- Although correlations of biomarkers to symptoms will be informative, very short symptomatic trials may be too short to reflect cartilage or bone biomarker level modifications.
- We lack information on the impact of therapy on biomarkers in generalized OA.
- One shortcoming of most biomarker studies is the failure to account for total body burden of disease.
- Proof of concept studies with serum COMP have shown that systemic concentrations in the serum report on burden of (systemic) disease while intra-articular concentrations report on local disease features [47].
- Little to date is known about markers specific for a particular joint site.
- The use of systemic biomarkers to report on local disease at a specific joint site tends to be confounded by high background from turnover in other cartilage tissues including the spine [48].
- There is a validated measure to evaluate spine OA structural changes [117] that could serve as an endpoint on which to qualify a biomarker for spine OA as exemplified by one past study [48].
- There is no definitive “gold standard” for assessing structural changes in all joint tissues with imaging techniques thus hampering the ability to qualify a biomarker for structural endpoints; sampling of fluid from a given joint will circumvent this problem in a trial.
- Statistically significant biomarker differences may not correlate with clinically meaningful differences in symptomatic or imaging endpoints.
- The interpretation of the biomarker values in urine and blood must take into account the possible confounders such as age, gender, body mass index, ethnicity, diurnal changes, food intake, physical activity and post-menopausal status.
- These confounders require that the biological fluids be collected at well defined times, with standardized procedures, accounting for all known confounders.
- Levels of biomarkers measured in blood and urine provide information on systemic skeletal tissue turnover and are not necessarily specific for the alterations occurring in a single affected joint.

- The clearance of the biomarker may also be affected to different extents by physical activity, time of day, and liver and kidney function. At the joint level, biochemical marker clearance may also vary with synovial inflammation.
- The use of multiple biomarkers that represent various components of the complex OA disease pathway, such as tissue synthesis, destruction and inflammation, may yield intermediate endpoints that offer a more comprehensive assessment of treatment effects such as impact on catabolism and anabolism.
- Because clinical decisions can depend on the quality of biomarker data, appropriate analytical validation of biomarker assays is essential to ensure high-quality data to maximize the value of such decisions [39].
- Many promising OA-related biomarkers have never been tested in appropriate clinical trial samples, often because of lack of access to samples by those developing assays, so it is premature to finalize the choice of the optimal biomarker(s) for OA trials.
- No single biomarker will be representative of all aspects of the biological changes in the complex organ represented by the joint.
- To encourage the application of biochemical and genomic biomarkers in drug development, a consensus on how to interpret results from these measurements is needed for regulatory submissions.

6.2.3. Difficulties encountered

Historically, much work on biomarkers has suffered from a number of limitations and obstacles. First among these include difficulties encountered in translating new biomarker assays developed in the laboratory into preclinical animal models and human clinical trials. Often scientists working independently with animal models and in clinics have had difficulties accessing appropriate collaborative opportunities for biomarker application and assessment. Researchers and companies developing biomarker assays continue to have serious problems evaluating assays due to inability to gain access to clinical samples, especially those from clinical trials. Also many assays do not cross-react between human and other species requiring the development of multiple assays. Second, although early events of the OA-process should be optimal for intervention, clinical studies focused on early disease have been very limited (the Cibere et al, 2009 study being a notable example [15]). Diagnosis of OA is typically made late in the disease process and no DMOADS are currently available for treatment, patients are often missed during the early phases of OA. A third obstacle hampering the application of biomarkers has been the lack of understanding of how the processes leading to tissue destruction also lead to symptoms and other clinical parameters and whether there are molecular indicators that correlate with these parameters. Another unknown is whether or how the processes vary over the progression of the disease. Soluble biomarkers are potentially as complex and varied as the biology they model but have the challenge of being qualified based on relatively generic symptomatic and structural outcomes. These obstacles form the basis of a research agenda for the study of OA biomarkers informed by the recommendations below. Fourth, investigators tend to study a single biomarker or a limited set of biomarkers at the exclusion of others. This trend is beginning to change with increased understanding of the need to evaluate many different biomarkers together within a given study. Information can be learned from these biomarkers both individually and in combination as well as in combination with imaging markers.

6.2.4. Critical needs

- To develop better structural endpoints for biomarker qualification;
- To develop biomarkers for various stages of disease;
- To develop biomarkers reporting on specific joint sites and to elucidate the specific joint site contributions to the systemic concentrations of existing biomarkers;
- To determine the clearance of biomarkers from the joint, from the lymphatics, and from the blood as well as the renal processing and elimination via the urine and the effect on their correlation with disease progression;
- To assess if there is a circadian rhythm in the level of a biomarker in a particular matrix to better design the sample collection schedule and the interpretation of the results;
- To assess if there are covariates that affect the concentration of a biomarker in the selected matrix such as age, gender, BMI, concomitant diseases/medications, or joint site involvement;
- To study a wide-variety of patient types with varied clinical characteristics and joint-site involvement;
- To develop biomarkers fit-to-purpose;
- To establish an ongoing critical assessment of the value of existing biomarkers in clinical trials;
- To establish minimal clinically important differences in biomarkers once the minimal clinical important differences are defined for the qualifying endpoints for biomarkers, namely in symptomatic and structural endpoints.
- To be able to gain easier access to body fluids from past, present and future clinical trials to enable more comprehensive and critical head to head evaluations of existing and new biomarkers for use in clinical trials.
- To develop multiplex assays incorporating existing promising biomarkers to provide efficient, cost-effective assays informing on multiple domains of joint biology and response to therapy while minimizing demands for sample.
- To increase the available knowledge of biomarker responses in clinical trials for biomarker qualification and clearance by FDA through public release by companies, of information related to use of biomarkers in their preclinical and clinical trials.

6.3. Recommendations to Advance the Science of Biomarkers

The availability of an expanding number of biomarkers provides increasing opportunities to combine biomarkers to study disease-subsets and to correlate these to clinical parameters and disease outcome. We recommend measurement of a broad set of biomarkers in available and future sample sets, and analysis of biomarkers singly and in combination, to provide a more comprehensive assessment of ongoing disease and efficacy of treatment. We recommend that a panel of biomarkers be used to examine the same samples and preferably in multiple past and future clinical trials. The most appropriate biomarkers would be those related to the proposed mechanism of drug action. The following commercially available biomarkers, some often studied and others less frequently, are nevertheless recommended for inclusion to provide comparative data and biological insights from which to continue to assess the utility and relevance of an array of established OA-related biomarkers: urinary CTXII, serum COMP, serum

Hyaluronan, serum and urine C1,2C, serum and urine C2C, serum and urine Coll2-1 and Coll2-1NO2, serum CPII, Serum PIIANP, urine/serum NTX-1, urine/serum CTX-1, serum CS846, and serum MMP-3. This panel is considered an initial starting point for a process in evolution. As knowledge is gained and additional OA-qualified biomarkers become either commercially or readily available to the OA community of investigators, it is anticipated that this will be revised.

- Recommendations should be developed for biomarker data presentation in publications from research studies and clinical trials. For clinical trials, this should include, at a minimum, reporting of the mean and standard deviations (in all groups before and after treatment) of biomarker concentrations and inter- and intra-assay variation.

- Minimal meaningful differences for biomarkers need to be defined and established and this can be done even in the absence of a treatment study in a longitudinal trial. A critical component for the success of this aim will be the establishment of clinically meaningful endpoints related to imaging and symptom-related outcomes which serve as the qualifying endpoints for biomarker studies.

- For clinical trials, consideration should be given to listing intended biomarker analyses at clinical trials.gov in addition to primary clinical endpoints; alternatively, a separate website could be considered to serve the purpose of tracking and reporting this information, results (both positive and negative apropos of next recommendation), and stimulating advances in the field.

- Biomarker data, both positive and negative, ideally should be released in a timely manner into the public domain, preferably by peer-reviewed publication. This will ensure the optimal development and use of important biomarker tools as exemplified in this guidance document. It will also serve to maintain the momentum generated by a recent increase in collaborative research on biomarkers of OA, ensuring that this continues as a concerted effort to serve the broader stakeholder community to solve common problems. This information could and should be summarized and included in a public database that is managed and regularly updated on a monthly basis.

- Resources should be made available to encourage, through a carefully controlled peer review process, access to body fluids from cohorts such as those harvested from studies of OA onset, progression and OA clinical trials. Many such cohorts are presently available for study (see proceedings of OARSI Biomarker Workshop, Bethesda, MD, 2009) (see http://www.oarsi.org/index2.cfm?section=Meetings_Events&content=OABiomarker). In addition, an effort also needs to be made to obtain cohorts depicting early events, including sample sets for investigation of risk groups after joint trauma, and past and future clinical trial sample sets.

- We note that in existing clinical trials, there has been no standardized method of sample collection. We call for a consensus regarding collection methods and recommend practices in Appendix A.

- We recommend body fluid collection and sample banking in future human (in particular all future prospective OA clinical trials) and animal studies to include serum and plasma, RNA and DNA isolated from whole blood, urine, and where possible, synovial fluid (SF). Synovial

fluid is included since it represents the most proximal fluid to the joint and can provide the most direct insight into joint metabolism in the case of biochemical and molecular biomarkers. Peripheral white blood cells exhibit changes in gene expression in OA that are detectable by microarray and PCR analyses [118, 119]. The process of cell isolation may be associated with artifactual gene expression changes so the collection of whole blood (via PaxGene or Tempus tubes), in lieu of cell isolation, may be preferable for studies of gene expression. Just as the FDA has encouraged voluntary submission of pharmacogenomic data in an effort to increase the knowledge base for therapeutic candidates (see http://www.fda.gov/oc/initiatives/criticalpath/Lesko/Lesko.html) [39], and in view of encouraging successful biomarker developments of this kind in other fields (described in section 5.3), we recommend collection of whole blood for future genomic analyses of gene expression in OA clinical trials.

- Since patterns of fragments may vary in different body fluids due to processing in the kidney, we recommend that both urine and serum samples be collected and analysed when biomarker assays are available for use with both these body fluids.

- Protocols enrolling patients with knee or hip OA (the so-called signal joints) have made measuring and interpreting treatment effects easier, and the development of specific OA measurements has paralleled, and in some ways guided, this signal joint approach. However, exclusive focus on the signal joint will miss what is happening at other OA sites that could affect systemic biomarker concentrations. For this reason it is recommended that clinical trials for OA that include systemic (serum, urine) biomarkers, collect information about other joints in addition to the target joint, such as by using a patient global assessment, or taking specific non-signal-joint measurements. Future developments may demonstrate that the status of particular joints can be distinguished even in the setting of generalized OA.

- Immunoassays based on monoclonal antibodies are preferred (or similar highly specific reactive agents such as those produced by phage libraries). The ability to accurately and quantitatively measure the concentration of epitopes in body fluids is a primary requisite for all assays. Competition immunoassays using a single antibody are often subject to higher assay variability than sandwich assays in which intra- and inter-assay variability can be minimized by use of, ideally, two monoclonal antibodies with different epitope specificities. Sandwich assays however may be problematic with small fragments when these do not span two epitopes. The reliance on polyclonal antisera makes it difficult to ensure continuing assay standardization when new antisera must be raised to replace depleted supplies. The incorporation of an appropriate standard is also an essential requirement for all immunoassays.

- In cases where multiple assays are available for the same analyte, these assays should be compared against each other as different information may be generated according to epitope recognition.

- Although technically challenging, for all existing and future assays, validation of assay specificity should include epitope identification of protein epitopes consisting of sequence verification of the epitope(s) being measured by their isolation and characterization from the sample under investigation using the antibodies that constitute the assay in combination with methods such as mass spectrometry. An example is provided by Nemirovskiy *et. al.* [40] who

examined the peptides in urine generated by collagenase cleavage of type II collagen and bound by the uTIINE antibody.

- For an improved understanding of a biomarker, the principle tissue source(s) of a given biomarker should be identified as accurately as possible, so that the origin(s) of the epitope(s) is/are clearly understood. These requirements are essential for a clear understanding of what the assay results represent and for the interpretation of data when biochemical and molecular biomarkers are used in preclinical or clinical studies.

- Assays developed in independent laboratories should be made available either commercially or through collaborative agreements.

- For the most effective assessment of existing and new biomarkers, strong collaborations involving both the academic and commercial sectors are essential so that accessibility to body fluids and different biomarker assays in past, present and future clinical trials is ensured. It is possible to envision a time when an expert advisory group could manage this and that one or more central reference laboratories perform assays in a standardized manner in both biomarker assessment/validation and in preclinical and clinical trials.

- Data on epitope stability with storage, and freezing and thawing, should be standardized and available in the public domain.

Figure Legends

Figure 1. Continuum of OA stages as paradigms for structure modifying OA trials.

Figure 2. Conceptual framework for biomarker qualification.

Declaration of Funding and Role of Funding Source

The OARSI FDA OA Initiative received financial support from the following professional organization:

American College of Rheumatology

Additionally the OARSI FDA OA Initiative received financial support from the following companies:

Amgen

ArthroLab

AstraZeneca

Bayer Healthcare

Chondrometrics

CombinatoRx

Cypress BioScience

DePuy Mitek

Expanscience

4QImaging

Genevrier/IBSA

Genzyme

King (Alpharma)

Merck

Merck Serono

NicOx

Pfizer

Rottapharm

Smith & Nephew

Wyeth

While individuals from pharmaceutical, biotechnology and device companies actively participated in on-going working group discussions, due to the conflict of interest policy enacted by OARSI, these individuals were not allowed to vote on the final recommendations made by OARSI to the Food and Drug Administration.

Author Contributions

All authors contributed to the writing and revision of the manuscript and approved the final version.

Conflict of Interest Statement

Virginia Byers Kraus was funded by 5P30AG028716 and AR50245 from the NIA/NIH and NIAMS/NIH.

Bruce K Burnett was funded by CTSA grant 1 UL1 RR024128 from NCRR/NIH.

JH, DE, JCR: The authors have nothing to disclose

RM: has received research support or compensation from sanofi-aventis, Pfizer, Bioiberica, Ferring, Endo, Adolor, Merck

SL: has received research support or compensation from AstraZeneca, Carbylan, MerckSerono, Pfizer, Tigenix, Wyeth

YH: has received research support or compensation from Danone, Kitozyme, Nestec, Expanscience, Pierre Fabre, Bioiberica, BioXtract, Intervet, Pfizer; Ownership interest or patent in Wellcare product, Medic all design, Zentech

DH: Patent and stock ownership Anamar Medical; research support or compensation from Anamar Medical, Novo Nordisk

PG: patent with Synarc

RP: has received research support or compensation from sanofi-aventis, Xceed, XOMA, DePuy-Mitek, Centocor, IBEX, AO Foundation, Ethicon; ownership interest or patent Transition Therapeutics, Biosyntech, GeneNews

JC: employee of Novartis Pharmaceuticals Corp

SC: employee of Amgen

MG: employee of Cypress Biosciences

JG: previous employee of Wyeth

AK, JM: employee of Merck & Company, Inc

GM: employee of Genzyme Corp

SP: employee of Rottapharm SpA

MT: employee of Smith & Nephew UK, Ltd

7. APPENDIX A: SAMPLE ACQUISITION AND HANDLING

Introduction to sample handing and considerations relevant to all samples: The measurement of biomarkers in biological samples has the potential to provide information on diagnosis, evaluation of risk, assessment of prognosis, monitoring treatment, prediction of response to treatment and as a surrogate response marker [120]. Biomarkers can be evaluated in a wide array of fluids or tissues depending on the pathology to be monitored. In this Appendix, we focus on the collection and storage of blood, urine, and synovial fluid for the assessment of onset and disease activity in osteoarthritis (OA). For all samples intended for biomarker analysis, the sample quality is dependent on two major factors: the pre-analytical parameters (methods used for sample collection, handling, processing) and the storage conditions (duration of storage, storage temperature, number of thaws) [120]. The time from body fluid sampling to storage should be reduced to a minimum to avoid degradation. There is a consensus that a temperature below minus 70°C is required for long-term stability of protein epitopes, although good prospective data on stability in frozen samples are missing for most assays or have not been published. Likewise, it is ideal to perform biomarker analyses as soon after sampling as possible. However, to avoid assay batch effects, it is best, when possible, to run all samples from a particular study at one time and to examine samples from the same patient on the same plate. For a clinical trial, the storage time prior to analyses may be kept to a minimum but this is less readily implemented for routine use. Epitope stability for each assay should be clearly established with respect to duration of storage and effects of freezing and thawing to ensure validity of the measurements. Samples for immunoassays should be aliquoted on isolation into volumes suitable for at least a single immunoassay in triplicate. A volume of 175 µl is generally recommended to be sufficient to accommodate most assays run in triplicate. With technical improvements in assay design in future, much smaller volumes should be able to be accommodated. The time of day of sample collection should be standardized and noted (recommend AM or PM at least 2 hours after rising and/or any meals for blood and 2nd morning void for urines). The body fluid collection for human and animal studies should include serum, plasma (to avoid the proteolysis that may be activated in blood coagulation), urine, and where possible, synovial fluid samples. Although collection of synovial fluid presents some unique challenges in both patients and animal studies, this sample represents the most proximal fluid to the joint and can provide the most direct insight into joint metabolism in the case of biochemical and molecular biomarkers. Whole blood should be collected in appropriate tubes to permit biomarker studies of gene expression and genetic polymorphisms. Robust standardized protocols for sample collection, handling, and storage should be developed and adhered to for high quality biomarker analyses.

Blood collection, handling, and recommendations: Blood should be collected and stored separately as serum, plasma, and whole blood. Some assays work better in serum and some only in plasma so the acquisition of both provides for maximal possible assays, as exemplified by assays for the matrix metalloproteinases (MMPs) [121]. The specific needs of the assay should be carefully checked in advance. For instance, consideration must be given to the potential for altered protein conformations and immunoreactivity in an assay upon chelation of divalent cations. Patients can be fasted overnight prior to blood collection but this is often found not necessary. Plasma should be collected into a commercial collection tube with anticoagulant added (commonly EDTA, citrate, or heparin) followed by centrifugation. The effect of different anticoagulants on the analyte should be examined [120] as the requirements differ for different

assays. For instance, EDTA and citrate plasma are unsuitable for MMP activity assays as these anticoagulants chelate calcium required for MMP activity. For serum, blood should be collected in a red top tube without additives. Serum separator (SST) tubes are particularly easy to use and minimize contamination of serum by clot; they have been successfully used for several years by some researchers (VBK). Upon blood collection, the plasma or serum tube is immediately gently inverted 3-5 times, and allowed to clot at room temperature for at least 30 minutes (maximum 60 minutes to avoid subsequent fibrinolysis), followed by centrifugation at approximately 1300g (~3500 rpm) for 10 minutes to separate the plasma from the buffy coat and red blood cells (anticoagulant tube), or to separate the serum from the clot (tube without additives). The supernatant from both plasma and serum collection tubes should be aliquoted into small fresh cryotubes (recommended 100 or 175µl aliquots according to the assay noting the volume to monitor for potential subsequent dessication of sample) and frozen below -70C. Depending on the intended use of the sample, mixed protease inhibitors can be added to blood sample collections to avoid degradation of specific analytes of interest. The material composition of the tube can affect measurement of analytes so it is recommended to use identical tubes for all samples within a study.

For total RNA and genomic DNA isolation from whole blood: The PaxGene blood collection tube (Becton Dickinson) can be used to obtain RNA and DNA from whole blood. RNA extraction can be performed using Qiagen's PaxGene 96 blood RNA kit. RNA amplification can be achieved using Ambion Illumina AMIL1791 Total Prep RNA amplification kit. DNA isolation can also be achieved from these tubes as described [122]. An alternative but similar system is provided by Tempus tubes (Applied Biosystems). A successful example of blood-derived gene expression analysis in OA is provided by Marshall 2005 [118] although *in vitro* manipulation of cells is ideally to be avoided in favor of direct RNA isolation with PAXGene or Tempus tubes.

Urine collection, handing and recommendations: A second morning void urine specimen is recommended as the standard for biomarker assays. Prior to aliquoting (1 ml aliquots recommended for urine), samples should be centrifuged at approximately 1300g for 10 minutes to remove any debris. As with blood samples, collection of urine samples should use a standardized tube and aliquoted supernatants should be stored, as for serum and plasma, in cryotubes below -70°C until measurements are made. Biomarker levels in urine are subject to dilutional variances due to varying hydration level and urine flow rate (volume produced/time) or total volume. This requires adjustment for differences in flow rate or volume to allow comparison of samples collected from different patients or from the same patient over time and is most commonly achieved through normalization of urine biomarker values with urinary creatinine, although urinary creatinine is influenced by age, diet, exercise, muscle mass, medications, tubular secretion and glomerular filtration rate [123].

Synovial fluid collection, handling, and recommendations: Synovial fluids can be aspirated directly in many cases, but if necessary, a small volume (10 mls) of sterile saline can be injected into the knee followed by aspiration of all obtainable fluid [47]. Using this technique, only one needle insertion is required for human studies. For animal studies (usually performed under anaesthesia except in rabbits), it is recommended that the needle be withdrawn after saline injection, and the knee flexed and extended 10 times to ensure mixing; this procedure can also potentially increase the yield of fluid aspirated. To obtain a total white blood cell count in the sample, 25 µl of synovial fluid can be mixed with 25 µl of trypan blue and the cell count

performed with a hemocytometer. Synovial fluid samples should be cleared by centrifugation (approximately 1300 g for 10 minutes) and the remaining supernatant fluid aliquoted (100 µl) and stored in cryotubes at -80°C for future assays. In cases where 10 ml lavage samples were obtained, a nearly simultaneous serum sample should also be obtained in order to determine the dilutional factor of the synovial fluid for subsequent correction of biomarker concentrations for this dilutional effect. The dilution factor can be determined as described by Kraus et al [124] based on measuring urea concentrations in the synovial fluid and serum. Synovial fluid up to 2.5 fold diluted shows a similar mass spectroscopic profile as synovial fluid aspirated directly (V. Kraus unpublished data); beyond this level of dilution there may be some specimen heterogeneity introduced by lack of mixing.

In the case of small animals, such as mice, a published methodology is available for obtaining synovial fluid at the time of sacrifice [125]. This utilizes an alginate product with high absorbancy that wicks the fluid from the joint. The method has to be tested for each biomarker or analyte of interest to insure that the alginate or the buffer components do not interfere with the assay but to date has been shown to be compatible with synovial fluid COMP [125] and IL-1 (VBK unpublished). For rabbits, Poole et al 1978 [126] have used a 1-2 ml saline injection, containing the tissue culture dye neutral red, into the stifle (knee) joint. Dilution and hence original synovial fluid volume is determined by spectrophotometric examination of dye concentration. Intra-articular injection volumes, determined by the relative size of the animal, should be used for other species in relation to the rabbit. However, this weakly cationic dye penetrates cell membranes by nonionic diffusion and binds intracellularly to sites of the lysosomal matrix [127]; these dye properties may confound the determinations of dilution factor by this method. In the past, Evans blue and indocyanine green dyes were shown to be inappropriate for monitoring dilutional effects of lavage because of their absorption and metabolism by intraarticular cells and precipitation upon exposure to synovial fluid [124]. Another useful approach for small animals (rats, guinea pigs, rabbits) at sacrifice is to blot the surfaces of the opened joint with a pre-weighed filter paper and then immediately record the weight with the synovial fluid blotted. This will provide a measure of the amount of non-diluted fluid. Biomarkers can be readily eluted from the paper as described and validated previously [125]. This method works well in the guinea pig [128, 129]. Biomarker ratios can also be calculated in joint fluids where dilutions cannot be determined. These are independent of dilution and provide useful data (RP, unpublished).

8. APPENDIX B: DEFINITIONS OF BIOMARKER TERMS

EXPLORATORY BIOMARKER: research and development tools accompanied by in vitro and/or preclinical evidence, but there is no consistent information linking the biomarker to clinical outcomes in humans. Used for hypothesis generation. First level of surrogacy based on Wagner et al [1].

DEMONSTRATION BIOMARKER: associated with adequate preclinical sensitivity and specificity and linked with clinical outcomes, but have not been reproducibly demonstrated in clinical studies. This category corresponds to ‘‘probable valid biomarkers’’ in nomenclature suggested in draft guidance from FDA. Used in decision-making; provides supporting evidence for primary clinical evidence. Second level of surrogacy based on Wagner et al [1].

CHARACTERIZATION BIOMARKER: associated with adequate preclinical sensitivity and specificity and reproducibly linked to clinical outcomes in more than one prospective clinical study in humans. This category corresponds to ‘‘known valid biomarkers’’ in nomenclature suggested in guidance by FDA. Used in decision-making, and dose finding, for secondary/tertiary claims. Third level of surrogacy based on Wagner et al [1].

SURROGATE BIOMARKER: A holistic evaluation of the available data demonstrates that the biomarker can substitute for a clinical endpoint. The designation of ‘‘surrogate end point’’ requires agreement with regulatory authorities. Used for drug registration. Fourth level of surrogacy based on Wagner et al [1].

VALID BIOMARKER: has been defined in the “Guidance for Industry: Pharmacogenomic Data Submissions”. Therein, a valid biomarker is described as a “biomarker that is measured in an analytical test system with well established performance characteristics and for which there is an established scientific framework or body of evidence that elucidates the physiologic, toxicologic, pharmacologic, or clinical significance of the test results.” The classification of biomarkers is context specific.
(http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm)

ANALYTIC VALIDITY: test’s ability to accurately and reliably detect the epitope of interest.

FORMAT: commercial availability, single or multiplex , type of assay (ELISA or mass spectroscopy, etc)

QUALIFICATION ENDPOINTS: symptoms; structure: radiographic OA, preradiographic OA; molecular OA

The following are summarized from the International Conference on Harmonisation of Pharmaceuticals for Human Use [111]

ACCURACY: expresses the closeness of agreement between the value that is accepted either as a conventional true value or an accepted reference value and the value found.

DETECTION LIMIT: is the lowest amount of analyte in a sample that can be detected but not necessarily quantified as an exact value.

LINEARITY: is the ability of an analytical procedure (within a given range) to obtain test results which are directly proportional to the concentration (amount) of analyte in the sample.

PRECISION: expresses the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions. Precision may be considered at three levels: repeatability, intermediate precision and reproducibility.

> **Repeatability:** expresses the precision under the same operating conditions over a short interval of time and is also termed intra-assay precision.
>
> **Intermediate precision:** expresses within-laboratories variations: different days, different analysts, different equipment, etc.
>
> **Reproducibility**: expresses the precision between laboratories (collaborative studies, usually applied to standardization of methodology).

QUANTITATION LIMIT: is the lowest amount of analyte in a sample that can be quantitatively determined with suitable precision and accuracy.

RANGE: the interval between the upper and lower concentration (amounts) of analyte in the sample (including these concentrations) for which it has been demonstrated that the analytical procedure has a suitable level of precision, accuracy and linearity.

ROBUSTNESS: a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters and provides an indication of its reliability during normal usage.

SPECIFICITY: the ability to assess unequivocally the analyte in the presence of components that may be expected to be present.

9. REFERENCES

1. Wagner JA, Williams SA, Webster CJ. Biomarkers and surrogate end points for fit-for-purpose development and regulatory evaluation of new drugs. Clin Pharmacol Ther 2007; 81: 104-107.
2. Biomarkers Definitions Working G. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001; 69: 89-95.
3. FDA FaDA. Guidance for Industry: E15 Definitions for Genomic Biomarkers, Pharmacogenomics, Pharmacogenetics, Genomic Data and Sample Coding Categories. 2008.
4. Goodsaid F, Frueh F. Biomarker qualification pilot process at the US Food and Drug Administration. AAPS J 2007; 9: E105-108.
5. Bauer DC, Hunter DJ, Abramson SB, Attur M, Corr M, Felson D, et al. Classification of osteoarthritis biomarkers: a proposed approach. Osteoarthritis Cartilage 2006; 14: 723-727.
6. FDA FaDA. Guidance for Industry: Pharmacogenomic Data Submissions 2005.
7. FDA FDA. E16 Genomic Biomarkers Related to Drug Response: Context, Structure, and Format of Qualification Submissions. 2009.
8. Colman EG, Food and Drug A. The Food and Drug Administration's Osteoporosis Guidance Document: past, present, and future. J Bone Miner Res 2003; 18: 1125-1128.
9. Information OoFo. Guidelines for the Clinical Evaluation of Drugs Used in the Treatment of Osteoporosis. In: Services DoHaH Ed. Rockville, MD, USA 1979.
10. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. Arthritis Rheum 1990; 33: 1601-1610.
11. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis & Rheumatism 1991; 34: 505-514.
12. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum 1986; 29: 1039-1049.
13. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957; 16: 494-502.
14. Hart DJ, Spector TD. Kellgren & Lawrence grade 1 osteophytes in the knee-doubtful or definite? Osteo Cartilage 2003; 11: 149-150.
15. Cibere J, Zhang H, Garnero P, Poole AR, Lobanok T, Saxne T, et al. Association of biomarkers with pre-radiographically defined and radiographically defined knee osteoarthritis in a population-based study. Arthritis Rheum 2009; 60: 1372-1380.
16. Ling SM, Patel DD, Garnero P, Zhan M, Vaduganathan M, Muller D, et al. Serum protein signatures detect early radiographic osteoarthritis. Osteoarthritis Cartilage 2009; 17: 43-48.
17. Mullan RH, Matthews C, Bresnihan B, FitzGerald O, King L, Poole AR, et al. Early changes in serum type II collagen biomarkers predict radiographic progression at one year in inflammatory arthritis patients after biologic therapy. Arthritis Rheum 2007; 56: 2919-2928.
18. Christgau S, Henrotin Y, Tanko LB, Rovati LC, Collette J, Bruyere O, et al. Osteoarthritic patients with high cartilage turnover show increased responsiveness to the cartilage

protecting effects of glucosamine sulphate. Clinical & Experimental Rheumatology 2004; 22: 36-42.
19. Garnero P, Aronstein WS, Cohen SB, Conaghan PG, Cline GA, Christiansen C, et al. Relationships between biochemical markers of bone and cartilage degradation with radiological progression in patients with knee osteoarthritis receiving risedronate: the Knee Osteoarthritis Structural Arthritis randomized clinical trial. Osteoarthritis Cartilage 2008; 16: 660-666.
20. 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: 3711-3722.
21. Eckstein F, Maschek S, Wirth W, Hudelmaier M, Hitzl W, Wyman B, et al. One year change of knee cartilage morphology in the first release of participants from the Osteoarthritis Initiative progression subcohort: association with sex, body mass index, symptoms and radiographic osteoarthritis status. Ann Rheum Dis 2009; 68: 674-679.
22. Qvist P, Bay-Jensen AC, Christiansen C, Dam EB, Pastoureau P, Karsdal MA. The disease modifying osteoarthritis drug (DMOAD): Is it in the horizon? Pharmacol Res 2008; 58: 1-7.
23. Sharif M, Kirwan JR, Elson CJ, Granell R, Clarke S. Suggestion of nonlinear or phasic progression of knee osteoarthritis based on measurements of serum cartilage oligomeric matrix protein levels over five years. Arthritis Rheum 2004; 50: 2479-2488.
24. Garnero P, Ayral X, Rousseau JC, Christgau S, Sandell LJ, Dougados M, et al. Uncoupling of type II collagen synthesis and degradation predicts progression of joint damage in patients with knee osteoarthritis. Arthritis & Rheumatism 2002; 46: 2613-2624.
25. Cahue S, Sharma L, Dunlop D, Ionescu M, Song J, Lobanok T, et al. The ratio of type II collagen breakdown to synthesis and its relationship with the progression of knee osteoarthritis. Osteoarthritis Cartilage 2007; 15: 819-823.
26. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997; 24: 799-802.
27. Pham T, van der Heijde D, Altman RD, Anderson JJ, Bellamy N, Hochberg M, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. Osteoarthritis Cartilage 2004; 12: 389-399.
28. Abadie E, Ethgen D, Avouac B, Bouvenot G, Branco J, Bruyere O, et al. Recommendations for the use of new methods to assess the efficacy of disease-modifying drugs in the treatment of osteoarthritis. Osteoarthritis Cartilage 2004; 12: 263-268.
29. Hunter DJ, Zhang YQ, Tu X, Lavalley M, Niu JB, Amin S, et al. Change in joint space width: hyaline articular cartilage loss or alteration in meniscus? Arthritis Rheum 2006; 54: 2488-2495.
30. Emrani PS, Katz JN, Kessler CL, Reichmann WM, Wright EA, McAlindon TE, et al. Joint space narrowing and Kellgren-Lawrence progression in knee osteoarthritis: an analytic literature synthesis. Osteoarthritis Cartilage 2008; 16: 873-882.
31. Charles HC, Kraus VB, Ainslie M, Hellio Le Graverand-Gastineau MP. Optimization of the fixed-flexion knee radiograph. Osteoarthritis Cartilage 2007; 15: 1221-1224.
32. King KB, Lindsey CT, Dunn TC, Ries MD, Steinbach LS, Majumdar S. A study of the relationship between molecular biomarkers of joint degeneration and the magnetic resonance-measured characteristics of cartilage in 16 symptomatic knees. Magnetic Resonance Imaging 2004; 22: 1117-1123.

33. Dam EB, Loog M, Christiansen C, Byrjalsen I, Folkesson J, Nielsen M, et al. Identification of progressors in osteoarthritis by combining biochemical and MRI-based markers. Arthritis Res Ther 2009; 11: R115.
34. Lukaszkiewicz J, Karczmarewicz E, Pludowski P, Jaworski M, Czerwinski E, Lewinski A, et al. Feasibility of simultaneous measurement of bone formation and bone resorption markers to assess bone turnover rate in postmenopausal women: an EPOLOS study. Med Sci Monit 2008; 14: PH65-70.
35. Karsdal MA, Henriksen K, Leeming DJ, Mitchell P, Duffin K, Barascuk N, et al. Biochemical markers and the FDA Critical Path: how biomarkers may contribute to the understanding of pathophysiology and provide unique and necessary tools for drug development. Biomarkers 2009; 14: 181-202.
36. Halasz K, Kassner A, Morgelin M, Heinegard D. Comp acts as a catalyst in collagen fibrillogenesis. J Biol Chem 2007.
37. Milan AM, Sugars RV, Embery G, Waddington RJ. Modulation of collagen fibrillogenesis by dentinal proteoglycans. Calcif Tissue Int 2005; 76: 127-135.
38. Brune K, Katus HA, Moecks J, Spanuth E, Jaffe AS, Giannitsis E. N-terminal pro-B-type natriuretic peptide concentrations predict the risk of cardiovascular adverse events from antiinflammatory drugs: a pilot trial. Clin Chem 2008; 54: 1149-1157.
39. Lee JW, Weiner RS, Sailstad JM, Bowsher RR, Knuth DW, O'Brien PJ, et al. Method validation and measurement of biomarkers in nonclinical and clinical samples in drug development: a conference report. Pharm Res 2005; 22: 499-511.
40. Nemirovskiy OV, Dufield DR, Sunyer T, Aggarwal P, Welsch DJ, Mathews WR. Discovery and development of a type II collagen neoepitope (TIINE) biomarker for matrix metalloproteinase activity: from in vitro to in vivo. Anal Biochem 2007; 361: 93-101.
41. Eyre DR, Weis MA. The Helix-II epitope: a cautionary tale from a cartilage biomarker based on an invalid collagen sequence. Osteoarthritis Cartilage 2009; 17: 423-426.
42. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003; 138: 40-44.
43. Felson DT, Lohmander LS. Whither osteoarthritis biomarkers? Osteoarthritis Cartilage 2009; 17: 419-422.
44. Goodsaid FM, Frueh FW, Mattes W. Strategic paths for biomarker qualification. Toxicology 2008; 245: 219-223.
45. Lassere MN, Johnson KR, Boers M, Tugwell P, Brooks P, Simon L, et al. Definitions and validation criteria for biomarkers and surrogate endpoints: development and testing of a quantitative hierarchical levels of evidence schema. J Rheumatol 2007; 34: 607-615.
46. Aurich M, Squires GR, Reiner A, Mollenhauer JA, Kuettner KE, Poole AR, et al. Differential matrix degradation and turnover in early cartilage lesions of human knee and ankle joints. Arthritis Rheum 2005; 52: 112-119.
47. Addison S, Coleman RE, Feng S, McDaniel G, Kraus VB. Whole-body bone scintigraphy provides a measure of the total-body burden of osteoarthritis for the purpose of systemic biomarker validation. Arthritis Rheum 2009; 60: 3366-3373.
48. Garnero P, Sornay-Rendu E, Arlot M, Christiansen C, Delmas PD. Association between spine disc degeneration and type II collagen degradation in postmenopausal women: the OFELY study. Arthritis & Rheumatism 2004; 50: 3137-3144.
49. Meulenbelt I, Kloppenburg M, Kroon HM, Houwing-Duistermaat JJ, Garnero P, Hellio Le Graverand MP, et al. Urinary CTX-II levels are associated with radiographic subtypes of

osteoarthritis in hip, knee, hand, and facet joints in subject with familial osteoarthritis at multiple sites: the GARP study. Ann Rheum Dis 2006; 65: 360-365.
50. Kraus VB, Kepler TB, Stabler T, Renner J, Jordan J. First qualification study of serum biomarkers as indicators of total body burden of osteoarthritis. PLoS One 2010; 5: e9739.
51. Clark AG, Jordan JM, Vilim V, Renner JB, Dragomir AD, Luta G, et al. Serum cartilage oligomeric matrix protein reflects osteoarthritis presence and severity: the Johnston County Osteoarthritis Project. Arthritis Rheum 1999; 42: 2356-2364.
52. Jordan JM, Luta G, Stabler T, Renner JB, Dragomir AD, Vilim V, et al. Ethnic and sex differences in serum levels of cartilage oligomeric matrix protein: the Johnston County Osteoarthritis Project. Arthritis Rheum 2003; 48: 675-681.
53. Kobayashi T, Yoshihara Y, Samura A, Yamada H, Shinmei M, Roos H, et al. Synovial fluid concentrations of the C-propeptide of type II collagen correlate with body mass index in primary knee osteoarthritis. Annals of the Rheumatic Diseases 1997; 56: 500-503.
54. Heinegard D, Saxne T. Molecular markers of processes in cartilage in joint disease. Br J Rheumatol 1991; 30 Suppl 1: 21-24.
55. Young-Min SA, Cawston TE, Griffiths ID. Markers of joint destruction: principles, problems, and potential. Ann Rheum Dis 2001; 60: 545-548.
56. Kong SY, Stabler TV, Criscione LG, Elliott AL, Jordan JM, Kraus VB. Diurnal variation of serum and urine biomarkers in patients with radiographic knee osteoarthritis. Arthritis Rheum 2006; 54: 2496-2504.
57. Gordon CD, Stabler TV, Kraus VB. Variation in osteoarthritis biomarkers from activity not food consumption. Clin Chim Acta 2008; 398: 21-26.
58. Quintana DJ, Garnero P, Huebner JL, Charni-Ben Tabassi N, Kraus VB. PIIANP and HELIXII diurnal variation. Osteoarthritis Cartilage 2008; 16: 1192-1195.
59. Garnero P, Gourley I, Mareau E, Durn B, Hickey L, Cohen S. Biological variability of biochemical markers of bone, cartilage, and synovial metabolism in patients with knee osteoarthritis: effect of food intake Osteo Cartilage 2007; 15: C68.
60. Rossler A, Laszlo Z, Kvas E, Hinghofer-Szalkay HG. Plasma hyaluronan concentration: no circadian rhythm but large effect of food intake in humans. Eur J Appl Physiol Occup Physiol 1998; 78: 573-577.
61. Christgau S. Circadian variation in serum CrossLaps concentration is reduced in fasting individuals.[comment]. Clinical Chemistry 2000; 46: 431.
62. van Spil WE, Degroot J, Lems WF, Oostveen JC, Lafeber FP. Serum and urinary biochemical markers for knee and hip-osteoarthritis: a systematic review applying the consensus BIPED criteria. Osteoarthritis Cartilage 2010; 18: 605-612.
63. Lohmander L, Eyre D. Biochemical markers as surrogate end points of joint disease. In: Clinical Trials in Rheumatoid Arthritis and Osteoarthritis, Reid D, Miller C Eds. New York: Springer 2008:249-274.
64. Charni-Ben Tabassi N, Garnero P. Monitoring cartilage turnover. Curr Rheumatol Rep 2007; 9: 16-24.
65. Rousseau JC, Delmas PD. Biological markers in osteoarthritis. Nat Clin Pract Rheumatol 2007; 3: 346-356.
66. Henrotin Y, Addison S, Kraus V, Deberg M. Type II collagen markers in osteoarthritis: what do they indicate? Curr Opin Rheumatol 2007; 19: 444-450.
67. Birmingham J, Vilim V, Kraus V. Collagen biomarkers for arthritis applications. Biomarker Insights 2006; 2: 61-76.

68. Lohmander L, Poole A. Defining and validating the clinical role of molecular markers in osteoarthritis. In: Osteoarthritis, Brandt K, Doherty M, Lohmander L Eds. Oxford: Oxford University Press 2003:468-477.
69. Hellio Le Graverand MP, Brandt KD, Mazzuca SA, Katz BP, Buck R, Lane KA, et al. Association between concentrations of urinary type II collagen neoepitope (uTIINE) and joint space narrowing in patients with knee osteoarthritis. Osteoarthritis Cartilage 2006.
70. Gineyts E, Mo JA, Ko A, Henriksen DB, Curtis SP, Gertz BJ, et al. Effects of ibuprofen on molecular markers of cartilage and synovium turnover in patients with knee osteoarthritis. Annals of the Rheumatic Diseases 2004; 63: 857-861.
71. Christgau S, Garnero P, Fledelius C, Moniz C, Ensig M, Gineyts E, et al. Collagen type II C-telopeptide fragments as an index of cartilage degradation. Bone 2001; 29: 209-215.
72. Gineyts E, Garnero P, Delmas PD. Urinary excretion of glucosyl-galactosyl pyridinoline: a specific biochemical marker of synovium degradation. Rheumatology (Oxford) 2001; 40: 315-323.
73. Manicourt DH, Azria M, Mindeholm L, Thonar EJ, Devogelaer JP. Oral salmon calcitonin reduces Lequesne's algofunctional index scores and decreases urinary and serum levels of biomarkers of joint metabolism in knee osteoarthritis. Arthritis Rheum 2006; 54: 3205-3211.
74. Spector TD, Conaghan PG, Buckland-Wright JC, Garnero P, Cline GA, Beary JF, et al. Effect of risedronate on joint structure and symptoms of knee osteoarthritis: results of the BRISK randomized, controlled trial [ISRCTN01928173]. Arthritis Res Ther 2005; 7: R625-633.
75. Bingham CO, 3rd, Buckland-Wright JC, Garnero P, Cohen SB, Dougados M, Adami S, et al. Risedronate decreases biochemical markers of cartilage degradation but does not decrease symptoms or slow radiographic progression in patients with medial compartment osteoarthritis of the knee: results of the two-year multinational knee osteoarthritis structural arthritis study. Arthritis Rheum 2006; 54: 3494-3507.
76. Mazieres B, Hucher M, Zaim M, Garnero P. Effect of chondroitin sulphate in symptomatic knee osteoarthritis: a multicentre, randomised, double-blind, placebo-controlled study. Ann Rheum Dis 2007; 66: 639-645.
77. Andersson ML, Thorstensson CA, Roos EM, Petersson IF, Heinegard D, Saxne T. Serum levels of cartilage oligomeric matrix protein (COMP) increase temporarily after physical exercise in patients with knee osteoarthritis. BMC Musculoskelet Disord 2006; 7: 98.
78. Chua SD, Jr., Messier SP, Legault C, Lenz ME, Thonar EJ, Loeser RF. Effect of an exercise and dietary intervention on serum biomarkers in overweight and obese adults with osteoarthritis of the knee. Osteoarthritis Cartilage 2008.
79. Li XQ, Thonar EJ, Knudson W. Accumulation of hyaluronate in human lung carcinoma as measured by a new hyaluronate ELISA. Connect Tissue Res 1989; 19: 243-253.
80. Thonar EJ, Lenz ME, Klintworth GK, Caterson B, Pachman LM, Glickman P, et al. Quantification of keratan sulfate in blood as a marker of cartilage catabolism. Arthritis Rheum 1985; 28: 1367-1376.
81. Cibere J, Thorne A, Kopec JA, Singer J, Canvin J, Robinson DB, et al. Glucosamine sulfate and cartilage type II collagen degradation in patients with knee osteoarthritis: randomized discontinuation trial results employing biomarkers. Journal of Rheumatology 2005; 32: 896-902.
82. Poole AR, Ionescu M, Fitzcharles MA, Billinghurst RC. The assessment of cartilage degradation in vivo: development of an immunoassay for the measurement in body fluids of type II collagen cleaved by collagenases. J Immunol Methods 2004; 294: 145-153.

83. Billinghurst RC, Dahlberg L, Ionescu M, Reiner A, Bourne R, Rorabeck C, et al. Enhanced cleavage of type II collagen by collagenases in osteoarthritic articular cartilage. J Clin Invest 1997; 99: 1534-1545.
84. Leff RL, Elias I, Ionescu M, Reiner A, Poole AR. Molecular changes in human osteoarthritic cartilage after 3 weeks of oral administration of BAY 12-9566, a matrix metalloproteinase inhibitor. J Rheumatol 2003; 30: 544-549.
85. Nelson F, Dahlberg L, Laverty S, Reiner A, Pidoux I, Ionescu M, et al. Evidence for altered synthesis of type II collagen in patients with osteoarthritis. J Clin Invest 1998; 102: 2115-2125.
86. Hollander AP, Heathfield TF, Webber C, Iwata Y, Bourne R, Rorabeck C, et al. Increased damage to type II collagen in osteoarthritic articular cartilage detected by a new immunoassay. J Clin Invest 1994; 93: 1722-1732.
87. Rizkalla G, Reiner A, Bogoch E, Poole AR. Studies of the articular cartilage proteoglycan aggrecan in health and osteoarthritis. Evidence for molecular heterogeneity and extensive molecular changes in disease. J Clin Invest 1992; 90: 2268-2277.
88. Lohmander LS, Brandt KD, Mazzuca SA, Katz BP, Larsson S, Struglics A, et al. Use of the plasma stromelysin (matrix metalloproteinase 3) concentration to predict joint space narrowing in knee osteoarthritis. Arthritis Rheum 2005; 52: 3160-3167.
89. Walakovits LA, Moore VL, Bhardwaj N, Gallick GS, Lark MW. Detection of stromelysin and collagenase in synovial fluid from patients with rheumatoid arthritis and posttraumatic knee injury. Arthritis Rheum 1992; 35: 35-42.
90. Mazzuca SA, Brandt KD, Eyre DR, Katz BP, Askew J, Lane KA. Urinary levels of type II collagen C-telopeptide crosslink are unrelated to joint space narrowing in patients with knee osteoarthritis. Ann Rheum Dis 2006; 65: 1055-1059.
91. Mazzuca SA, Poole AR, Brandt KD, Katz BP, Lane KA, Lobanok T. Associations between joint space narrowing and molecular markers of collagen and proteoglycan turnover in patients with knee osteoarthritis. J Rheumatol 2006; 33: 1147-1151.
92. Otterness IG, Brandt KD, Le Graverand MP, Mazzuca SA. Urinary TIINE concentrations in a randomized controlled trial of doxycycline in knee osteoarthritis: implications of the lack of association between TIINE levels and joint space narrowing. Arthritis Rheum 2007; 56: 3644-3649.
93. Uebelhart D, Thonar EJ, Delmas PD, Chantraine A, Vignon E. Effects of oral chondroitin sulfate on the progression of knee osteoarthritis: a pilot study. Osteoarthritis Cartilage 1998; 6 Suppl A: 39-46.
94. Uebelhart D, Gineyts E, Chapuy MC, Delmas PD. Urinary excretion of pyridinium crosslinks: a new marker of bone resorption in metabolic bone disease. Bone Miner 1990; 8: 87-96.
95. Hasegawa M, Nakoshi Y, Tsujii M, Sudo A, Masuda H, Yoshida T, et al. Changes in biochemical markers and prediction of effectiveness of intra-articular hyaluronan in patients with knee osteoarthritis. Osteoarthritis Cartilage 2008; 16: 526-529.
96. Yamada H, Miyauchi S, Morita M, Yoshida Y, Yoshihara Y, Kikuchi T, et al. Content and sulfation pattern of keratan sulfate in hip osteoarthritis using high performance liquid chromatography. J Rheumatol 2000; 27: 1721-1724.
97. Yoshida K, Miyauchi S, Kikuchi H, Tawada A, Tokuyasu K. Analysis of unsaturated disaccharides from glycosaminoglycuronan by high-performance liquid chromatography. Anal Biochem 1989; 177: 327-332.
98. Shinmei M, Miyauchi S, Machida A, Miyazaki K. Quantitation of chondroitin 4-sulfate and chondroitin 6-sulfate in pathologic joint fluid. Arthritis Rheum 1992; 35: 1304-1308.

99. Arjmandi BH, Khalil DA, Lucas EA, Smith BJ, Sinichi N, Hodges SB, et al. Soy protein may alleviate osteoarthritis symptoms. Phytomedicine 2004; 11: 567-575.
100. Buckland-Wright JC, Messent EA, Bingham CO, 3rd, Ward RJ, Tonkin C. A 2 yr longitudinal radiographic study examining the effect of a bisphosphonate (risedronate) upon subchondral bone loss in osteoarthritic knee patients. Rheumatology (Oxford) 2007; 46: 257-264.
101. Conrozier T, Poole AR, Ferrand F, Mathieu P, Vincent F, Piperno M, et al. Serum concentrations of type II collagen biomarkers (C2C, C1, 2C and CPII) suggest different pathophysiologies in patients with hip osteoarthritis. Clin Exp Rheumatol 2008; 26: 430-435.
102. Hellio Le Graverand MP, Brandt KD, Mazzuca SA, Katz BP, Buck R, Lane KA, et al. Association between concentrations of urinary type II collagen neoepitope (uTIINE) and joint space narrowing in patients with knee osteoarthritis. Osteoarthritis Cartilage 2006; 14: 1189-1195.
103. Atley L, Shao P, Shaffer K, Eyre D. Matrix metalloproteinase-mediated release of immunoreactive telopeptides from cartilage type II collagen [abstract]. . Trans Orthop Res Soc 1998; 23: 850.
104. Charni-Ben Tabassi N, Desmarais S, Bay-Jensen AC, Delaisse JM, Percival MD, Garnero P. The type II collagen fragments Helix-II and CTX-II reveal different enzymatic pathways of human cartilage collagen degradation. Osteoarthritis Cartilage 2008.
105. Oestergaard S, Sondergaard BC, Hoegh-Andersen P, Henriksen K, Qvist P, Christiansen C, et al. Effects of ovariectomy and estrogen therapy on type II collagen degradation and structural integrity of articular cartilage in rats: implications of the time of initiation. Arthritis Rheum 2006; 54: 2441-2451.
106. Norlund L, Shao P, Yoshihara P, Eyre D. Markers of bone type I and cartilage type II collagen degradation in the Hartley guinea pig model of osteoarthritis. Trans Orthop Res 1997; 22: 313.
107. Eyre D, Atley L, Wu J-J. Collagen cross-links as markers of bone and cartilage degradation. In: The Many Faces of Osteoarthritis, Hascall V, Kuettner K Eds. Basel: Birkhauser Verlag 2002:275-284.
108. Huebner JL, Williams JM, Deberg M, Henrotin Y, Kraus VB. Collagen fibril disruption occurs early in primary guinea pig knee osteoarthritis. Osteoarthritis Cartilage 2009.
109. Bay-Jensen AC, Andersen TL, Charni-Ben Tabassi N, Kristensen PW, Kjaersgaard-Andersen P, Sandell L, et al. Biochemical markers of type II collagen breakdown and synthesis are positioned at specific sites in human osteoarthritic knee cartilage. Osteoarthritis Cartilage 2008; 16: 615-623.
110. FDA UFDA. 21 CFR Part 211: Current Good Manufacturing Practice for Finished Pharmaceuticals. In, vol. 2009 2009.
111. ICH EWG. Validation of analytical procedures: text and methodology Q2(R1), ICH Harmonised Tripartite Guideline. In: International conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use 2005.
112. Goodsaid F, Frueh F. Process map proposal for the validation of genomic biomarkers. Pharmacogenomics 2006; 7: 773-782.
113. 314.510 C. Approval based upon a surrogate endpoint. In.
114. FDA UFDA. Table of valid genomic biomarkers in teh context of approved drug labels. In, vol. 2009: US Dept Health Human Services 2009.

115. Lesko LJ, Atkinson AJ, Jr. Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies. Annu Rev Pharmacol Toxicol 2001; 41: 347-366.
116. FDA UFDA. Drug-diagnostic co-development concept paper. 2005.
117. Lane NE, Nevitt MC, Genant HK, Hochberg MC. Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. J Rheumatol 1993; 20: 1911-1918.
118. Marshall KW, Zhang H, Yager TD, Nossova N, Dempsey A, Zheng R, et al. Blood-based biomarkers for detecting mild osteoarthritis in the human knee. Osteoarthritis Cartilage 2005; 13: 861-871.
119. Krasnokutsky S, Attur M, Belitskaya-Levy I, Greenberg J, Samuels J, Smiles S, et al. Peripheral blood leukocyte gene expression as biomarkers of disease in kneee osteoarthritis (OA). Arthritis Rheum 2007; 56: S355:846.
120. Schrohl AS, Wurtz S, Kohn E, Banks RE, Nielsen HJ, Sweep FC, et al. Banking of biological fluids for studies of disease-associated protein biomarkers. Mol Cell Proteomics 2008; 7: 2061-2066.
121. Gerlach RF, Uzuelli JA, Souza-Tarla CD, Tanus-Santos JE. Effect of anticoagulants on the determination of plasma matrix metalloproteinase (MMP)-2 and MMP-9 activities. Anal Biochem 2005; 344: 147-149.
122. Kruhoffer M, Dyrskjot L, Voss T, Lindberg RL, Wyrich R, Thykjaer T, et al. Isolation of microarray-grade total RNA, microRNA, and DNA from a single PAXgene blood RNA tube. J Mol Diagn 2007; 9: 452-458.
123. Jantos-Siwy J, Schiffer E, Brand K, Schumann G, Rossing K, Delles C, et al. Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. J Proteome Res 2009; 8: 268-281.
124. Kraus VB, Huebner JL, Fink C, King JB, Brown S, Vail TP, et al. Urea as a passive transport marker for arthritis biomarker studies. Arthritis Rheum 2002; 46: 420-427.
125. Seifer DR, Furman BD, Guilak F, Olson SA, Brooks SC, 3rd, Kraus VB. Novel synovial fluid recovery method allows for quantification of a marker of arthritis in mice. Osteoarthritis Cartilage 2008; 16: 1532-1538.
126. Poole AR, Oldham G, Coombs RR. Early rheumatoid-like lesions in rabbits injected with foreign serum: relationship to localization of immune complexes in the lining tissues of joints and cellular content of synovial fluid. Int Arch Allergy Appl Immunol 1978; 57: 135-145.
127. Babich H, Borenfreund E. Cytotoxicity of T-2 toxin and its metabolites determined with the neutral red cell viability assay. Appl Environ Microbiol 1991; 57: 2101-2103.
128. Huebner JL, Hanes MA, Beekman B, TeKoppele JM, Kraus VB. A comparative analysis of bone and cartilage metabolism in two strains of guinea-pig with varying degrees of naturally occurring osteoarthritis. Osteoarthritis Cartilage 2002; 10: 758-767.
129. Huebner JL, Kraus VB. Assessment of the utility of biomarkers of osteoarthritis in the guinea pig. Osteoarthritis Cartilage 2006; 14: 923-930.