Seminars in Arthritis and Rheumatism ▮ (2017) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

Choice of knee cartilage thickness change metric for different treatment goals in efficacy studies

Robert J. Buck, PhD[a,*], Marie-Pierre Hellio Le Graverand, MD[b], Wolfgang Wirth, PhD[c,d], Felix Eckstein, MD[c,d]

[a] StatAnswers Consulting LLC, 408 N 1st St 706, Minneapolis, MN 55401, USA
[b] Pfizer Japan Inc, Tokyo, Japan
[c] Institute of Anatomy, Paracelsus Medical University, Salzburg, Austria
[d] Chondrometrics GmbH, Ainring, Germany

ARTICLE INFO

Keywords:
Knee cartilage
Biomarker
Clinical trial

ABSTRACT

Introduction: In knee osteoarthritis, local increase and decrease in cartilage thickness has been observed over short time intervals. Hence, averaging cartilage change across large regions may not capture the complexity of structural alterations in disease progression. This study aims to examine the relative performance of different metrics of cartilage thickness change for different clinical studies scenarios.
Materials and methods: Metrics for assessing cartilage thickness change were characterized by conventional measures of change versus absolute values (the magnitude) of change, and by different methods of summarizing change over (sub-) regions. Sample sizes for these metrics were derived for 6–24-month observation periods, and for different treatment efficacies. Treatment effects were derived from an observational trial with 6-, 12-, and 24-month follow-up, ranging from slowing cartilage loss to stimulating cartilage growth.
Results: Projected sample sizes ranged from 10 to $>$ 10,000 patients/arm (median = 164), depending on metric choice, treatment efficacy, and observation period. The smallest sample sizes for metrics using magnitude of change typically were half the size of those using conventional measures of change. Extreme values, e.g., minimum change or average of last four-ordered values of absolute change, required smaller sample sizes than metrics averaging over one or more regions.
Conclusions: Metrics using extreme magnitudes of change were most efficient in detecting differences between treatment and placebo, i.e., involved the smallest sample sizes across different DMOAD study lengths and treatment efficacies. Ancillary metrics can be used to clarify whether differences between treatment and placebo indicate structural benefit when needed.

Introduction

Osteoarthritis is the most common form of arthritis, and has a large socio-economic impact on health care and society [1]. There exists a large unmet medical need in developing drugs that can prevent the progression of structural pathology in knee osteoarthritis (OA) in conjunction with providing clinically relevant benefits. However, such disease-modifying OA drugs (DMOAD) have not yet been approved by the regulatory agencies.

Cartilage loss is an important characteristic of knee osteoarthritis (OA) progression and is often viewed as a slow chronic monotonic process. While cartilage loss appears to be the overall long-term trend in disease progression, it has been previously shown that knee cartilage not only experiences periods of (rapid) loss, but also stasis or even (rapid) thickening [2–5].These findings appear to agree with observations made in animal models of knee OA [6–9] in which cartilage thickening has been described as hypertrophy or swelling. In clinical studies, thickening and thinning have been observed in subregions over time courses as short as 3–6 months in some participants [5]. Thickening and thinning have been observed to be spatially localized, i.e., to occur in certain subregions only, with locations varying greatly between individuals [3,4]. Hence it appears that, in the short term, cartilage change is highly heterogeneous, both temporally and spatially, and cannot be described as a simple monotonic process.

The presence of thinning and thickening in the disease process creates challenges in developing imaging biomarker metrics for cartilage thickness change. At any given time, different study

* Corresponding author.
E-mail address: robert.j.buck@gmail.com (R.J. Buck).

http://dx.doi.org/10.1016/j.semarthrit.2017.05.012
0049-0172/

participants are likely to be at different stages of disease, so clinical studies are likely to include subjects undergoing cartilage thickening as well as subjects undergoing cartilage thinning in various subregions of the knee. Disease-related cartilage thickening (e.g., edema or pathologic hypertrophy) impacts the ability to observe treatment benefit either by direct confounding with cartilage growth, or by masking cartilage thinning in other subjects or knee subregions when calculating average thickness (or volume) change. Hence, heterogeneous subregional cartilage thickening and thinning across patients has a significant impact on designing a clinical study with the goal of assessing cartilage thickness change with and without administration of a disease-modifying OA drug (DMOAD). Two options may be considered when trying to increase the efficiency of clinical studies: (a) find predictors of immediate progression patterns, so recruited patients will mostly be at the same stage of progression, or (b) consider other metrics to improve efficiency in detecting a structural treatment benefit. As the first option has proven challenging [10], the latter objective is the focus of the current study.

When both cartilage thinning and thickening are present, alternative measures to averaging thickness (or volume) change over all subregions and subjects should be considered [11]. One option is to analyze thinning and thickening separately. The use of ordered values has been proposed to organize observed changes by direction and magnitude, i.e., subregional cartilage changes are sorted based on the amount (negative or positive) of change independent of where they occur in an individual joint [11,12]. Ordered values were shown to more efficiently discriminate between rates of cartilage thickness change in knees at different radiographic stages of osteoarthritis than region-specific analysis [12,13]; they also were shown to be more efficient in detecting risk factors of structural progression than region-specific measurement [14], and were able to detect structural differences in cartilage loss between DMOAD-treated and placebo-treated study participants that were not detected by either radiography (JSW loss) or by region-based analysis of cartilage thickness loss from MRI [11,15]. Determining whether analysis of cartilage thinning or thickening is more efficient can be made by assessing which metric leads to more efficient clinical trial design, i.e., a smaller sample size. Alternatively, if both thickening and thinning are part of the disease process, treatment optimally would slow both thinning and thickening; although ultimately cartilage response to treatment depends on the biological mechanism of the drug candidate. Choosing the absolute value of change, $|\Delta Th|$, which examines the magnitude of change without considering the direction of change, as an endpoint, the hypothesis would be to see whether smaller changes occurred in the treated group than in the placebo/best of care group. Hence a metric such as the absolute value of ordered values may be a more efficient metric of disease progression than average change or ordered values of thinning or thickening separately, particularly if treatment only slows or stops progression and does not stimulate cartilage hypertrophy.

In the context of clinical studies, the following are 3 important characteristics of a good biomarker:

1. It measures the treatment effect as directly as possible.
2. It is efficient, i.e., requires small sample sizes.
3. It is universal, i.e., it works for all treatment effects and observation periods.

While this last characteristic is not necessary, it helps make an ideal biomarker as it implies that one does not need to be confident of the exact nature of the treatment effect before the clinical trial.

The goal of the current study was thus to examine the efficiency of different metrics for cartilage thickness change in detecting the efficacy of DMOADs under different treatment scenarios in a clinical trial, in particular with regard to regrowth versus slowing/stopping disease progression, and with regard to shorter versus longer observation periods. Specific questions of interest examined were as follows:

- Which specific metrics for assessing cartilage thickness change provide the smallest sample sizes under different treatment scenarios and observation periods?
- Are these metrics robust across treatment scenarios and observation periods?

Materials and methods

The parameter sample size per arm (N) of a simple clinical trial (treatment—control in a two-sample t-test with equal sample sizes for each arm) was used to assess the efficiency of different treatment effects, observation periods, and various biomarker choices for cartilage thickness change that will be described below. All cartilage thickness change biomarkers considered were univariate endpoints, i.e., they summarized change from multiple regions, if necessary, into a single value per knee.

The cartilage thickness change observed in an ROA sample at 6, 12, and 24 months follow-up was derived from coronal 3 Tesla MR images of 71 female study participants with medial radiographic osteoarthritis (ROA) with Kellgren-Lawrence Grades (KLG) of 0 (n = 1), 1, (n = 7), 2 (n = 26), and 3 (n = 37) adjudicated after recruitment, and average WOMAC scores for pain = 5.8 (of 20), stiffness = 2.8 (of 8), and function = 20.4 (of 68) [16,17]. Projected effects for a treatment cohort were derived from the differences of cartilage thickness change observed between the ROA and an asymptomatic cohort in the same study consisting of 74 female study participants without knee pain or any sign of radiographic knee OA (all adjudicated KLG = 0, average WOMAC scores for pain = 0.2, stiffness = 0.1, and function = 0.7). In both cohorts, knee cartilage thickness was determined for 5 tibial and 3 femoral subregions in both the medial and lateral compartments for 16 subregions in total [18]. Cartilage thickness change was annualized for each time interval and then normalized by dividing by the standard deviation of annualized change found in the asymptomatic cohort for each subregion [19]. The difference in normalized cartilage thickness change between the ROA (e.g., placebo treated) vs. asymptomatic (e.g., simulated DMOAD treated) cohort for different knee subregions ranged from, −0.23 to 0.38 at month 6, −0.48 to 0.35 at month 12, and −0.80 to 0.32 at Month 24, with more details having been reported previously [12–14].

The observed change in cartilage thickness in the ROA cohort was defined as the Placebo response. The sample standard deviation that was assumed to characterize population variability in cartilage thickness change in the projected clinical trial, i.e., the standard deviation used in calculating the effect size for the presumed study, was the observed sample standard deviation in the ROA cohort. Four treatment effects were considered:

1) A 50% reduction in disease progression, i.e., half the difference in average metric value between ROA and asymptomatic cohorts.
2) 100% reduction, i.e., the full difference in average metric value between ROA and asymptomatic cohorts.
3) A uniform increase in cartilage thickness of 0.5%/y (see below).
4) A uniform increase in cartilage thickness of 2%/y (see below).

Cartilage thickness increase in a subregion was defined as the sum of the observed normalized rate of change and the percent cartilage thickness increase, i.e., 0.5%/y or 2%/y, relative to baseline thickness in asymptomatic subjects. All subregions were