OsteoArthritis and Cartilage (2006) **14,** A112–A115

doi:10.1016/j.joca.2006.02.027

Osteoarthritis and Cartilage

International Cartilage Repair Society

OARSI OSTEOARTHRITIS RESEARCH SOCIETY INTERNATIONAL

Responsiveness, effect size, and smallest detectable difference of Magnetic Resonance Imaging in knee osteoarthritis

D. J. Hunter M.B.B.S., Ph.D., F.R.A.C.P.†*, P. G. Conaghan M.B.B.S., Ph.D., F.R.A.C.P., F.R.C.P.‡, C. G. Peterfy M.D., Ph.D.§, D. Bloch Ph.D.‖, A. Guermazi M.D.§, T. Woodworth M.D.¶, R. Stevens M.D.# and H. K. Genant M.D.§††

† *Clinical Epidemiology Research and Training Unit, Boston University School of Medicine, Boston, MA, USA*
‡ *Academic Unit of Musculoskeletal Disease, University of Leeds, Leeds, W Yorkshire, UK*
§ *Synarc Inc., San Francisco, CA, USA*
‖ *Department of Health Research and Policy, Stanford University Medical Center, Stanford, CA, USA*
¶ *Novartis, Basel, Switzerland*
# *Hoffman-LaRoche Inc., Nutley, NJ, USA*
†† *Department of Radiology, University of California, San Francisco, USA*

Summary

Objective: The aim of this study was to determine the responsiveness, effect size (ES) and smallest detectable difference (SDD) of two Magnetic Resonance Imaging (MRI) measures for osteoarthritis (OA) of the knee: a whole-organ semiquantitative evaluation and cartilage volume.

Design: This analysis was performed on a dataset from a randomized, double-blind trial (Roche NI-15713) conducted in 1998 of a novel therapy in subjects with mild-moderate knee OA, with MRI at baseline and 6-month follow-up. The trial measurements included (1) cartilage volume measured using a proprietary software method; and (2) semiquantitative scoring of other parameters important for ''whole organ'' evaluation of OA knee joint pathology, using the Whole-Organ MRI Score (WORMS). The analysis initially examined the distributional characteristics of WORMS items, such as cartilage morphology. Standardized response mean (SRM), ES, and SDD between baseline and 6-month follow-up were then calculated in the whole group and the placebo group alone.

Results: In general, the differences were small and this was reflected in the small ESs and SRMs. There was also a suggestion of a treatment effect with reduction in differences between baseline and follow-up in the treatment group.

Conclusion: Of the MRI semiquantitative measures, cartilage morphology, synovitis and osteophytes appeared to be responsive to change and the focus of repeat measures should highlight these articular features. In general, the ESs and SRMs were small.

Key words: Osteoarthritis, Knee, MRI, Whole organ, Responsiveness, Effect size.

Magnetic Resonance Imaging (MRI) is being developed as a measurement tool in knee osteoarthritis (OA) for both longitudinal studies and clinical trials. There is increasing information on how the measures obtained are responsive alone or in combination, how much change one might expect per unit of time, and the smallest amount of change that is real, as distinct from measurement error. On December 5 and 6, 2002, OMERACT (Outcome Measures in Rheumatology Clinical Trials) and OARSI (Osteoarthritis Research Society International), with support from various pharmaceutical companies listed at the beginning of this supplement, held a Workshop for Consensus on Osteoarthritis Imaging in Bethesda, MD. The overall aim of the workshop was to provide a state-of-the-art review of imaging outcome measurement in OA to help guide scientists and pharmaceutical companies who want to use MRI in multi-site studies of knee OA. Applications of MRI were initially reviewed by a multidisciplinary, international panel of expert scientists and physicians from academia, the pharmaceutical industry and regulatory agencies. The panel was co-chaired by Charles Peterfy, M.D., Ph.D. (Synarc, Inc.) and Roy Altman, M.D. (University of Miami, Miami, FL, USA[a]) and also included Deborah Burstein, Ph.D. (Harvard-MIT, Cambridge, MA, USA), Flavia Cicuttini (Epidemiology, Monash University, Prahran, Australia), Gary Cline, Ph.D. (Biostatistics, Proctor & Gamble), Philip Conaghan, M.B.B.S., F.R.A.C.P. (Rheumatology, Leeds University, Leeds, UK), Bernard Dardzinski, Ph.D. (MRI Physics, University of Cincinnati, Cincinnati, USA), Felix Eckstein, Ph.D. (MR image analysis, Ludwig-Maximilians-Universität, München, Germany[b]), David Felson, M.D., M.P.H. (Rheumatology, Boston University, Boston, MA, USA), Garry Gold, M.D.,

*Address correspondence and reprint requests to: David J. Hunter, Clinical Epidemiology Research and Training Unit, Boston University School of Medicine, 715 Albany Street, Room A203, Boston, MA 02118-2526, USA. Tel: 1-617-638-5180; Fax: 1-617-638-5239; E-mail: djhunter@bu.edu
Received 26 February 2006; revision accepted 26 February 2006.

[a]Dr Altman is currently at UCLA, CA, USA.
[b]Dr Eckstein is currently at Paracelsus Private Medical University Salzburg, Austria & Chondrometrics GmbH, Munich, Germany.

Ph.D. (Radiology, Stanford University, Stanford, CA, USA), Benjamin Hsu, Ph.D. (GlaxoSmithKlein[c]), Marissa Lassere, M.B.B.S., Ph.D., F.R.A.C.P. (Epidemiology, St George Hospital, Kogarah, Australia), Stefan Lohmander, M.D., Ph.D. (Orthopaedics, University of Lund, Lund, Sweden), Jean-Pierre Raynauld, M.D. (Rheumatology, University of Montreal and Arthrovision, Montreal, PQ, Canada), Randall Stevens, M.D. (Hoffman-LaRoche Inc., Nutley, NJ, USA), Saara Totterman, M.D., Ph.D. (VirtualScopics, Pittsford, NY, USA), James Witter, M.D. (Food and Drug Administration (FDA), Washington, DC, USA), and Thasia Woodworth, M.D. (Pfizer, Groton, CT, USA[d]). The panel met in New Orleans, LA on October 29, 2002 prior to the Workshop in Bethesda to define a preliminary set of MRI features to include in whole-organ assessment of the knee[1] and to review the relative strengths and weaknesses of various imaging protocols for multi-feature, multi-site MRI. The findings of the panel were presented to the participants of the Workshop in Bethesda for open discussion. In addition, data sets from previous clinical trials and epidemiological studies of OA were analyzed with respect to the metrological properties of the methods employed. One of these analyses examined the responsiveness, effect size (ES) and smallest detectable difference (SDD) of two increasingly used methods for assessing OA knee abnormalities and progression: a whole-organ semiquantitative scoring system and cartilage volume using data from a clinical trial of OA. The results of this analysis were presented to the participants of the Workshop in Bethesda for discussion. This report summarizes this analysis.

Methods

This analysis was conducted on data kindly made available by the investigators and management of Roche, from a multicenter, prospective, randomized, placebo-controlled, five arm, parallel group, dose-ranging trial (NI-15713) of 24 weeks duration including methods development for assessing OA joints. A total of 504 subjects recruited from 58 clinical sites in 1998 were randomly assigned to receive one of four doses of test drug or matching placebo tablets. Ambulatory subjects of either sex with knee OA who met the following clinical criteria were eligible: age 45–85 years, at least moderate knee pain on the Western Ontario and McMaster Universities (WOMAC) scale on at least 15 days in the previous month, and either stiffness <30 min or crepitus on motion of the target knee. At screening the target knee must have met all of the following radiographic criteria by the OARSI Atlas[1] using fixed-flexion radiography[2] with a positioning frame: grade 1 or 2 joint-space narrowing (JSN) of the medial tibiofemoral compartment, joint-space width (JSW) ≥ 1.5 mm of both the medial and lateral tibiofemoral compartments, and grade 1, 2 or 3 osteophytes. The mean age of the subjects was 60.6 years, with 72% 65 years of age or younger, and 72% were women. The mean body mass index (BMI) was 30.0, with 35% of the patients having a BMI less than 28. Sixty percent of the patients entering the study used nonsteroidal anti-inflammatory drugs (NSAIDs) on a regular basis for the treatment of OA.

MRI of a target knee was performed at baseline in the 504 patients originally enrolled. A total of 150 of these patients also received a 24-week follow-up MRI examination of the same knee before the study was terminated. Only data from baseline images of these 150 patients were included in this analysis. MRI was performed with 1.5 T whole body scanners using a circumferential extremity coil. Imaging included axial T2-weighted fast spin echo (FSE) through the patella (repetition time (TR) = 3500 ms; echo time (TE) = 60 ms, echo-train length (ETL) = 16, field of view (FOV) = 12 cm, slice thickness = 3 mm with no interslice gap, matrix = 256 × 256 pixels, anterior–posterior frequency encoding, two excitations); coronal T2-weighted FSE with spectral fat suppression (TR = 3500 ms, TE = 60 ms, ETL = 8, FOV = 12 cm, slice thickness = 3 mm with no interslice gap, matrix = 256 × 256 pixels, superior–inferior frequency encoding, two excitations); sagittal dual-echo FSE (TR = 3500 ms, TE = 20 ms and 60 ms, ETL = 8, FOV = 12 cm, slice thickness = 3 mm with no interslice gap, matrix = 256 × 256 pixels, superior–inferior frequency encoding, two excitations, wide superior and inferior external saturation bands to limit vascular pulsation artifacts); sagittal multi-echo spin-echo with spectral fat suppression (TR = 2500 ms, TE = 15 ms, 30 ms, 45 ms and 60 ms, FOV = 12 cm, slice thickness = 4 mm with no interslice gap, matrix = 256 × 160 pixels, 16 kHz bandwidth, anterior–posterior frequency encoding, one excitation, wide superior and inferior external saturation bands to limit vascular pulsation artifacts) for measuring T2 relaxation time of the articular cartilage; and sagittal T1-weighted three dimensional (3D) spoiled gradient echo with spectral fat suppression (TR = 58 ms, TE = 6 ms, flip angle = 40°, FOV = 12 cm, 60 contiguous 2-mm slices, matrix = 256 × 192 pixels, superior–inferior frequency encoding, one excitation, wide superior and inferior external saturation bands to limit vascular pulsation artifacts).

Radiographic and MRI protocol design, site training, imaging quality control and image analysis were performed by a central radiology service (Synarc, Inc.). The volumes of articular cartilage over the patella, femur, medial tibia and lateral tibia were quantified using a previously described method [add reference:] employing semiautomated cartilage segmentation with a seed-growing algorithm[3]. Semiquantitative scoring of the MR images using the Whole-Organ MRI Score (WORMS)[4] was performed by radiologists experienced in the method and a Sun Workstation equipped with MRVision software (MRVision Inc., Menlo Park, CA, USA). WORMS examines 14 features at multiple intra-joint sites. The cartilage, bone (edema, attrition and cysts) and osteophyte abnormalities are assessed in 15 different regions including the medial and lateral patellar facets, the medial and lateral femoral condyles and tibial plateaus (each divided into anterior, central and posterior sections) and the subspinous region of the tibial plateaus (the latter is not included in cartilage assessment). Synovial cavity distension is scored without distinction between synovial hypertrophy and effusion. Periarticular cysts and bursae (popliteal, meniscal, tibiofibular, anserine, infra-patellar and pre-patella) are also recorded. These features are scored with semiquantitative scales varying from 0–1 to 0–7. For example, a 0–1 scale indicates normal or abnormal whereas a 0–3 scale represents normal, mild,

Table I
Demographic characteristics of study population

	Mean (SD)	Range
Age (years)	58.9 (8.6)	44–81
BMI	30.7 (4.6)	21–39.8
Female (%)	72	

[c]Dr Hsu is currently at Centocor, Philadelphia, PA, USA.
[d]Dr Woodworth is currently at Novartis, Basel, Switzerland.

Table II
Cartilage morphology

	Lateral FTJ		Medial FTJ		PFJ		SUM	
Possible range	0–30		0–30		0–24		0–84	
Baseline mean (SD)	1.63 (3.78)		9.29 (6.21)		7.58 (6.05)		19.45 (11.82)	
Baseline range	0–30		0–30		0–22		0–74	
	Plac	Treat	Plac	Treat	Plac	Treat	Plac	Treat
Mean difference	−0.2	−0.2	−2.1	−0.9	−1.4	−0.4	−3.7	−2.3
$SD_{Difference}$	1.3	1.5	4.1	2.6	3.2	1.9	8.1	6.3
SRM	−0.18	−0.12	−0.50	−0.36	−0.43	−0.23	−0.46	−0.37
ES	−0.09	−0.04	−0.28	−0.16	−0.22	−0.07	−0.28	−0.21

FTJ = femorotibial joint and PFJ = patellofemoral joint.

Table III
Bone marrow edema

	Lateral FTJ		Medial FTJ		SS		PFJ		SUM	
Possible Range	0–15		0–15		0–3		0–12		0–45	
Baseline mean (SD)	0.12 (0.72)		0.48 (1.21)		0.19 (0.61)		0.78 (1.47)		1.57 (2.31)	
Baseline range	0–7		0–7		0–4		0–12		0–16	
	Plac	Treat	Plac	Treat	Plac	Treat	Plac	Treat	Plac	Treat
Mean difference	−0.03	0.01	−0.10	−0.07	0.04	0.03	−0.20	0.05	−0.29	0.02
$SD_{Difference}$	0.30	0.39	0.51	0.66	0.59	0.54	1.14	0.61	1.58	1.19
SRM	−0.10	0.03	−0.20	−0.10	−0.07	0.05	−0.18	0.08	−0.18	0.02
ES	−0.08	0.01	−0.09	−0.05	−0.10	0.04	−0.09	0.05	−0.10	0.01

FTJ = femorotibial joint, PFJ = patellofemoral joint, and SS = subspinous region.

Table IV
Marginal osteophytes

	Lateral FTJ		Medial FTJ		PFJ		SUM	
Possible range	0–35		0–35		0–28		0–98	
Baseline mean (SD)	3.5 (4.8)		7.2 (6.4)		6.4 (6.5)		17.5 (15.9)	
Baseline range	0–26		0–30		0–25		0–77	
	Plac	Treat	Plac	Treat	Plac	Treat	Plac	Treat
Mean difference	−0.7	−0.3	−1.3	−0.8	−0.4	−0.5	−2.0	−1.9
$SD_{Difference}$	2.7	1.1	3.3	2.5	0.9	1.8	6.2	6.0
SRM	−0.27	−0.27	−0.38	−0.32	−0.40	−0.27	−0.33	−0.32
ES	−0.11	−0.07	−0.17	−0.13	−0.07	−0.07	−0.12	−0.12

FTJ = femorotibial joint and PFJ = patellofemoral joint.

moderate or severe abnormality. These scores can been aggregated by feature or compartment to produce medial tibiofemoral, lateral tibiofemoral and patellofemoral compartment scores as well as total knee scores[1].

STATISTICAL METHODS

The analysis initially examined the distributional characteristics of WORMS items, e.g. cartilage morphology. The following measures were then calculated in the treated group and the placebo group alone:

1. Standardized response mean (SRM) − responsiveness (relative ease of detecting change). Difference/Standard Deviation (SD) Difference;
2. ES − standardized measure of the amount of change. Difference/SD Baseline;
3. SDD − minimum amount of an observed change in a single measure that represents a real change, as distinct from noise. (1.96 × SD of observation difference).

The difference was between baseline and 6-month follow-up.

Table V
Synovitis/effusion

Baseline mean (SD)	1.25 (1.01)	
Baseline range	0–6	
	Plac	Treat
Mean difference	−0.28	−0.17
$SD_{Difference}$	0.77	0.73
SRM	−0.36	−0.24
ES	−0.24	−0.18

Table VI
Cartilage volume (mL)

	Femur		Patella		Lateral Tibia		Medial Tibia	
Baseline mean (SD)	7.4 (1.8)		2.9 (0.9)		2.4 (0.7)		1.8 (0.5)	
Baseline range	3.9–13.9		0.1–6.3		0.7–4.6		0.8–3.7	
	Plac	Treat	Plac	Treat	Plac	Treat	Plac	Treat
Mean difference	0.02	0.07	−0.04	−0.02	−0.04	0.01	−0.01	−0.00
$SD_{Difference}$	0.41	0.45	0.21	0.34	0.14	0.18	0.21	0.22
SRM	0.06	0.16	−0.20	−0.05	−0.27	0.05	−0.05	−0.02
ES	0.01	0.04	−0.04	−0.02	−0.06	0.01	−0.02	−0.01
SDD	1.3	0.8	1.1	0.7	0.5	0.6	0.7	0.4

Results

The subject characteristics are displayed in Table I.

The following series of tables displays the SRM, ES and SDD for cartilage morphology (Table II), bone marrow edema (Table III), marginal osteophytes (Table IV), synovitis/effusion (Table V) and cartilage volume (Table VI). The data from the remaining WORMS parameters (cartilage signal, subarticular cysts, attrition, meniscus and ligament scores) are not included here, as these measures were less responsive than the other WORMS items in this study. The SDD for all WORMS features is not included, as there was insufficient variability in the baseline measures to make meaningful conclusions from this study.

In general the differences are small and this is reflected in the small ESs and SRMs. There is also a suggestion of a treatment effect with reduction in differences between baseline and follow-up in the treatment group.

Discussion

Of the MRI measures cartilage morphology, synovitis and osteophytes appeared to be responsive to change, and the focus of repeat measures should highlight these articular features. In general, the ESs and SRMs were small. This may reflect a small real difference, given the short duration of the study, and/or limitations in responsiveness of the measures used. The limitations of this study include the short duration, which limits our ability to detect change, and that the treatment may have had a measurable influence on the amount of change. Additionally, the limited number of repeat observations at baseline inhibited our ability to make meaningful conclusions about the SDDs.

Acknowledgments

We are grateful to Roche and the MMPI study investigators for supplying the raw data for this exercise, and to the team at Boston University Clinical Epidemiology and Research Training Unit for assistance in analysis and result synthesis.

References

1. Altman RD, Hochberg M, Murphy WA, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthritis Cartilage 1995;3(Suppl A):3–70.
2. Peterfy C, Li J, Zaim S, Duryea J, Lynch L, Miaux Y, *et al*. Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-space width in the knee: test–retest reproducibility. Skeletal Radiol 2003;32:128–32.
3. Peterfy CG, van Dijke CF, Janzen DL, Glüer C, Namba R, Majumdar S, *et al*. Quantification of articular cartilage in the knee by pulsed saturation transfer and fat-suppressed MRI: optimization and validation. Radiology 1994;192:485–91.
4. Peterfy CG, Guermazi A, Zaim S, Tirman PFJ, Miaux Y, White D, *et al*. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. Osteoarthritis Cartilage 2004;12:177–90.