failure, three moved to distant areas and three lost contact. Twelve other patients who showed symptoms in the lower extremities due to cerebral infarction, myelopathy or dementia during the postoperative follow-up period were excluded. Symptoms in the lower extremities of the remaining 89 patients (56 males and 33 females; mean ± SD, 66.3 ± 11.2 years) were surveyed 2 years after surgery. There was no complication in the surgical procedure except for slight dural tears in four patients, which were repaired without additional treatment. During the follow-up period, a superficial infection, a pseudomembranous enteritis, a disc herniation at the operated level and a compression vertebral fracture occurred, all of which were cured with conservative therapies. None of the patients underwent spinal re-operation because of progression of stenosis or instability.

The background data of the 89 patients are shown in Table 1. Comparison of preoperative and postoperative JOA scores on symptoms in the lower extremities of all the patients revealed that both leg pain/numbness (1.0–2.0) and gait disturbance (0.7–2.4) were significantly improved by surgery (Table 1). The stratified comparisons by gender, preoperative presence of the above findings and the number of decompressed levels showed that the JOA scores of both symptoms were significantly improved by the surgery in all subgroups. However, the subgroup with preoperative drop foot showed somewhat less improvement in both leg pain/numbness ($P = 0.009$) and gait disturbance ($P = 0.007$) than other subgroups ($P < 0.0001$).

Predictors of the residual symptoms in lower extremities

According to the definition of residual symptoms as above, 27 (30.3%) and 13 (14.6%) patients showed residual leg pain/numbness and gait disturbance, respectively (Table 2). To identify the predictors of residual symptoms in the lower extremities, we compared the number (percentage) of patients with and without residual symptoms in the stratified subgroup according to the variables. Among the variables, preoperative resting numbness was positively associated with both residual leg pain/numbness ($P = 0.03$) and residual gait disturbance ($P = 0.02$). Furthermore, preoperative drop foot was more strongly associated with residual gait disturbance ($P = 0.0002$), although not with residual leg pain/numbness. Age, gender, preoperative presence of cauda equina syndrome, degenerative spinal deformity, myelographic complete filling defect or the number of decompressed levels was not associated with either of the residual symptoms in the lower extremities.

To further identify the principal predictors, we further performed logistic regression analysis after adjustment for age and gender to estimate OR and 95% CI. We confirmed the significant association of resting numbness with residual leg pain/numbness and gait disturbance, as well as the significant association of drop foot with residual gait disturbance (Table 3).

Table 1 Comparison of preoperative and postoperative JOA scores on symptoms in lower extremities

		n	Leg pain/numbness				Gait disturbance			
			Preop. (SD)	Postop. (SD)	Change (SD)	*P* value	Preop. (SD)	Postop. (SD)	Change (SD)	*P* value
All		89	1.0 (0.5)	2.0 (0.8)	10 (0.8)	<0.0001	0.7 (0.8)	2.4 (0.8)	1.8 (1.0)	<0.0001
Gender	Male	56	1.0 (0.6)	2.1 (0.8)	1.1 (0.8)	<0.0001	0.7 (0.8)	2.4 (0.8)	1.7 (0.9)	<0.0001
	Female	33	0.9 (0.4)	1.8 (0.8)	0.9 (0.8)	<0.0001	0.6 (0.8)	2.4 (0.8)	1.8 (1.0)	<0.0001
Resting numbness	(+)	40	0.9 (0.5)	1.4 (0.7)	0.6 (0.7)	<0.0001	0.6 (0.7)	2.2 (0.9)	1.6 (1.0)	<0.0001
	(−)	49	1.1 (0.6)	2.5 (0.5)	1.4 (0.6)	<0.0001	0.8 (0.8)	2.7 (0.6)	1.9 (0.9)	<0.0001
Drop foot	(+)	9	0.9 (0.8)	1.9 (0.6)	1.0 (0.9)	0.009	0.6 (0.9)	1.9 (1.2)	1.3 (1.1)	0.007
	(−)	80	0.9 (0.5)	2.0 (0.8)	1.0 (0.8)	<0.0001	0.7 (0.8)	2.5 (0.7)	1.8 (0.9)	<0.0001
Cauda equina syndrome	(+)	66	0.9 (0.5)	1.9 (2.2)	1.0 (0.8)	<0.0001	0.6 (0.7)	2.4 (0.8)	1.8 (0.9)	<0.0001
	(−)	23	1.0 (0.7)	2.2 (0.9)	1.0 (0.8)	<0.0001	0.9 (0.9)	2.4 (0.7)	1.5 (1.0)	<0.0001
Degenerative spinal deformity	(+)	47	1.0 (0.6)	2.0 (0.8)	1.0 (0.8)	<0.0001	0.7 (0.8)	2.4 (0.7)	1.7 (0.9)	<0.0001
	(−)	42	0.9 (0.4)	2.0 (0.9)	1.1 (0.9)	<0.0001	0.6 (0.7)	2.5 (1.0)	1.9 (1.0)	<0.0001
Complete filling defect	(+)	56	0.9 (0.6)	1.9 (0.8)	1.0 (0.8)	<0.0001	0.7 (0.8)	2.4 (0.8)	1.7 (1.0)	<0.0001
	(−)	33	1.0 (0.5)	2.2 (0.9)	1.2 (0.7)	<0.0001	0.7 (0.8)	2.5 (0.8)	1.8 (0.9)	<0.0001
Number of decompressed levels	1	50	1.0 (0.5)	2.1 (0.7)	1.1 (0.7)	<0.0001	0.9 (0.8)	2.6 (0.6)	1.7 (0.8)	<0.0001
	≥2	39	0.9 (0.6)	1.9 (0.8)	0.9 (0.9)	<0.0001	0.4 (0.7)	2.2 (0.9)	1.8 (1.2)	<0.0001

P value was determined by the paired *t* test

Table 2 Number (percentage) of patients with and without residual symptoms in the lower extremities

		Leg pain/numbness			Gait disturbance		
		(+) $n = 27$	(−) $n = 62$	P value	(+) $n = 13$	(−) $n = 76$	P value
Mean age (years)		72.0	68.1	0.10	64.6	69.1	0.10
Gender	Male	17 (30.4)	39 (69.6)	0.26	10 (17.9)	46 (82.1)	0.26
	Female	10 (30.3)	23 (69.7)		3 (9.0)	30 (91.0)	
Resting numbness	(+)	26 (65.0)	14 (35.0)	0.03*	9 (22.5)	31 (77.5)	0.02*
	(−)	1 (2.0)	48 (98.0)		4 (8.2)	45 (91.8)	
Drop foot	(+)	4 (44.4)	5 (55.6)	0.33	5 (55.6)	4 (44.4)	0.0002*
	(−)	23 (28.8)	57 (61.2)		8 (10.0)	72 (90.0)	
Cauda equina syndrome	(+)	23 (34.8)	43 (65.2)	0.11	9 (13.6)	57 (86.4)	0.66
	(−)	4 (17.4)	19 (82.6)		4 (17.4)	19 (82.6)	
Degenerative spinal deformity	(+)	17 (36.2)	30 (63.8)	0.60	7 (14.9)	40 (55.1)	0.60
	(−)	10 (23.8)	32 (76.2)		6 (14.3)	36 (55.7)	
Complete filling defect	(+)	21 (37.5)	35 (62.5)	0.06	8 (14.3)	48 (85.7)	0.91
	(−)	6 (18.2)	27 (81.8)		5 (15.2)	28 (84.8)	
Number of decompressed levels	1	11 (22.0)	39 (78.0)	0.05	4 (8.0)	46 (92.0)	0.05
	≥2	16 (41.0)	23 (59.0)		9 (23.1)	30 (76.9)	

P value was determined by the chi-square test

Table 3 Logistic regression analyses for odds ratio (OR) and 95% confidence interval (CI) of the variables for residual symptoms in the lower extremities

	Leg pain/numbness		Gait disturbance	
	OR	(95% CI)	OR	(95% CI)
Resting numbness	85.6*	(15.9–1603.1)	4.5*	(1.2–23.2)
Drop foot	2.1	(0.5–9.0)	11.6*	(2.5–59.1)
Cauda equina syndrome	2.6	(0.09–4.1)	1.3	(0.006–2.5)
Degenerative spinal deformity	0.7	(0.1–4.9)	0.6	(0.1–2.2)
Complete filling defect	2.2	(0.9–2.4)	1.4	(0.004–2.3)
Number of decompressed levels	2.5	(1.0–2.7)	4.2	(0.06–9.8)

Data were calculated by logistic regression analysis after adjustment for age and gender, *$P < 0.01$

Discussion

This prospective observational study for the first time identified the specific predictors for the remaining major symptoms of LSS after decompression surgery: leg pain/numbness and gait disturbance. Preoperative resting numbness was found to be a predictor of both residual leg pain/numbness and gait disturbance, and preoperative drop foot was a predictor of residual gait disturbance. It would seem to be natural that preoperative resting numbness eventually leads to postoperative leg pain/numbness. In fact, 65.0% (26 of 40 patients) of patients with preoperative resting numbness still showed residual leg pain/numbness 2 years after the operation (Table 2). Numbness caused by LSS has been reported to be more difficult to alleviate by surgery than other neurological symptoms such as muscle weakness or pain [2, 6]. Also, it is not surprising that preoperative drop foot eventually leads to postoperative gait disturbance. More than half (55.6%; 5 of 9 patients) the patients with preoperative drop foot showed residual gait disturbance (Table 2). A previous study on the surgical outcome of LSS patients with drop foot revealed that especially those with a preoperative MMT score of 0 or 1 for ankle dorsiflexion exhibited poor alleviation of this disorder [1]. In the present study as well, there were three patients with an MMT score of 0 or 1, and all of them showed residual gait disturbance due to the unchanged drop foot (data not shown). Furthermore, comparison of preoperative and postoperative JOA scores revealed that the subgroup with preoperative drop foot showed less improvement of symptoms in the lower extremities than other subgroups (Table 1). Taken together, resting numbness and drop foot may be derived from less reversible neurological disorders, so that they are difficult to restore by decompression. The preoperative durations of these symptoms may influence the surgical outcomes, which we should examine as a next task. More interesting is that preoperative resting numbness, a sensory disorder, was identified as a predictor of residual gait disturbance, which is a motor disorder. Although the underlying mechanism remains unclear, we speculate a possible involvement of irreversible peripheral neural damage that is related to both resting numbness and gait disturbance.

In addition to resting numbness and drop foot, the number of decompressed levels also showed a trend toward positive association with residual leg pain/numbness and residual gait disturbance, although not statistically significant (Table 2, $P = 0.05$ in both symptoms; Table 3, OR = 2.5 and 4.2, respectively). Although the present comparison was performed between a single level of decompression and two or more levels of decompression, a comparison between one or two levels ($n = 80$) and three or more levels ($n = 9$) showed a significant association of the number of decompressed levels with residual leg pain/numbness ($P = 0.01$, OR = 7.5, 95% CI = 1.6–50.0), but not with residual gait disturbance ($P = 0.50$, OR = 1.3, 95% CI = 0.1–6.3) (data not shown in the tables). Indeed, there is greater possibility of multiple levels of decompression to cause surgical invasion, which may eventually result in the residual symptoms. Alternatively, independently of the surgery itself, residual symptoms may be derived from irreversible symptoms of the preoperative disorders, since the multi-level canal stenosis, which is an indication of multi-level decompression, may cause more damages to the nerve root, cauda equina and the blood circulation [13, 14, 18]. Hence, unlike preoperative symptoms such as resting numbness and drop foot, the number of decompressed levels may not be suitable for the predictor. In fact, previous reports on the relationship between the number of decompressed levels and the surgical outcome have been controversial, depending on the outcome measures including standardized instruments and self-reported satisfaction by patients [7, 9].

As the decompression surgery, the present study utilized our original technique called modified fenestration with restorative spinoplasty. Since this technique was developed to achieve good visibility of the spinal canal and safe decompression even in patients with narrow and steep facet joints, the ability to decompress the spinal canal and nerve roots is sufficient, similar to conventional laminectomy/foraminotomy techniques [11]. Hence, we believe that the present results obtained using this unique technique are applicable generally to typical decompression surgeries.

Although approximately 30 and 15% of patients were shown to have residual symptoms after the decompression surgery, respectively (Table 2), the present residual symptoms were defined according to our original criteria based on the JOA score and were not completely or directly linked with the dissatisfaction of the patients. Indeed, there are other factors such as back pain and psychological status to be considererd for the recommendation of the surgery. Furthermore, since the JOA scores of leg pain/numbness and gait disturbance were significantly improved after the surgery, regardless of the presence or absence of these predictors (Table 1), decompression surgery is definitely worth performing to decrease the severity of these symptoms. Hence, the present study suggests that it is desirable for this surgery to be performed before the onset of resting numbness or drop foot at least to prevent residual symptoms in the lower extremities. Even in the presence of these preoperative predictors, however, we encourage this surgery with sufficient informed consent, including the findings obtained from this study, to avoid misunderstanding or over-expectation of the patient with regard to the surgical outcome.

References

1. Aono H, Iwasaki M, Ohwada T et al (2007) Surgical outcome of drop foot caused by degenerative lumbar diseases. Spine 32:E262–E266
2. Babb A, Carlson WO (2006) Spinal stenosis. S D Med 59(3): 103–105
3. Berthelot JM, Bertrand VA, Rodet D et al (1997) Lumbar spinal stenosis: a review. Rev Rheum Engl Ed 64:315–325
4. Bolender NF, Schonstrom NSR, Spengler DM (1985) Role of computed tomography and myelography in the diagnosis of central spinal stenosis. J Bone Joint Surg 67A:240–246
5. Fujiwara A, Kobayashi N, Saiki K et al (2003) Association of the Japanese Orthopaedic Association score with the oswestry disability index, roland–morris disability questionnare, and short-Form 36. Spine 28:1601–1607
6. Goh KJ, Khalifa W, Anslow P et al (2004) The clinical syndrome associated with lumbar spinal stenosis. Eur Neurol 52:242–249
7. Gunzburg R, Keller TS, Szpalski M et al (2003) Clinical and psychofunctional measures of conservative decompression surgery for lumbar spinal stenosis: a prospective cohort study. Eur Spine J 12:197–204
8. Katz J, Stucki G, Lipson S et al (1999) Predictors of surgical outcome in degenerative lumbar spinal stenosis. Spine 24:2229–2233
9. Katz JN, Lipson SJ, Brick GW et al (1995) Clinical correlates of patient satisfaction after laminectomy for degenerative lumbar spinal stenosis. Spine 20:1155–1160
10. Matsudaira K, Yamazaki T, Seichi A et al (2005) Spinal stenosis in grade I degenerative lumbar spondylolisthesis: a comparative study of outcomes following laminoplasty and laminectomy with instrumented spinal fusion. J Orthop Sci 10:270–276
11. Matsudaira K, Yamazaki T, Seichi A et al (2009) Modified fenestration with restorative spinoplasty for lumbar spinal stenosis. J Neurosurg Spine 10:587–594
12. Mazanec DJ, Podichetty VK, Hsia A (2002) Lumbar canal stenosis: start with nonsurgical therapy. Cleve Clin J Med 69:909–917
13. Olmarker K, Rydevik B (1992) Single- versus double-level nerve root compression: An experimental study on the porcine cauda equina with analysis of nerve impulse conduction properties. Clin Orthop 279:35–39
14. Porter RW, Ward D (1992) Cauda equina dysfunction. The significance of two-level pathology. Spine 17:9–15
15. Spratt KF, Keller TS, Szpalski M et al (2004) A predictive model for outcome after conservative decompression surgery for lumbar spinal stenosis. Eur Spine J 13:14–21
16. Taillard W (1954) Spondylolisthesis in children and adolescents. Acta Orthop Scand 24:115–144

17. Takahashi K, Miyazaki T, Takino T et al (1995) Epidural pressure measurements: relationship between epidural pressure and posture in patients with lumbar spinal stenosis. Spine 20:650–653
18. Takahashi K, Olmarker K, Holm S et al (1993) Double-level cauda equina compression: an experimental study with continuous monitoring of intraneural blood flow in the porcine cauda equina. J Orthop Res 11:104–109
19. Voorhies RM, Jiang X, Thomas N (2007) Predicting outcome in the surgical treatment of lumbar radiculopathy using the pain drawing score, McGill short form pain questionnaire, and risk factors including psychosocial issues and axial joint pain. Spine J 7:516–524
20. Weinstein JN, Tosteson TD, Lurie JD et al (2008) Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 358:794–810

JBMR

A Local Application of Recombinant Human Fibroblast Growth Factor 2 for Tibial Shaft Fractures: A Randomized, Placebo-Controlled Trial

Hiroshi Kawaguchi,[1] Hiroyuki Oka,[1] Seiya Jingushi,[2] Toshihiro Izumi,[3] Masao Fukunaga,[4] Katsumi Sato,[5] Takashi Matsushita,[6] and Kozo Nakamura[1] for the TESK Group

[1]Sensory and Motor System Medicine, University of Tokyo, Tokyo, Japan
[2]Department of Orthopaedic Surgery, Kyushu Rosai Hospital, Fukuoka, Japan
[3]Department of Orthopaedic Surgery, Saihaku Hospital, Tottori, Japan
[4]Department of Nuclear Medicine, Kawasaki Medical School, Okayama, Japan
[5]Department of Orthopaedic Surgery, Tohoku Rosai Hospital, Miyagi, Japan
[6]Department of Orthopaedic Surgery, Teikyo University, Tokyo, Japan

ABSTRACT

Fibroblast growth factor 2 (FGF-2) is a potent mitogen for mesenchymal cells, and a local application of recombinant human FGF-2 (rhFGF-2) in a gelatin hydrogel has been reported to accelerate bone union in our animal studies and preparatory dose-escalation trial on patients with surgical osteotomy. We have performed a randomized, double-blind, placebo-controlled trial in which patients with fresh tibial shaft fractures of transverse or short oblique type were randomly assigned to three groups receiving a single injection of the gelatin hydrogel containing either placebo or 0.8 mg (low-dosage group) or 2.4 mg (high-dosage group) of rhFGF-2 into the fracture gap at the end of an intramedullary nailing surgery. Of 194 consecutive patients over 2 years, 85 met the eligibility criteria, and 70 (24 in the placebo group and 23 each in low- and high-dosage groups) completed the 24-week study. The cumulative percentages of patients with radiographic bone union were higher in the rhFGF-2-treated groups ($p = .031$ and .009 in low- and high-dosage group, respectively) compared with the placebo group, although there was no significant difference between low- and high-dosage groups ($p = .776$). At 24 weeks, 4, 1, and 0 patients in the placebo, low-dosage, and high-dosage groups, respectively, continued to show delayed union. No patient underwent a secondary intervention, and the time to full weight bearing without pain was not significantly different among the three groups ($p = .567$). There also was no significant difference in the profiles of adverse events among the groups. In conclusion, a local application of the rhFGF-2 hydrogel accelerated healing of tibial shaft fractures with a safety profile. © 2010 American Society for Bone and Mineral Research.

KEY WORDS: FRACTURE HEALING; FIBROBLAST GROWTH FACTOR; CLINICAL TRIAL

Introduction

Tibial shaft fractures, common fractures that generally are caused by traumatic injuries, are sometimes associated with delayed union or nonunion, even after conventional intramedullary nailing surgery.[1,2] To prevent this, the development of synthetic materials with osteogenic properties that overcome the limitations of the conventional autologous or allogeneic bone graft has been desired.[3,4] Because several growth factors are expressed during various phases of fracture healing, there has been considerable interest in their use as therapeutic agents to enhance bone union.[5–7]

Fibroblast growth factors (FGFs), constituting a family of structurally related polypeptides characterized by their affinity for the glycosaminoglycan heparin-binding sites on cells, are known to play crucial roles in mitogenesis of mesenchymal cells.[8] Since mutations in genes of FGFs or their receptors cause genetic diseases with severe skeletal abnormalities, such as achondroplasia and thanatophoric dysplasia, the FGF signal has been suggested to play a pivotal role in osteogenesis.[9] Among the FGF family members, FGF-2 (basic FGF) is most abundantly accumulated in bone matrix and expressed from the early stages of bone formation.[10,11] We and others have reported osteogenic properties of recombinant human FGF-2 (rhFGF-2) by

Received in original form March 11, 2010; revised form May 1, 2010; accepted May 27, 2010. Published online June 7, 2010.
Address correspondence to: Hiroshi Kawaguchi, MD, Sensory and Motor System Medicine, Faculty of Medicine, University of Tokyo, Hongo 7-3-1, Bunkyo, Tokyo 113-8655, Japan. E-mail: kawaguchi-ort@h.u-tokyo.ac.jp
Members of the Therapeutic Exploratory Study of KCB-1B (TESK) Group are listed in the Supplemental Text S1. Additional Supporting Information may be found in the online version of this article.

Journal of Bone and Mineral Research, Vol. 25, No. 12, December 2010, pp 2735–2743
DOI: 10.1002/jbmr.146

The December 2010 issue of *Journal of Bone and Mineral Research* was published online on 23 Nov 2010. A pagination error was subsequently identified. This notice is included to indicate that the pagination is now correct and authoritative [20 January 2011].

a single local application for bone fracture, bone defect, and osteoporotic bone using several animal models, and the effects were enhanced when a biodegradable gelatin hydrogel prepared through glutaraldehyde cross-linking of gelatin was used as the carrier.[12–17] Aiming at clinical application, we recently performed a dose-escalation trial of a local injection of rhFGF-2 in the gelatin hydrogel into the surgical high tibial osteotomy site of knee osteoarthritis patients and found that rhFGF-2 (0.2, 0.4, or 0.8 mg) dose dependently accelerated bone union of the osteotomy with a safety profile.[18] Considering that delayed union or nonunion generally can be determined on radiographs in less than 6 months after the fracture,[19] we performed a prospective, randomized, double-blind, placebo-controlled multicenter trial based on the hypothesis that a local application of the rhFGF-2 hydrogel increases the number of patients who achieve bone union of tibial shaft fractures within 24 weeks of intramedullary nailing surgery.

Materials and Methods

Study design

The protocol of this trial was designed collaboratively by the academic authors and the sponsor, Kaken Pharmaceutical Co., Ltd. The sponsor was responsible for gathering data from participating institutions to create the clinical database. The data analyses were conducted by an independent faculty member at the University of Tokyo who was not employed by the sponsor or by a commercial contract research organization. The corresponding author wrote the article with full access to all data and the final responsibility for the decision to submit for publication. The protocol for this study and the CONSORT checklist are available as Supplemental Text S2 and S3.

All patients provided written informed consent, and the protocol was approved by the institutional review boards of all institutions involved. The trial was undertaken in compliance with the *ICH Good Clinical Practice Guideline* (www.ich.org/LOB/media/MEDIA482.pdf) and the ethics principles set out in the Declaration of Helsinki.

Patients

A succession of 194 patients with tibial fractures who visited 48 institutions in Japan from March 2006 through April 2008 were recruited and assessed for eligibility on the basis of the following criteria. Inclusion criteria were males and females aged $\geq$20 and $<$75 years with transverse or short oblique fractures (fracture-line length less than twice the diameter of the diaphyseal shaft on anteroposterior and lateral radiographs) that were considered to be indications for closed intramedullary nailing surgery. Although the fracture type was limited principally to closed fractures, Gustilo-Anderson type I open fractures with a clean minimum pinhole skin injury were included. Exclusion criteria were pathologic fractures, intraarticular fractures, spiral fractures, and comminuted fractures with large fragments; complications of other fractures influencing weight bearing; use of medications that may affect bone and cartilage metabolism, such as bisphosphonates, sex hormones (ie, estrogen or androgen) and their derivatives such as selective estrogen-receptor modulators (SERMs), vitamin D and its derivatives, calcitonin, vitamin K_2, ipriflavone, and calcium supplements; and complications of osteoporosis, malignancy, rheumatoid arthritis, hematologic, immunologic, or neurologic disorders and other serious disorders in lung, heart, liver, kidney, brain, etc., based on the criteria of the Pharmaceuticals and Medical Devices Agency in the Japanese Ministry of Health, Labor and Welfare; history of anaphylactic shock or asthma; allergy to gelatin; pregnant or nursing women; and administration of other clinical trials within 24 weeks.

Randomization

Patients who met the inclusion and exclusion criteria were randomly assigned to one of the three groups: receiving a single administration of the biodegradable gelatin hydrogel containing 0.8 or 2.4 mg of rhFGF-2 or the hydrogel alone (placebo) according to stratification by three factors (closed versus open fracture, age $<$30 years versus $\geq$30 years, and with versus without fibula fracture). To keep treatment assignments inaccessible to all individuals, including surgeons, institutional investigators, assessors, and patients who were involved in the conduct of the trial, the randomization was performed by an independent organization, the Registration Center (Adjust Co., Ltd., Sapporo, Japan). Assignments were made centrally by the Registration Center via an automated telephone system containing the pregenerated randomization schedule programmed via algorithm. The difference in patient numbers among the groups in one institution was programmed to be no more than 2. Institutional investigators were required to call directly into the telephone system to randomly assign eligible patients (which occurred at the appropriate visit on the basis of demographic data entered into the system by the institution) and to receive assignments of a masked study agent at each visit according to the randomization algorithm. The external organization examined the appearances of the three kinds of hydrogel before and after the trial and guaranteed that they were indistinguishable by surgeons, institutional investigators, or patients. The blindness to all individuals involved in the conduct of the trial was not broken until the clinical trial was completely finished.

Interventions

At the end of the closed intramedullary nailing surgery performed by experienced orthopedic surgeons with proximal and distal static locking by transverse screws under general anesthesia, all patients received two percutaneous injections of the hydrogel (0.5 mL each, containing 0, 0.4, or 1.2 mg of rhFGF-2) into the fracture gap from medial and lateral sides under fluoroscopic guidance. After the intervention, all patients received conventional care according to up-to-date guidelines on the management of tibial fractures. Neither external fixation, cast, or brace was used, and weight bearing was allowed unless the patient felt pain while walking. Patients who had had events with the potential to affect the bone union, such as a fall or a screw or nail breakage, were excluded from the study.

Outcome assessment

The primary outcome was radiographic bone union according to the criteria established by a previous fracture-healing study.[20]

For the assessment, anteroposterior and lateral radiographs were taken immediately after the surgery and at least every 2 weeks thereafter up to 24 weeks postoperatively under standardized conditions with the same exposure setting. After completion of the 24-week trial, all radiographs of 71 patients were randomly mixed to be blinded to the treatment group, institution, and time after intervention. A panel of three blinded assessors, two orthopedic surgeons and one musculoskeletal radiologist, who did not belong to participating institutions, read the radiographs. The assessors independently evaluated the number of extracortical bridging calluses on the four cortices (medial, lateral, anterior, and posterior) on the anteroposterior and lateral radiographs. The number was determined when at least two of the three assessors agreed. If not, the three assessors consulted and determined the number. Bridging callus formation on all four of the cortices was defined as bone union, and that on at least three of the four cortices was defined as primary bone union. The kappa values for interobserver reliability on the diagnosis of bone union and primary bone union were 0.804 and 0.759, respectively. Absence of primary bone union at 24 weeks was defined as delayed union.

The secondary outcomes were performance of a secondary intervention to promote bone union and the time to full weight bearing without pain. Although nonsteroidal anti-inflammatory drugs (NSAIDs) were used freely for analgesia at the discretion of the investigators and patients, *pain-free* was defined as no pain without such medication.

Adverse events were defined as any local, systemic, or laboratory disorders that occurred or worsened after surgery regardless of causality to the intervention. All adverse events were classified as serious or nonserious according to the ICH guidelines (www.ich.org/LOB/media/MEDIA436.pdf). Adverse drug reactions were defined as adverse events that were judged by the institutional investigator to have a possibility of relationship to the intervention.

The serum concentration of FGF-2 was measured by ELISA (Kaken Pharmaceutical Co., Ltd., Tokyo, Japan) 1, 2, 4, 6, and 24 hours postoperatively, and the maximum concentration was shown. The transition rate was calculated as the percentage of the time-concentration curve of the serum FGF-2 concentration above compared with our preparatory data of that for 24 hours after the intravenous injection of respective concentrations of rhFGF-2. The serum concentrations of the antibody to rhFGF-2 preoperatively and 2 and 4 weeks postoperatively were measured by ELISA (Kaken Pharmaceutical Co., Ltd).

Statistical analysis

Using the incidence of bone union after closed intramedullary nailing surgery for closed tibial fractures (76% in 6 months),[2] we calculated that a sample of 23 participants per group would be required for the study to have 90% power to show at least 20% advantage of the rhFGF-2 over placebo with respect to bone union, with a hazard ratio of 3.0, based on a two-sided type 1 error rate, of 5%.[21]

All analyses were performed according to the intention-to-treat principle. Comparisons of the baseline characteristics among the three groups were made using the chi-square test for categorical variables and the one-factor analysis of variance for numerical variables. Cumulative rates for bone union, primary bone union, and clinical outcomes were measured using the Kaplan-Meier method, and the comparisons among the three groups were made by the generalized Wilcoxon test. A subject who was lost to follow-up before 24 weeks was considered as a censored case in the time-to-event analysis. A weighted kappa statistic was used to determine the interobserver agreement on the diagnosis of bone union and primary bone union. Hazard ratios of bone union or primary bone union were calculated based on the Cox proportional-hazards model. Differences in the number of adverse events and adverse drug reactions were evaluated by Fisher's exact test. All statistical tests were performed at a significance level of .05 (two sides) and have not been adjusted for multiple testing. Data analyses were performed using SAS Version 9.1.3 (SAS Institute, Inc., Cary, NC, USA).

Results

Study population

Of 194 consecutive patients with tibial fracture who were assessed for eligibility, 109 did not meet the eligibility criteria, and 14 declined to participate; so 71 patients from 33 institutions were randomly assigned to the three groups [24 to the placebo group, 23 to the 0.8 mg of rhFGF-2 (low-dosage) group, and 24 to the 2.4 mg of rhFGF-2 (high-dosage) group) (Fig. 1). One patient in the high-dosage group withdrew from the trial after the 14-week assessment because he did not visit the institution every 2 weeks, so 70 patients completed the 24-week study.

None of the baseline characteristics of the patients (age, sex, body mass index, diabetes comorbidity, current smoker, complication of other fractures, transverse or oblique fracture, closed or open fracture, use of NSAIDs after fracture, days from injury to surgery, and days from surgery to discharge) was significantly different among the three groups (all $p > .05$; Table 1).

Efficacy

The cumulative percentage of patients with bone union, defined as bridging callus formation on all four of the cortices on radiographs, increased as a function of time without refracture

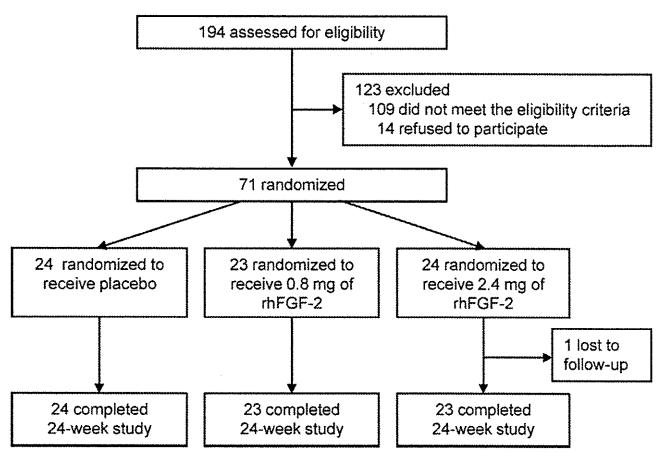

Fig. 1. Patient flow diagram.

Table 1. Baseline Characteristics of Study Patients

		rhFGF-2		
	Placebo ($n=24$)	0.8 mg ($n=23$)	2.4 mg ($n=24$)	p Value
Age (years)	46.8(14.9)	46.1(15.8)	46.2(12.7)	0.88
<30 yr	3(12.5%)	5(21.7%)	3(12.5%)	0.42
30 - 59 yr	15(62.5%)	10(43.5%)	17(70.1%)	
>/=60 yr	6(25.0%)	8 (34.8%)	4(16.7%)	
Female	3(12.5%)	8 (34.8%)	10(41.7%)	0.07
Body-mass index*	24.3(4.1)	24.5(3.6)	22.4(3.0)	0.07
Diabetes comorbidity†	5 (20.8%)	1(4.3%)	2(8.3%)	0.17
Current smoker	10(41.7%)	9(39.1%)	13(54.2%)	0.54
Other fractures	20(83.3%)	19(82.6%)	21(87.5%)	0.88
Transverse fracture‡	7(29.2%)	3(13.0%)	5(20.8%)	0.40
Closed fracture§	23(95.8%)	23(100%)	23(95.8%)	0.61
Use of NSAIDs	18(75.0%)	17(73.9%)	22(91.7%)	0.23
Days from injury to surgery	6.4(3.9)	5.7(2.6)	6.5(6.3)	0.93
Days from surgery to discharge	29.5(14.9)	33.5(20.2)	26.3(20.2)	0.55

Data are mean (SD) or number of patients (%). There were no significant differences among the three groups for any of the measured variables (p>0.05).
*The body-mass index is the weight in kilograms divided by the square of the height in meters.
†Diabetes was determined by serum hemoglobin A1c>/=6.5%.
‡The other patients had oblique fractures whose length was less than double the midshaft diameter on anteroposterior and lateral radiographs. Neither intraarticular fracture nor spiral fracture was included.
§The other 2 patients (one in the placebo and the other in the high dosage group) had Gustilo-Anderson type-I open fracture with a clean pin-hole skin injury.

during the 24 weeks in all groups (Fig. 2, *left column*). The percentages in the rhFGF-2-treated groups were higher than that in the placebo ($p=.028$ among the three groups, $p=.031$ between the low-dosage and placebo groups, and $p=.009$ between the high-dosage and placebo groups) and were more than double that in the placebo group during the clinically critical period from 12 to 16 weeks. There was, however, no significant difference between the low- and high dosage groups ($p=.776$). The median time to bone union was 128 days [95% confidence interval (CI) 115–152 days), 100 days (95% CI 95–116 days), and 101 days (95% CI 92–116 days), in the placebo, low-dosage, and high-dosage groups, respectively.

The cumulative percentage of patients with primary bone union, defined as bridging callus formation on at least three of the four cortices, also was higher in the rhFGF-2 treated groups than in the placebo group, although the effect was not dose-dependent ($p=.038$ between low-dosage and placebo groups and $p>.05$ among the three groups and between the other two

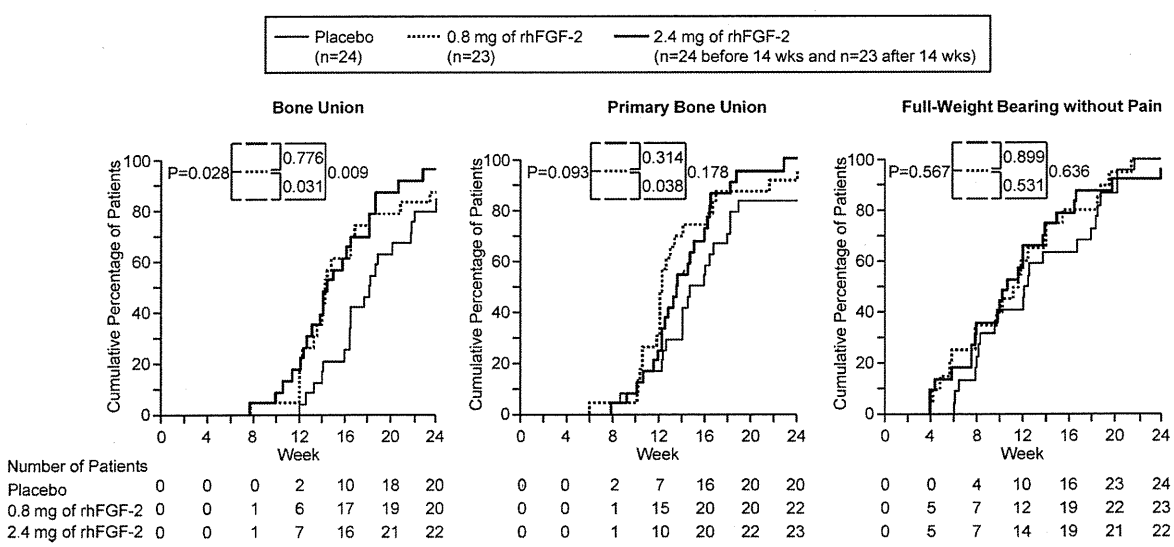

Fig. 2. Kaplan-Meier curves of the time to bone union, primary bone union, and full weight bearing without pain over 24 weeks after intervention. The extracortical bridging callus formation on all four of the cortices (medial, lateral, anterior, and posterior) on radiographs was defined as bone union and that on at least three of the four cortices was defined as primary bone union. As much weight bearing as possible was allowed unless the patient felt pain while walking, and the time to full weight bearing without pain was determined by the institutional investigators. p Values among the three groups (*left*) and between the two groups (*right*) were determined by the generalized Wilcoxon test.

groups) (Fig. 2, *middle column*). The median time to primary bone union was 107 days (95% CI 98–125 days), 85 days (95% CI 84–93 days), and 95 days (95% CI 85–111 days) in the placebo, low-dosage, and high-dosage groups, respectively. At the end of this study of 24 weeks, 4 patients in the placebo group and 1 patient in the low-dosage group remained in a state of delayed union without primary bone union, although all patients in the high-dosage group achieved primary union. Figure 3 shows the time course of radiographs of a patient in each group. In the patient from the placebo group, bridging callus formation was not apparent even at 24 weeks, indicating a delayed union; however, in patients from the rhFGF-2-treated groups, the calcified callus was observed at 16 weeks.

The Cox proportional-hazards model in all patients confirmed that the high and low dosages of rhFGF-2 were the only significant factors for bone union and primary bone union, respectively ($p = .03$ and .048), although none of the baseline characteristics was significantly associated with either of them (Table 2). Considering that there were trends, although not statistically significant, of a higher percentage of females and lower percentage of diabetics in the rhFGF-2-treated groups than in the placebo group (Table 1), we performed multivariate analyses adjusted by sex and diabetes comorbidity to avoid the biases and obtained reproducible results (Table 3). The high dosage of rhFGF-2 was significantly associated with bone union ($p = .03$) after adjustment for sex, and the high and low dosages were associated with bone union and primary bone union, respectively ($p = .04$ and .03) after adjustment for diabetes comorbidity.

During the study period, no patient experienced events with a potential to affect bone union, nor did any patient undergo a secondary intervention to promote bone union. The cumulative percentage of patients who achieved full weight bearing without pain showed a tendency to be higher in the rhFGF-2-treated groups than in the placebo group, although it was not significantly different among the three groups (Fig. 2, *right column*). The median time to full weight bearing without pain was 89 days (95% CI 68–129 days), 84 days (95% CI 69–109 days), and 79 days (95% CI 57–98 days) in the placebo, low-dosage, and high-dosage groups, respectively.

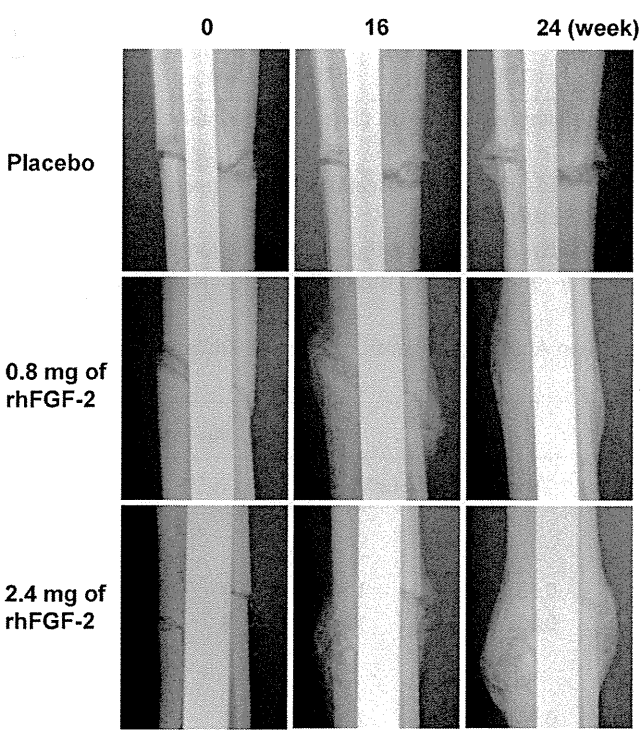

Fig. 3. Anteroposterior radiographs of fractured sites of a patient in each group 0, 16, and 24 weeks after intervention.

Adverse events

Although more than 50% of patients in all groups showed adverse events, no significant difference was observed in the profiles among the three groups (Table 4). The number of each serious adverse event or adverse drug reaction was only one in a group and showed no difference among the three groups. All had recovered before the last visit at 24 weeks. An infection in a patient in the high-dosage group was a superficial soft tissue infection without bacteriologic confirmation that was healed quickly with treatment with an antibiotic.

The transition rate of the injected rhFGF-2 into serum was less than 5% and unrelated to adverse events or adverse drug reactions (Table 4). No antibody to rhFGF-2 was detected in any patient preoperatively or postoperatively.

Discussion

Based on the evidence of animal studies and a preparatory dose-escalation clinical trial,[12–18] this prospective, randomized, double-blind, placebo-controlled multicenter trial for the first time revealed that local application of a synthetic material, rhFGF-2 hydrogel, accelerated healing of fresh fractures. The hydrogel alone was not likely to affect the healing process because the time to bone union in our placebo group was similar to that without any intervention but the nailing surgery in a previous report.[2] In our previous dose-escalation trial,[18] we selected 0.2, 0.4, and 0.8 mg as the dosages of rhFGF-2 based on results of animal experiments[12,14,17] and found dose-dependent effects on bone union of the tibial osteotomy.[18] To further determine the dosage with maximal efficacy, this study added a dosage of 2.4 mg because our preparatory toxicity tests in dogs with daily intramuscular administration of the rhFGF-2 hydrogel predicted the nontoxic dosage of rhFGF-2 to be 40 μg/kg per day (Supplemental Table S1). Since the effect of the 2.4-mg dosage on bone union was comparable with that of 0.8 mg (Fig. 2), we assume that the optimal concentration of rhFGF-2 for clinical use may be around 0.8 mg.

FGF-2 is known to stimulate proliferation of immature mesenchymal cells, but not their differentiation or matrix synthesis.[22–24] During bone fracture healing, expressions of endogenous FGF-2 and the principal receptor FGFR1 have been identified mainly at the early stage when immature osteoprogenitor cells are undergoing proliferation.[5,6,25] Considering that the local half-life of the injected ^{125}I-labeled rhFGF-2 in the hydrogel at the fractured site was less than 2 days in our animal experiments,[12,13] rhFGF-2 may exert its effect by affecting the earlier stage of bone union, probably through its potent

Table 2. Factors Influencing Bone Union and Primary Bone Union (Univariate Analysis)

		Bone union		Primary bone union	
Baseline factors	Number of patients	Hazard ratio (95% CI)*	*p* Value	Hazard ratio (95% CI)*	*p* Value
Age					
>/=30yr	60	1.44(0.71-2.94)	0.31	0.85(0.44-1.63)	0.63
<30yr	11	1.00		1.00	
Sex					
Female	21	1.21(0.71-2.07)	0.48	1.61(0.95-2.72)	0.08
Male	50	1.00		1.00	
Body-mass index					
>/=25	23	0.96(0.56-1.64)	0.87	0.89(0.53-1.50)	0.67
<25	48	1.00		1.00	
Diabetes comorbidity					
Yes	8	1.05(0.50-2.22)	0.89	1.36(0.65-2.87)	0.42
No	63	1.00		1.00	
Current smoker					
Yes	32	1.11(0.67-1.83)	0.70	0.80(0.49-1.31)	0.37
No	39	1.00		1.00	
Other fractures					
Yes	60	0.90(0.44-1.84)	0.78	0.89(0.44-1.83)	0.75
No	11	1.00		1.00	
Transverse fracture					
Yes	15	0.78(0.41-1.47)	0.44	0.74(0.40-1.39)	0.35
No	56	1.00		1.00	
Closed fracture					
Yes	69	0.83(0.20-3.43)	0.79	2.63(0.62-11.28)	0.19
No	2	1.00		1.00	
Use of NSAIDs					
Yes	57	0.75(0.41-1.36)	0.34	0.65(0.36-1.19)	0.16
No	14	1.00		1.00	
rhFGF-2					
Placebo	24	1.00		1.00	
0.8 mg	23	1.61(0.86-3.01)	0.13	1.85(1.01-3.41)	0.048
2.4 mg	24	1.94(1.05-3.57)	0.03	1.71(0.93-3.14)	0.09

*The hazard ratios of bone union and primary bone union associated with the listed factors are expressed. CI= confidence interval

mitogenic effect on immature cells. FGF-2 is also known to induce expressions of differentiation factors such as bone morphogenetic proteins (BMPs), transforming growth factor β (TGF-β), and prostaglandins.[12,24,26] Hence FGF-2 may initiate the cellular and molecular cascade of the osteogenesis process during fracture healing.

This study allowed investigators and patients to use NSAIDs freely for analgesia at their discretion. NSAIDs, especially

Table 3. Effects of rhFGF-2 Treatment on Bone Union and Primary Bone Union After Adjustments by Sex and Diabetes Comorbidity (Multivariate Analysis)

		Bone union		Primary bone union	
Characteristics	Number of patients	Adjusted hazard ratio (95% CI)*	*p* Value	Adjusted hazard ratio (95% CI)*	*p* Value
rhFGF-2 (adjusted by sex)					
Placebo	24	1.00		1.00	
0.8 mg	23	1.68(0.86-3.30)	0.13	1.85(0.98-3.33)	0.06
2.4 mg	24	1.99(1.06-3.75)	0.03	1.50(0.79-2.83)	0.22
rhFGF-2 (adjusted by diabetes comorbidity)					
Placebo	24	1.00		1.00	
0.8 mg	23	1.68(0.86-3.29)	0.14	2.07(1.06-4.04)	0.03
2.4 mg	24	1.96(1.02-3.77)	0.04	1.87(0.98-3.57)	0.06

*The hazard ratios of bone union and primary bone union adjusted by sex and diabetes comorbidity are expressed. CI= confidence interval

Table 4. Adverse Events, and Serum Concentrations of FGF-2 and Anti-FGF-2 Antibody

		rhFGF-2		
	Placebo ($n=24$)	0.8 mg ($n=23$)	2.4 mg ($n=24$)	p Value
All adverse events	18(75.0%)	17(73.9%)	14(58.3%)	0.43
Most common events: >5 % in either group				
Nasopharyngitis	5(20.1%)	8(34.8%)	3(12.5%)	
Back pain	2(8.3%)	3(13.4%)	1(4.2%)	
Insomnia	2(8.3%)	2(8.7%)	2(8.3%)	
Arthralgia	2(8.3%)	0	1(4.2%)	
Pyrexia	3(12.5%)	2(8.7%)	0	
Diarrhoea	0	2(8.7%)	2(8.3%)	
Pruritus	1(4.2%)	2(8.7%)	0	
Pain in extremity	0	0	2(8.3%)	
Serious adverse events	1(4.2%)	2(8.7%)	2(8.4%)	0.87
Skin necrosis	1(4.2%)	0	0	
Acute myocardial infarction	0	1(4.4%)	0	
Syncope	0	1(4.4%)	0	
Infection	0	0	1(4.2%)	
Peripheral artery aneurysm	0	0	1(4.2%)	
Adverse drug reactions	1(4.2%)	2(8.7%)	3(12.5%)	0.69
Hepatic dysfunction	0	1(4.4%)	0	
Hyperlipidaemia	0	1(4.4%)	0	
High serum alkaline phosphatase	1(4.2%)	0	1(4.2%)	
High serum uric acid	1(4.2%)	0	0	
Eosinophilia	0	0	1(4.2%)	
Infection	0	0	1(4.2%)	
Serum FGF-2	($n=10$)	($n=6$)	($n=8$)	
Maximum concentration (ng/mL)	0.031(0.030)	0.070(0.081)	0.080(0.082)	
Transitional rate (%)	—	4.9(3.7)	2.0(1.3)	
Serum antibody to FGF-2	ND	ND	ND	

Data are mean (SD) or number of patients (%). ND = not detected.

selective cyclooxygenase-2 inhibitors, are reported to have a negative impact on fracture healing[27,28]; however, this analysis showed no significant association of the NSAID use with bone union or primary bone union (Table 2). Although the number of patients with NSAID use was not significantly different among the three groups (Table 1), our further analysis revealed a higher number of patients with chronic use (>2 weeks) of NSAIDs in the high-dosage group than in the other two groups: 3, 5, and 10 for more than 2 weeks; 3, 3, and 8 for more than 4 weeks; 1, 1, and 6 for more than 8 weeks; and 0, 1, and 3 for more than 12 weeks in the placebo, low-dosage, and high-dosage groups, respectively. Considering that all patients in the high-dosage group have achieved bone union, NSAIDs may not influence fracture healing, at least under stimulation by rhFGF-2 treatment. We should note, however, that this might possibly be so because all NSAIDs used in this study were nonselective agents such as loxoprofen or diclofenac but not selective cyclooxygenase-2 inhibitors.

Of 194 patients, 104 did not meet the eligibility criteria, indicating a diversity of tibial fractures. Among the fracture types, the development of novel treatments for open, comminuted, or severe fractures is clinically more challenging and more desired than that of the simple fractures examined in this study. However, the healing of these complicated fractures is influenced by variable conditions, such as skin wound, contamination, soft tissue damage, bone displacement, and cortical defects, so it is difficult to perform an objective and accurate evaluation of the effect of interventions. In fact, although clinical trials of rhBMP-2, another synthetic protein for bone fractures, deal with open tibial fractures in a greater number of patients than this study, the primary outcome is the proportion of patients requiring secondary intervention.[29,30] Since the intervention is determined based on the surgeons' clinical and radiographic assessment, it may be considerably influenced by their discretion and subjective judgment. Hence, to evaluate the rhFGF-2 effect on fracture healing objectively and accurately, this study focused on transverse or short oblique fractures with more strict criteria whose background conditions are much more consistent and has successfully disclosed the significant effect on radiographic bone union. However, the increase in patients who achieved full weight bearing without pain by the treatment was not statistically significant, although there was a trend (Fig. 2, *right column*), probably because this outcome still depends on factors other than bone union, such as subjectiveness of patients and investigators, soft tissue inflammation, skin wound, and body weight. Taking the biologic nature of FGF-2 into consideration, we believe that rhFGF-2 hydrogel can exert its osteogenic action

even in more complicated fractures only if they are fresh and contain immature mesenchymal cells, which are targets of the FGF-2 mitogenic action.

Owing to its formulation, the rhFGF-2 hydrogel can be injected easily percutaneously under fluoroscopic guidance at the time of closed reduction even without surgery. In addition to fractures and surgical osteotomies, there are other potential clinical applications of this agent, such as acceleration of spinal fusion and prevention of loosening of total-joint arthroplasty. This agent also may act as one element of a comprehensive tissue-engineering strategy for bone and cartilage regeneration.[31] The extensive application of the osteogenic properties of FGF-2 may change clinical practice for various skeletal disorders, providing a substantial public health impact.

Disclosures

The role of the sponsor is stated under "Materials and Methods." None of the authors or investigators was employed by Kaken Pharmaceutical Co., Ltd., before, during, or after the study period, nor did any of them receive support grants or consultant/honorarium fees from the company. TM reports receiving a grant from Teijin Pharmaceutical Co., Ltd., and consultancy fees from Stryker Japan, Co., Ltd. All the other authors state that they have no conflicts of interest.

Acknowledgments

We thank the study participants and the TESK Group members (Supplementary Text S1). This study was funded by Kaken Pharmaceutical Co., Ltd.

Author contributions: *Study concept and design:* Kawaguchi, Jingushi, Izumi, Fukunaga, Sato, Matsushita, and Nakamura. *Acquisition of data:* Jingushi, Izumi, Fukunaga, Sato, and Matsushita. *Analysis and interpretation of data:* Kawaguchi, Oka, Matsushita, and Nakamura. *Drafting of the manuscript:* Kawaguchi. *Critical revision of the manuscript for important intellectual content:* Kawaguchi, Oka, Jingushi, Izumi, Fukunaga, Sato, Matsushita, and Nakamura. *Statistical analysis:* Oka. *Trial registration:* Umin.ac.jp, number 000000847.

References

1. Heckman JD, Sarasohn-Kahn J. The economics of treating tibia fractures: the cost of delayed unions. Bull Hosp Joint Dis. 1997;56:63–72.
2. Finkemeier CG, Schmidt AH, Kyle RF, Templeman DC, Varecka TF. A prospective, randomized study of intramedullary nails inserted with and without reaming for the treatment of open and closed fractures of the tibial shaft. J Orthop Trauma. 2000;14:187–193.
3. Younger EM, Chapman MW. Morbidity at bone graft donor sites. J Orthop Trauma. 1989;3:192–195.
4. Finkemeier CG. Bone-grafting and bone-graft substitutes. J Bone Joint Surg Am. 2002;84:454–464.
5. Bolander ME. Regulation of fracture repair by growth factors. Proc Soc Exp Biol Med. 1992;200:165–170.
6. Bourque WT, Gross M, Hall BK. Expression of four growth factors during fracture repair. Int J Dev Biol. 1993;37:573–579.
7. Lieberman JR, Daluiski A, Einhorn TA. The role of growth factors in the repair of bone biology and clinical applications. J Bone Joint Surg Am. 2002;84:1032–1044.
8. Itoh N, Ornitz DM. Evolution of the Fgf and Fgfr gene families. Trends Genet. 2004;20:563–569.
9. Ornitz DM. FGF signaling in the developing endochondral skeleton. Cytokine Growth Factor Rev. 2005;16:205–213.
10. Canalis E, Centrella M, McCarthy T. Effects of basic fibroblast growth factor on bone formation in vitro. J Clin Invest. 1988;81:1572–1577.
11. Khan SN, Bostrom MP, Lane JM. Bone growth factors. Orthop Clin North Am. 2000;31:375–388.
12. Kawaguchi H, Kurokawa T, Hanada K, et al. Stimulation of fracture repair by recombinant human basic fibroblast growth factor in normal and streptozotocin-diabetic rats. Endocrinology. 1994;135:774–781.
13. Nakamura K, Kawaguchi H, Aoyama I, et al. Stimulation of bone formation by intraosseous application of recombinant basic fibroblast growth factor in normal and ovariectomized rabbits. J Orthop Res. 1997;15:307–313.
14. Kato T, Kawaguchi H, Hanada K, et al. Single local injection of recombinant fibroblast growth factor-2 stimulates healing of segmental bone defects in rabbits. J Orthop Res. 1998;16:654–659.
15. Radomsky ML, Aufdemorte TB, Swain LD, et al. Novel formulation of fibroblast growth factor-2 in a hyaluronan gel accelerates fracture healing in nonhuman primates. J Orthop Res. 1999;17:607–614.
16. Tabata Y, Yamada K, Hong L, et al. Skull bone regeneration in primates in response to basic fibroblast growth factor. J Neurosurg. 1999;91:851–856.
17. Kawaguchi H, Nakamura K, Tabata Y, et al. Acceleration of fracture healing in non-human primates by fibroblast growth factor-2. J Clin Endocrinol Metab. 2001;86:875–880.
18. Kawaguchi H, Jingushi S, Izumi T, et al. Local application of recombinant human fibroblast growth factor-2 on bone repair: a dose-escalation prospective trial on patients with osteotomy. J Orthop Res. 2007;25:480–487.
19. Charles M, Court-Brown C. Fractures of the tibia and fibula. In: Bucholz RW, Heckman JD, Court-Brown C, eds. *Rockwood and Green's Fractures in Adults* 6th ed. Philadelphia, USA: Lippincott Williams & Wilkins, 2079–2146. 2006.
20. Heckman JD, Ryaby JP, McCabe J, Frey JJ, Kilcoyne RF. Acceleration of tibial fracture-healing by non-invasive, low-intensity pulsed ultrasound. J Bone Joint Surg Am. 1994;76:26–34.
21. Freedman LS. Tables of the number of patients required in clinical trials using the logrank test. Stat Med. 1982;1:121–129.
22. Rodan SB, Wesolowski G, Kyonggeun Y, Rodan GA. Opposing effects of fibroblast growth factor and pertussis toxin on alkaline phosphatase, osteopontin, osteocalcin and type I collagen mRNA levels in ROS 17/2.8 cells. J Biol Chem. 1989;264:19934–19941.
23. Hurley MM, Abreu C, Harrison JR, et al. Basic fibroblast growth factor inhibits type I collagen gene expression in osteoblastic MC3T3-E1 cells. J Biol Chem. 1993;268:5588–5593.
24. Kawaguchi H, Pilbeam CC, Gronowicz G, et al. Transcriptional induction of prostaglandin G/H synthase-2 by basic fibroblast growth factor. J Clin Invest. 1995;96:923–930.
25. Nakajima A, Nakajima F, Shimizu S, et al. Spatial and temporal gene expression for fibroblast growth factor type I receptor (FGFR1) during fracture healing in the rat. Bone. 2001;29:458–466.
26. Noda M, Vogel R. Fibroblast growth factor enhances type ß1 transforming growth factor gene expression in osteoblast-like cells. J Cell Biol. 1989;109:2529–2535.
27. Gerstenfeld LC, Einhorn TA. COX inhibitors and their effects on bone healing. Expert Opin Drug Saf. 2004;3:131–136.

28. Gerstenfeld LC, Al-Ghawas M, Alkhiary YM, et al. Selective and nonselective cyclooxygenase-2 inhibitors and experimental fracture-healing. Reversibility of effects after short-term treatment. J Bone Joint Surg Am. 2007;89:114–125.
29. Govender S, Csimma C, Genant HK, et al. Recombinant human bone morphogenetic protein-2 for treatment of open tibial fractures: a prospective, controlled, randomized study of four hundred and fifty patients. J Bone Joint Surg Am. 2002;84:2123–2134.
30. Swiontkowski MF, Aro HT, Donell S, et al. Recombinant human bone morphogenetic protein-2 in open tibial fractures. A subgroup analysis of data combined from two prospective randomized studies. J Bone Joint Surg Am. 2006;88:1258–1265.
31. Liu G, Kawaguchi H, Ogasawara T, et al. Optimal combination of soluble factors for tissue engineering of permanent cartilage from cultured human chondrocytes. J Biol Chem. 2007;282:20407–20415.

Osteoarthritis and Cartilage 18 (2010) 1227–1234

Osteoarthritis and Cartilage

Association of radiographic and symptomatic knee osteoarthritis with health-related quality of life in a population-based cohort study in Japan: the ROAD study

S. Muraki †*, T. Akune †, H. Oka ‡, Y. En-yo §, M. Yoshida §, A. Saika §, T. Suzuki ||, H. Yoshida ||, H. Ishibashi ||, F. Tokimura ||, S. Yamamoto ||, K. Nakamura ¶, H. Kawaguchi ¶, N. Yoshimura ‡

† *Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*
‡ *Department of Joint Disease Research, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*
§ *Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan*
|| *Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan*
¶ *Department of Sensory & Motor System Medicine, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*

ARTICLE INFO

Article history:
Received 4 August 2009
Accepted 17 June 2010

Keywords:
Osteoarthritis
Knee
Quality of life
Cohort
Epidemiology

SUMMARY

Objective: Knee osteoarthritis (OA) is a major public health issue causing chronic pain and disability. However, there is little information on the impact of this disease on quality of life (QOL) in Japanese men and women. The objective of the present study was to clarify the impact of radiographic and symptomatic knee OA on QOL in Japan.

Methods: This study examined the association of radiographic and symptomatic knee OA with QOL parameters such as the Medical Outcomes Study Short Form-8 (SF-8), EuroQOL (EQ-5D) and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). Radiographic knee OA was defined according to Kellgren/Lawrence (KL) grades, and symptomatic knee OA was defined as KL = 3 or 4 with knee pain. We also examined the independent association of symptomatic knee OA and grip strength with QOL.

Results: From the 3040 participants in the Research on Osteoarthritis Against Disability (ROAD) study, the present study analyzed 2126 subjects older than 40 years who completed the questionnaires (767 men and 1359 women; mean age, 68.9 ± 10.9 years). Subjects with KL = 3 or 4 had significantly lower physical QOL as measured by the physical component summary (PCS) score of the SF-8 and pain domains of the WOMAC, whereas mental QOL, as measured by the mental component summary (MCS) score of the SF-8, was higher in subjects with KL = 3 or 4 than KL = 0 or 1. Symptomatic knee OA was significantly more likely than radiographic knee OA without pain to be associated with physical QOL loss as measured by the PCS score and physical domains of the WOMAC. Symptomatic knee OA and grip strength were independently associated with physical QOL.

Conclusion: This cross-sectional study revealed that subjects with symptomatic knee OA had significantly lower physical QOL than subjects without it.

Introduction

Knee osteoarthritis (OA) is a major public health issue that causes chronic pain and disability[1–3]. The prevalence of radiographic knee OA is high in Japan[4], with 25,300,000 subjects aged 40 years and older estimated to experience radiographic knee OA[5]. According to the recent National Livelihood Survey of the Ministry of Health, Labour and Welfare in Japan, OA is ranked fourth among diseases that cause disabilities that subsequently require support with activities of daily living[6].

Quality of life (QOL) measurements in patients with chronic diseases are useful tools for estimating disease impact; these QOL scales may be generic or disease specific. Among the generic scales, the EuroQOL (EQ-5D) has been widely used to measure health-related QOL (HRQOL) in patients with OA[7,8], and several studies have used the Medical Outcomes Study Short Form-36 (SF-36) in

* Address correspondence and reprint requests to: Shigeyuki Muraki, Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. Tel: 81-3-5800-9178; Fax: 81-3-5800-9179.
E-mail address: murakis-ort@h.u-tokyo.ac.jp (S. Muraki).

1063-4584/$ – see front matter © 2010 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.joca.2010.06.001

Caucasian patients with OA[9–11]. However, almost all of these studies include only patients with knee OA, and there are few population-based studies regarding knee OA and QOL[11]. A previous population-based study in Caucasians showed that arthritis has a major impact on the HRQOL measured by the SF-36 in a community setting[11], although arthritis was examined by self-reported means and not by radiographs. In terms of disease-specific scales for knee OA, the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) has been used for Caucasians[12] and Asians[13,14], although these reports were not population-based studies. Furthermore, there is little information on the impact of knee OA with QOL in Japan, although a population survey suggests that the disease pattern differs among races[15–17]. In fact, the prevalence of knee OA in Japan[4] was much higher than that of previous epidemiologic studies in elderly Caucasians[16,18]. Furthermore, in terms of risk factors, studies in Caucasians have suggested that occupational activities that include kneeling and squatting were associated with knee OA[19], whereas these activities were not associated with Kellgren/Lawrence (KL) grades ≥ 2 OA in our previous study in Japan[20]. Therefore, the impact of knee OA on QOL also appears to differ in different populations. It would thus be of interest to clarify the impact of OA on QOL in a Japanese population.

The principal clinical symptom of knee OA is pain[21], but the correlation with the radiographic severity of knee OA is controversial[4,22–24]. Thus it would be interesting to determine whether the impact of radiographic knee OA on QOL differs according to the severity of OA. Furthermore, pain is strongly associated with QOL, so it would be of interest to clarify the impact of symptomatic OA as well as radiographic knee OA on QOL.

Gender differences have also been observed in knee OA. The prevalence of knee OA is higher in women than men[4], and the association of knee pain with knee OA also differs by gender[4]. Thus, the impact of these diseases on QOL may also differ between genders. However, to the best of our knowledge, there are no population-based studies that assess the association of knee OA with QOL in men and women separately.

Grip strength is a useful marker of muscle function and sarcopenia[25]. There is growing evidence that reduced grip strength is associated with adverse outcomes including morbidity[26], disability[27], falls[27], higher fracture rates[28], increased length of hospital stay[29], and mortality[27]. A previous study also showed that grip strength is related to total muscle strength[30]. Furthermore, there is increasing recognition that grip strength is a useful clinical marker of sarcopenia, and recent work has validated this approach, demonstrating that grip strength is more strongly associated with age and is a better predictor of poor mobility than other potential markers such as calf muscle area[31]. Previous reports have shown that low muscle mass was also associated with reduced QOL[32,33]; thus, the association of knee OA with QOL may be influenced by grip strength, but again, no studies have examined the association of knee OA and grip strength with QOL simultaneously in the same population.

The first objective of this study is to clarify the association of radiographic severity of knee OA with QOL among Japanese men and women using the large-scale, population-based cohort study called the Research on Osteoarthritis Against Disability (ROAD). Because pain is strongly associated with QOL, we also examined the association of symptomatic knee OA with QOL. Finally, we analyzed the independent associations of knee OA and grip strength with QOL.

Subjects and methods

Subjects

The ROAD study is a nationwide prospective study designed to establish epidemiologic indexes for evaluation of clinical evidence for the development of a disease-modifying treatment for bone and joint diseases (with OA and osteoporosis as the representative bone and joint diseases). It consists of population-based cohorts in several communities in Japan. A detailed profile of the ROAD study has been described in detail elsewhere[4,5,34]; a brief summary is provided here. To date, we have completed creation of a baseline database including clinical and genetic information for 3040 inhabitants (1061 men and 1979 women) ranging in age from 23 to 95 years (mean, 70.6 years), who were recruited from resident registration listings in three communities: an urban region in Itabashi, Tokyo, a mountainous region in Hidakagawa, Wakayama, and a seacoast region in Taiji, Wakayama. All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo and the Tokyo Metropolitan Institute of Gerontology. Anthropometric measurements included height and weight, and body mass index (BMI) (weight [kg]/height^2 [m^2]) was calculated. Grip strength was measured on bilateral sides using a TOEI LIGHT handgrip dynamometer (TOEI LIGHT Co., Ltd, Saitama, Japan), and the better measurement was used to characterize maximum muscle strength. Among 2995 subjects aged 40 years or older in the ROAD study, 2243 (74.9%), 2245 (75.0%) and 2222 (74.2%) subjects completed the SF-8, the EQ-5D and the WOMAC, respectively, and 2126 (71.0%) subjects completed all three questionnaires. The present study analyzed 2126 subjects (767 men and 1359 women) aged 40 years (mean, 68.9 $\pm$ 10.9 years) or older who had completed the SF-8, the EQ-5D, and the WOMAC.

Radiographic assessment

All participants had radiographic examination of both knees using anterior–posterior and lateral views with weight-bearing and foot map positioning. Knee radiographs were read without knowledge of participant clinical status by a single well-experienced orthopaedist (SM) using the KL radiographic atlas for overall knee radiographic grades[35]. In KL grade, radiographs are scored as grade 0 through 4, with higher grades being associated with more severe OA. The higher KL grade in both knees was designated as that of the participant. Symptomatic knee OA was defined as: (1) a subject reporting knee pain lasting at least 1 month with pain having last occurred within the current or previous year; and (2) KL = 3 or 4 OA in the painful knee. To evaluate the intra-observer variability of KL grading, 100 randomly selected radiographs of the knee were scored by the same observer more than 1 month after the first reading. One hundred other radiographs were also scored by two experienced orthopaedic surgeons (SM & HO) using the same atlas for inter-observer variability. The evaluated intra- and inter-observer variabilities were confirmed by kappa analysis to be sufficient for assessment (0.86 and 0.80, respectively).

Instruments

The SF-8 generates a health profile consisting of eight scales and two summary measures describing HRQOL. The SF-8 is an alternate form to the SF-36, which is the most widely used patient-based health status survey, translated into more than 40 languages; the Japanese version of the SF-36 has been well validated[36]. The SF-8 uses a single question to measure each of the eight SF-36 domains. In the SF-8, each of the eight items assesses a different dimension of health: General Health (GH), Physical Functioning (PF), Role Physical (RP), Bodily Pain (BP), Vitality (VT), Social Functioning (SF), Mental Health (MH) and Role Emotional (RE). The SF-8 was scored by assigning the mean SF-36 scale score from the 2002 general Japanese population to each response category of the SF-8 measuring the same concept, and then weighting each SF-8 item to

compute aggregate physical component summary (PCS) and mental component summary (MCS) scores. The SF-8 may be scored using a published algorithm for Japanese versions of the SF-8, which has been well validated[37]. The EQ-5D self-report questionnaire measures five domains of HRQOL, including mobility, self-care, usual activities, pain/discomfort, and anxiety/depression[38]. Each of the five domains is assessed by a single question with three response levels (no problem, some problems, and extreme problems), so the EQ-5D defines a total of 243 health states. These results were coded and converted to a score of utility using the tables of values[39]. The EQ-5D scoring algorithm was first developed using time trade off-based preference scores for a sample of these health states from a representative sample of the UK general population[38]; the Japanese version of the EQ-5D has been validated[39]. This EQ-5D algorithm is used worldwide and generates scores ranging from −0.111 to 1.000, with negative scores representing health states worse than being dead, 0 representing being dead, and 1.00 representing a state of full health. The WOMAC, a 24-item OA-specific index, consists of three domains: pain, stiffness, and physical function. Each of these 24 items is graded on either a five-point Likert scale or a 100-mm visual analogue scale[12,40]. In the present study, we used the Likert scale (version LK 3.0). The domain score ranges from 0 to 20 for pain, 0 to 8 for stiffness, and 0 to 68 for physical function. Japanese versions of the WOMAC have also been validated[41].

Statistical analysis

The differences in age, height, weight, BMI, grip strength, and QOL measurements between men and women were examined by the Student's *t* test. The prevalence of radiographic and symptomatic knee OA was compared between men and women using the chi-square test. We also used the chi-square test to analyze whether subjects with one symptomatic knee were likely to have symptomatic OA in the other knee. According to KL grade[35], KL = 2 was defined as definite osteophytosis but no definite joint space narrowing, and KL = 3 and 4 included definite joint space narrowing. We thus categorized KL grade in KL = 0 or 1, KL = 2, or KL = 3 or 4, and differences among each KL grade with QOL measurements were determined using the Tukey Honestly Significant Difference (HSD) test without adjustment and after adjustment for age, BMI, and grip strength in men and women. We further classified subjects into those with symptomatic knee OA, those with KL = 3 or 4 knee OA without pain, and those without KL = 3 or 4 knee OA, and compared their association with QOL using the Tukey HSD test after adjustment for age, BMI, and grip strength. To determine the independent association of symptomatic knee OA and grip strength with QOL, we used multiple regression analysis without adjustment and after adjustment for age and BMI. Data analyses were performed using SAS version 9.0 (SAS Institute Inc., Cary, NC).

Results

The characteristics of the 2126 participants in the present study are shown in Table I. The prevalence of knee OA was significantly higher in women than men. The prevalence of bilateral and unilateral symptomatic knee OA was 2.0% and 3.0% in men, and 5.6% and 5.8% in women, respectively. Chi-square test showed that when the right knee had symptomatic knee OA, the odds ratio for the left knee to have symptomatic knee OA was 86.3 and 59.7 in men and women, respectively. The PCS and MCS of the SF-8 and the EQ-5D utility scores were significantly higher and the all domains of WOMAC were significantly lower in men than women, indicating that the QOL scores were higher in men than women.

Table I
Characteristics of participants

	Overall	Men	Women	*P*-Values
Number of subjects	2126	767	1359	
Age, years	68.9 ± 10.9	69.7 ± 10.5	68.4 ± 11.1	0.006
Height, cm	154.6 ± 9.2	162.8 ± 6.7	150.0 ± 6.9	<0.0001
Weight, kg	55.0 ± 10.9	61.5 ± 10.8	51.4 ± 9.0	<0.0001
BMI, kg/m^2	22.9 ± 3.6	23.1 ± 3.4	22.8 ± 3.7	0.03
Grip strength, kg	25.5 ± 9.3	33.2 ± 8.9	21.2 ± 6.3	<0.0001
Radiographic knee OA, %	17.9	11.6	21.5	<0.0001
Symptomatic knee OA, %	9.0	5.0	11.3	<0.0001
SF-8				
PCS	47.0 ± 7.0	47.4 ± 6.8	46.8 ± 7.0	0.03
MCS	52.8 ± 5.9	53.4 ± 5.3	52.5 ± 6.1	0.0009
EQ-5D	0.90 ± 0.15	0.91 ± 0.14	0.90 ± 0.15	0.03
WOMAC				
Pain (0–20)	1.37 ± 2.44	1.13 ± 2.16	1.50 ± 2.57	0.0003
Stiffness (0–8)	0.71 ± 1.25	0.63 ± 1.09	0.77 ± 1.33	0.01
Function (0–68)	4.08 ± 7.93	3.35 ± 7.06	4.49 ± 8.37	0.001

Except where otherwise indicated, values are the mean ± SD.
The differences between men and women were examined by the Student's *t* test except for the prevalence of radiographic and symptomatic knee OA.
The prevalence of radiographic and symptomatic knee OA was compared between men and women using the chi-square test.
Radiographic knee OA was defined as KL grade 3 or 4.
Symptomatic knee OA was defined as KL grade 3 or 4 with knee pain.
SF-8, Medical Outcomes Study Short Form-8.

The scores for PCS and MCS in the SF-8, the EQ-5D utility scores, and all domains in the WOMAC by KL grade of knee OA in men and women are shown in Tables II and III. The associations of age, BMI, and grip strength with each QOL parameter were significant in men and women by linear regression analysis ($P < 0.01$), except for the association of age with the MCS of the SF-8. Thus, we used the Tukey HSD test after adjustment for age, BMI, and grip strength to determine the association of radiographic severity of knee OA with QOL. Men and women with KL = 3 or 4 had significantly lower QOL measured by PCS of the SF-8 and pain domains of the WOMAC than those with KL = 0 or 1 as well as KL = 2. In addition, the MCS scores were higher in men and women with KL = 3 or 4 compared with KL = 0 or 1. The EQ-5D utility scores were not significantly associated with the KL grade of the knee after adjustment for age, BMI and grip strength.

Next, to determine impact of symptoms of radiographic knee OA with QOL, we classified subjects into those with symptomatic knee OA, defined as KL = 3 or 4 with knee pain, those with KL = 3 or 4 without pain, and those without KL = 3 or 4 and compared the impact of each type of OA on QOL using the Tukey HSD test after adjustment for age, BMI, and grip strength (Fig. 1). In men and women, PCS of the SF-8 and physical function domain of the WOMAC were significantly lower in subjects with symptomatic knee OA compared with those without KL = 3 or 4 knee OA (men: difference in mean −5.9, 95% CI −8.6 to −3.2 and difference in mean 4.9, 95% CI 2.2 to 7.6, respectively; women: difference in mean −4.3, 95% CI −5.7 to −2.9 and difference in mean 3.9, 95% CI 2.3 to 5.5, respectively) as well as KL = 3 or 4 knee OA without pain (men: difference in mean −6.3, 95% CI −9.7 to −3.0 and difference in mean 5.7, 95% CI 2.3 to 9.1, respectively; women: difference in mean −4.9, 95% CI −6.7 to −3.1 and difference in mean 3.9, 95% CI 1.8 to 5.9, respectively), whereas among those with KL = 3 or 4 knee OA without pain and no KL = 3 or 4 knee OA, there were no significant differences in PCS of the SF-8 and physical function domain of the WOMAC. In women, MCS of the SF-8 was significantly higher in subjects with symptomatic knee OA compared with those without KL = 3 or 4 knee OA (difference in mean 2.6, 95% CI 1.3 to 4.0) as well as KL = 3 or 4 knee OA without pain (difference in mean 2.3, 95% CI 0.6 to 4.0). The EQ-5D utility score was

Table II
Mean scores of the SF-8, EQ-5D, and WOMAC scales by KL grade in men

		Severity of knee OA			Difference in means (95% CI)	
		KL = 0 or 1 (n = 444)	KL = 2 (n = 231)	KL = 3 or 4 (n = 92)	KL = 3 or 4 vs KL = 0 or 1	KL = 3 or 4 vs KL = 2
SF-8						
PCS	Crude	48.1 ± 0.3	47.1 ± 0.4	44.7 ± 0.7	−3.3 (−5.2, −1.5)	−2.3 (−4.3, −0.4)
	Adjusted	47.8 ± 0.3	47.4 ± 0.4	45.5 ± 0.7	−2.3 (−4.2, −0.5)	−1.9 (−3.9, 0.0)
MCS	Crude	52.8 ± 0.2	53.7 ± 0.3	55.3 ± 0.5	2.5 (1.1, 3.9)	1.6 (0.1, 3.1)
	Adjusted	52.9 ± 0.3	53.7 ± 0.4	55.2 ± 0.6	2.3 (0.8, 3.8)	1.5 (−0.02, 3.1)
EQ-5D	Crude	0.92 ± 0.01	0.91 ± 0.01	0.87 ± 0.01	−0.06 (−0.10, −0.02)	−0.04 (−0.08, 0.00)
	Adjusted	0.92 ± 0.01	0.91 ± 0.01	0.89 ± 0.01	−0.03 (−0.07, 0.01)	−0.03 (−0.07, 0.01)
WOMAC						
Pain	Crude	0.92 ± 0.10	1.13 ± 0.14	2.11 ± 0.22	1.19 (0.61, 1.76)	0.97 (0.36, 1.59)
	Adjusted	1.03 ± 0.10	1.02 ± 0.14	1.75 ± 0.22	0.72 (0.14, 1.30)	0.73 (0.12, 1.34)
Stiffness	Crude	0.57 ± 0.05	0.65 ± 0.07	0.91 ± 0.11	0.34 (0.05, 0.64)	0.26 (0.05, 0.58)
	Adjusted	0.60 ± 0.05	0.61 ± 0.07	0.80 ± 0.12	0.20 (−0.10, 0.50)	0.19 (0.13, 0.51)
Function	Crude	2.83 ± 0.33	3.38 ± 0.46	6.08 ± 0.73	3.24 (1.36, 5.12)	2.70 (0.67, 4.73)
	Adjusted	3.31 ± 0.32	2.88 ± 0.45	4.66 ± 0.72	1.35 (−0.53, 3.23)	1.77 (−0.19, 3.74)

Values are mean ± standard error (SE). SF-8, Medical Outcomes Study Short Form-8.
Adjusted differences in means were calculated by Tukey HSD test after adjustment for age, BMI and grip strength.

significantly lower in subjects with symptomatic knee OA compared with those without KL = 3 or 4 knee OA (difference in mean −0.08, 95% CI −0.13 to −0.02) as well as KL = 3 or 4 knee OA without pain in men (difference in mean −0.08, 95% CI −0.15 to −0.01), but not in women.

Next, to examine the independent association of symptomatic knee OA and grip strength on QOL, multiple regression analysis was used with age, BMI, grip strength, and the presence of symptomatic knee OA as independent variables (Table IV). In men and women, symptomatic knee OA and grip strength were independently associated with PCS of the SF-8 (R^2, 0.11 and 0.17, respectively), EQ-5D utility scores (R^2, 0.08 and 0.12, respectively), and pain (R^2, 0.12 and 0.16, respectively), stiffness (R^2, 0.06 and 0.09, respectively) and physical function domains (R^2, 0.13 and 0.21, respectively) of the WOMAC.

Discussion

This is the first study to examine the association of radiographic and symptomatic knee OA with QOL measured by generic scales such as the SF-8, which is an alternate form of the SF-36, and the EQ-5D, as well as a disease-specific scale such as WOMAC in Japanese men and women using a large-scale population-based cohort study. In the present study, subjects with KL = 3 or 4 had significantly lower physical QOL than those with KL = 0 or 1 as well as KL = 2. At the same time, the MCS scores were higher in KL = 3 or 4 than KL = 0 or 1 in men and women. Furthermore, symptomatic knee OA was significantly associated with lower physical QOL compared with radiographic knee OA without pain. We further clarified the independent associations with symptomatic knee OA and grip strength. Symptomatic knee OA and grip strength were independently associated with lower QOL.

In the present study, physical QOL was significantly lower in subjects with KL = 3 or 4 compared with KL = 0 or 1 as well as KL = 2 in men and women. Samsa *et al.* reviewed the existing literature and concluded that the Minimally Clinically Important Difference (MCID) for the SF-36 is typically in the range of 3–5 points[42], implying that differences in SF-36 scores of 1–2 points are not important, but differences in scores of 3 points or more are clinically important. In this study, differences of PCS scores between subjects with KL = 3 or 4 and those with KL = 0 or 1 were 3.4 and 4.6 in men and women, respectively. The differences were similar to MCID thresholds, indicating that KL = 3 or 4 knee OA may be clinically important for physical QOL. A previous study in China

Table III
Mean scores of the SF-8, EQ-5D, and WOMAC scales by KL grade in women

		Severity of knee OA			Difference in means (95% CI)	
		KL = 0 or 1 (N = 541)	KL = 2 (N = 526)	KL = 3 or 4 (N = 292)	KL = v3 or 4 vs KL = 0 or 1	KL = 3 or 4 vs KL = 2
SF-8						
PCS	Crude	48.4 ± 0.3	46.9 ± 0.3	43.8 ± 0.4	−4.5 (−5.7, −3.4)	−3.0 (−4.2, −1.9)
	Adjusted	47.1 ± 0.3	47.4 ± 0.3	45.5 ± 0.4	−1.6 (−2.9, −0.3)	−1.9 (−3.1, −0.7)
MCS	Crude	52.1 ± 0.3	52.3 ± 0.3	53.8 ± 0.4	1.7 (0.7, 2.7)	1.4 (0.4, 1.5)
	Adjusted	51.9 ± 0.3	52.5 ± 0.3	53.8 ± 0.4	1.9 (0.7, 3.1)	1.3 (0.2, 2.4)
EQ-5D	Crude	0.92 ± 0.01	0.89 ± 0.01	0.85 ± 0.01	−0.07 (−0.09, −0.04)	−0.04 (−0.07, −0.02)
	Adjusted	0.89 ± 0.01	0.91 ± 0.01	0.89 ± 0.01	−0.003 (−0.04, 0.03)	−0.02 (−0.04, 0.01)
WOMAC						
Pain	Crude	0.96 ± 0.11	1.45 ± 0.10	2.62 ± 0.15	1.65 (1.23, 2.08)	1.16 (0.74, 1.59)
	Adjusted	1.45 ± 0.11	1.19 ± 0.11	1.99 ± 0.15	0.53 (0.07, 1.00)	0.80 (0.38, 1.21)
Stiffness	Crude	0.55 ± 0.06	0.79 ± 0.06	1.14 ± 0.08	0.59 (0.37, 0.81)	0.35 (0.12, 0.57)
	Adjusted	0.75 ± 0.06	0.68 ± 0.06	0.85 ± 0.08	0.10 (−0.15, 0.34)	0.16 (0.06, 0.39)
Function	Crude	2.41 ± 0.34	4.54 ± 0.35	8.32 ± 0.47	5.91 (4.54, 7.28)	3.78 (2.40, 5.16)
	Adjusted	4.37 ± 0.35	3.62 ± 0.33	5.79 ± 0.47	1.42 (−0.04, 2.88)	2.17 (0.85, 3.50)

Values are mean ± SE. SF-8, Medical Outcomes Study Short Form-8.
Adjusted differences in means were calculated by Tukey HSD test after adjustment for age, BMI and grip strength.

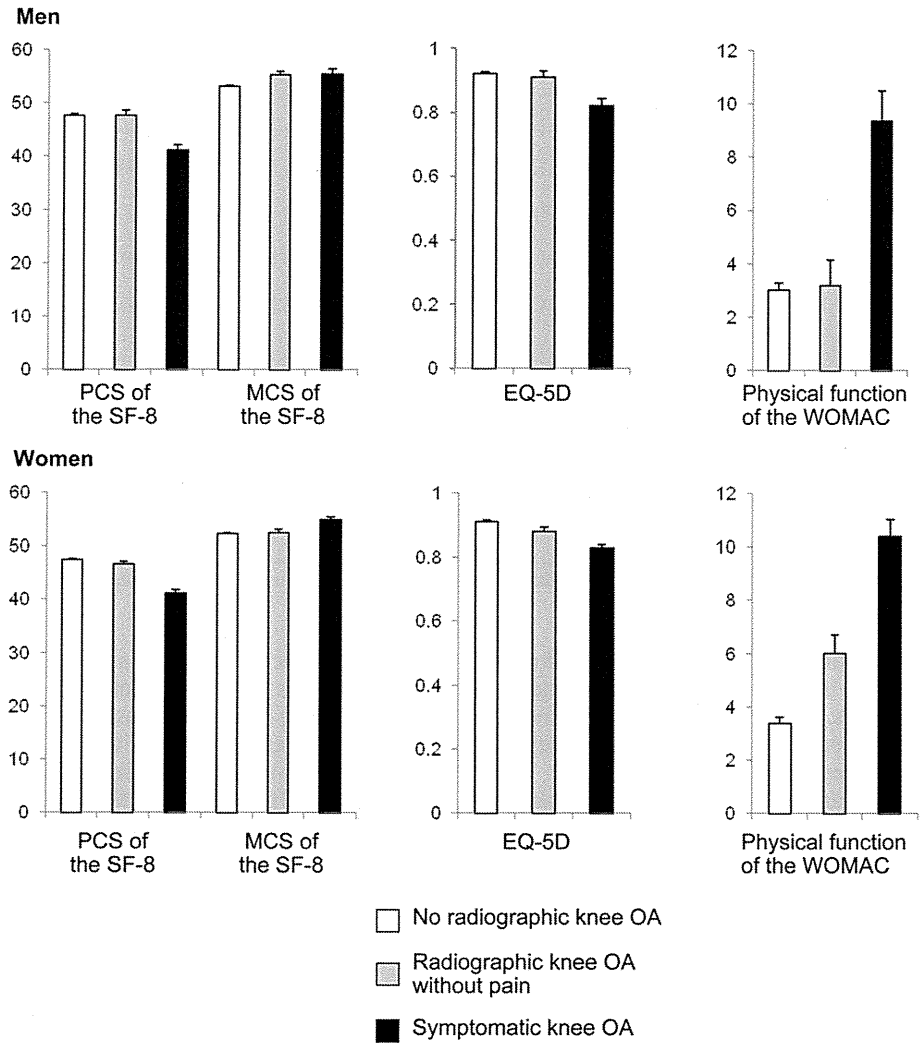

Fig. 1. Mean scores and SE of the SF-8, EQ-5D, and WOMAC scales in men and women with symptomatic knee OA ($N = 38$ and 154, respectively), radiographic knee OA without pain ($N = 53$ and 140, respectively), and no radiographic knee OA ($N = 676$ and 1065, respectively). Symptomatic knee OA was defined as KL = 3 or 4 with knee pain, radiographic knee OA without pain was defined as KL = 3 or 4 without knee pain, and no radiographic knee OA was defined as KL = 0, 1 or 2.

also showed that subjects with severe knee OA had lower QOL than those with mild knee OA[14], although their subjects were recruited from hospitals, so QOL parameters were not compared between subjects with mild knee OA and those without knee OA. The present study showed that there were no significant differences between subjects with KL = 2 and those with KL = 0 or 1. Considering the definitions of the KL grade, our findings may indicate that osteophytosis and joint space narrowing, which are representative features of knee OA, have a different impact on QOL. In other words, osteophytosis may have a weak impact on QOL, whereas joint space narrowing may have a strong impact.

Because QOL was shown to be strongly associated with pain, we next compared the impact of radiographic knee OA with and without pain on QOL. The present study showed that symptomatic knee OA was significantly associated with lower physical QOL than radiographic knee OA without pain. Differences in PCS scores among subjects with symptomatic knee OA and those without radiographic knee OA without pain were 6.6 and 6.5 in men and women, respectively. The differences were higher than the MCID; thus, symptomatic knee OA is considered clinically important for physical QOL. In addition, there were no significant differences in physical QOL between subjects with radiographic knee OA without pain and those without radiographic knee OA. This finding indicates that loss of physical QOL was more strongly associated with symptoms such as pain due to radiographic knee OA rather than radiographic changes of the knee itself. In other words, QOL may improve when pain is relieved by medical care, even if subjects have radiographic knee OA.

As measured by MCS of the SF-8, knee OA was associated with higher QOL scores in men and women, although it was also associated with lower PCS. Past studies also showed the dissociation between PCS and MCS in knee OA[43]. Several factors may contribute to this phenomenon. First, the MCS questions within the SF-8 include generic questions about energy levels, feelings of being "downhearted and blue," and interference in daily activities as a result of emotional problems. These questions are less sensitive to the presence of mental health issues than disease-specific scales such as the Kessler psychological distress scale[44]. In fact, psychological distress has been shown to be significantly more frequent in those with arthritis than those without it, although scores on the MCS were not significantly different between these two groups[45]. Second, the dissociation may be due to a disability paradox[46], which suggests that people with chronic disabilities report serious limitations in activities of daily living and problems in performing social roles, yet

Table IV
Correlations of symptomatic knee OA and grip strength with scores of the SF-8, EQ-5D, and WOMAC scales

		SF-8		EQ-5D	WOMAC		
		PCS	MCS		Pain	Stiffness	Function
Men							
Symptomatic knee OA	Crude regression coefficient	−6.64 (−8.82, −4.46)	2.49 (0.77, 4.21)	−0.10 (−0.14, −0.05)	2.46 (1.78, 3.13)	0.83 (0.48, 1.18)	6.19 (3.95, 8.42)
($N=38$)	Adjusted regression coefficient	−6.00 (−8.17, −3.81)	2.10 (0.33, 3.88)	−0.08 (−0.12, −0.03)	2.18 (1.51, 2.86)	0.75 (0.39, 1.10)	4.88 (2.67, 7.10)
Grip strength	Crude regression coefficient	0.20 (0.15, 0.25)	−0.03 (−0.07, 0.01)	0.003 (0.002, 0.004)	−0.06 (−0.07, −0.04)	−0.02 (−0.03, −0.01)	−0.23 (−0.28, −0.17)
	Adjusted regression coefficient	0.19 (0.12, 0.26)	−0.02 (−0.08, 0.03)	0.003 (0.001, 0.004)	−0.04 (−0.06, −0.02)	−0.01 (−0.02, 0.00)	−0.19 (−0.26, −0.12)
Women							
Symptomatic knee OA	Crude regression coefficient	−6.29 (−7.42, −5.16)	2.66 (1.64, 3.69)	−0.07 (−0.10, −0.05)	2.05 (1.64, 2.47)	0.80 (0.59, 1.02)	6.74 (5.40, 8.08)
($N=154$)	Adjusted regression coefficient	−4.36 (−5.52, −3.21)	2.52 (1.43, 3.61)	−0.03 (−0.06, −0.01)	1.44 (1.02, 1.85)	0.51 (0.29, 0.74)	3.97 (2.68, 5.27)
Grip strength	Crude regression coefficient	0.34 (0.28, 0.41)	0.06 (0.01, 0.12)	0.007 (0.006, 0.009)	−0.11 (−0.13, −0.08)	−0.04 (−0.05, −0.03)	−0.46 (−0.53, −0.39)
	Adjusted regression coefficient	0.20 (0.13, 0.27)	0.08 (0.01, 0.15)	0.004 (0.003, 0.006)	−0.04 (−0.07, −0.02)	−0.01 (−0.03, 0.00)	−0.21 (−0.30, −0.13)

Adjusted regression coefficient is calculated by multiple regression analysis with age, BMI, grip strength, and the presence of symptomatic knee OA as independent variables.
SF-8, Medical Outcomes Study Short Form-8.

state that they have excellent or good QOL. Many subjects with knee OA had knee pain, which may lead to functional impairment. Particularly in elderly individuals, pain or functional impairment may be considered a natural consequence of being elderly. Knee OA was thus not associated with lower scores for MCS in the SF-8.

In the present study, grip strength was independently associated with QOL measured by almost all domains of the three scales. Previous reports showed that low muscle mass was associated with reduced QOL[32,33]. There is increasing recognition that grip strength is a useful clinical marker of sarcopenia, and recent work has validated this approach, demonstrating that grip strength is more strongly associated with age and is a better predictor of poor mobility than other potential markers such as calf muscle area[31]. The independent association of grip strength with QOL suggests that QOL may improve with increase of muscle power in subjects with symptomatic knee OA, although longitudinal studies will be required to clarify this finding.

The present study showed that the association of radiographic and symptomatic knee OA with QOL differed among the SF-8, the WOMAC, and the EQ-5D. Radiographic and symptomatic knee OA were significantly associated with physical QOL in men and women, but not with EQ-5D utility scores. The reason for this difference may be explained by the fact that in the EQ-5D, all five domains are combined to analyze the association with knee OA, whereas the PCS and MCS of the SF-8 are analyzed separately. In fact, associations of knee OA differed between PCS and MCS of the SF-8, so when all domains were combined, the results may differ. For WOMAC, previous studies have found that WOMAC discriminates better among individuals with knee OA, whereas the SF-36 discriminates better among individuals with varying levels of self-reported general health status and comorbidities[47]. In addition, WOMAC was shown to be a more responsive measure than SF-36 in documenting changes after surgery[7,10]. Although our survey is not strictly comparable in design, it would appear that in our Japanese population, the PCS of the SF-8 and physical function domains of the WOMAC are able to discriminate among individuals with knee OA. It has been suggested that these two scales provide complementary information and may be useful in assessing both generic and disease-specific aspects of OA. However, this was a cross-sectional study, so the efficacy of these scales for knee OA in a longitudinal analysis could not be clarified. In longitudinal studies, generic measures such as the SF-8 may be much less useful than disease-specific measures such as the WOMAC because the generic measures pick up a lot of "noise" from comorbidities and may therefore be relatively unresponsive.

There are several limitations to the present study. First, this is a large-scale, population-based study, with a cross-sectional study of baseline data. Thus, causal relationships could not be determined. The ROAD study is a longitudinal survey, so further progress may help elucidate any causal relationships. Second, we did not include other weight-bearing OAs, such as hip OA, in the analysis, although this disorder may also affect QOL. However, the prevalence of KL = 3 or 4 hip was 1.4% and 3.5% in Japanese men and women[48], respectively, which was smaller compared with KL = 3 or 4 knee in the present study. Thus it is possible that hip OA would not strongly affect the results in the present study. Third, among the 2995 subjects ≥40 years old in the ROAD study, 2126 subjects had completed questionnaires for the SF-8, the EQ-5D, and the WOMAC, for a response rate of 71.0%. Subjects who completed questionnaires may have had better QOL than those who did not, so our results regarding QOL may have represented overestimations of QOL.

In conclusion, the present cross-sectional study using a large-scale population from the ROAD study revealed that KL = 3 or 4 OA was significantly associated with lower physical QOL scores, whereas KL = 2 OA was not. Symptomatic knee OA was more strongly associated with QOL than radiographic knee OA without pain. Further studies, along with continued longitudinal surveys in the ROAD study, will help to elucidate the background of knee OA and relations with QOL.

Author contributions

All authors have made substantial contributions to all three of sections (1), (2) and (3) below;

(1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data
(2) drafting the article or revising it critically for important intellectual content
(3) final approval of the version to be submitted.

Conflicts of interest

There are no conflicts of interest.

Acknowledgements

This study was supported by a Grant-in-Aid for Scientific Research (B20390182, C20591737, C20591774), for Young Scientists (A18689031), and for Exploratory Research (19659305) from the Japanese Ministry of Education, Culture, Sports, Science and Technology, H17-Men-eki-009, H18-Choujyu-037, and H20-Choujyu-009 from the Ministry of Health, Labor and Welfare, the Research Aid from the Japanese Orthopaedic Association and Grant of Japan Orthopedics and Traumatology Foundation, Inc., No. 166. The sponsors had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

The authors wish to thank Dr Anamizu and members of the Department of Orthopedics; Mr Kutsuma and other members of the Department of Radiology at Tokyo Metropolitan Geriatric Medical Center; Mrs Tomoko Takijiri and other members of the Public Office in Hidakagawa Town; and Mrs Tamako Tsutsumi, Mrs Kanami Maeda, and other members of the Public Office in Taiji Town, for their assistance in the location and scheduling of participants for examinations.

References

1. Sharma L, Kapoor D. Epidemiology of osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM, Eds. Osteoarthritis: Diagnosis and Medical/Surgical Management. 4th edn. Philadelphia: Lippincott Williams & Wilkins; 2007:3–26.
2. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, *et al.* The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. Am J Public Health 1994;84:351–8.
3. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41:1343–55.
4. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, *et al.* Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthritis Cartilage 2009;17:1137–43.
5. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-yo Y, Yoshida M, *et al.* Prevalence of knee osteoarthritis, lumbar spondylosis and osteoporosis in Japanese men and women: the Research on Osteoarthritis/osteoporosis Against Disability (ROAD). J Bone Miner Metab 2009;27:620–8.
6. Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey 2007. Available at: http://www.mhlw.go.jp/toukei/list/20-19-1.html.
7. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology (Oxford) 1999;38:870–7.
8. Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. Rheumatology (Oxford) 1999;38:807–13.
9. March LM, Cross MJ, Lapsley H, Brnabic AJ, Tribe KL, Bachmeier CJ, *et al.* Outcomes after hip or knee replacement surgery. Med J Aust 1999;171:235–8.
10. Bachmeier CJ, March LM, Cross MJ, Lapsley HM, Tribe KL, Courtenay BG, *et al.* A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery. Osteoarthritis Cartilage 2001;9:137–46.
11. Hill CL, Parsons J, Taylor A, Leach G. Health related quality of life in a population sample with arthritis. J Rheumatol 1999;26:2029–35.
12. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
13. Thumboo J, Chew LH, Soh CH. Validation of the Western Ontario and McMaster University osteoarthritis index in Asians with osteoarthritis in Singapore. Osteoarthritis Cartilage 2001;9:440–6.
14. Woo J, Lau E, Lee P, Kwok T, Lau WC, Chan C, *et al.* Impact of osteoarthritis on quality of life in a Hong Kong Chinese population. J Rheumatol 2004;31:2433–8.
15. Dillon CF, Rasch EK, Gu Q, Hirsch R. Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991–94. J Rheumatol 2006;33:2271–9.
16. Jordan JM, Helmick CG, Renner JB, Luta G, Dragomir AD, Woodard J, *et al.* Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol 2007;34:172–80.
17. Yoshida S, Aoyagi K, Felson DT, Aliabadi P, Shindo H, Takemoto T. Comparison of the prevalence of radiographic osteoarthritis of the knee and hand between Japan and the United States. J Rheumatol 2002;29:1454–8.
18. Felson DT, Naimark A, Anderson J, Kazis L, Castelli W, Meenan RF. The prevalence of knee osteoarthritis in the elderly. The Framingham osteoarthritis study. Arthritis Rheum 1987;30:914–8.
19. Coggon D, Croft P, Kellingray S, Barrett D, McLaren M, Cooper C. Occupational physical activities and osteoarthritis of the knee. Arthritis Rheum 2000;43:1443–9.
20. Muraki S, Akune T, Oka H, Mabuchi A, En-yo Y, Yoshida M, *et al.* Association of occupational activity with radiographic knee osteoarthritis and lumbar spondylosis in elderly patients of population-based cohorts: a large-scale population-based study. Arthritis Rheum 2009;61:779–86.
21. Linaker CH, Walker-Bone K, Palmer K, Cooper C. Frequency and impact of regional musculoskeletal disorders. Baillieres Clin Rheumatol 1999;13:197–215.
22. Duncan R, Peat G, Thomas E, Hay E, McCall I. Croft P symptoms and radiographic osteoarthritis: not as discordant as they are made out to be? Ann Rheum Dis 2007;66:86–91.
23. Hannan MT, Felson DT, Pincus T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. J Rheumatol 2000;27:1513–7.
24. Neogi T, Felson D, Niu J, Nevitt M, Lewis CE, Aliabadi P, *et al.* Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. BMJ 2009;339:b2844.
25. Roubenoff R. Sarcopenia: a major modifiable cause of frailty in the elderly. J Nutr Health Aging 2000;4:4140–2.
26. Sayer AA, Syddall HE, Dennison EM, Martin HJ, Phillips DI, Cooper C, *et al.* Grip strength and the metabolic syndrome: findings from the Hertfordshire Cohort Study. QJM 2007;100:707–13.
27. Bohannon RW. Hand-grip dynamometry predicts future outcomes in aging adults. J Geriatr Phys Ther 2008;31:3–10.
28. Sirola J, Rikkonen T, Tuppurainen M, Jurvelin JS, Kroger H. Association of grip strength change with menopausal bone loss and related fractures: a population based follow-up study. Calcif Tissue Int 2006;78:218–26.
29. Kerr A, Syddall HE, Cooper C, Turner GF, Briggs RS, Sayer AA. Does admission grip strength predict length of stay in hospitalised older patients? Age Ageing 2006;35:82–4.