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.

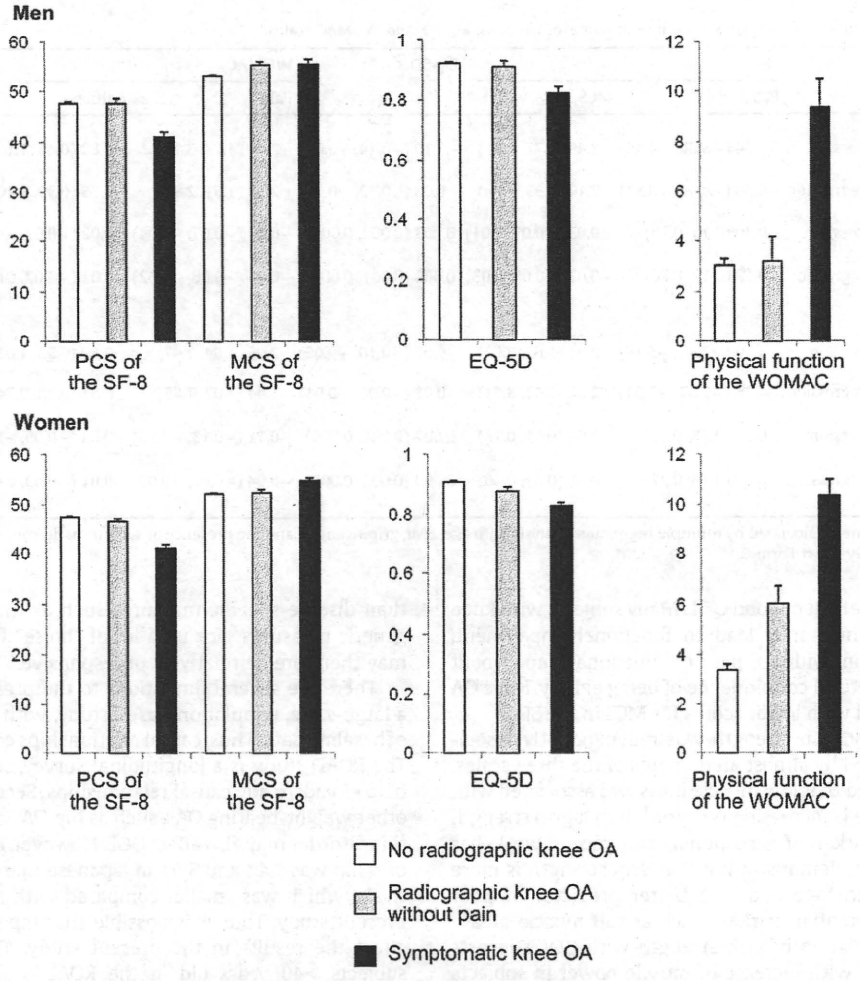

Fig. 1. Mean scores and SE of the SF-8, EQ-5D, and WOMAC scales in men and women with symptomatic knee OA ($N = 38$ and 154, respectively), radiographic knee OA without pain ($N = 53$ and 140, respectively), and no radiographic knee OA ($N = 676$ and 1065, respectively). Symptomatic knee OA was defined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

30. Wind AE, Takken T, Helders PJ, Engelbert RH. Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? Eur J Pediatr 2010;169:281–7.
31. Lauretani F, Russo CR, Bandinelli S, Bartali B, Cavazzini C, Di Iorio A, *et al.* Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia. J Appl Physiol 2003;95:1851–60.
32. Iannuzzi-Sucich M, Prestwood KM, Kenny AM. Prevalence of sarcopenia and predictors of skeletal muscle mass in healthy, older men and women. J Gerontol A Biol Sci Med Sci 2002;57:M772–7.
33. Sayer AA, Syddall HE, Martin HJ, Dennison EM, Roberts HC, Cooper C. Is grip strength associated with health-related quality of life? Findings from the Hertfordshire Cohort Study. Age Ageing 2006;35:409–15.
34. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability study. Int J Epidemiol (In press).
35. Kellgren JH, Lawrence JS, Eds. The Epidemiology of Chronic Rheumatism: Atlas of Standard Radiographs of Arthritis. Oxford: Blackwell Scientific; 1963.
36. Fukuhara S, Bito S, Hsiano A, Green J, Kurokawa K. Translation, adaptation, and validation of the SF-36 health survey for use in Japan. J Clin Epidemiol 1998;51:1045–53.
37. Fukuhara S, Suzukamo Y. Manual of the SF-8 Japanese Version (In Japanese). Kyoto: Institute for Health Outcomes & Process Evaluation Research; 2004.
38. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35:1095–108.
39. Japanese EuroQOL Translation Team. The development of the Japanese EuroQOL instrument. J Health Care Soc 1998;8:109–1123.
40. Barr S, Bellamy N, Buchanan WW, Chalmers A, Ford PM, Kean WF, *et al.* A comparative study of signal versus aggregate methods of outcome measurement based on the WOMAC osteoarthritis index. J Rheumatol 1994;21:2106–12.
41. Hashimoto H, Hanyu T, Sledge CB, Lingard EA. Validation of a Japanese patient-derived outcome scale for assessing total knee arthroplasty: comparison with Western Ontario and McMaster Universities osteoarthritis index (WOMAC). J Orthop Sci 2003;8:288–93.
42. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics 1999;15:141–55.
43. Boutron I, Rannou F, Jardinaud-Lopez M, Meric G, Revel M, Poiraudeau S. Disability and quality of life of patients with knee or hip osteoarthritis in the primary care setting and factors associated with general practitioners' indication for prosthetic replacement within 1 year. Osteoarthritis Cartilage 2008;16:1024–31.
44. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, *et al.* Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002;32:959–76.
45. Hill CL, Gill T, Taylor AW, Daly A, Grande ED, Adams RJ. Psychological factors and quality of life in arthritis: a population-based study. Clin Rheumatol 2007;26:1049–54.
46. Albrecht G, Devlieger P. The disability paradox: high quality of life against all odds. Soc Sci Med 1999;48:977–88.
47. Hawker G, Melfi C, Paul J, Green R, Bombardier C. Comparison of a generic (SF-36) and a disease specific (WOMAC) (Western Ontario and McMaster Universities Osteoarthritis Index) instrument in the measurement of outcomes after knee replacement surgery. J Rheumatol 1995;22:1193–6.
48. Inoue K, Wicart P, Kawasaki T, Huang J, Ushiyama T, Hukuda S, *et al.* Prevalence of hip osteoarthritis and acetabular dysplasia in french and japanese adults. Rheumatology (Oxford) 2000;39:745–8.

J Orthop Sci (2010) 15:781–789
DOI 10.1007/s00776-010-1545-2

JOURNAL OF
ORTHOPAEDIC SCIENCE

Original article

Normal and threshold values of radiographic parameters for knee osteoarthritis using a computer-assisted measuring system (KOACAD): the ROAD study

Hiroyuki Oka[1], Shigeyuki Muraki[2], Toru Akune[2], Kozo Nakamura[3], Hiroshi Kawaguchi[3], and Noriko Yoshimura[1]

[1]Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan
[2]Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Tokyo, Japan
[3]Department of Orthopaedic Surgery, Sensory and Motor System Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Abstract

Background. Although radiographic severity of the knee is commonly determined by the Kellgren-Lawrence (KL) grading scale, it does not separately assess joint space narrowing or osteophyte formation. The present study aimed to establish normal and threshold values of radiographic parameters for knee osteoarthritis (OA) using the knee osteoarthritis computer-aided diagnosis (KOACAD) measuring system on a large-scale population-based cohort of the Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) population.

Methods. From a total of 3040 participants in the ROAD study, standing anteroposterior radiographs of the knee were obtained from 2975 subjects (1041 men, 1934 women) in the ROAD cohort, and 5950 knees were evaluated using the KOACAD system to obtain the medial and lateral minimum joint space width (mJSW), medial and lateral joint space area (JSA), osteophyte area (OPA), and femorotibial angle (FTA). These indices were compared with the KL scores, and cutoff values for radiographic knee OA were determined by receiver operating characteristic (ROC) curve analysis.

Results. The mean KOACAD parameters for KL = 0 were as follows: medial mJSW 3.70 mm; lateral mJSW 4.77 mm, medial JSA 125.0 mm^2, lateral JSA 140.0 mm^2, OPA 0, and FTA 176.1° in men; for women they were medial mJSW 3.26 mm, lateral mJSW 4.22 mm, medial JSA 100.9 mm^2, lateral JSA 111.0 mm^2, OPA 0, and FTA 174.9°. Threshold values for KL ≥ 2 provided by ROC curve analysis with area under the curve (AUC) > 0.7 were medial mJSW 2.8 mm and medial JSA 107.3 mm^2 in men and medial mJSW 2.7 mm in women. Those for KL ≥ 3 were medial mJSW 2.1 mm, medial JSA 81.1 mm^2, OPA 2.4 mm^2, and FTA 179.6° in men; and medial mJSW 2.1 mm, medial JSA 66.6 mm^2, OPA 2.5 mm^2, and FTA 178.1° in women. We then determined the cutoff values for medial knee OA and lateral knee OA.

Conclusions. The present study established normal and threshold values of parameters for knee OA using an automated computer-assisted program on plain radiographs.

Offprint requests to: N. Yoshimura
Received: January 28, 2010 / Accepted: August 10, 2010

Introduction

Osteoarthritis (OA) is a major public health problem among the elderly that affects activities of daily living (ADL) and quality of life (QOL), leading to increased morbidity and mortality. According to the recent National Livelihood Survey by the Ministry of Health, Labor, and Welfare in Japan, OA is ranked first and fourth among the diseases that cause disabilities requiring support and long-term care, respectively.[1]

Given the increasing proportion of the elderly among the Japanese population, a comprehensive, evidence-based prevention strategy for OA is urgently needed. However, few indices have been identified that can predict the future occurrence and progression of OA. Risk factors for knee OA are known to include age, heavy weight, previous knee injury, and history of work involving overloading the knees.[2–4]

One of the reasons for the scarcity of epidemiological studies on OA might be the diagnostic criteria. Most cases of OA are radiographically determined based on a rating of grade 2 or more on the Kellgren-Lawrence (KL) grading scale.[5] The KL scale is the most conventional grading system for determining the radiographic severity of knee OA, but this categorical scale does not assess joint space narrowing or osteophyte formation separately. Accumulating evidence has shown that osteophytosis and joint space narrowing display distinct etiological mechanisms, and their progression is neither constant nor proportional.[6–8] When examining factors associated with knee OA, these two features should thus be assessed separately. In addition, other indices

of knee OA on plain radiographs, such as the femorotibial angle (FTA) and joint space area, should be determined for evaluation of the severity of knee OA. However, reference values of these indices have yet to be established for a general population.

In the present study, we obtained values for medial and lateral minimum joint space width (mJSW), medial and lateral joint space area (JSA), osteophyte area (OPA), and FTA using the knee osteoarthritis computer-aided diagnosis (KOACAD) measuring system, which we developed and reported elsewhere.[9] We used it in a large-scale population-based cohort study called the Research on Osteoarthritis/Osteoporosis Against Disability (ROAD)[10,11] to establish normal reference and threshold values of radiographic parameters for knee OA.

Participants and methods

Participants

Reference values were obtained based on the results of measurements from the participants of the ROAD study, a nationwide prospective study comprising population-based cohorts established in several communities in Japan. Recruitment methods for this study have been described in detail elsewhere.[9,10] To date, we have completed creation of a baseline database that includes clinical and genetic information of 3040 inhabitants (1061 men, 1979 women) in the age range of 23–95 years (mean 70.6 years), recruited from resident registrations in three communities. All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the participating institutions.

In the present study, from among a total cohort of 3040 participants, 2975 individuals (1041 men, 1934 women) with knee radiographs that could be evaluated by the KOACAD system were selected as subjects for analysis. Among 65 dropouts in the present analysis, 29 underwent total knee arthroplasty (TKA), and 1 had unilateral knee arthroplasty (UKA). The radiographic conditions for the remaining 35 subjects were insufficient for automated analysis due to severe flexion contracture, so we excluded them from the overall analysis.

A summary of the characteristics of subjects are shown in Table 1. No significant differences in baseline characteristics were seen between the 3040 participants in the whole cohort and the 2975 subjects in the present analysis.

Table 1. Summary characteristics of participants

Characteristic	Men	Women
No. of participants	1041	1934
Age (yrs)	71.0 (10.7)	69.8 (11.3)
Height (cm)	162.5 (6.7)	149.8 (6.5)
Weight (kg)	61.3 (10.0)	51.5 (8.6)
Body mass index (kg/m^2)	23.2 (3.0)	22.9 (3.5)
Current smokers	25.7%	3.5%
Current drinkers	64.2%	25.9%

Results are the mean (SD) unless otherwise specified

Radiographic assessment

All participants underwent radiographic examination of both knees using an anteroposterior (AP) view with weight-bearing and foot-map positioning. Fluoroscopic guidance with a horizontal AP X-ray beam was used to visualize the joint space properly, and images were downloaded into digital imaging and communication in medicine (DICOM) format files. The KOACAD system has been described in detail elsewhere[9] and is only briefly summarized here. The KOACAD was programmed to measure mJSW and JSA in the medial and lateral compartments, OPA at the medial tibia, and FTA using digitized knee radiographs. Initially, correction for radiographic magnification was performed based on the image size of a rectangular metal plate.

Next, to determine the region of interest (ROI) including the tibiofemoral joint space, a vertical neighborhood difference filter was applied to identify points with high absolute values for difference of scales. The center of all points was then calculated, and the ROI was chosen. Within the ROI, the outline of the femoral condyle was designated as the upper rim of the joint space. The two ends were determined, and vertical lines from the ends were designated as the outside rims of the joint space. Outlines of anterior and posterior margins of the tibial plateau were drawn similarly to that of the femoral condyle, and the middle line between the two outlines was designated as the lower rim of the joint space (Fig. 1A). A straight regression line for the lower rim outline was then drawn, and the intersections of the lower rim outline and the regression line were designated as the inside rims.

Medial and lateral JSAs were determined as areas surrounded by the upper, lower, inside, and outside rims, as defined above. Medial and lateral mJSWs were further determined as the minimum vertical distances in the respective JSA (Fig. 1B).

To measure the osteophyte area and FTA, medial and lateral outlines of the femur and tibia were drawn. Inflection points for these outlines were then calculated. The medial outline of the tibia from the inflection point was drawn upward to the joint level, and the area that was medially prominent over the smoothly extended outline was designated the osteophyte area (Fig. 1C).

For FTA, a middle line between the medial and lateral outlines of the femur from the top of the image

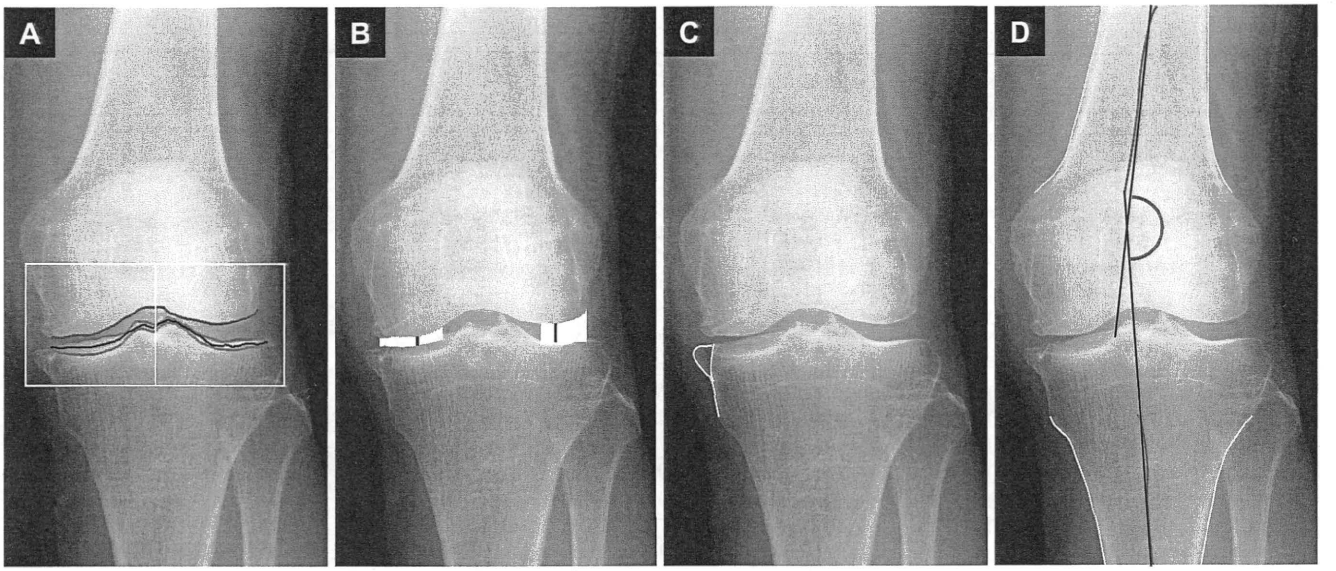

Fig. 1. Schema of image processing by the knee osteoarthritis computer-aided diagnosis (KOACAD) measuring system. **A** Outlines of anterior and posterior margins of the tibial plateau. The *middle line* between these two outlines represents the lower rim of the joint space. **B** Medial and lateral minimum joint space widths (mJSWs) as minimum vertical distances in joint space areas (JSAs). **C** Osteophyte area (area surrounded by *white lines*) determined as the medial prominence over the smoothly extended outline of the tibia. **D** Tibiofemoral angle as the lateral angle between straight regression lines (*black lines*) of the *middle lines* above in the femur and tibia. (From Oka et al.[9], with permission)

to the inflection points was drawn, and the straight regression line was determined as the axis of the femur. Similarly, the straight regression line of the middle line of the tibia from the bottom to the inflection points was designated as the axis of the tibia. The lateral angle between the two axis lines was calculated as the FTA (Fig. 1D).

This system can thus quantify the major features of knee OA on standard radiographs. Moreover, it allows objective, accurate, simple, easy assessment of the structural severity of knee OA without any manual operation in general clinical practice.

The severity of OA was radiographically determined according to the KL grading scale as follows[5]: KL0, normal joint; KL1, slight osteophytes; KL2, definite osteophytes; KL3, disc-space narrowing and large osteophytes; and KL4, bone sclerosis, joint space narrowing, and large osteophytes. In the ROAD study, joints that exhibited only joint space narrowing and no large osteophytes were graded as KL3. All radiographs were examined by a single, experienced orthopedic surgeon (S.M.) who was blinded to the clinical status of the participants.

Establishment of normal values and threshold values for parameters for knee OA

Each index determined using the KOACAD system — medial mJSW, lateral mJSW, medial JSA, lateral JSA, OPA, FTA — was compared with KL scores. First, we established normal values of these parameters using mean values for knees with KL grade 0. Cutoff values for radiographic knee OA for KL ≥ 2 and KL ≥ 3 were then determined using receiver operating characteristic (ROC) curve analysis. The present study adopted cutoff points that maximized the area under the ROC curve (AUC) as threshold values. Threshold values were determined for total OA including medial and lateral OA, and then separate threshold values were determined for medial and lateral OA. Although parameters for medial OA could be obtained separately for each sex, those for lateral OA were evaluated for the combined number of men and women, given the small number of cases with lateral OA.

Statistical analysis

All statistical analyses were performed using STATA statistical software (STATA, College Station, TX, USA). First, the Kolmogrov-Smirnov Lilliefors test was used to evaluate the normality of distribution of each variable. All parameters were confirmed to fit normal distributions (*P* values after Kolmogrov-Smirnov Lilliefors test: medial mJSW 0.40; lateral mJSW 0.6; medial JSA 0.37; lateral JSA 0.76; OPA 0.10; FTA 0.49). Differences in values of the parameters were tested for significance using analysis of variance (ANOVA) for comparisons among multiple groups. Trends of values

Table 2. Various parameters classified by age and sex

Age strata (years)	No. of knees	Medial mJSW (mm)	Lateral mJSW (mm)	Medial JSA (mm^2)	Lateral JSA (mm^2)	OPA (mm^2)	FTA (°)
Men							
<40	28	4.12 (0.92)	4.91 (1.22)	141.6 (27.3)	153.2 (30.1)	0	176.5 (2.2)
40–49	82	3.67 (0.75)	5.06 (1.10)	132.9 (27.9)	156.2 (34.4)	0.40 (1.98)	177.5 (2.4)
50–59	214	3.63 (0.75)	4.88 (1.06)	124.1 (26.9)	148.5 (33.1)	0.30 (1.51)	176.5 (3.2)
60–69	334	3.37 (0.93)[a]	4.59 (0.96)[bc]	113.9 (28.4)[abc]	138.9 (28.8)[bc]	0.96 (3.60)	177.0 (3.0)
70–79	1052	3.13 (0.96)[abcd]	4.40 (1.02)[bc]	102.5 (26.7)[abcd]	125.7 (30.3)[abcd]	1.35 (4.68)[c]	177.1 (3.3)
80+	372	2.94 (0.98)[abcd]	4.22 (0.87)[abcd]	97.2 (27.4)[abcd]	121.3 (28.0)[abcd]	1.31 (4.06)	177.0 (4.0)
Total	2082	3.22 (0.96)	4.48 (1.02)	107.3 (29.0)	130.6 (31.7)	1.12 (4.07)	177.0 (3.3)
Women							
<40	62	3.37 (0.61)	4.31 (1.23)	106.3 (24.0)	116.8 (32.6)	0.18 (1.25)	175.6 (3.0)
40–49	210	3.22 (0.64)	4.36 (1.01)	104.0 (22.2)	116.9 (25.8)	0.46 (2.09)	175.5 (2.7)
50–59	418	3.03 (0.78)	4.23 (1.15)	97.5 (26.2)	112.3 (28.3)	0.96 (2.87)	175.8 (4.0)
60–69	762	2.80 (0.98)[abc]	4.03 (1.09)[b]	89.8 (28.7)[abc]	106.6 (28.7)[bc]	2.33 (6.39)	176.4 (3.8)
70–79	1764	2.52 (0.92)[abcd]	3.87 (1.00)[bcd]	79.5 (26.3)[abcd]	99.4 (26.8)[abcd]	4.60 (11.2)[abcd]	176.9 (4.2)[bc]
80+	652	2.32 (0.95)[abcde]	3.80 (1.08)[abcd]	77.4 (26.9)[abcd]	97.9 (28.0)[abcd]	6.39 (12.70)[abcde]	177.1 (4.6)[bc]
Total	3868	2.65 (0.95)[#]	3.96 (1.07)[#]	84.9 (28.0)[#]	103.2 (28.2)[#]	3.76 (9.87)[#]	176.6 (4.1)[#]

Results are the mean (SD)
mJSW, minimal joint space width; JSA, joint space area; OPA, osteophyte area; FTA, femoro-tibial angle; KL grade: Kellgren-Lawrence grade
[a] Significantly different from those of the group <40 years ($P < 0.05$)
[b] Significantly different from those of the group in their 40s ($P < 0.05$)
[c] Significantly different from those of the group in their 50s ($P < 0.05$)
[d] Significantly different from those of the group in their 60s ($P < 0.05$)
[e] Significantly different from those of the group in their 70s ($P < 0.05$)
[#] Significantly different from those of men ($P < 0.001$)

according to the KL grade were tested using the Jonckheere-Terpstra trend test. Scheffe's least significant difference test was then used for pairs of groups.

Results

The mean values for mJSW, JSA, OPA, and FTA, classified by sex, are shown in Table 2. The values for medial and lateral mJSW, medial and lateral JSA, OPA, and FTA all differed significantly between the sexes ($P < 0.001$). The mean values for medial mJSW, JSA, and FTA were significantly greater in men than in women ($P < 0.001$). By contrast, the OPA values in both knees were significantly lower in men than in women ($P < 0.001$).

The mean values for mJSW, JSA, OPA, and FTA classified by age are also shown in Table 2. The medial mJSW for men in their sixties was significantly smaller than that for men <40 years; and that for men ≥70 years was significantly smaller than that of men <70 years ($P < 0.05$). Lateral mJSW for subjects in their sixties and seventies was significantly smaller than that for subjects in their forties and fifties ($P < 0.05$). Medial JSA for subjects in their sixties was significantly smaller than that for subjects <60 years old ($P < 0.05$); and that for subjects ≥70 years was significantly smaller than that for subjects <70 years ($P < 0.05$). Lateral JSA for subjects in their sixties was significantly smaller than that for subjects in their forties and fifties ($P < 0.05$); and that for subjects ≥70 years was significantly smaller than that for subjects <70 years ($P < 0.05$). No significant differences in OPA or FTA were seen according to age in men except for OPA in subjects between their fifties and seventies. In women, the medial mJSW for subjects in their sixties was significantly smaller than that for subjects <60 years ($P < 0.05$); that for women in their seventies was significantly smaller than that for subjects <70 years ($P < 0.05$); and that for women ≥80 years was significantly smaller than that of subjects <80 years ($P < 0.05$). The lateral mJSW for subjects in their sixties was significantly smaller than that for subjects in their forties ($P < 0.05$); that for subjects in their seventies was significantly smaller than that for subjects in their forties, fifties, and sixties ($P < 0.05$); and that for subjects ≥80 years was significantly smaller than that for subjects <70 years ($P < 0.05$). The medial JSA and lateral JSA showed trends with age similar to those of men. However, the mean OPA was significantly larger in women ≥70 years than in younger women ($P < 0.05$), and the FTA was significantly larger for subjects ≥70 years than for women in their forties and fifties ($P < 0.05$).

Table 3 shows mean values for mJSW, JSA, OPA, and FTA classified by KL grade. All values increased significantly according to the severity of the KL grade (P for trends < 0.0001). Regarding differences in the above-mentioned indices, values for medial mJSW and

Table 3. Various parameters classified by Kellgren-Lawrence grade

KL grade	Proportion of knees (%)	Medial mJSW (mm)	Lateral mJSW (mm)	Medial JSA (mm^2)	Lateral JSA (mm^2)	OPA (mm^2)	FTA (°)
Men							
KL0	24.4	3.70 (0.77)	4.77 (1.01)	125.0 (27.1)	140.0 (33.6)	0	176.1 (2.6)
KL1	38.4	3.40 (0.76)[a]	4.50 (0.93)[a]	109.8 (23.5)[a]	128.9 (29.0)[a]	0.48 (2.24)	176.6 (2.7)
KL2	28.5	3.02 (0.78)[ab]	4.38 (1.02)[a]	99.3 (22.5)[ab]	125.1 (29.8)[a]	1.08 (3.25)[ab]	177.5 (3.1)[ab]
KL3	6.3	2.10 (1.00)[abc]	4.06 (1.40)[abc]	84.1 (31.3)[abc]	129.5 (38.2)[a]	5.37 (8.70)[abc]	178.1 (4.5)[ab]
KL4	2.4	0.79 (0.84)[abcd]	4.04 (1.12)[ab]	44.7 (32.7)[abcd]	137.3 (39.5)	12.05 (10.36)[abcd]	184.2 (6.2)[abcd]
Total	100.0	3.22 (0.96)	4.48 (1.02)	107.3 (29.1)	130.8 (31.8)	1.12 (4.08)	177.0 (3.3)
Women							
KL0	13.9	3.26 (0.65)	4.22 (1.08)	100.9 (23.7)	111.0 (29.4)	0	174.9 (2.9)
KL1	30.6	2.95 (0.73)[a]	3.95 (0.99)[a]	89.7 (24.3)[a]	101.3 (26.0)[a]	0.68 (2.26)	175.6 (3.0)[a]
KL2	38.3	2.66 (0.73)[ab]	3.93 (0.96)[a]	84.5 (23.5)[ab]	100.3 (25.5)[a]	3.39 (6.67)[ab]	176.6 (3.3)[ab]
KL3	13.1	1.85 (0.92)[abc]	3.91 (1.20)[a]	73.3 (27.4)[abc]	106.5 (30.2)[bc]	11.15 (17.54)[abc]	178.7 (4.8)[abc]
KL4	4.1	0.67 (1.02)[abcd]	3.83 (1.68)[a]	34.6 (34.8)[abcd]	112.1 (43.7)[bc]	19.70 (20.65)[abcd]	183.8 (7.1)[abcd]
Total	100.0	2.65 (0.94)[#]	3.97 (1.06)[#]	84.9 (27.9)[#]	103.4 (28.1)[#]	3.76 (9.90)[#]	176.6 (4.1)[#]

Results are the mean (SD)
[a] Significantly different from those of KL0 ($P < 0.05$)
[b] Significantly different from those of KL1 ($P < 0.05$)
[c] Significantly different from those of KL2 ($P < 0.05$)
[d] Significantly different from those of KL3 ($P < 0.05$)
[#] Significantly different from those of men ($P < 0.05$)

medial JSA in both sexes tended to be smaller with increasing KL grade ($p < 0.05$). Values for OPA and FTA in both sexes were significantly larger in the KL 2–4 group than in the KL 0–1 group ($P < 0.05$). Age differences in values of lateral mJSW and JSA were smaller than those for medial mJSW and JSA.

We performed ROC curve analysis to identify threshold values of these indices to determine the knee OA defined by KL ≥ 2 and KL ≥ 3. ROC curve analysis provided threshold values of KL ≥ 2 and KL ≥ 3 in OA parameters for the two knees (Table 4). Threshold values of KOACAD parameters for KL ≥ 2 with AUC > 0.7 were medial mJSW 2.8 mm and medial JSA 107.3 mm^2 in men and medial mJSW 2.7 mm in women. Those for KL ≥ 3 were medial mJSW 2.1 mm, medial JSA 81.1 mm^2, OPA 2.4 mm^2, and FTA 179.6° in men; and they were medial mJSW 2.1 mm, medial JSA 66.6 mm^2, OPA 2.5 mm^2, and FTA 178.1° in women. In contrast, the AUC of the lateral mJSW and lateral JSA for KL ≥ 2 and KL ≥ 3 in OA parameters was near 0.5, meaning that the capacity of these parameters to distinguish diseased knees from normal knees was low.

In addition, we provided threshold values for parameters for both the medial and lateral knee OA using ROC curve analysis (Table 4). Medial OA comprised 97.8% of total OA cases, with the lateral type making up the remaining 2.2%. Although most threshold values for medial OA were similar to those for total OA, the AUC values for parameters of medial OA (e.g., medial mJSW, medial JSA) were higher than for total OA. In contrast, for lateral OA, the AUC values for lateral mJSW and lateral JSA for KL ≥ 2 and KL ≥ 3 in OA parameters were higher than those for total OA, which were near 0.99, meaning that the capacity of these parameters to distinguish disease states from the normal population was high.

Discussion

We have reported elsewhere the automated computer-assisted program KOACAD, which can accurately measure values of mJSW, JSA, OPA, and FTA.[9] In the previous report,[9] we clarified that KOACAD allows accurate, easy assessment of the structural severity of knee OA without any manual operation. The present study applied this system to baseline data from the ROAD study, obtaining normal and threshold values of the above-mentioned indices for objective diagnosis of knee OA.

In the present study, we first established normal values for mJSW, JSA, OPA, and FTA using mean values of these parameters for knees with KL grade 0. The mean values were medial mJSW 3.70 mm, lateral mJSW 4.77 mm, medial JSA 125.0 mm^2, lateral JSA 140.0 mm^2, OPA 0, and FTA 176.1° in men; and medial mJSW 3.26 mm, lateral mJSW 4.22 mm, medial JSA 100.9 mm^2, lateral JSA 111.0 mm^2, OPA 0, and FTA 174.9° in women. All these indices except OPA were significantly lower in women than in men, suggesting that the values are influenced by differences in stature. We concluded that normal and threshold values for knee OA should be established for each sex.

The JSW has been recommended as a candidate index for progression of knee OA,[12] but few data

Table 4. Threshold values of various parameters, by Kellgren-Lawrence grades 2 and 3

Parameter	Threshold value	AUC	Sensitivity	Specificity (%)
Total				
KL ≥ 2				
Men				
Medial mJSW (mm)	2.8	0.726	58.4	76.8
Lateral mJSW(mm)	4.3	0.566	52.3	59.0
Medial JSA (mm^2)	107.3	0.715	71.0	60.3
Lateral JSA (mm^2)	115.5	0.551	39.5	68.2
OPA (mm^2)	1.0	0.599	23.9	95.5
FTA (°)	178.5	0.633	42.7	79.3
Women				
Medial mJSW (mm)	2.7	0.730	63.7	72.5
Lateral mJSW(mm)	4.3	0.521	66.4	38.5
Medial JSA (mm^2)	85.9	0.654	64.5	59.9
Lateral JSA (mm^2)	79.2	0.509	19.8	83.4
OPA (mm^2)	1.0	0.691	44.3	92.4
FTA (°)	177.4	0.664	48.6	77.0
KL ≥ 3				
Men				
Medial mJSW (mm)	2.1	0.875	73.6	92.1
Lateral mJSW(mm)	4.3	0.608	65.2	54.3
Medial JSA (mm^2)	81.1	0.800	58.4	88.9
Lateral JSA (mm^2)	135.7	0.522	50.0	60.1
OPA (mm^2)	2.4	0.739	52.8	93.5
FTA (°)	179.6	0.702	52.5	85.5
Women				
Medial mJSW (mm)	2.1	0.842	65.3	92.0
Lateral mJSW(mm)	2.5	0.507	15.7	93.0
Medial JSA (mm^2)	66.6	0.717	48.7	83.2
Lateral JSA (mm^2)	116.4	0.562	38.8	73.0
OPA (mm^2)	2.5	0.768	66.1	82.2
FTA (°)	178.1	0.744	64.6	76.3
Medial OA				
KL ≥ 2				
Men				
Medial mJSW (mm)	2.8	0.728	58.5	76.8
Lateral mJSW(mm)	4.3	0.560	51.7	59.0
Medial JSA (mm^2)	107.3	0.717	71.3	60.3
Lateral JSA (mm^2)	115.5	0.545	38.8	68.2
OPA (mm^2)	1.2	0.599	23.9	95.5
FTA (°)	178.5	0.639	43.2	79.3
Women				
Medial mJSW (mm)	2.7	0.732	63.9	72.5
Lateral mJSW(mm)	5.4	0.505	92.9	10.9
Medial JSA (mm^2)	85.9	0.655	64.7	59.9
Lateral JSA (mm^2)	97.9	0.505	56.1	46.3
OPA (mm^2)	1.0	0.693	44.7	92.4
FTA (°)	177.4	0.677	49.9	77.0
KL ≥ 3				
Men				
Medial mJSW (mm)	2.1	0.888	76.3	90.4
Lateral mJSW(mm)	4.3	0.598	64.2	54.4
Medial JSA (mm^2)	81.1	0.809	59.0	89.0
Lateral JSA (mm^2)	135.3	0.536	52.6	59.7
OPA (mm^2)	2.4	0.741	53.2	93.4
FTA (°)	179.6	0.719	54.0	85.5
Women				
Medial mJSW (mm)	2.1	0.854	66.6	92.2
Lateral mJSW(mm)	4.6	0.512	29.7	75.8
Medial JSA (mm^2)	66.6	0.727	49.4	83.4
Lateral JSA (mm^2)	116.5	0.587	40.8	72.8
OPA (mm^2)	2.5	0.774	67.3	82.1
FTA (°)	178.1	0.771	67.9	76.0

Table 4. *Continued*

Parameter	Threshold value	AUC	Sensitivity	Specificity (%)
Lateral OA				
KL ≥ 2				
Men and women				
Medial mJSW (mm)	2.1	0.683	43.1	92.4
Lateral mJSW(mm)	2.2	0.995	100.0	98.1
Medial JSA (mm^2)	75.7	0.664	50.0	84.2
Lateral JSA (mm^2)	69.7	0.990	100.0	95.4
OPA (mm^2)	1.2	0.626	30.6	93.8
FTA (°)	173.3	0.795	65.3	81.5
KL ≥ 3				
Men and women				
Medial mJSW (mm)	2.1	0.680	46.0	92.0
Lateral mJSW(mm)	2.2	0.992	100.0	97.0
Medial JSA (mm^2)	75.1	0.638	48.7	84.5
Lateral JSA (mm^2)	69.1	0.987	100.0	95.6
OPA (mm^2)	4.8	0.706	43.2	96.5
FTA (°)	173.3	0.805	64.9	80.8

AUC, area under the curve

regarding normal values have been accumulated.[13] Gensburger et al. showed that the mean medial and lateral JSW in women were 5.1 mm and 6.0 mm, respectively,[13] suggesting that those values in Caucasian populations may be larger than our results in women; no normal values for men were available. In addition, although evaluations of knee alignment are known to be useful for diagnosing arthritic conditions affecting the knee joint and also serve as a guide for conservative management and surgical planning,[14,15] few reports have shown normal values of FTA along with JSA and OPA.

Koshino measured the FTA of 85 knees in men and 97 knees in women aged 25–35 years and reported normal FTA values of 178° in men and 176° in women.[16] These results seem broadly consistent with our results, although no sex differences were apparent in our study, with values of 176° for both men and women. In any case, this represents the first report of reference values for the above-mentioned parameters using a population-based cohort. The results may thus be useful for diagnosing knee OA. Furthermore, by a longitudinal follow-up of the present cohort, these parameters would be expected to predict the progress of knee OA.

We then determined the threshold values for knee OA using ROC curve analysis. In this analysis, we regarded parameters with AUC > 0.7 as good indices for features of knee OA according to a previous report.[17] For KL ≥ 2, threshold values of KOACAD parameters with AUC > 0.7 were only the mJSW in men and women and the medial JSA in men. AUCs > 0.7 on ROC curve analysis means that the threshold of parameters might show good capacity for accurate diagnosis of the disorder in question. In contrast, AUCs of threshold values of parameters regarding the lateral region (i.e., KL ≥ 2; 0.566 for lateral mJSW 4.3 mm, 0.551 for lateral JSA 115.5 mm^2 in men; 0.521 for lateral mJSW 4.3 mm, 0.509 for lateral JSA 79.2 mm^2 in women) seem insufficient as indicators for knee OA. In contrast, for KL ≥ 3, OPA and FTA seem to represent good predictors with satisfactory AUCs. These results suggest that such parameters are more useful in severe knee OA than in mild knee OA.

We also tried to determine threshold values for medial knee OA and lateral knee OA. Because most cases of knee OA were medial knee OA (97.8%), the above-mentioned threshold values were considered applicable for medial OA. Conversely, in the diagnosis for lateral OA, for both KL ≥ 2 and KL ≥ 3, threshold values for medial mJSW and medial JSA were no longer parameters with good predictive capacity. By contrast, AUCs of threshold values for parameters of the lateral region (KL ≥ 2: 0.995 for lateral mJSW 2.2 mm, 0.990 for lateral JSA 69.7 mm^2; KL ≥ 3: 0.992 for lateral mJSW 2.2 mm, 0.987 for lateral JSA 69.1 mm^2) were preferable as good predictors. Similar to medial knee OA, for KL ≥ 3 the OPA and FTA seem to represent good predictors with satisfactory AUC. These results suggest that parameters at the medial side are useful in medial knee OA, and parameters at the lateral side are useful in lateral knee OA. However, evaluation of lateral OA was performed in only 65 participants (2.2%), so we could not analyze data for men and women separately. Regarding lateral OA and threshold KOACAD parameters, further investigation is warranted.

On the other hand, discrepancies between continuous values obtained from the KOACAD system and categorical scales such as the KL scale might add to the limitations of the KL grading scale. Most previous

studies have been performed in patients with knee OA defined by a KL score; but utilizing this categorical scale at the diagnosis of OA seems to result in the loss of a considerable amount of information, as the contribution of joint space narrowing and osteophytes is relatively small. Even though these indices are linear and constant in number, joint space narrowing is simply categorized as mild or severe and osteophytes as slight, definite, or large. In addition, the optimal method for handling joints with severe joint space narrowing but no osteophyte formation is unclear.

One solution to such problems might be found in a radiographic atlas of individual features published by the Osteoarthritis Research Society International (OARSI).[18] OARSI proposed a new grading scale in which joint space narrowing and osteophyte formation at the medial and lateral tibiofemoral compartments on radiographs should be evaluated separately. Several studies have evaluated the severity of joint space narrowing and osteophytes in the osteoarthritic knee utilizing the OARSI scale,[19] although these studies did not assess distinct features of knee OA such as joint space narrowing, osteophyte formation, or joint angulation in one sitting. To the best of our knowledge, no quantitative assessment systems for osteophytes have been described other than in the KOACAD,[9] so the present study is the first to assess threshold values for knee OA in a population-based cohort.

Unlike categorical methods for grading the severity of knee OA (e.g., KL or OARSI scales), KOACAD enables measurement of independent parameters for knee OA. We have already confirmed that low medial mJSW and high FTA are associated with the presence of knee pain, unlike lateral mJSW or an osteophyte area.[9] These accurate and continuous parameters obtained by KOACAD might be candidates for predictors of rapid progress from mild knee OA. These parameters might also be helpful for assessing risk factors for the occurrence of OA. We assumed that 25.3 million people (8.6 million men, 16.7 million women) ≥40 years would be affected by radiographic knee OA, and 7.8 million people (2.2 million men, 5.6 million women) ≥40 years would be affected by knee OA with knee pain.[10] Preventive strategies for OA are certainly in urgent demand. At the planning stage for the strategies against knee OA, the provision of accurate, objective, quantitative indices to measure outcomes seems highly important.

However, some limitation might apply to automated systems for all knee OA. First, as we stated, the number of cases with lateral knee OA was small for accurate determination of thresholds. Second, with radiographs of cases showing severe flexion contracture of the knee (>20°), the KOACAD system failed to measure parameters automatically. However, the system includes a manual mode, and in such cases orthopedic specialists can obtain values by manual measurement.

We believe this system may not only be useful for objective evaluation of knee OA in daily clinical practice or population-based epidemiological studies, it also acts as a proper surrogate measure for the development of disease-modifying drugs for OA. We hope in the future that this system will be applied worldwide to develop international criteria and for the diagnosis and treatment of knee OA.

Conclusion

We have established normal and threshold values of parameters for knee OA using an automated computer-assisted program, KOACAD, on plain radiographs.

Acknowledgments. This work was supported by the following Grants-in-Aid for Scientific Research: B20390182 (Noriko Yoshimura); C20591737 (Toru Akune); C20591774 (Shigeyuki Muraki); Young Scientists A18689031 (Hiroyuki Oka); and Collaborating Research with NSF 08033011-00262 (Director, Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science, and Technology. H17-Men-eki-009 (Director, Kozo Nakamura); H18-Choujyu-037 (Director, Toshitaka Nakamura); and H20-Choujyu-009 (Director, Noriko Yoshimura) from the Ministry of Health, Labor, and Welfare in Japan. This study was also supported by grants from the Japan Osteoporosis Society (Noriko Yoshimura), Nakatomi Foundation (Noriko Yoshimura), and research aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1) (Director, Hiroshi Kawaguchi). The sponsors had no role in the study design, data collection, data analysis, data interpretation, or writing of this report. The authors thank Mrs. Tomoko Takijiri and other members of the Public Office in Hidakagawa Town; Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, and other members of the Public Office in Taiji Town; Dr. Seizo Yamamoto and other members of the Department of Orthopaedics; Mr. Kutsuma and other members of Department of Radiology at Tokyo Metropolitan Geriatric Medical Center for their assistance in the location and scheduling of participants for examinations.

No conflict of interest has been declared by the authors.

References

1. Ministry of Health, Labour and Welfare (2007) The outline of the results of National Livelihood Survey 2007 available at http://www.mhlw.go.jp/toukei/list/20-19-1.html.
2. Du H, Chen SL, Bao CD, Wang XD, Lu Y, Gu YY, et al. Prevalence and risk factors of knee osteoarthritis in Huang-Pu District, Shanghai, China. Rheumatol Int 2005;25:585–90.
3. Hart DJ, Spector TD. The relationship of obesity, fat distribution and osteoarthritis in women in the general population: the Chingford Study. J Rheumatol 1993;20:331–5.
4. Anderson JJ, Felson DT. Factors associated with osteoarthritis of the knee in the first national Health and Nutrition Examination Survey (HANES I): evidence for an association with overweight, race, and physical demands of work. Am J Epidemiol 1988;128:179–89.

5. Kellgren JH, Lawrence LS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
6. Yamada T, Kawano H, Koshizuka Y, Fukuda T, Yoshimura K, Kamekura S, et al. Carminerin contributes to chondrocyte calcification during endochondral ossification. Nat Med 2006;12: 665–70.
7. Kamekura S, Kawasaki Y, Hoshi K, Shimoaka T, Chikuda H, Maruyama Z, et al. Contribution of runt-related transcription factor 2 to the pathogenesis of osteoarthritis in mice after induction of knee joint instability. Arthritis Rheum 2006;54:2462–70.
8. Jones G, Ding C, Scott F, Glisson M, Cicuttini F. Early radiographic osteoarthritis is associated with substantial changes in cartilage volume and tibial bone surface area in both males and females. Osteoarthritis Cartilage 2004;12:169–74.
9. Oka H, Muraki S, Akune T, Mabuchi A, Suzuki T, Yoshida H, et al. Fully automatic quantification of knee osteoarthritis severity on standard radiographs. Osteoarthritis Cartilage 2008;16: 1300–6.
10. 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 study. J Bone Miner Metab 2009;27:620–8.
11. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: Research on Osteoarthritis/ Osteoporosis Against Disability (ROAD) study. Int J Epidemiol (in press, 2010).
12. Altman R, Brandt K, Hochberg M, Moskowitz R, Bellamy N, Bloch DA, et al. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society: results from a workshop. Osteoarthritis Cartilage 1996;4:217–43.
13. Gensburger D, Arlot M, Sornay-Rendu E, Roux JP, Delmas P. Radiologic assessment of age-related knee joint space changes in women: a 4-year longitudinal study. Arthritis Rheum 2009;61: 336–43.
14. Coventry MB. Upper tibial osteotomy for osteoarthritis. J Bone Joint Surg Am 1985;67:1136–40.
15. Cerejo R, Dunlop DD, Cahue S, Channin D, Song J, Sharma L. The influence of alignment on risk of knee osteoarthritis progression according to baseline stage of disease. Arthritis Rheum 2002;46:2632–6.
16. Koshino T. Surgical treatment of arthrosis deformans of the knee: high tibial osteotomy. Nippon Seikeigeka Gakkai Zasshi 1971;45: 1121–33.
17. Hosmer DW, Lemeshow S. Assessing the fit of the model. In: Hosmer DW, Lemeshow S, editors. Applied logistic regression, 2nd edn. New York: Wiley; 2000. p. 143–202.
18. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage 2007;15(suppl A):A1–56.
19. 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.

Evidence for Geographical and Racial Variation in Serum Sex Steroid Levels in Older Men

Eric S. Orwoll, Carrie M. Nielson, Fernand Labrie, Elizabeth Barrett-Connor, Jane A. Cauley, Steven R. Cummings, Kristine Ensrud, Magnus Karlsson, Edith Lau, P. C. Leung, Östen Lunggren, Dan Mellström, Alan L. Patrick, Marcia L. Stefanick, Kozo Nakamura, Noriko Yoshimura, Joseph Zmuda, Liesbeth Vandenput, and Claes Ohlsson, for the Osteoporotic Fractures in Men (MrOS) Research Group

Oregon Health & Science University (E.S.O., C.M.N.), Portland, Oregon 97239; Le Centre hospitalier de l'Université Laval (F.L.), Québec City, Canada G1K 7P4; University of California, San Diego (E.B.-C.), La Jolla, California 92161; University of Pittsburgh (J.A.C., A.L.P., J.Z.), Pittsburgh, Pennsylvania 15260; San Francisco Coordinating Center (S.R.C.), San Francisco, California 94107; University of Minnesota (K.E.), Minneapolis, Minnesota 55455; Lund University (M.K.), 221 00 Malmö, Sweden; The Chinese University of Hong Kong (E.L., P.C.L.), Hong Kong 1, China; University of Uppsala (O.L.), 753 12 Uppsala, Sweden; University of Gothenburg (D.M., L.V., C.O.), 411 24 Göteborg, Sweden; Stanford University (M.L.S.), Palo Alto, California 94305; and University of Tokyo (K.N., N.Y.), Tokyo 113-8655, Japan

Background: Despite considerable racial and geographical differences in human phenotypes and in the incidence of diseases that may be associated with sex steroid action, there are few data concerning variation in sex steroid levels among populations. We designed an international study to determine the degree to which geography and race influence sex steroid levels in older men.

Methods: Using mass spectrometry, concentrations of serum androgens, estrogens, and sex steroid precursors/metabolites were measured in 5003 older men from five countries. SHBG levels were assessed using radioimmunoassay.

Results: There was substantial geographical variation in the levels of sex steroids, precursors, and metabolites, as well as SHBG. For instance, Asian men in Hong Kong and Japan, but not in the United States, had levels of total testosterone approximately 20% higher than in other groups. Even greater variation was present in levels of estradiol, SHBG, and dihydrotestosterone. Group differences in body mass index did not explain most geographical differences. In addition, body mass index-independent racial differences were present; Black men had higher levels of estrogens (estradiol, estrone), and Asian men had lower levels of glucuronidated androgen metabolites.

Conclusions: On a global scale, there are important geographical and racial differences in the concentrations of serum sex steroids and SHBG in older men. ***(J Clin Endocrinol Metab* 95: E151–E160, 2010)**

Sex steroids have pleiotropic actions in men. In addition to their effects on reproductive tissues, they influence numerous biological systems including those that determine body composition, immunological function, glucose and lipid metabolism, and cardiovascular health. Aging is well known to result in a decline in sex hormone levels, and those changes have been linked to disorders such as frailty, muscle and skin atrophy, fat accumulation, osteoporosis, and prostate disease, as well as sexual and cognitive dysfunction.

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.

doi: 10.1210/jc.2009-2435 Received November 16, 2009. Accepted June 18, 2010.
First Published Online July 28, 2010

Abbreviations: ADT, Androsterone; ADT-G, ADT-glucuronide; CI, confidence interval; CV, coefficient of variation; DHEA, dehydroepiandrosterone; DHEA-S, DHEA sulfate; DHT, dihydrotestosterone; 3α-diol-3G, androstane-3α,17β-diol-3-glucuronide; 3α-diol-17G, androstane-3α,17β-diol-17-glucuronide; 4-dione, androstenedione.

Racial and geographical differences exist in some human phenotypes that are influenced by sex steroids (*e.g.* lean mass, bone density) and in the incidence of age-related disorders that are linked to sex steroid actions. For instance, the risk of developing prostate cancer is higher in Black men, whereas Asian men living in the Far East have rates considerably lower than those living in Western countries (1, 2). These differences have stimulated interest in whether serum sex steroid concentrations or metabolism are affected by race and geography. Although some comparisons have suggested that there are racial differences in sex hormone levels (3–5), those findings are not consistent and are difficult to interpret because of small study sizes and the use of some sex steroid assay methods (RIA) with questionable specificity in lower concentration ranges. There are fewer studies of geographical variation, but some indicate that there may be environmental factors that affect sex hormone physiology (6, 7).

To examine racial and geographic differences in sex hormone levels, we assembled a large international cohort of older men and measured the serum levels of the most biologically active estrogens and androgens as well as their precursors and metabolites.

Subjects and Methods

Study populations

We evaluated 5003 ambulatory, community-dwelling men at least 65 yr of age from five countries (Japan, Hong Kong, Sweden, Tobago, and the United States). All were participants in ongoing observational studies that used similar enrollment criteria and evaluation methods (Supplemental Table 1, published on The Endocrine Society's Journals Online web site at http://jcem.endojournals.org).

Men in the United States, Sweden, and Hong Kong were participants in the Osteoporotic Fractures in Men Study (MrOS). Details of the U.S. MrOS cohort have been described (8). Self-defined racial/ethnic ancestry (Caucasian, Black, or Asian) was ascertained through questionnaire; fasting morning serum was collected. A random sample of MrOS Caucasian men and all Black and Asian men with sufficient stored serum were included. Samples were collected between 0800 and 1200 h, and serum was prepared and frozen for 6–7 yr in 1.0-ml cryovials at −80 C until assay. MrOS participants in Sweden (9) were Caucasian; all participants with sufficient morning, fasting serum were included. After phlebotomy, serum was prepared and stored for 2–4 yr in 0.5-ml microtubes at −70 C until assay. The MrOS Study in Hong Kong included Chinese men (10). All participants with sufficient stored serum from a morning, fasting blood draw were included. Serum was prepared and stored for 4–5 yr at −20 to −85 C in 0.6-ml microtubes. Vials were briefly thawed and refrozen once before assay.

In Tobago, the parent study was the Tobago Prostate Cancer Survey (11), a study of men of West African descent. A random sample of 500 participants at least 65 yr old was included. Fasting morning serum was obtained. After phlebotomy, serum was prepared and stored for 0.5–2.5 yr in 2.0-ml tubes at −80 C until assay.

In Japan, the parent study was the Research on Osteoarthritis and Osteoporosis Against Disability (ROAD) Study. Japanese participants were recruited from resident registration lists in two communities in the Wakayama District: a mountainous region in Hidakagawa and a coastal region in Taiji (12). Samples from all 367 men who were at least 65 yr old were analyzed; serum was collected between 0900 and 1500 h and stored for 1–2 yr in 2.0-ml vials at −80 C until assay.

The Institutional Review Board at each study center approved the protocol, and written informed consent was obtained from all participants.

Laboratory methods

The steroids measured are depicted in Fig. 1. All hormone assays were performed at the Centre de Recherché du CHUL (University of Laval, Quebec, Canada) using gas chromatography/liquid chromatography/mass spectrometry (13). Serum was stored at −80 C until assay. Standard samples of several concentrations were included in all assay runs to ensure precision and accuracy. Moreover, blinded aliquots of a single serum pool included in all assay runs showed no longitudinal trend in any measurement. Ranges of detection for each steroid and mean coefficients of variation (CV) for pooled serum aliquots are in Supplemental Table 2.

SHBG assays for Japan, the United States, and Tobago samples were performed at the Oregon Clinical and Translational Research Institute (Diagnostic Products Corp., Los Angeles, CA; CV, 7.1%), and Hong Kong and Sweden samples at the Clinical Pharmacology Lab, Sahlgrenska University Hospital, Gothenburg (Orion Diagnostics, Espoo, Finland; CV, 10.7%). Duplicate measures of 50 serum samples in each laboratory resulted in nearly identical mean values (40.3 *vs.* 40.4 nM) and SD values (14.2 *vs.* 13.2) and a very high correlation ($r = 0.98$). A Bland-Altman plot showed that 48 of 50 (96%) pairs fell within the 95% confidence limits of agreement (Supplemental Fig. 1). In addition, we tested a term for "laboratory" in the adjusted SHBG models and found that the term was not statistically significant ($P = 0.25$) and did not alter our estimates of mean SHBG levels across country or race. For these reasons, we assume any small differences in assay performance between the labs are not influential. Free fractions of testosterone and estradiol were calculated as described by Sodergard *et al.* (14). These calculated values appear to accurately reflect directly measured results (15–18).

Statistical methods

Men who reported the use of androgen or antiandrogen therapy had been orchidectomized or had testosterone or dihydrotestosterone (DHT) below the limit of detection (Supplemental Table 2) were excluded. Seasonal variation was not present in any steroid, and adjustment for season did not reduce group differences. Square root or natural log transformations were made to produce normally distributed variables. Age-adjusted least-squares means and 95% confidence intervals (CIs) were calculated for each group, and a P for ANOVA was presented. Age-group and BMI-group interactions were tested for each analyte; when significant ($P < 0.10$), the interaction term was retained to allow age and BMI adjustments to be group-specific. Means and CIs were back-transformed to facilitate the presentation of results in the original units of measure. Evaluation of patterns in mean concentrations across groups led to *post hoc* tests of differ-

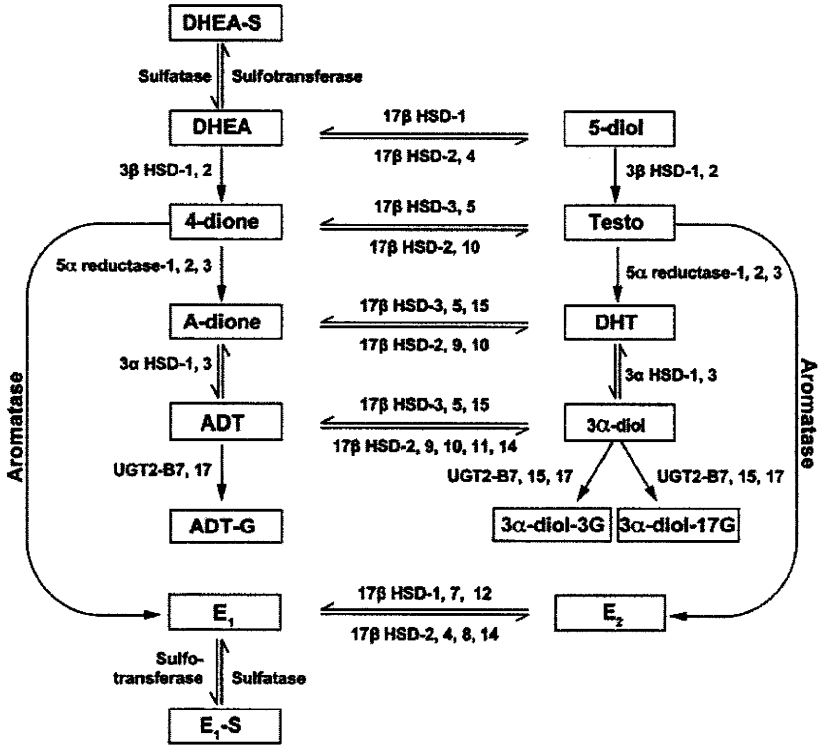

FIG. 1. Measured steroids and their metabolic relationships. 5-diol, Androst-5-ene-3β,17β-diol; A-dione, androstanedione; E2, estradiol; E1, estrone; E1-S, estrone sulfate; HSD, hydroxysteroid dehydrogenase. [Adapted from Refs. 41–43.]

ences after collapse of groups with similar means. Spearman correlations among untransformed sex steroid variables are shown in Supplemental Table 3.

For racial comparisons, we used three major groups (Caucasian, Black, and Asian). Country means were weighted equally within racial group to account for heterogeneity of sample sizes across countries. All analyses were performed using SAS version 9.1.3 (SAS Institute Inc., Cary, NC).

Results

The mean ages of the participants in the seven geographical/racial groups were very similar, but there were differences in indices of body size (Table 1).

Geographical variation in sex steroid and SHBG concentrations

Androgen and estrogen levels varied among populations (Table 1). For example, whereas age-adjusted mean total testosterone levels were similar in men from Sweden, Tobago, and the United States, they were higher (16%; $P < 0.0001$) in men living in Hong Kong and Japan (Fig. 2A); Asian men living in the United States had levels similar to other U.S. residents. As a result of geographical heterogeneity, the proportion of men with total testosterone levels suggestive of hypogonadism (<230 ng/dl) (19) differed between populations (*e.g.* 5–6% in U.S. Caucasian and Swedish men; 3% in Hong Kong and Japanese men) (Fig. 3). There was also geographical variation in DHT concentrations, but the interpopulation pattern was distinct from that of testosterone. When compared with men in the United States and Sweden, age-adjusted levels were higher (33%; $P < 0.0001$) in Tobago and in Asia (Fig. 2B). Total estradiol levels were higher in men from Hong Kong and Tobago (22%) and among Black men in the United States (20%) (Fig. 2C) (Table 1). Free estradiol showed a similar pattern. We recently identified a level of free estradiol (0.3 pg/ml) below which fractures (including hip fracture) are more likely in older Caucasian men (20). The proportion of men with levels below that threshold varied considerably between cohorts (4 to 28%) (Fig. 3).

Major geographical variation was also present in sex steroid precursors and metabolites (Table 1). For example, the age-adjusted mean dehydroepiandrosterone (DHEA) level was 44% higher, and the mean DHEA sulfate (DHEA-S) level was 20% higher in Sweden than in U.S. Caucasians. The concentrations of androgen metabolites varied across groups, primarily because androgen glucuronide levels were lower in Asian men (see *Racial variation*).

Whereas SHBG levels were similar in most groups, they were slightly lower (12%) in Swedish men and markedly higher (47%) in Japanese (Fig. 2D). As a result, mean free testosterone levels were significantly lower in Japanese (Fig. 2E). Mean free estradiol levels were higher in Hong Kong, U.S. Black, and Tobagonian men, but were lower in Japanese (Fig. 2F).

Adjustment for BMI somewhat altered the population differences in serum concentrations of steroids and SHBG, but significant heterogeneity generally remained or was enhanced (Supplemental Table 4). For instance, BMI adjustment had little effect on the cohort differences in SHBG, total estradiol, free estradiol, or free testosterone levels. On the other hand, total testosterone levels were similar among populations after BMI adjustment (Fig. 2G).

Racial variation

In addition to the effects of geography, two consistent racial patterns were found. First, Black men (from the United

TABLE 1. Cohort characteristics (mean ± SD) and age-adjusted mean (95% CI) sex steroid concentrations

	Sweden	U.S. White	U.S. Asian	Hong Kong	Japan	Tobago	U.S. Black	P[a]
n	1874	427	156	1479	364	482	221	
Age (yr)	75 ± 3	74 ± 6	73 ± 5	72 ± 5	74 ± 6	72 ± 6	72 ± 5	
BMI (kg/m^2)	26 ± 4	27 ± 4	25 ± 3	23 ± 3	23 ± 3	27 ± 4	28 ± 4	
SHBG (nmol/liter)	42.95 (42.1–43.38)	47.47 (45.6–49.4)	47.47 (44.7–50.91)	48.42 (47.47–49.40)	72.97 (70.11–75.94)	51.42 (49.9–53.52)	50.40 (47.94–52.98)	<0.0001
Precursors								
DHEA-S (μg/ml)[b]	0.67 (0.66–0.71)	0.55 (0.52–0.59)	0.76 (0.69–0.83)	0.85 (0.83–0.88)	0.81 (0.76–0.86)	0.94 (0.88–1.00)	0.58 (0.52–0.62)	0.003
DHEA (ng/ml)[b]	1.90 (1.88–1.96)	1.21 (1.14–1.3)	1.51 (1.37–1.66)	1.72 (1.66–1.77)	1.35 (1.25–1.44)	1.39 (1.32–1.46)	1.35 (1.23–1.46)	<0.0001
4-dione (ng/ml)	0.82 (0.8–0.82)	0.68 (0.65–0.7)	0.68 (0.65–0.73)	0.72 (0.7–0.72)	0.68 (0.65–0.7)	0.63 (0.6–0.65)	0.72 (0.68–0.75)	<0.0001
5-diol (ng/ml)[b]	0.59 (0.58–0.61)	0.50 (0.48–0.53)	0.53 (0.48–0.58)	0.62 (0.59–0.64)	0.71 (0.67–0.74)	0.58 (0.55–0.61)	0.58 (0.53–0.62)	<0.0001
Androgens								
Testosterone (ng/ml)	4.54 (4.45–4.62)	4.54 (4.37–4.71)	4.41 (4.12–4.67)	5.29 (5.15–5.38)	5.20 (4.97–5.38)	4.54 (4.37–4.67)	4.71 (4.45–4.93)	<0.0001
Free T (ng/ml)	9.52 (9.37–9.66)	9.14 (8.86–9.42)	8.82 (8.35–9.29)	10.49 (10.34–10.65)	8.17 (7.86–8.47)	8.76 (8.49–9.03)	9.06 (8.67–9.46)	<0.0001
DHT (ng/ml)	0.36 (0.36–0.37)	0.36 (0.34–0.37)	0.34 (0.31–0.36)	0.45 (0.44–0.46)	0.52 (0.49–0.53)	0.46 (0.45–0.49)	0.38 (0.36–0.41)	<0.0001
Estrogens								
E1 (pg/ml)[b]	35.28 (34.57–36)	30.36 (29.27–31.58)	29.38 (27.46–31.25)	33.41 (32.72–34.11)	25.00 (23.81–26.21)	42.51 (41.22–43.82)	39.82 (37.82–41.73)	0.0002
E1-S (ng/ml)[b]	0.48 (0.45–0.49)	0.51 (0.48–0.54)	0.55 (0.51–0.62)	0.57 (0.55–0.58)	0.45 (0.42–0.48)	0.7 (0.67–0.73)	0.48 (0.43–0.52)	0.07
E2 (pg/ml)[b]	20.7 (20.29–21.12)	20.09 (19.49–20.91)	18.73 (17.64–19.89)	22.87 (22.2–23.34)	18.17 (17.46–18.92)	24.53 (23.57–25.28)	23.10 (21.98–24.53)	0.05
Free E2 (pg/ml)[b]	0.52 (0.51–0.54)	0.49 (0.48–0.51)	0.46 (0.43–0.49)	0.55 (0.54–0.57)	0.38 (0.36–0.39)	0.58 (0.55–0.60)	0.55 (0.52–0.57)	0.02
Metabolites								
ADT (ng/ml)	0.17 (0.17–0.19)	0.16 (0.15–0.16)	0.17 (0.16–0.19)	0.21 (0.21–0.22)	0.16 (0.16–0.17)	0.17 (0.17–0.19)	0.16 (0.15–0.17)	<0.0001
ADT-G (ng/ml)	28.37 (27.79–29.27)	24.28 (22.81–25.58)	19.09 (17.54–20.98)	22.10 (21.42–22.57)	17.36 (16.29–18.49)	28.37 (26.94–29.88)	27.22 (25.31–29.27)	<0.0001
3α-Diol-3G (ng/ml)	1.29 (1.27–1.34)	1.27 (1.18–1.34)	0.86 (0.75–0.95)	0.82 (0.79–0.86)	0.58 (0.54–0.65)	1.23 (1.16–1.29)	1.16 (1.08–1.27)	<0.0001
3α-Diol-17G (ng/ml)	2.46 (2.39–2.53)	2.60 (2.46–2.74)	1.86 (1.66–2.06)	1.92 (1.83–1.97)	1.25 (1.14–1.36)	2.42 (2.32–2.56)	2.56 (2.35–2.78)	<0.0001

T, Testosterone; 5-diol, androst-5-ene-3β,17β-diol; E1, estrone; E1-S, estrone sulfate; E2, estradiol.

[a] *P* for ANOVA.

[b] Model includes age × group interaction to allow age adjustment to be group-specific.