

Canadian normative data for the SF-36 health survey

Wilma M. Hopman,^{*} Tanveer Towheed,[†] Tassos Anastassiades,[†] Alan Tenenhouse,[‡] Suzette Poliquin,[‡] Claudie Berger,[§] Lawrence Joseph,[¶] Jacques P. Brown,^{**} Timothy M. Murray,^{††} Jonathan D. Adachi,^{‡‡} David A. Hanley,^{§§} Emmanuel Papadimitropoulos,^{¶¶} and the Canadian Multicentre Osteoporosis Study Research Group^{***}

Abstract

Background: The Medical Outcomes Study 36-item Short Form (SF-36) is a widely used measure of health-related quality of life. Normative data are the key to determining whether a group or an individual scores above or below the average for their country, age or sex. Published norms for the SF-36 exist for other countries but have not been previously published for Canada.

Methods: The Canadian Multicentre Osteoporosis Study is a prospective cohort study involving 9423 randomly selected Canadian men and women aged 25 years or more living in the community. The sample was drawn within a 50-km radius of 9 Canadian cities, and the information collected included the SF-36 as a measure of health-related quality of life. This provided a unique opportunity to develop age- and sex-adjusted normative data for the Canadian population.

Results: Canadian men scored substantially higher than women on all 8 domains and the 2 summary component scales of the SF-36. Canadians scored higher than their US counterparts on all SF-36 domains and both summary component scales and scored higher than their UK counterparts on 4 domains, although many of the differences are not large.

Interpretation: The differences in the SF-36 scores between age groups, sexes and countries confirm that these Canadian norms are necessary for comparative purposes. The data will be useful for assessing the health status of the general population and of patient populations, and the effect of interventions on health-related quality of life.

Over the past 20 years, there has been an increased recognition of the patient's point of view as an important component in the assessment of health care outcomes. This has resulted in the development of several instruments to measure health-related quality of life. One of the most widely used and psychometrically sound instruments is the Medical Outcomes Study 36-item Short Form (SF-36). This relatively brief and simple questionnaire contains 36 items covering 8 health concepts chosen on the basis of reliability, validity and frequency of measurement in health surveys.^{1,2} Two summary scores have also been developed for the SF-36.³

The reliability and validity of the SF-36 have been well documented by the developers of the instrument.⁴⁻⁷ A comparison of a series of generic health status measures indicated that the SF-36 is not only psychometrically sound but is also more responsive to clinical improvement than the other instruments tested.^{8,9} Moreover, health functioning changed in the hypothesized direction with increased age, socioeconomic status and disease status in a population-based longitudinal study of the SF-36, which suggests that the instrument is sensitive to changes in the health of the general population.¹⁰

Normative data are the key to determining whether a group or an individual scores below or above the average for their country, age or sex. Published norms

Research

Recherche

From ^{*}the MacKenzie Health Services Research Group and [†]the Division of Rheumatology, Queen's University, Kingston, Ont.; [‡]the Canadian Multicentre Osteoporosis Study (CaMos) National Coordinating Centre and [§]the CaMos Analysis Centre, McGill University, Montreal, Que.; [¶]the Department of Epidemiology and Biostatistics, McGill University, Montreal, Que.; the Departments of Medicine at ^{**}Laval University, Sainte-Foy, Que., ^{††}the University of Toronto, Toronto, Ont., ^{‡‡}McMaster University, Hamilton, Ont., and ^{§§}the University of Calgary, Calgary, Alta.; and ^{¶¶}Eli Lilly Canada Inc., Toronto, Ont.

****The list of CaMos investigators appears at the end of this article.*

This article has been peer reviewed.

CMAJ 2000;163(3):265-71

‡ See related article page 283

now exist for the United States,¹ the Queensland region of Australia,¹¹ the United Kingdom,^{12,13} certain regions of the United Kingdom,¹⁴ Australian women¹⁵ and US residents with a variety of medical conditions.^{1,4} Comparable norms do not yet exist for Canadians. This forces researchers and policy-makers to compare data from Canadian studies to those from other countries. The initiation of the Canadian Multicentre Osteoporosis Study (CaMos) in 1995 provided a unique opportunity to incorporate the SF-36 into a population-based survey and develop age- and sex-adjusted norms for Canadians.

Methods

CaMos is a prospective cohort study of 9423 randomly selected women and men aged 25 years or more living in the community. The sample was drawn within a 50-km radius of 9 Canadian cities (Vancouver, Calgary, Saskatoon, Hamilton, Toronto, Kingston, Quebec, Halifax and St. John's). The study was designed to provide estimates of the prevalence and incidence of osteoporosis and osteoporotic fractures among Canadian men and women and of regional variation in the rates of these conditions. Baseline data were collected by means of an interviewer-administered questionnaire and included sociodemographic information, medical, fracture, re-

Table 1: Mean age- and sex-standardized scores for the 8 domains of the Medical Outcomes Study 36-item Short Form (SF-36) and for the 2 summary scales (physical and mental component) for Canadians

Age, yr	Physical functioning	Role physical	Bodily pain	General health perceptions	Energy/vitality	Social functioning	Role emotional	Mental health	Physical component scale	Mental component scale
25-34	<i>n</i> = 399	<i>n</i> = 399	<i>n</i> = 399	<i>n</i> = 399	<i>n</i> = 398	<i>n</i> = 399	<i>n</i> = 399	<i>n</i> = 399	<i>n</i> = 398	<i>n</i> = 398
Mean score	92.4	87.1	77.0	79.0	64.9	86.3	82.9	75.9	53.0	50.1
SD	14.6	29.3	21.8	16.1	17.7	20.3	32.3	15.7	7.2	9.6
95% CI	91.0-93.9	84.3-90.0	74.9-79.1	77.4-80.6	63.2-66.6	84.3-88.3	79.8-86.1	74.3-77.4	52.2-53.7	49.2-51.1
% at floor*	0.5	7.2	0.3	0.04	0.2	1.0	8.2	0.3	0.7	0.2
% at ceiling*	55.4	80.7	29.5	10.5	0.6	57.0	75.2	1.1	0.3	0.2
35-44	<i>n</i> = 499	<i>n</i> = 499	<i>n</i> = 499	<i>n</i> = 499	<i>n</i> = 497	<i>n</i> = 499	<i>n</i> = 499	<i>n</i> = 498	<i>n</i> = 497	<i>n</i> = 497
Mean score	90.9	83.4	76.2	78.9	66.1	85.5	83.2	77.3	52.0	50.9
SD	15.1	31.6	22.1	16.9	17.4	18.4	32.5	14.7	8.0	9.0
95% CI	89.6-92.2	80.6-86.2	74.3-78.2	77.4-80.3	64.6-67.7	83.9-87.1	80.3-86.0	76.0-78.5	51.3-52.7	50.1-51.7
% at floor	0.1	7.8	0.2	0.2	0.3	0.3	9.8	0.1	0.1	0.1
% at ceiling	47.8	72.9	33.3	12.6	0.1	48.0	74.6	1.2	0.1	0.4
45-54	<i>n</i> = 1690	<i>n</i> = 1690	<i>n</i> = 1690	<i>n</i> = 1689	<i>n</i> = 1690	<i>n</i> = 1690	<i>n</i> = 1690	<i>n</i> = 1689	<i>n</i> = 1688	<i>n</i> = 1688
Mean score	88.0	84.9	76.2	77.3	65.5	86.4	85.6	76.8	51.3	51.4
SD	16.9	31.9	23.4	18.4	18.2	20.3	30.1	15.8	9.0	9.2
95% CI	87.2-88.8	83.3-86.4	75.1-77.3	76.4-78.2	64.6-66.4	85.5-87.4	84.2-87.1	76.0-77.5	50.9-51.7	51.0-51.8
% at floor	0.1	9.0	0.5	0.1	0.1	0.3	7.6	0.01	0.01	0.01
% at ceiling	36.1	77.9	34.5	11.0	0.7	56.6	77.8	3.2	0.1	0.01
55-64	<i>n</i> = 2282	<i>n</i> = 2282	<i>n</i> = 2282	<i>n</i> = 2276	<i>n</i> = 2280	<i>n</i> = 2282	<i>n</i> = 2282	<i>n</i> = 2279	<i>n</i> = 2271	<i>n</i> = 2271
Mean score	82.3	81.3	74.9	74.8	68.3	88.1	87.8	79.5	49.0	53.7
SD	19.3	33.1	23.7	19.4	17.7	18.8	28.3	14.7	9.2	8.2
95% CI	81.5-83.0	80.0-82.7	73.9-75.9	74.0-75.6	67.6-69.0	87.3-88.9	86.6-88.9	78.9-80.1	48.6-49.3	53.4-54.0
% at floor	0.2	9.2	0.2	0.1	0.4	0.2	6.3	0.03	0.02	0.0
% at ceiling	19.1	70.3	32.5	8.3	2.0	60.4	81.4	4.7	0.1	0.1
65-74	<i>n</i> = 2925	<i>n</i> = 2925	<i>n</i> = 2927	<i>n</i> = 2921	<i>n</i> = 2921	<i>n</i> = 2926	<i>n</i> = 2924	<i>n</i> = 2922	<i>n</i> = 2910	<i>n</i> = 2910
Mean score	75.7	76.2	74.0	73.5	67.7	87.0	83.4	79.3	47.2	53.7
SD	22.2	36.5	23.9	18.4	18.1	19.8	32.8	15.0	9.7	8.3
95% CI	74.9-76.5	74.9-77.5	73.1-74.8	72.8-74.1	67.0-68.3	86.2-87.7	82.2-84.6	78.8-79.8	46.8-47.6	53.4-54.0
% at floor	0.6	13.3	0.2	0.01	0.2	0.04	9.9	0.01	0.0	0.0
% at ceiling	9.1	63.8	31.9	4.6	2.7	58.8	76.3	7.3	0.0	0.01
≥ 75	<i>n</i> = 1613	<i>n</i> = 1609	<i>n</i> = 1614	<i>n</i> = 1611	<i>n</i> = 1613	<i>n</i> = 1612	<i>n</i> = 1612	<i>n</i> = 1613	<i>n</i> = 1603	<i>n</i> = 1603
Mean score	59.1	62.6	69.8	71.2	61.1	83.2	80.3	79.4	42.0	54.5
SD	27.4	41.9	25.1	17.9	19.6	22.5	34.3	15.1	10.3	8.6
95% CI	57.8-60.4	60.5-64.6	68.6-71.0	70.3-72.1	60.2-62.1	82.1-84.2	78.6-81.9	78.6-80.1	41.5-42.5	54.1-54.9
% at floor	1.3	23.7	0.3	0.2	0.4	0.7	11.3	0.01	0.03	0.1
% at ceiling	2.9	47.7	27.0	3.9	2.0	50.2	70.4	6.4	0.01	0.1
All ages	<i>n</i> = 9408	<i>n</i> = 9404	<i>n</i> = 9411	<i>n</i> = 9395	<i>n</i> = 9399	<i>n</i> = 9408	<i>n</i> = 9406	<i>n</i> = 9400	<i>n</i> = 9367	<i>n</i> = 9367
Mean score	85.8	82.1	75.6	77.0	65.8	86.2	84.0	77.5	50.5	51.7
SD	20.0	33.2	23.0	17.7	18.0	19.8	31.7	15.3	9.0	9.1
95% CI	85.4-86.2	81.5-82.8	75.1-76.0	76.6-77.3	65.4-66.1	85.8-86.6	83.3-84.6	77.2-77.8	50.3-50.7	51.5-51.9
% at floor	0.4	9.8	0.3	0.1	0.2	0.2	8.6	0.01	0.0	0.0
% at ceiling	36.7	72.7	31.8	9.7	1.0	54.9	76.1	3.0	0.01	0.0

Note: SD = standard deviation, CI = confidence interval.

*% at floor and % at ceiling refer to the proportion of respondents achieving the minimum (0) and maximum (100) possible scores respectively.

productive and family history, medication use, diet, alcohol and tobacco use, and physical activity. The health status instrument was self-administered at the end of the interview. Ethical approval for CaMos was obtained through the review boards of each participating centre as well as at the coordinating centre in Montreal.

Health status was assessed with the SF-36, which contains 36 items that, when scored, yield 8 domains. Physical functioning (10 items) assesses limitations in physical activities, such as walking and climbing stairs. The role physical (4 items) and role emotional (3 items) domains measure problems with work or other daily activities as a result of physical health or emotional problems. Bodily pain (2 items) assesses limitations due to pain, and vitality (4 items) measures energy and tiredness. The social functioning domain (2 items) examines the effect of physical and emotional health on normal social activities, and mental health (5 items) assesses happiness, nervousness and depression. The general health perceptions domain (5 items) evaluates personal health and the expectation of changes in health.¹ All domains are scored on a scale from 0 to 100, with 100 representing the best possible health state. One additional, unscored item compares the respondent's assessment of her or his current health with that 1 year earlier. Summary scores for a physical component (physical functioning, role physical, bodily pain and general health perceptions) and a mental component (vitality, social functioning, mental health and role emotional) can also be derived.³

The US English-language version of the SF-36 was used because the Canadian English-language version had not been finalized at the time the study began. However, the only difference between the 2 is the use of the word "kilometre" rather than "mile" in one item; the developers of the Canadian English-language version agree that the concept — being able to walk some distance — is the same.¹⁶ For Quebec, the Canadian French-language version was obtained from the International Quality of Life Assessment Project Group.¹⁷ The data were scored by means of the Medical Outcomes Trust scoring method.¹³

The sample was identified through the use of all postal codes within 50 km of the study centres. This list was provided to InfoDirect (Bell Canada), who in turn provided a random sample of listed residential telephone numbers in these areas. This method was selected because it was the only one available at all centres. Sample size calculations were completed for each of 12 age and sex stratifications at each of the 9 centres. Because the underlying purpose of the CaMos is to study osteoporosis, fracture and bone density, sample size calculations were based on these features. The prevalence of osteoporosis and fracture increases with increasing age and is believed to be greater among women than among men, so the largest strata are for older women. The data were therefore age- and sex-standardized using simple direct standardization to the Canadian population by weighting the total means based on the underlying population characteristics using Statistics Canada data.^{18,19}

An introductory letter and information brochure were sent to all sampled households. Trained CaMos interviewers telephoned each selected household about 2 weeks after the introductory material was mailed. Telephone screening identified all eligible members of a household, and a random number table was used if more than one person was eligible. Eligibility was determined on the basis of predefined age, sex, region and calendar period (quarterly) to ensure that each centre obtained the necessary number of participants in each stratum and to eliminate seasonal bias. Up to 12 contact attempts were made. If the first few attempts were unsuccessful, the interviewer telephoned the household again after 2 weeks, at various times of the day, to allow for absences such as vacation.

Not all of the people who were invited to participate agreed to participate fully in all aspects of the study. In most of these cases we collected such information as age, sex, smoking status and number of household members. To evaluate selection bias, we compared these data for subjects for whom we had SF-36 values with the data for those who did not participate fully. We created regression models that predicted SF-36 values from these poten-

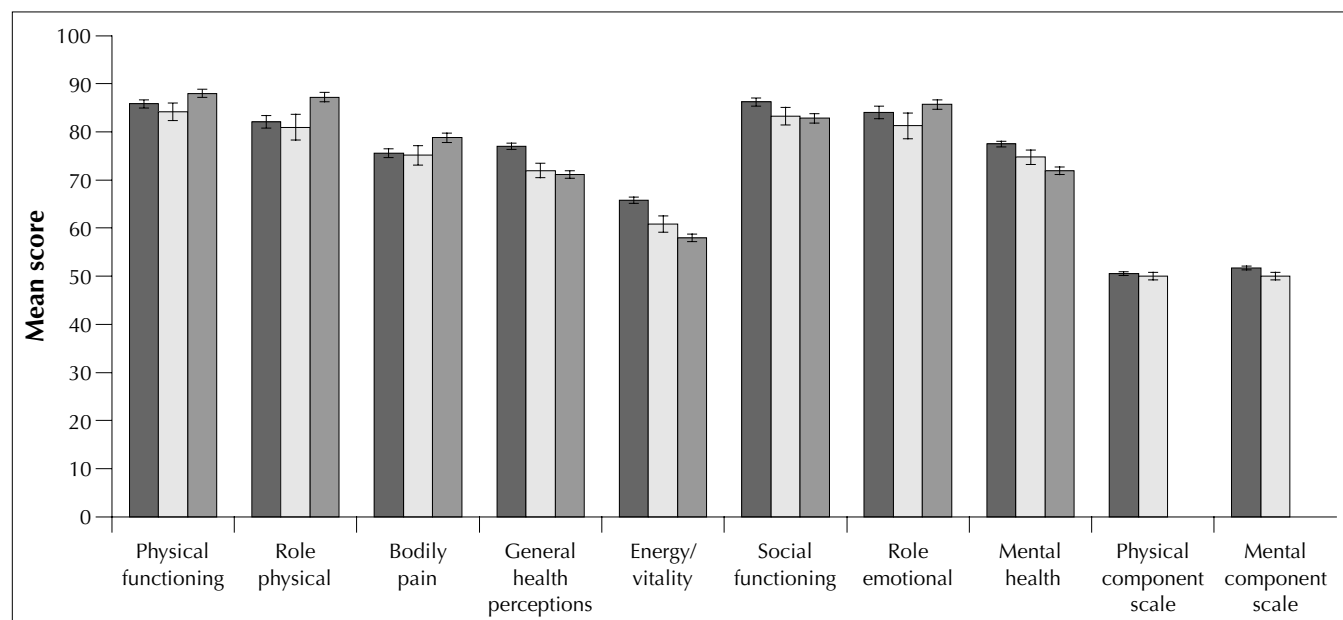


Fig. 1: Mean age- and sex-standardized scores for the 8 domains of the Medical Outcomes Study 36-item Short Form (SF-36) and for the 2 summary scales (physical component and mental component) for Canada (dark grey bars), the United States¹³ (light grey bars) and the United Kingdom¹³ (medium grey bars). Error bars indicate 95% confidence intervals.

tial predictors and applied the models to predict, through multiple imputation,²⁰ what the SF-36 values would have been for subjects who did not participate fully. We were thus able to predict whether the means we observed were likely to be different from those we would have observed had we been able to collect data from everyone. We found no differences that could have substantially changed the results reported here.

Results

Data were collected between February 1996 and September 1997. Of the 80 163 households sampled, 59.0% were ineligible, primarily because the age, sex or calendar

period stratum was already filled. In addition, 7.8% were invalid or wrong numbers, and 5.2% were unreachable after 12 attempts. Of the remaining households, 28.4% declined to participate, 29.6% completed a short questionnaire only, and 9423 (42.0%) went on to participate fully in the study and complete the SF-36.

The mean age of the sample was 62.1 (standard deviation [SD] 13.4) years. Just over 30% of the sample were men (mean age 59.9 [SD 14.5] years, range 25–97 years), and 69.4% were women (mean age 63.1 [SD 12.8] years, range 25–101 years). The age distribution was similar across the centres, as was the sex distribution (proportion of women ranged from 70.6% in Quebec to 67.2% in Toronto).

Table 2: Mean age- and sex-standardized SF-36 scores for Canadian men

Age, yr	Physical functioning	Role physical	Bodily pain	General health perceptions	Energy/vitality	Social functioning	Role emotional	Mental health	Physical component scale	Mental component scale
25–34	<i>n</i> = 200	<i>n</i> = 200	<i>n</i> = 200	<i>n</i> = 200	<i>n</i> = 200	<i>n</i> = 200	<i>n</i> = 200	<i>n</i> = 200	<i>n</i> = 200	<i>n</i> = 200
Mean score	94.0	90.6	79.1	80.1	68.7	88.9	88.3	77.7	53.5	51.7
SD	14.2	26.5	21.8	16.6	17.3	18.9	28.0	14.9	7.3	8.4
95% CI	92.1–96.0	87.0–94.3	76.1–82.1	77.8–82.4	66.3–71.1	86.3–91.5	84.4–92.2	75.6–79.7	52.5–54.6	50.5–52.8
% at floor*	0.3	6.8	3.2	0.2	1.3	1.3	5.9	3.2	1.3	0.5
% at ceiling*	62.8	86.4	32.3	14.0	0.6	62.2	82.9	1.4	0.5	0.5
35–44	<i>n</i> = 212	<i>n</i> = 212	<i>n</i> = 212	<i>n</i> = 212	<i>n</i> = 211	<i>n</i> = 212	<i>n</i> = 212	<i>n</i> = 211	<i>n</i> = 211	<i>n</i> = 211
Mean score	91.7	85.8	77.3	79.6	69.4	87.5	84.3	78.0	52.6	51.7
SD	13.9	28.8	22.2	16.8	16.8	17.1	32.2	15.1	7.1	8.8
95% CI	89.8–93.6	81.9–89.6	74.3–80.3	77.4–81.9	67.1–71.7	85.2–89.8	80.0–88.6	76.0–80.0	51.6–53.5	50.5–52.9
% at floor	0.3	5.3	0.3	2.6	0.6	0.6	9.9	0.6	0.1	0.6
% at ceiling	52.6	74.7	36.6	12.9	4.2	50.8	77.0	1.6	0.05	0.05
45–54	<i>n</i> = 583	<i>n</i> = 583	<i>n</i> = 583	<i>n</i> = 583	<i>n</i> = 583	<i>n</i> = 583	<i>n</i> = 583	<i>n</i> = 582	<i>n</i> = 582	<i>n</i> = 582
Mean score	89.3	87.7	79.6	77.4	67.7	88.5	87.1	77.9	52.1	52.0
SD	17.4	28.9	21.8	17.7	16.8	19.5	28.8	15.1	8.7	8.9
95% CI	87.9–90.8	85.4–90.1	77.8–81.4	75.9–78.8	66.3–69.0	87.0–90.1	84.8–89.4	76.7–79.2	51.4–52.8	51.3–52.7
% at floor	0.2	6.8	0.1	0.2	0.1	0.2	6.8	0.02	0.02	0.02
% at ceiling	40.6	80.9	39.8	11.3	0.7	62.4	79.8	3.9	0.2	0.1
55–64	<i>n</i> = 645	<i>n</i> = 645	<i>n</i> = 645	<i>n</i> = 644	<i>n</i> = 645	<i>n</i> = 645	<i>n</i> = 645	<i>n</i> = 644	<i>n</i> = 643	<i>n</i> = 643
Mean score	84.7	85.4	77.8	74.4	70.8	89.8	91.9	81.7	49.7	54.8
SD	17.7	29.6	22.5	18.9	15.9	17.2	22.9	13.4	8.5	7.2
95% CI	83.3–86.1	83.1–87.7	76.0–79.5	73.0–75.9	69.5–72.0	88.5–91.2	90.1–93.7	80.7–82.8	49.0–50.4	54.3–55.4
% at floor	0.1	6.2	0.2	0.1	0.2	0.2	3.5	0.1	0.1	0.1
% at ceiling	21.3	75.8	35.6	6.8	2.2	64.0	86.6	5.9	0.1	0.2
65–74	<i>n</i> = 799	<i>n</i> = 799	<i>n</i> = 800	<i>n</i> = 799	<i>n</i> = 799	<i>n</i> = 800	<i>n</i> = 798	<i>n</i> = 798	<i>n</i> = 796	<i>n</i> = 796
Mean score	78.6	78.7	77.2	73.7	70.2	87.7	85.4	82.1	48.1	54.6
SD	20.5	35.1	22.9	18.6	17.2	19.0	30.8	13.9	9.1	7.7
95% CI	77.2–80.0	76.3–81.2	75.6–78.8	72.4–75.0	69.0–71.4	86.4–89.0	83.3–87.6	81.2–83.1	47.4–48.7	54.1–55.2
% at floor	0.3	11.8	0.1	0.3	0.1	0.5	8.1	0.1	0.4	0.2
% at ceiling	10.5	67.0	37.4	4.8	3.8	60.0	78.4	9.9	0.2	0.1
≥ 75	<i>n</i> = 442	<i>n</i> = 440	<i>n</i> = 442	<i>n</i> = 442	<i>n</i> = 441	<i>n</i> = 442	<i>n</i> = 442	<i>n</i> = 442	<i>n</i> = 440	<i>n</i> = 440
Mean score	65.1	67.7	73.1	70.9	64.9	84.8	82.4	81.0	43.7	54.9
SD	27.2	40.2	24.7	18.1	19.2	21.7	32.4	14.1	10.3	8.0
95% CI	62.5–67.6	63.9–71.5	70.8–75.4	69.2–72.5	63.1–66.6	82.8–86.8	79.3–85.4	79.7–82.3	42.8–44.7	54.2–55.7
% at floor	1.8	18.9	0.3	0.3	0.1	0.6	9.7	0.2	0.02	0.3
% at ceiling	4.8	53.0	31.6	6.3	3.9	51.8	72.2	8.0	0.4	0.2
All ages	<i>n</i> = 2881	<i>n</i> = 2879	<i>n</i> = 2882	<i>n</i> = 2880	<i>n</i> = 2879	<i>n</i> = 2882	<i>n</i> = 2880	<i>n</i> = 2877	<i>n</i> = 2872	<i>n</i> = 2872
Mean score	88.2	85.7	78.0	77.6	68.9	88.3	87.0	79.0	51.4	52.6
SD	18.4	30.2	22.3	17.7	17.1	18.6	29.3	14.7	8.5	8.5
95% CI	87.5–88.9	84.6–86.8	77.2–78.8	76.9–78.2	68.3–69.5	87.6–88.9	85.9–88.0	78.4–79.5	51.1–51.7	52.3–52.9
% at floor	0.2	7.5	0.1	0.02	0.05	0.3	7.2	0.0	0.0	0.0
% at ceiling	42.3	77.2	35.6	10.9	1.2	58.8	80.3	3.7	0.04	0.02

The age- and sex-standardized scores for the 8 domains and the 2 summary scales (physical component and mental component) of the SF-36 varied by age (Table 1). Although several domains exhibited somewhat of a ceiling effect (proportion of subjects receiving the maximum possible score) (76.1% in the role emotional domain and 72.7% in the role physical domain), there did not appear to be a strong floor effect (proportion of subjects receiving the minimum possible score) (8.6% in the role emotional domain and 9.8% in the role physical domain).

Fig. 1 shows the Canadian, US and UK normative data for the 8 domains and the 2 summary scales (summary scores are not available for the UK data). Australian norma-

tive data are not included as they are provided by age and sex stratification only.¹¹ The Canadian norms are higher than the US norms in every domain and are higher than the UK norms in 4 domains. However, the magnitude of the differences is small, even though the confidence intervals do not overlap for several domains. For example, when comparing the Canadian and US norms, only the vitality domain (difference of 4.9) and general health perceptions domain (difference of 5.1) are close to the difference of 5 points considered to be clinically and socially meaningful.¹

The age- and sex-standardized scores for Canadian men and women varied by age and by sex (Tables 2 and 3). As in the entire sample, several domains exhibited a ceiling effect

Table 3: Mean age- and sex-standardized SF-36 scores for Canadian women

Age, yr	Physical functioning	Role physical	Bodily pain	General health perceptions	Energy/vitality	Social functioning	Role emotional	Mental health	Physical component scale	Mental component scale
25-34	<i>n</i> = 199	<i>n</i> = 199	<i>n</i> = 199	<i>n</i> = 199	<i>n</i> = 198	<i>n</i> = 199	<i>n</i> = 199	<i>n</i> = 199	<i>n</i> = 198	<i>n</i> = 198
Mean score	90.9	83.7	75.0	77.9	61.2	83.7	77.6	74.1	52.4	48.6
SD	14.8	31.4	21.6	15.5	17.3	21.3	35.3	16.3	7.0	10.4
95% CI	88.8-92.9	79.3-88.0	72.0-77.9	75.8-80.1	58.7-63.6	80.7-86.6	72.7-82.5	71.8-76.4	51.4-53.4	47.1-50.0
% at floor*	0.9	7.6	0.6	0.1	0.3	0.7	10.5	0.6	0.1	0.3
% at ceiling*	48.0	75.1	26.8	7.0	0.7	51.8	67.5	0.9	0.4	0.1
35-44	<i>n</i> = 287	<i>n</i> = 287	<i>n</i> = 287	<i>n</i> = 287	<i>n</i> = 286	<i>n</i> = 287	<i>n</i> = 287	<i>n</i> = 287	<i>n</i> = 286	<i>n</i> = 286
Mean score	90.1	81.0	75.1	78.1	62.9	83.5	82.1	76.6	51.5	50.2
SD	16.1	34.0	22.0	16.9	17.5	19.5	32.7	14.3	8.7	9.2
95% CI	88.2-92.0	77.1-84.9	72.6-77.7	76.1-80.0	60.9-64.9	81.3-85.8	78.3-85.8	74.9-78.2	50.5-52.5	49.1-51.2
% at floor	0.2	10.3	0.2	0.3	0.6	0.3	9.7	0.3	0.2	0.3
% at ceiling	43.1	71.0	30.0	12.4	0.3	45.2	72.1	0.9	0.2	0.8
45-54	<i>n</i> = 1107	<i>n</i> = 1107	<i>n</i> = 1107	<i>n</i> = 1106	<i>n</i> = 1107	<i>n</i> = 1107	<i>n</i> = 1107	<i>n</i> = 1107	<i>n</i> = 1106	<i>n</i> = 1106
Mean score	86.6	82.0	72.9	77.2	63.3	84.3	84.2	75.6	50.5	50.8
SD	16.3	34.5	24.5	19.1	19.2	20.8	31.4	16.4	9.2	9.5
95% CI	85.7-87.6	79.9-84.0	71.4-74.3	76.1-78.3	62.2-64.5	83.1-85.6	82.3-86.0	74.6-76.6	49.9-51.0	50.2-51.3
% at floor	0.03	11.2	0.9	0.3	0.1	0.4	8.4	0.04	0.3	0.2
% at ceiling	31.6	74.8	29.1	10.8	0.6	50.7	75.8	2.5	0.2	0.01
55-64	<i>n</i> = 1637	<i>n</i> = 1637	<i>n</i> = 1637	<i>n</i> = 1632	<i>n</i> = 1635	<i>n</i> = 1637	<i>n</i> = 1637	<i>n</i> = 1635	<i>n</i> = 1628	<i>n</i> = 1628
Mean score	79.9	77.4	72.1	75.3	65.9	86.4	83.8	77.4	48.3	52.6
SD	20.4	35.7	24.5	19.8	19.0	20.0	32.2	15.5	9.7	9.0
95% CI	78.9-80.9	75.6-79.1	70.9-73.3	74.3-76.2	65.0-66.8	85.4-87.4	82.3-85.4	76.6-78.1	47.8-48.7	52.2-53.1
% at floor	0.3	12.1	0.3	0.2	0.6	0.2	9.0	0.1	0.04	0.01
% at ceiling	17.0	65.0	29.5	9.9	1.7	57.0	76.4	3.6	0.2	0.04
65-74	<i>n</i> = 2126	<i>n</i> = 2126	<i>n</i> = 2127	<i>n</i> = 2122	<i>n</i> = 2122	<i>n</i> = 2126	<i>n</i> = 2126	<i>n</i> = 2124	<i>n</i> = 2114	<i>n</i> = 2114
Mean score	73.3	74.1	71.4	73.3	65.6	86.4	81.8	77.0	46.5	53.0
SD	23.3	37.6	24.4	18.3	18.5	20.4	34.3	15.6	10.2	8.8
95% CI	72.3-74.3	72.5-75.7	70.3-72.4	72.5-74.0	64.8-66.4	85.5-87.2	80.3-83.2	76.3-77.7	46.1-46.9	52.7-53.4
% at floor	0.8	14.5	0.4	0.01	0.3	0.1	11.3	0.01	0.01	0.01
% at ceiling	8.0	61.2	27.5	4.5	1.9	57.8	74.6	5.2	0.01	0.02
≥ 75	<i>n</i> = 1171	<i>n</i> = 1169	<i>n</i> = 1172	<i>n</i> = 1169	<i>n</i> = 1172	<i>n</i> = 1170	<i>n</i> = 1170	<i>n</i> = 1171	<i>n</i> = 1163	<i>n</i> = 1163
Mean score	55.5	59.5	67.8	71.4	58.9	82.2	79.0	78.4	40.9	54.3
SD	26.9	42.6	25.1	17.8	19.6	23.0	35.4	15.6	10.2	8.9
95% CI	54.0-57.0	57.1-62.0	66.3-69.2	70.4-72.4	57.7-60.0	80.8-83.5	77.0-81.0	77.5-79.3	40.3-41.5	53.8-54.8
% at floor	1.0	26.6	0.3	0.3	0.6	0.8	12.3	0.01	0.05	0.2
% at ceiling	1.8	44.5	24.2	2.4	0.9	49.3	69.3	5.4	0.01	0.10
All ages	<i>n</i> = 6527	<i>n</i> = 6525	<i>n</i> = 6529	<i>n</i> = 6515	<i>n</i> = 6520	<i>n</i> = 6526	<i>n</i> = 6526	<i>n</i> = 6523	<i>n</i> = 6495	<i>n</i> = 6495
Mean score	83.5	78.7	73.3	76.4	62.9	84.3	81.2	76.1	49.7	50.9
SD	21.2	35.5	23.3	17.7	18.4	20.7	33.7	15.7	9.4	9.6
95% CI	83.0-84.0	77.9-79.6	72.7-73.8	76.0-76.8	62.4-63.3	83.8-84.8	80.3-82.0	75.7-76.4	49.4-49.9	50.6-51.1
% at floor	0.5	11.9	0.5	0.1	0.4	0.2	10.0	0.01	0.05	0.0
% at ceiling	31.5	68.5	28.1	8.6	0.9	51.2	72.1	2.4	0.02	0.0

(80.3% for men and 72.1% for women in the role emotional domain), but there did not appear to be a strong floor effect (11.9% for women and 7.5% for men in the role physical domain).

The mean scores for Canadian women and men are shown in Fig. 2. Men had higher scores than women for all domains and the 2 summary scales. Although the confidence intervals did not overlap for any of the domains or summary scales, the magnitude of the difference needs to be taken into account. Only 3 of the domains (role physical, role emotional and vitality) had between-sex differences greater than 5 points, and 2 (physical functioning and bodily pain) had differences of just under 5 points.

Interpretation

The Canadian scores for the 8 domains and 2 summary scales of the SF-36 are similar to those from the United States and the United Kingdom, but there is a pattern of higher scores in the Canadian sample for all domains when compared with the US data and for 4 domains when compared with the UK data. This finding is consistent with those of other researchers¹ and underscores the importance of Canadian norms for comparative purposes. The variability of the scores by age underscores the need to use the appropriate age-specific normative data whenever possible.

The differences between countries could be due to methodologic differences rather than representing true differences. For example, the US normative data are based not on a random sample¹ but, rather, on the responses of 2474 participants in the National Survey of Functional Status, who were selected to receive a mailed version on the basis

of previous participation in a General Social Survey. The UK norms were based on the responses from 8889 people to a postal survey mailed to randomly selected households.¹³ These differences in methodology will introduce variation in the normative data for international comparisons. Thus, a clear description of methods is a vital part of the interpretation of normative data.

There are also sex differences within the Canadian sample, with men scoring higher than women on all domains and summary scales. These results are consistent with the data from the United States, where men scored higher than women on all domains,¹ and with those from the United Kingdom, where men scored higher than women on all but 1 domain.¹³ Although the differences are not large, there is evidence that some may be clinically and socially relevant, as a 5-point difference between groups or a 5-point change over time is considered clinically relevant.¹

For normative data to be valid, they must be based on a well-defined and representative sample of the population of interest.¹ The Canadian data are based on a sample of 9423 participants from 9 centres across Canada, which included an area within a 50-km radius of the cities in order to include the rural population. The complex sampling framework further increases the likelihood that the sample is representative.

The CaMos subjects were invited to participate, and there is evidence that there may be systematic differences between those who are and those who are not willing to participate in a study.²¹ However, we found no evidence that selection bias could have changed our reported mean values substantially, as determined through multiple imputation methods.²⁰ Moreover, because both the US and the UK data

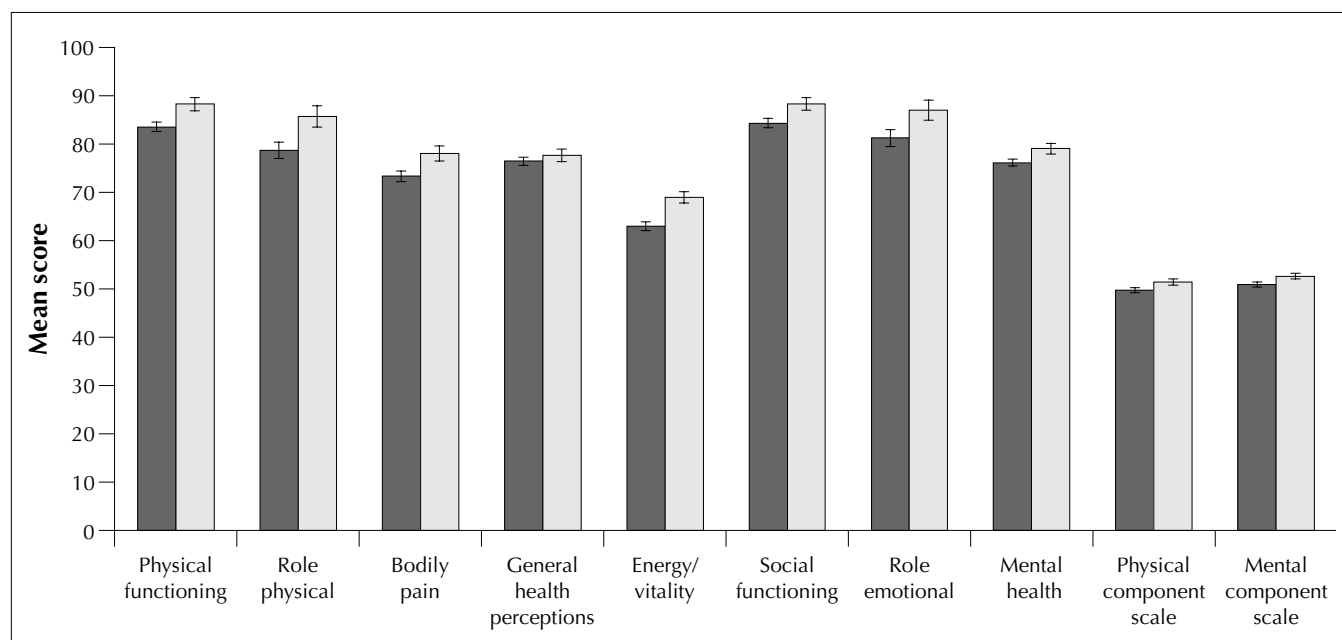


Fig. 2: Mean age- and sex-standardized SF-36 and summary scale scores for Canadian women (dark grey bars) and men (light grey bars). Error bars indicate 95% confidence intervals.

are also based on voluntary participation, this limitation applies to all the studies reported so far. We therefore conclude that the normative data that we present are valid and are based on a representative sample of residents of Canada.

We thank all the participants in the Canadian Multicentre Osteoporosis Study (CaMos). We also acknowledge the early contributions of Dr. Thomas MacKenzie, who died in October 1997.

CaMos was funded by the Senior's Independence Research Program through the National Health Research and Development Program (project no. 6605-4003-OS), the Medical Research Council of Canada-Pharmaceutical Manufacturers Association of Canada (MRC-PMAC) Health Program, Merck Frosst Canada Inc., Eli Lilly Canada Inc., Procter & Gamble Pharmaceuticals Canada, Inc. and the Dairy Farmers of Canada.

Competing interests: None declared.

References

- Ware JE Jr. *SF-36 Health Survey manual and interpretation guide*. Boston: The Health Institute, New England Medical Centre; 1993.
- Ware JE Jr. The SF-36 Health Survey. In: Spilker B, editor. *Quality of life and pharmaco-economics in clinical trials*. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1996. p. 337-45.
- Ware JE Jr, Kosinski M, Keller SD. *SF-36 physical and mental health summary scales: a user's manual*. Boston: The Health Institute, New England Medical Centre; 1994.
- Ware JE Jr, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
- McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993;31:247-63.
- McHorney CA, Ware JE Jr, Lu JFR, Sherbourne CD. The MOS 36-Item Short Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;32:40-66.
- Ware JE Jr, Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. *Med Care* 1995;33(4 Suppl):AS264-79.
- Beaton DE, Hogg-Johnson S, Bombardier C. Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *J Clin Epidemiol* 1997;50:79-93.
- Essink-Bot ML, Krabbe PFM, Bonsel GJ, Aaronson NK. An empirical comparison of four generic health status measures. *Med Care* 1997;35:522-37.
- Hemingway H, Stafford M, Stansfield S, Shipley M, Marmot M. Is the SF-36 a valid measure of change in population health? Results from the Whitehall II study. *BMJ* 1997;315:1273-9.
- Watson EK, Firman DW, Baade PD, Ring I. Telephone administration of the SF-36 Health Survey: validation studies and population norms for adults in Queensland. *Aust N Z J Public Health* 1996;20:359-63.
- Jenkinson C, Coulter C, Wright L. Short Form 36 (SF-36) Health Survey questionnaire: normative data for adults of working age. *BMJ* 1993;306:1437-40.
- Jenkinson C, Stewart-Brown S, Petersen S, Paice C. Assessment of the SF-36 version 2 in the United Kingdom. *J Epidemiol Community Health* 1999;53:46-50.
- Lyons RA, Fielder H, Littlepage NC. Measuring health status with the SF-36: the need for regional norms. *J Public Health Med* 1995;17:46-50.
- Mishra G, Schofield MJ. Norms for the physical and mental health component summary scores of the SF-36 for young, middle-aged and older Australian women. *Qual Life Res* 1998;7:215-20.
- Wood-Dauphinee SW, Gauthier L, Gandek B, Magnan L, Pierre U. Readying a US measure of health status, the SF-36, for use in Canada. *Clin Invest Med* 1997;20:224-38.
- Ware JE Jr, Keller SD, Gandek B, Brazier JE, Sullivan M, and the International Quality of Life Assessment Project Group. Evaluating translations of health status questionnaires. Methods from the IQOLA project. *Int J Technol Assess Health Care* 1995;11:525-51.
- Statistics Canada provincial census data, 1991. Age, sex and marital status: the nation*. Ottawa: Statistics Canada; 1993. Cat no 93-310.
- Statistics Canada provincial census data, 1991. Profile of census metropolitan areas and census agglomeration, part A*. Ottawa: Statistics Canada; 1993. Cat no 93-337.
- Rubin D. *Multiple imputation for non-response in surveys*. New York: John Wiley & Sons; 1987.
- Meltzoff J. *Critical thinking about research: psychology and related fields*. Washington: American Psychological Association; 1998.

Reprint requests to: Wilma M. Hopman, Director, MacKenzie Health Services Research Group, Department of Community Health and Epidemiology, 3rd floor, Abramsky Hall, Queen's University, Kingston ON K7L 3N6; hopmanw@post.queensu.ca

Investigators in the Canadian Multicentre Osteoporosis Study Research Group: McGill University, Montreal General Hospital, Montreal (coordinating centre): Alan Tenenhouse (principal investigator), Suzette Poliquin (national coordinator), Lawrence Joseph (study statistician), Lucie Blondeau (statistician), Claudie Berger (statistician), Suzanne Lefebvre (administrative assistant); University of British Columbia, Vancouver: Jerilynn C. Prior (centre director), Brian Lentle (Radiology Study consultant), Yvette Vigna (centre coordinator); University of Alberta, Edmonton: Stuart Jackson (medical physicist), Loralee Robertson (research assistant); University of Calgary: David A. Hanley (centre director), Jane Allan (centre coordinator); University of Saskatchewan, Saskatoon: Wojciech P. Olszynski (centre director), and Pat Krutzen and Jola Kedra (centre coordinators); McMaster University, Hamilton, Ont.: Jonathan D. Adachi (centre director), Laura Pickard (centre coordinator); University of Toronto: Nancy Kreiger (study epidemiologist), Timothy M. Murray (centre director), Barbara Gardner-Bray (centre coordinator); Queen's University, Kingston, Ont.: Tassos Anastassiades (centre director), and Pamela Hartman and Barbara Matthews (centre coordinators); Laval University, Sainte-Foy, Que.: Jacques P. Brown (centre director), and Nathalie Migneault-Roy and Evelyne Lejeune (centre coordinators); Dalhousie University, Halifax: Roger S. Rittmaster (centre director), Susan Kirkland (epidemiologist), Barbara Stanfield (centre coordinator); and Memorial University, St. John's: Carol Joyce (centre director), Minnie Parsons (centre coordinator)

Change of address

We require 6 to 8 weeks' notice to ensure uninterrupted service. Please send your current mailing label, new address and the effective date of change to:

CMA Member Service Centre

1867 Alta Vista Dr.
Ottawa ON K1G 3Y6

tel 888 855-2555 or
613 731-8610 x2307
fax 613 236-8864
cmamsc@cma.ca

ASSOCIATION
MÉDICALE
CANADIENNE  CANADIAN
MEDICAL
ASSOCIATION