Qual Life Res
DOI 10.1007/s11136-014-0733-5

Prospectively measured 10-year changes in health-related quality of life and comparison with cross-sectional estimates in a population-based cohort of adult women and men

Wilma M. Hopman · Claudie Berger · Lawrence Joseph · Wei Zhou · Jerilynn C. Prior · Tanveer Towheed · Tassos Anastassiades · Jonathan D. Adachi · David A. Hanley · Emmanuel A. Papadimitropoulos · Susan Kirkland · Stephanie M. Kaiser · Robert G. Josse · David Goltzman · and the CaMos Research Group

Accepted: 6 June 2014

Abstract
Purpose To prospectively assess changes in health-related quality of life (HRQOL) over 10 years, by age and sex, and to compare measured within-person change to estimates of change based on cross-sectional data.
Methods Participants in the Canadian Multicentre Osteoporosis Study completed the 36-item short form (SF-36) in 1995/1997 and 2005/2007. Mean within-person changes for domain and summary components were calculated for men and women separately, stratified by 10-year age groups. Projected changes based on published age- and sex-stratified cross-sectional data were also calculated. Mean differences between the two methods were then estimated, along with the 95 % credible intervals of the differences.
Results Data were available for 5,569/9,423 (59.1 %) of the original cohort. Prospectively collected 10-year changes suggested that the four physically oriented domains declined in all but the youngest group of men and women, with declines in the elderly men exceeding 25 points. The four mentally oriented domains tended to improve over time, only showing substantial declines in vitality and role

Please see the "Appendix" section for CaMos Research Group members.

W. M. Hopman (✉)
Clinical Research Centre, Kingston General Hospital, Kingston, ON, Canada
e-mail: Hopmanw@kgh.kari.net

W. M. Hopman · T. Towheed
Department of Public Health Sciences, Queen's University, Kingston, ON, Canada

C. Berger · W. Zhou
CaMos Methods Centre, McGill University, Montreal, QC, Canada

L. Joseph
Department of Epidemiology and Biostatistics, McGill University, Montreal, QC, Canada

J. C. Prior
Department of Medicine/Endocrinology, University of British Columbia, Vancouver, BC, Canada

T. Towheed · T. Anastassiades
Division of Rheumatology, Department of Medicine, Queen's University, Kingston, ON, Canada

J. D. Adachi
Department of Medicine, McMaster University, Hamilton, ON, Canada

D. A. Hanley
University of Calgary, Calgary, AB, Canada

E. A. Papadimitropoulos
Eli Lilly Canada, Inc., Toronto, ON, Canada

E. A. Papadimitropoulos
Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada

S. Kirkland
Departments of Community Health and Epidemiology and Medicine, Dalhousie University, Halifax, NS, Canada

S. M. Kaiser
Department of Endocrinology and Metabolism, Dalhousie University, Halifax, NS, Canada

R. G. Josse
Department of Medicine, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada

Published online: 13 June 2014

emotional in older women, and all four domains in older men. Cross-sectional estimates identified a similar pattern of change but with a smaller magnitude, particularly in men. Correspondence between the two methods was generally high.

Conclusions Changes in HRQOL may be minimal over much of the life span, but physically oriented HRQOL can decline substantially after middle age. Although clinically relevant declines were more evident in prospectively collected data, differences in 10-year age increments of cross-sectional data may be a reasonable proxy for longitudinal changes, at least in those under 65 years of age. Results provide additional insight into the natural progression of HRQOL in the general population.

Keywords SF-36 · Normative · Prospective · Quality of life · HRQOL

Rationale/background

The Medical Outcomes Trust 36-item health survey (SF-36) is widely used to assess the health-related quality of life (HRQOL) of general and specific adult populations, estimate the relative burden of different diseases and examine the impact of a wide range of treatment interventions [1–6]. Numerous studies have demonstrated the reliability (test–retest and internal consistency) and validity (content, concurrent, criterion, construct and predictive) of the measure [3, 7–9]. Age- and sex-standardized normative data have been developed for a number of countries including both the USA and Canada, based on 10-year age increments [1, 2, 10]. However, there are few reports of *change* in HRQOL over time in randomly selected population samples. Knowledge of the lifecycle evolution of HRQOL is important, because any measure of change following a diagnosis of disease, or the effect of an intervention over time, could be confounded by the natural changes in HRQOL due to aging.

To date, few studies have prospectively documented changes in HRQOL in an adult population-based sample. The Whitehall II study (London, UK) assessed HRQOL in office-based civil servants, including 5,070 men and 2,197 women between the ages of 39 and 63 years, with a mean follow-up of 36 months. They found that the SF-36 domains changed in the hypothesized direction with increased age (directly for physical domains and inversely for mental domains), socioeconomic status and the presence of chronic, progressive or recurrent disease at baseline [8, 11].

Between 1972 and 1974, the Rancho Bernardo Study enrolled 6,339 individuals into a study of heart disease risk factors; in 1995 and again in 2000, all surviving members of the original cohort completed the SF-36 for a final sample of 1,570 participants who completed both surveys (complete case analysis). They found that there were significant declines in physical health of all four of their age groups (50–59, 60–69, 70–79 and 80+ years), modest improvements in mental health of the younger age groups, but declines in men >70 and women >80 years [12].

Data from the Canadian Multicentre Osteoporosis Study (CaMos), an ongoing prospective cohort study of 9,423 non-institutionalized, randomly selected men and women 25 years of age and older at baseline, were used to examine changes over a 3-year period in a subset of 1974 women and 975 men between the ages of 40 and 59 years at baseline. Results suggested that mean scores changed only slightly in this sample, but there was considerable variation between individuals [13]. An analysis of the full cohort over a 5-year period noted that there was an overall trend toward decreasing HRQOL over time, particularly in the older participants and in the physical domains, while there were small mean improvements in some of the younger groups, particularly in the mentally oriented domains [14].

These three past studies suggest that the HRQOL of the general population is relatively stable over the short term, but that changes accumulated over time could become quite large [8, 11–14]. However, no prospectively collected data exist to evaluate the changes over a period beyond 3–5 years. Moreover, there is some preliminary evidence from the Whitehall II study that the common practice of using the differences between 10-year age increments of cross-sectional data as a proxy for longitudinal changes over 10 years may be misleading [8]. Estimates of within-person worsening of health with increased age were larger than those estimated by cross-sectional data alone, and sex differences also showed a different pattern in cross-sectional and longitudinal results [8]. There is therefore a clear need for repeated measures of HRQOL in longitudinal population-based studies [7, 8, 14].

The 10-year follow-up of the CaMos participants included administration of the same version of the SF-36 utilized at baseline and at 5 years. This provided the opportunity to assess changes over 10 years for all age and gender stratifications within the cohort. In addition, the changes over time could also be compared to the estimates of change based on cross-sectional data alone, providing an evaluation of the common practice of using the differences between 10-year age increments of cross-sectional data as a proxy for within-person changes over 10 years.

D. Goltzman
Division of Endocrinology and Metabolism, Department of Medicine, McGill University, Montreal, QC, Canada

Methods

The 9423 CaMos participants were drawn from a 50-km radius of nine Canadian cities (St John's, Halifax, Quebec City, Toronto, Hamilton, Kingston, Saskatoon, Calgary and Vancouver). Details regarding the study purpose, methodology and sampling framework are available elsewhere [10, 15]. In brief, households within each region were selected by random draws of listed telephone numbers, and one randomly selected household member ≥25 years of age was asked to participate. Ethics approval was obtained through the Review Boards of each of the participating centers and the coordinating center in Montreal.

Baseline interviews took place between September 1995 and September 1997, with 10-year follow-ups taking place between September 2005 and September 2007. The interviews were conducted in person at both time points and included detailed sociodemographic, family history and lifestyle questions; the Medical Outcomes Trust 36-item short form (SF-36) [1, 2] was self-administered at each session. The SF-36 yields 8 domain scores, including physical functioning (PF), role physical (RP), role emotional (RE), bodily pain (BP), general health perceptions (GH), vitality (VT), social function (SF) and mental health (MH). These eight domains are scored from 0 to 100, with a score of 100 representing excellent HRQOL and no pain [1]. In addition, a Physical Component Summary (PCS) and a Mental Component Summary (MCS) can also be derived. The PCS and MCS are standardized to a mean of 50, with scores above and below 50 representing better than average and poorer than average scores, respectively [2].

All data were age- and sex-standardized to the Canadian population, using 1996 Census data to maximize the comparability of the data from the two time points. All analyses were done for each of the domain and summary scores, stratified by sex and within the 10-year age increments used for the original normative data (e.g., 25–34, 35–44, 45–54, 55–64, 65–74, and 75+ years) [10]. Participants who died were omitted from all analysis as change scores are not defined for this population [16, 17]. Prospective individual changes were directly estimated for each domain and component summary using within subject differences for all those who contributed SF-36 data at both the start of the study (baseline) and 10 years later (year 10). Baseline scores were subtracted from year 10 scores so that positive values represent improvements in HRQOL and vice versa. Results included the mean differences, as well as the 95 % credible intervals (CI) of the mean differences (CI are the Bayesian analogue to frequentist confidence intervals).

Projected cross-sectional changes were calculated using aggregate information based on the published normative data [10]. The mean of each age/sex domain or component summary was subtracted from that of the adjacent, older age/sex domain or component summary, representing information available to the research community to estimate longitudinal change in the absence of prospectively collected data. For example, the change in PF for women aged 25–34 at baseline involved subtracting the mean of the 25–34 age group from that of the 35–44 age group to obtain a projected change over 10 years. As such, positive values would again represent improvements and vice versa. However, there are no projected cross-sectional changes for those in the 75+ age group at baseline as that was the oldest age group, and therefore, no corresponding 10-year change data are available. Finally, in order to compare the prospectively measured change and the projected cross-sectional change, the mean differences between the two were assessed with a Bayesian approach using highly diffuse prior distributions across all parameters to estimate all parameters. WinBUGS software (version 1.4.3, MRC Biostatistics Unit, Cambridge, UK) was used for these analyses.

Missing data

The issue of missing data cannot be underestimated in any assessment of change over time, and a considerable amount of research has been expended in determining ways of defining and handling missing data. Missing data can be defined as values that are not available, but that would be meaningful if they had been observed [18]. High rates of missing data can not only affect conclusions, but are a particularly serious problem if the data are not missing at random (non-ignorable) [16], as is the case when it is the outcome itself—poorer HRQOL—that contributes to the missing data. There is also no universally accepted method for handling missing data [18, 19], and it is often ignored or handled inadequately [20].

For the longitudinal component of the 10-year changes, we elected to do a complete case analysis [18]. Multiple imputation was initially considered and would have merit for correcting at least some of the bias [21, 22]. However, it was not utilized since it too has limitations within the context of non-ignorable data. Moreover, when multiple imputation was completed for the assessment of 5-year changes [14], the differences between the existing data and the existing + imputed data were negligible. Instead, we compared the baseline data of those who did and did not have complete data at year 10 to obtain some insight into the extent and direction of the differences between the two. Our approach is therefore to be as transparent as possible, attempt to measure and describe bias in either direction and provide the reader with the information required to make an informed decision regarding the validity of the findings. Within-person changes of 5 points in domain scores or 2–3

points in component summaries were considered clinically important [1, 2].

Results

Complete data were available for 5,569 or 59.1 % of the original cohort. Table 1 outlines the participants by sex and for each age group at baseline and year 10, as well as the number who died, only completed a short questionnaire that did not include the SF-36, or were true non-responders (unable to contact, no longer interested, too sick, moved away, canceled, no time or no reason provided). Attrition was slightly higher among men (45.6 %) than women (38.9 %) and was much higher in the advanced age groups due to mortality.

Comparison of baseline scores, those with and without year 10 data

Baseline domain and summary component scores of those who did and did not complete the 10-year assessment were compared using the mean scores and corresponding 95 % CIs. One comparison was substantially higher in the non-respondents (men, PF, baseline age 25–34). For women, 15/48 age-domain cells and 5/12 age-component summary cells were substantially lower in non-respondents and respondents. For men, 18/48 age-domain cells and 3/12 age-component summary cells were substantially lower in non-respondents. This was evident across all age groups and domains except for bodily pain and was somewhat more pronounced in the older groups.

Comparison of change over 10 years, lowest quartile versus highest quartile at baseline

The changes in domain and summary component scores were assessed for those in the lowest and the highest quartile at baseline to gain insight into the pattern of change in these groups. This is relevant as it may at least partially address the question of bias arising from the finding above that those without year 10 data tended to have lower HRQOL scores. For those in the highest quartile, all domain and summary component scores declined over the 10 years; this was true for all age groups, for both men and women. However, for those in the lowest quartile, the opposite was true, quite possibly representing simple regression toward the mean. For both men and women, 46/48 age-domain cells and 11/12 age-component summary cells improved over the 10 years. The only exceptions were for men age 75+, where the lowest quartile saw declines of −12.1, −12.2 and −1.9 in SF, PF and the PCS, respectively. The same domains saw declines of −17.0, −30.3 and −12.4 for the highest quartile, indicating that the highest quartile had a more substantial decline. For women, the three exceptions were in PF age 65–74, PF age 75+ and PCS 65–74, where declines of −4.5, −1.9 and −0.7 were noted in the lowest quartile. The highest quartile saw much larger declines of −17.8, −28.6 and −7.7 in these same age-domain or age-component summary cells. Detailed results and 95 % CIs are not presented for these data since it is the pattern of change, rather than the actual values, that are of interest.

Table 1 Numbers of participants by gender, age group and respondent status at 10 years in the Canadian Multicentre Osteoporosis Study population-based cohort

Age at baseline	All participants at baseline *N* Women	Responders at year 10 *N* (%) Women	Died by year 10 *N* (%) Women	Short questionnaire *N* (%) Women	Lost to follow-up *N* (%) Women
25–34	200	135 (67.5)	0 (0.0)	11 (5.5)	54 (27.0)
35–44	286	230 (80.4)	2 (0.7)	12 (4.2)	42 (14.7)
45–54	1,111	886 (79.7)	24 (2.2)	55 (5.0)	146 (13.1)
55–64	1,639	1,182 (72.1)	86 (5.2)	118 (7.2)	253 (15.4)
65–74	2,134	1,251 (58.6)	278 (13.0)	178 (8.3)	427 (20.0)
75+	1,169	315 (26.9)	377 (32.2)	105 (9.0)	372 (31.8)
All ages	6,539	3,999 (61.1)	767 (11.7)	479 (7.3)	1,294 (19.8)
	Men	Men	Men	Men	Men
25–34	200	115 (57.5)	0 (0.0)	15 (7.5)	70 (35.0)
35–44	212	137 (64.6)	1 (0.5)	12 (5.7)	62 (29.2)
45–54	587	413 (70.4)	22 (3.7)	36 (6.1)	116 (19.8)
55–64	640	426 (66.6)	56 (8.8)	43 (6.7)	115 (18.0)
65–74	802	396 (49.4)	166 (20.7)	66 (8.2)	174 (21.7)
75+	443	83 (18.7)	199 (44.9)	38 (8.6)	123 (27.8)
All ages	2,884	1,570 (54.4)	444 (15.4)	210 (7.3)	660 (22.9)

Prospective mean change over 10 years

The estimated mean changes over the 10 years are presented in Tables 2 (women) and 3 (men), along with the corresponding 95 % CI of the change. Although the mean changes tended to be relatively small in the younger age groups, the standard deviations of the observations (not shown) tended to be quite high, suggesting that some participants saw substantial declines while others saw substantial improvements, resulting in relatively small overall change. In general, small improvements or small declines that fell short of the 5-point criteria for clinical relevance [1] (or 2–3 points for the summary components) [2] were noted in the younger groups and mentally oriented domains, with much larger declines in the older groups and in the physically oriented domains.

For women, one domain (RE, baseline age 25–34 years) saw a clinically relevant mean improvement of 11.5 points. Three of the MCS age groups also saw relevant improvement (baseline age 25–34, 35–44 and 75+ years). An additional 13/48 (27 %) age-domain cells and 5/12 (42 %) age-component summary cells saw clinically relevant declines, highlighted in Table 2, particularly in the older age groups and the physical domains. The declines tended to be much smaller in magnitude than those seen in the men, with the largest decline of 20.2 points seen in the PF domain in the women aged 75+ years at baseline.

The results for men were similar. Two age-domain cells (SF, baseline age 25–34 years and RE, baseline age 35–44 years) saw clinically relevant improvements with a mean of 5.7 and 9.2 points, respectively. Two of the MCS groups (baseline age 35–44, and 45–54 years) also saw clinically relevant improvement. An additional 17/48 (35 %) age-domain cells and 6/12 (50 %) age-component summary cells saw clinically relevant declines, as highlighted in Table 3. The pattern was similar to the women in that older age groups and physically oriented domains were particularly affected. In the oldest age group, these declines could be quite substantial, exceeding 25 points for the domains of PF, RP and RE for the men aged 75+ years at baseline.

Cross-sectional estimates of mean change over 10 years

The estimates of change obtained from subtracting the cross-sectional baseline data from the cross-sectional 10-year data are presented in Tables 4 (women) and 5 (men), along with the 95 % CI of the estimates. For these results, there are 40 age-domain and 10 age-component summary cells rather than 48 and 12, as there are no comparison data available for the 75+ group, the oldest age group collected at baseline. The pattern of change is remarkably similar to that of the prospectively collected change, although the magnitude of the changes generally tended to be smaller. For the women, only 5/40 (13 %) of age-domain cells and 2/10 (20 %) age-component summary cells saw clinically relevant declines, highlighted in Table 4. Of these 7, 6 were physically oriented (3 PF, 1 RP, 2 PCS), with an additional one in the VT domain. Four of these were in the oldest age group (65–74 years at baseline) for which comparison data were available.

The cross-sectional estimates of age-stratified change for men were also similar to that of the prospectively measured change. Six of 40 (15 %) age-domain cells and 2/10 (20 %) age-component summary cells saw clinically relevant declines, as highlighted in Table 5. They were also physically oriented (2 PF, 2 RP, 2 PCS), with the remaining 2 in the VT and RE domains. They were again more pronounced in the older age groups in that 4/8 of these were in the oldest age group of 65–74 years at baseline.

For both women and men, the magnitude of the age-stratified differences when using the cross-sectional estimates tended to be much smaller than in the prospectively collected data. This is likely because the oldest group (75+ years at baseline) could not be represented within these results.

Comparison of mean change, prospectively measured versus projected cross-sectional estimates

Figures 1 (PCS) and 2 (MCS) provide a visual comparison of the longitudinally measured change and the projected age-stratified cross-sectional change, demonstrating the remarkable correspondence between the two methods. Tables 6 (women) and 7 (men) provide the comparison of the differences between the means of the longitudinally measured change and the means of the projected cross-sectional estimates, along with the 95 % CIs of the difference. For women, the differences did not attain clinical relevance for any of the component summary comparisons (all mean differences were <2–3 points) and only one age-domain cell (RP age group 25–34 to 35–44) was above the 5-point threshold for clinical relevance for domain scores. There were no clinically relevant differences for men for any of the component summary or domain comparisons. Age-stratified 10-year cross-sectional estimates were more likely to underestimate 10-y HRQOL changes as compared to the prospectively collected data, as a much larger proportion of the age-domain and age-component summary comparisons saw clinically relevant declines using the prospectively collected data.

Discussion

Changes in HRQOL tend to be minimal over much of the life span, but physically oriented HRQOL declines substantially

Table 2 Mean prospectively measured changes in SF-36 scores from baseline to year 10, by age group in women in the Canadian Multicentre Osteoporosis Study population-based cohort

Age group	Physical functioning	Role physical	Bodily pain	General health	Vitality	Social functioning	Role emotional	Mental health	Physical component scale	Mental component scale
From 25–34 to 35–44										
Mean	0.6	0.4	1.4	0.3	1.2	2.1	**11.5**	3.0	−0.6	**2.6**
95 % CI	−1.1, 2.2	−5.0, 5.9	−2.3, 5.1	−2.1, 2.7	−1.8, 4.2	−1.6, 5.8	**5.5**, **17.6**	0.1, 5.9	−1.7, 0.5	**0.9**, **4.4**
From 35–44 to 45–54										
Mean	−4.9	−2.9	−4.8	−1.1	0.9	3.5	4.7	1.9	*−2.5*	**2.5**
95 % CI	−7.2, −2.6	−8.8, 3.1	−8.2, −1.4	−3.4, 1.3	−1.1, 2.9	0.8, 6.3	0.2, 9.2	−0.1, 3.9	*−3.9*, *−1.1*	**1.2**, **3.7**
From 45–54 to 55–64										
Mean	*−5.4*	−4.7	−2.8	−1.9	1.3	0.5	2.2	2.0	*−2.4*	1.9
95 % CI	*−6.6*, *−4.3*	−7.3, −2.0	−4.3, −1.2	−2.9, −0.8	0.2, 2.5	−1.0, 2.1	−0.3, 4.7	1.0, 3.1	*−3.0*, *−1.8*	1.3, 2.6
From 55–64 to 65–74										
Mean	*−5.7*	−1.5	−4.1	−1.2	−0.4	1.2	1.6	1.6	*−2.2*	1.5
95 % CI	*−6.9*, *−4.5*	−3.9, 0.8	−5.5, −2.6	−2.2, −0.3	−1.4, 0.6	−0.1, 2.5	−0.6, 3.8	0.7, 2.4	*−2.8*, *−1.7*	1.0, 2.1
From 65–74 to 75+										
Mean	*−14.9*	*−16.0*	*−8.5*	*−5.5*	*−7.1*	−3.7	*−6.2*	−0.1	*−6.0*	0.4
95 % CI	*−16.1*, *−13.7*	*−18.4*, *−13.6*	*−9.9*, *−7.1*	*−6.4*, *−4.5*	*−8.1*, *−6.2*	−5.0, −2.4	*−8.4*, *−3.9*	−0.9, 0.7	*−6.5*, *−5.4*	−0.1, 0.9
75+ at baseline										
Mean	*−20.2*	*−9.4*	*−6.1*	*−6.8*	*−6.6*	−0.9	2.4	0.1	*−6.7*	**2.1**
95 % CI	*−23.1*, *−17.3*	*−15.0*, *−3.8*	*−9.0*, *−3.2*	*−8.8*, *−4.9*	*−8.7*, *−4.4*	−3.6, 1.8	−2.6, 7.4	−1.5, 1.7	*−7.8*, *−5.6*	**1.2**, **3.1**
All ages										
Mean	−4.7	−3.3	−2.7	−1.3	−0.1	1.4	5.0	2.0	−2.3	2.1
95 % CI	−5.3, −4.2	−4.5, −2.0	−3.5, −2.0	−1.9, −0.8	−0.6, 0.5	0.8, 2.1	3.8, 6.1	1.5, 2.5	−2.6, −2.0	1.8, 2.4

Emphasized cells represent clinically relevant decline (italics) or improvement (bold) (>5 points for domains and >2–3 points for component summaries) [1, 2]

CI credible interval

Table 3 Mean prospectively measured changes in SF-36 scores from baseline to year 10, by age group, in men in the Canadian Multicentre Osteoporosis Study population-based cohort

Age group	Physical functioning	Role physical	Bodily pain	General health	Vitality	Social functioning	Role emotional	Mental health	Physical component scale	Mental component scale
From 25–34 to 35–44										
Mean	1.5	1.0	−2.5	*−6.4*	−1.8	**5.7**	2.6	2.1	−1.2	1.3
95 % CI	−0.4, 3.3	−5.3, 7.2	−6.3, 1.2	*−9.3*, *−3.6*	−4.5, 0.8	**1.9**, **9.5**	−4.2, 9.5	−0.4, 4.6	−2.5, 0.0	−0.4, 3.1
From 35–44 to 45–54										
Mean	−0.9	*−5.3*	−1.4	0.0	0.5	−1.3	**9.2**	4.5	*−2.0*	**2.7**
95 % CI	−2.7, 0.9	*−11.5*, *0.8*	−5.6, 2.9	−2.1, 2.2	−1.7, 2.8	−6.1, 3.5	**3.0**, **15.5**	2.4, 6.5	*−3.3*, *−0.7*	**1.1**, **4.2**
From 45–54 to 55–64										
Mean	−2.5	−2.5	−3.4	−1.8	2.8	1.2	3.1	3.7	*−2.0*	2.4
95 % CI	−4.0, −1.0	−5.7, 0.6	−5.7, −1.1	−3.2, −0.4	1.3, 4.3	−0.6, 2.9	0.5, 5.7	2.2, 5.2	*−2.8*, *−1.2*	1.5, 3.2
From 55–64 to 65–74										
Mean	*−6.1*	*−8.3*	−2.3	−2.4	−2.4	−0.0	−3.8	0.8	*−2.5*	0.5
95 % CI	*−7.9*, *−4.4*	*−12.3*, *−4.2*	−4.5, −0.1	−3.9, −1.0	−3.9, −0.9	−1.8, 1.7	−6.8, −0.7	−0.4, 2.0	*−3.3*, *−1.7*	−0.2, 1.2
From 65–74 to 75+										
Mean	*−11.8*	*−13.7*	*−5.5*	−4.7	*−7.7*	−4.1	*−8.1*	−1.7	*−4.4*	−0.9
95 % CI	*−14.2*, *−9.5*	*−17.7*, *−9.8*	*−8.0*, *−2.9*	−6.5, −2.9	*−9.2*, *−6.1*	−6.1, −2.2	*−11.6*, *−4.6*	−3.0, −0.3	*−5.3*, *−3.4*	−1.8, −0.1
75+ at baseline										
Mean	*−26.3*	*−36.2*	*−10.5*	*−11.1*	*−16.6*	*−17.5*	*−25.8*	*−5.3*	*−9.4*	*−4.3*
95 % CI	*−31.5*, *−21.2*	*−46.4*, *−25.9*	*−16.9*, *−4.1*	*−15.3*, *−6.9*	*−21.5*, *−11.7*	*−22.3*, *−12.6*	*−36.2*, *−15.3*	*−8.6*, *−2.0*	*−11.6*, *−7.2*	*−6.4*, *−2.2*
All ages										
Mean	−2.1	−4.0	−2.6	−3.3	−1.1	1.3	2.5	2.5	−2.1	1.5
95 % CI	−2.9, −1.4	−5.8, −2.2	−3.8, −1.5	−4.1, −2.6	−1.8, −0.3	0.2, 2.4	0.8, 4.3	1.8, 3.2	−2.4, −1.7	1.1, 2.0

See footer of Table 2

Table 4 Mean differences in 10-year cross-sectional SF-36 scores by age group for women in the Canadian Multicentre Osteoporosis Study population-based cohort

For 10-year period	Physical functioning	Role physical	Bodily pain	General health	Vitality	Social functioning	Role emotional	Mental health	Physical component scale	Mental component scale
From 25–34 to 35–44										
Mean	−0.9	−2.5	0.2	0.4	1.6	−0.1	3.9	2.5	−0.9	1.5
95 % CI	−2.7, 1.0	−6.0, 1.0	−2.2, 2.5	−1.4, 2.1	−0.2, 3.5	−2.2, 2.0	0.6, 7.3	1.0, 4.1	−1.8, 0.0	0.6, 2.5
From 35–44 to 45–54										
Mean	−3.4	0.7	−2.4	−0.8	0.5	0.8	2.2	−0.8	−1.0	0.6
95 % CI	−5.4, −1.5	−3.0, 4.4	−4.8, 0.1	−2.7, 1.0	−1.5, 2.4	−1.4, 3.0	−1.3, 5.7	−2.5, 0.8	−2.0, −0.1	−0.4, 1.7
From 45–54 to 55–64										
Mean	*−6.8*	−4.6	−0.8	−2.0	2.7	2.2	−0.3	1.6	*−2.2*	1.9
95 % CI	*−9.1*, *−4.5*	−9.0, −0.2	−3.7, 2.1	−4.3, 0.2	0.4, 5.0	−0.4, 4.8	−4.5, 3.9	−0.3, 3.6	*−3.3*, *−1.1*	0.7, 3.0
From 55–64 to 65–74										
Mean	*−6.7*	−3.4	−1.2	−1.9	−0.2	−0.3	−2.2	−0.4	−1.8	0.4
95 % CI	*−9.3*, *−4.1*	−8.4, 1.6	−4.5, 2.2	−4.5, 0.6	−2.9, 2.4	−3.3, 2.7	−7.0, 2.6	−2.7, 1.8	−3.1, −0.5	−1.0, 1.7
From 65–74 to 75+										
Mean	*−17.4*	*−15.0*	−2.9	−1.6	*−7.1*	−4.1	−3.1	1.8	*−5.5*	1.3
95 % CI	*−20.5*, *−14.3*	*−20.9*, *−9.1*	−6.9, 1.0	−4.6, 1.4	*−10.2*, *−4.0*	−7.6, −0.5	−8.7, 2.6	−0.8, 4.5	*−7.0*, *−3.9*	−0.3, 2.9

Emphasized cells represent estimated clinically relevant decline (>5 points for domains and >2–3 points for component summaries) [1, 2]

CI credible interval

Table 5 Mean differences in 10-year cross-sectional sf-36 scores by age group for men in the Canadian Multicentre Osteoporosis Study population-based cohort

For 10-year period	Physical functioning	Role physical	Bodily pain	General health	Vitality	Social functioning	Role emotional	Mental health	Physical component scale	Mental component scale
From 25–34 to 35–44										
Mean	−2.3	−4.7	−1.9	−0.6	0.7	−1.7	−4.5	0.1	−0.9	−0.2
95 % CI	−4.9, 0.2	−9.1, −0.2	−5.2, 1.5	−3.2, 2.0	−1.8, 3.3	−4.5, 1.1	−8.9, −0.1	−2.1, 2.3	−2.1, 0.3	−1.4, 1.1
From 35–44 to 45–54										
Mean	−2.4	2.1	2.4	−2.4	−1.8	1.4	3.3	0.0	−0.5	0.5
95 % CI	−5.1, 0.3	−2.6, 6.7	−1.1, 5.9	−5.2, 0.4	−4.5, 0.9	−1.5, 4.4	−1.3, 8.0	−2.3, 2.3	−1.8, 0.8	−0.9, 1.8
From 45–54 to 55–54										
Mean	−4.6	−2.7	−2.0	−2.9	3.1	1.1	4.7	3.7	*−2.5*	2.8
95 % CI	−7.8, −1.5	−8.3, 2.9	−6.1, 2.2	−6.2, 0.4	−0.1, 6.3	−2.4, 4.6	−0.8, 10.2	1.0, 6.4	*−4.0*, *−0.9*	1.2, 4.3
From 55–64 to 65–74										
Mean	*−6.1*	*−6.0*	−0.5	−0.5	−0.2	−2.2	*−6.7*	0.3	−1.5	−0.2
95 % CI	*−9.8*, *−2.5*	*−12.4*, *0.4*	−5.3, 4.2	−4.2, 3.3	−3.8, 3.4	−6.2, 1.8	*−13.0*, *−0.4*	−2.8, 3.5	−3.2, 0.2	−2.0, 1.6
From 65–74 to 75+										
Mean	*−13.5*	*−11.3*	−4.2	−3.3	*−5.9*	−2.8	−2.8	−1.1	*−4.5*	0.3
95 % CI	*−18.4*, *−8.6*	*−19.9*, *−2.8*	−10.6, 2.2	−8.3, 1.7	*−10.8*, *−1.0*	−8.1, 2.6	−11.3, 5.6	−5.3, 3.1	*−6.8*, *−2.1*	−2.1, 2.7

See footer of Table 4

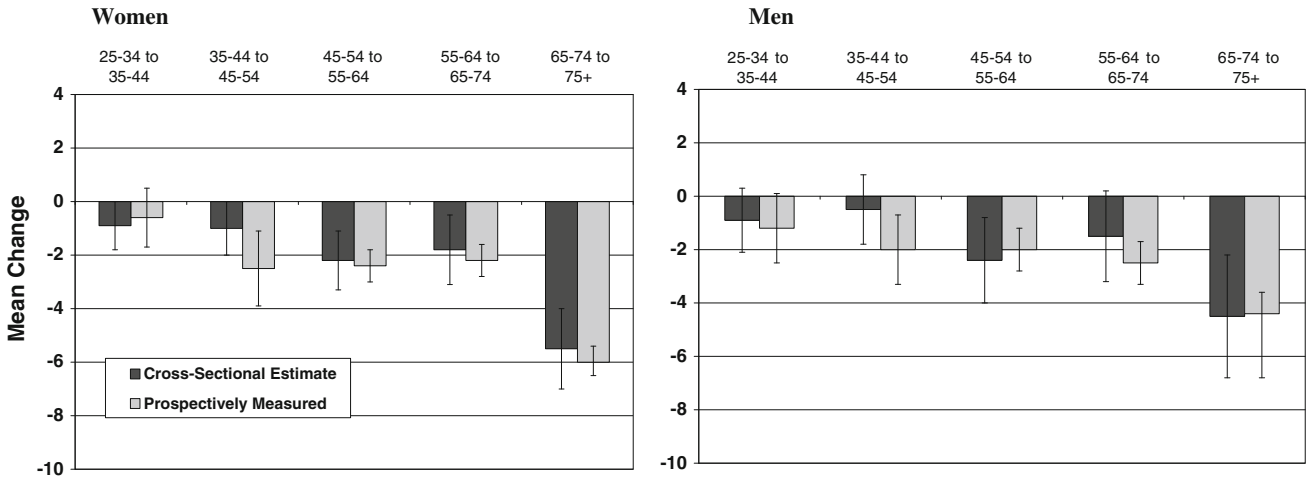

Fig. 1 Comparison of cross-sectional estimates of change and prospectively measured change in the physical component summary of the SF-36 in the population-based Canadian Multicentre Osteoporosis Study cohort

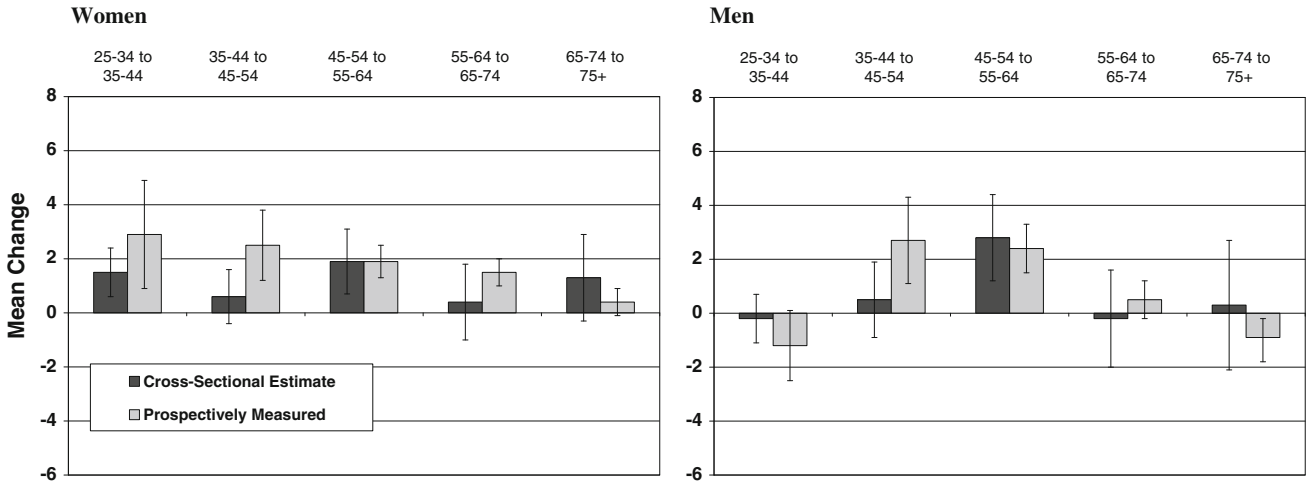

Fig. 2 Comparison of cross-sectional estimates of change and prospectively measured change in the mental component summary of the SF-36 in the population-based Canadian Multicentre Osteoporosis Study cohort

after middle age. Patterns of change in these prospective population-based data support the findings of previous research, suggesting that trends seen over 3–5 years continue in a similar direction [8, 11–14]. Physically oriented domains tended to decline in all but the youngest 10-year age groups of women and men, with declines becoming substantial in the oldest groups, particularly the oldest men where some declines were in excess of 25 points. Mentally oriented domains tended to improve over time, only showing substantial declines in the VT and RE domains in older women, and all domains in the two oldest groups of men. The larger declines in men may be related to the fact that at baseline, they had higher HRQOL on all domains and summary components compared to the women [10], and therefore, they had more room for decline. However, this pattern was also seen in the Rancho Bernardo study, where the mental health of men began to show decline in those over 70 years, while the mental health of women began to decline in those over 80 years [12].

The data also suggest that using the differences between 10-year age increments of cross-sectional data as an estimate of prospective change may be reasonable at least for the younger age groups, e.g., those under 65 years, where the estimates based on cross-sectional data are generally quite similar to those based on prospectively measured change. Both methods showed small improvements or non-relevant declines in virtually all domains and component summary comparisons. The two figures demonstrate the remarkable correspondence between the prospective and cross-sectional estimates of change in the PCS and MCS.

Table 6 Difference between prospectively measured change and 10-year cross-sectional estimates of change based on normative data, by age group for women in the Canadian Multicentre Osteoporosis Study population-based cohort

	Physical functioning	Role physical	Bodily pain	General health	Vitality	Social functioning	Role emotional	Mental health	Physical component scale	Mental component scale
From 25–34 to 35–44										
Mean	1.2	*7.6*	1.2	1.4	0.5	3.3	4.8	2.0	0.9	1.3
95 % CI	−2.6, 4.9	*−1.1, 16.5*	−4.6, 7.1	−2.8, 5.8	−4.1, 5.1	−2.2, 8.9	−3.6, 13.2	−2.3, 6.4	−1.2, 3.0	−1.4, 4.0
From 35–44 to 45–54										
Mean	−0.2	−3.6	−3.6	0.2	0.4	2.2	0.4	2.1	−1.1	1.2
95 % CI	−3.5, 3.1	−10.8, 3.8	−8.4, 1.2	−3.1, 3.5	−2.9, 3.8	−1.7, 6.0	−5.9, 6.5	−0.7, 4.8	−2.9, 0.7	−0.5, 3.0
From 45–54 to 55–54										
Mean	0.2	−0.4	−1.8	−0.6	−0.7	−1.0	1.9	0.8	−0.5	0.3
95 % CI	−1.7, 2.2	−4.2, 3.5	−4.4, 0.6	−2.5, 1.2	−2.6, 1.2	−3.3, 1.2	−1.3, 5.2	−0.8, 2.4	−1.5, 0.4	−0.6, 1.3
From 55–64 to 65–74										
Mean	1.1	0.8	−2.7	0.7	0.4	1.4	1.6	1.5	−0.2	0.8
95 % CI	−0.7, 3.0	−2.7, 4.2	−4.9, −0.6	−0.9, 2.3	−1.2, 2.0	−0.5, 3.3	−1.4, 4.5	0.2, 2.8	−1.1, 0.6	0.0, 1.6
From 65–74 to 75+										
Mean	−1.0	−0.8	−3.8	−2.9	−0.5	1.4	−2.1	−0.3	−1.1	0.1
95 % CI	−3.5, 1.4	−4.9, 3.4	−6.3, −1.3	−4.7, −1.1	−2.4, 1.4	−0.9, 3.7	−5.6, 1.4	−1.8, 1.2	−2.1, −0.1	−0.8, 1.0

Positive values indicate that the prospectively collected data saw an improvement (or smaller mean decline) than the cross-sectional estimate, while negative values indicate that the prospectively collected data showed a greater decline (or smaller mean improvement) than the cross-sectional estimate. The emphasized cell represents the only one that represents a clinically relevant difference

CI Credible interval

Table 7 Difference between prospectively measured change and 10-year cross-sectional estimates of change based on normative data, by age group for men in the Canadian Multicentre Osteoporosis Study population-based cohort

	Physical functioning	Role physical	Bodily pain	General health	Vitality	Social functioning	Role emotional	Mental health	Physical component scale	Mental component scale
From 25–34 to 35–44										
Mean	2.9	4.0	−0.9	−2.5	−2.7	3.3	3.4	0.6	0.1	0.3
95 % CI	−0.6, 6.3	−3.1, 11.2	−6.8, 5.0	−6.8, 1.8	−6.9, 1.5	−1.7, 8.3	−4.9, 11.7	−3.2, 4.3	−1.7, 2.0	−2.0, 2.6
From 35–44 to 45–54										
Mean	1.2	0.2	0.1	1.9	1.3	1.5	−0.5	2.1	0.3	0.7
95 % CI	−1.8, 4.2	−6.4, 6.9	−4.9, 5.1	−1.5, 5.3	−2.1, 4.7	−2.9, 5.9	−7.3, 6.2	−1.1, 5.2	−1.3, 1.9	−1.2, 2.6
From 45–54 to 55–54										
Mean	1.6	0.4	−2.4	−0.7	−1.3	0.1	−0.6	0.7	−0.1	−0.1
95 % CI	−1.0, 4.1	−4.4, 5.0	−5.8, 1.0	−3.1, 1.8	−3.9, 1.2	−2.9, 3.1	−4.8, 3.5	−1.5, 2.9	−1.4, 1.1	−1.4, 1.2
From 55–64 to 65–74										
Mean	−0.8	−2.2	−1.7	−2.0	−2.0	−0.2	1.2	0.3	−1.0	0.3
95 % CI	−3.6, 2.0	−7.6, 3.1	−5.1, 1.7	−4.6, 0.5	−4.4, 0.4	−3.0, 2.5	−2.8, 5.1	−1.6, 2.3	−2.3, 0.3	−0.8, 1.3
From 65–74 to 75+										
Mean	−2.3	−4.3	−4.7	−3.5	−2.6	−2.0	−2.1	−0.6	−1.7	−0.4
95 % CI	−6.6, 2.2	−11.1, 2.8	−8.9, −0.4	−6.5, −0.4	−5.9, 0.7	−5.5, 1.5	−7.8, 3.5	−3.0, 1.9	−3.4, 0.0	−1.8, 1.0

See footer of Table 6

We found some support for the Hemingway et al. [8] finding that estimates of within-person change are larger in prospectively collected data, indicating that the use of cross-sectional data may underestimate change, but this was only evident in men. The greatest within-domain inconsistency was in the BP domain for both men and women, where prospectively collected data suggested increased levels of pain compared with the cross-sectional projections in all sex and age group strata. It should be noted that while the direction of the findings was consistent, none of the comparisons attained clinical relevance.

Examining changes in HRQOL, both prospectively and cross-sectionally, adds to the body of literature for establishing meaningful change over time, information which is needed to assess changes due to aging versus those related to interventions or development of disease [8, 23]. This is particularly important since a change in health state can result in a change in internal standards, values or conceptualization known as a "response shift" [24], which can occur when an individual experiences a catalyst such as a cancer diagnosis. This can be particularly important in prospective studies such as this one, where many participants may have health changes over the 10-year period. However, a meta-analysis of the effect size of response shifts in studies that assessed quality of life found that the magnitude was small and that positive and negative values tended to cancel each other out [25]. The authors indicated that a definite conclusion regarding the clinical significance of response shift cannot currently be drawn on the basis of existing studies. Notably, an evaluation of response shift within the CaM*os* database is currently underway.

There are clear limitations to this study, the most significant one being the loss to follow-up. Only 61.1 and 54.4 % of the original cohort of women and men, respectively, completed the SF-36 at year 10. However, multiple imputation was not done since previous work assessing the 5-year changes produced negligible differences in the estimates of change [14]. Moreover, although multiple imputation may have merit for correcting at least some of the bias, it has limitations of its own in the context of non-ignorable data, which may explain why it was also not utilized in the other two prospective assessments of HRQOL [8, 12].

There is little doubt that a complete case analysis, while relatively common, will produce bias because those who have died or are lost to follow-up are not accounted for. Use of a measure that does incorporate death and life states worse than death would be a useful approach, but the current study incorporated the SF-36 from the beginning (1995) and was utilized at subsequent follow-ups to ensure consistency. A comparison of the baseline scores of those who did and did not have year 10 data demonstrated that with one exception, the baseline scores of those without year 10 data were consistently lower than those with complete data. This would suggest that had these participants been included in some way, the prospectively measured change would likely have shown greater decline.

These results are based on a random sample of the Canadian population, but we believe that the results are generalizable to populations beyond Canada. The concept of HRQOL transcends issues of different health care delivery and contextual situations, as is evidenced by the similarity of scores across different countries and different cultures [14, 26]. Likewise, the refinement of the SF-36 for different populations such as Veterans, since the items, the domains and the general concepts measured remain the same.

Conclusions

Changes in HRQOL may be minimal over much of an individual's life span. This is similar to what was noted in this population over a 5-year follow-up [14] and to the conclusions of the two other prospective but shorter duration studies of HRQOL change [8, 12]. It should be noted that these are general patterns and large standard deviations for the individual scores suggest that many participants saw either large declines or large improvements, functionally canceling each other out and resulting in relatively small overall changes. Care also needs to be taken when examining change in physical function by middle age, and particularly in the elderly, as changes accumulated over time may become substantial. These data also suggest that differences in 10-year age increments of cross-sectional data may be a reasonable proxy for longitudinal changes but that accuracy can decrease for men in particular and within assessment of pain over time. Cross-sectional estimates tended to underestimate change, as a much larger proportion of comparisons saw clinically relevant declines with the prospectively collected data.

Ongoing work into the prospective assessment of HRQOL, as well as an evaluation of response shift within such longitudinal databases, will provide additional insight into the natural progression of HRQOL across the lifespan.

Appendix: CaMos Research Group

David Goltzman (co-principal investigator, McGill University), Nancy Kreiger (co-principal investigator, Toronto), Alan Tenenhouse (principal investigator emeritus, Toronto).

CaMos Coordinating Centre, McGill University, Montreal, Quebec: Suzanne Godmaire (research assistant), Silvia Dumont (administrative assistant), Claudie Berger

(study statistician), Lisa Langsetmo (Fellow), Wei Zhou (statistician).
Memorial University, St. John's Newfoundland: Carol Joyce (director), Christopher Kovacs (co-director), Emma Sheppard (coordinator).
Dalhousie University, Halifax, Nova Scotia: Susan Kirkland, Stephanie Kaiser (co-directors), Barbara Stanfield (coordinator).
Laval University, Quebec City, Quebec: Jacques P. Brown (director), Louis Bessette (co-director).
Queen's University, Kingston, Ontario: Tassos Anastassiades (director), Tanveer Towheed (co-director), Wilma Hopman (research associate), Karen Rees-Milton (coordinator).
University of Toronto, Toronto, Ontario: Bob Josse (director), Sophie Jamal (co-director), Barbara Gardner-Bray (coordinator).
McMaster University, Hamilton, Ontario: Jonathan D. Adachi (director), Alexandra Papaioannou (co-director), Laura Pickard (coordinator).
University of Saskatchewan, Saskatoon, Saskatchewan: Wojciech P. Olszynski (director), K. Shawn Davison (co-director), Jola Thingvold (coordinator).
University of Calgary, Calgary, Alberta: David A. Hanley (director), Jane Allan (coordinator).
University British Columbia, Vancouver, British Columbia: Jerilynn C. Prior (director), Millan Patel (co-director), Brian Lentle (radiologist), Nerkeza Andjelic (coordinator).
McGill University, Montreal, Quebec: Elham Rahme (biostatistician), Brent Richards (research associate).
University of Alberta, Edmonton, Alberta: Stuart Jackson (medical physicist).
University of Manitoba, Winnipeg, Manitoba: William Leslie (nuclear medicine physician). CaMos is currently funded by Canadian Institutes of Health Research (CIHR), Amgen Canada Inc., Dairy Farmers of Canada, Merck Canada, Eli Lilly Canada, Novartis Canada.

References

1. Ware, J. E., Jr., Snow, K. K., Kosinski, M., & Gandek, B. (1993). *SF-36 health survey: Manual and interpretation guide*. Boston: The Health Institute, New England Medical Center.
2. Ware, J. E., Jr., Kosinski, M., & Keller, S. D. (1994). *SF-36 physical and mental summary scales: A user's manual*. Boston: The Health Institute, New England Medical Center.
3. Ware, J. E., Jr. (2000). SF-36 health survey updated. *Spine, 25*(24), 3130–3139.
4. Garratt, A., Schmidt, L., Mackintosh, A., & Fitzpatrick, R. (2002). Quality of life assessment: Bibliographic study of patient assessed health outcome measures. *British Medical Journal, 324*, 1417–1421.
5. Hopman, W. M., Harrison, M., Coo, H., Friedberg, E., Buchanan, M., & Vandenkerkhof, E. (2009). The impact of chronic disease and increased age on physical and mental health. *Chronic Diseases in Canada, 29*(3), 108–116.
6. Bayliss, E. A., Bayliss, M. S., Ware, J. E., Jr., & Steiner, J. F. (2004). Predicting declines in physical function in persons with multiple chronic medical conditions: What we can learn from the medical problem list. *Health and Quality of Life Outcomes, 2*, 47–55.
7. Andreson, E. M., & Meyers, A. R. (2000). Health-related quality of life outcomes measures. *Archives of Physical Medicine and Rehabilitation, 81*(12 Suppl. 2), S30–S45.
8. Hemingway, H., Stafford, M., Stansfield, S., Shipley, M., & Marmot, M. (1997). Is the SF-36 a valid measure of change in population health? Results from the Whitehall II study. *British Medical Journal, 315*, 1273–1279.
9. Beaton, D. E., Hogg-Johnson, S., & Bombardier, C. (1997). Evaluating changes in health status: Reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *Journal of Clinical Epidemiology, 50*(1), 79–93.
10. Hopman, W. M., Towheed, T., Anastassiades, T., Tenenhouse, A., Poliquin, S., Berger, C., et al. (2000). Canadian normative data for the SF-36 health survey. *Canadian Medical Association Journal, 163*, 265–271.
11. Mein, G., Martikainen, P., Hemingway, H., Stansfield, S., & Marmot, M. (2003). Is retirement good or bad for mental and physical health functioning? Whitehall II longitudinal study of civil servants. *Journal of Epidemiology and Community Health, 57*(1), 46–49.
12. Der-Martirosian, C., Kritz-Silverstein, D., & Barret-Connor, E. (2010). Five-year stability in associations of health-related quality of life measures in community-dwelling older adults: The Rancho Bernardo Study. *Quality of Life Research, 19*(9), 1333–1341.
13. Hopman, W. M., Berger, C., Joseph, L., Towheed, T., vanden-Kerkhof, E., Anastassiades, T., et al. (2004). Stability of normative data for the SF-36: Results of a three-year prospective study in middle-aged Canadians. *Canadian Journal of Public Health, 95*(3), 387–391.
14. Hopman, W. M., Berger, C., Joseph, L., Towheed, T., vanden-Kerkhof, E., Anastassiades, T., et al. (2006). The natural progression of health-related quality of life: Results of a five-year prospective study of SF-36 in a normative population. *Quality of Life Research, 15*(3), 527–536.
15. Kreiger, N., Tenenhouse, A., Joseph, L., MacKenzie, T., Poliquin, S., Brown, J., et al. (1999). Research Notes: The Canadian Multicentre Osteoporosis Study (CaMos): Background, rationale, methods. *Canadian Journal on Aging, 18*, 376–387.
16. Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581–592.
17. Diehr, P., Patrick, D. L., Spertus, J., Kiefe, C. I., McDonell, M., & Fihn, S. D. (2001). Transforming self-rated health and the SF-36 scales to include death and improve interpretability. *Medical Care, 39*(7), 670–680.
18. Little, R. J., D'Agostino, R., Cohen, M. L., Dickersin, K., Emerson, S. S., et al. (2012). Special Report: The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine, 367*(14), 1355–1360.
19. Shih, W. J. (2002). Problems in dealing with missing data and informative censoring in clinical trials (review). *Current Controlled Trials in Cardiovascular Medicine, 3*(4), 7.
20. Wood, A. M., White, A. R., & Thompson, S. G. (2004). Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clinical Trials, 1*(4), 368–376.

21. Kmetic, A., Joseph, L., Berger, C., & Tenenhouse, A. (2002). Multiple imputation to account for missing data in a survey: Estimating the prevalence of osteoporosis. *Epidemiology, 13*(4), 437–444.
22. Rubin, D. (1987). *Multiple imputation for non-response in surveys*. New York: Wiley.
23. Testa, M. A. (2000). Interpretation of quality of life outcomes: Issues that affect magnitude and meaning. *Medical Care, 38*(Suppl. II), 166–174.
24. Oort, F. J., Visser, M. R., & Sprangers, M. A. (2009). Formal definitions and measurement of bias and explanation bias clarify measurement and conceptual perspectives on response shift. *Journal of Clinical Epidemiology, 62*, 1126–1137.
25. Schwartz, C. E., Bode, R., Repucci, N., Becker, J., Sprangers, M. A., & Fayers, P. M. (2006). The clinical significance of adaptation to changing health: A meta-analysis of response shift. *Quality of Life Research, 15*(9), 1533–1550.
26. Wagner, A. K., Gandek, B., Aaronson, N. K., Acquadro, C., Alonzo, J., Apolone, G., et al. (1998). Cross-cultural comparisons of the content of SF-36 translations across 10 countries: Results from the IQOLA Project International Quality of Life Assessment. *Journal of Clinical Epidemiology, 51*(11), 925–932.