Qual Life Res (2014) 23:1515–1521
DOI 10.1007/s11136-013-0586-3

A comparison of SF-36 summary measures of physical and mental health for women across the life course

Gita D. Mishra · Richard Hockey · Annette J. Dobson

Accepted: 19 April 2013 / Published online: 3 December 2013

Abstract

Background Physical and mental component summary scores (PCS and MCS, respectively) are often used to summarise SF-36 quality of life subscales. This paper investigates PCS and MCS across the life course and compares the trajectories obtained from two different methods of calculation.

Methods The Australian Longitudinal Study on Women's Health is a population-based study with three cohorts of women and SF-36 surveys taken at multiple time points. Scoring coefficients for each component score were determined using factor analysis with uncorrelated (orthogonal) and correlated (oblique) rotation at the baseline survey, which were then used to compute correlated and uncorrelated PCS and MCS scores at each survey (scaled to have mean of 50 and standard deviation of 10 at baseline).

Results For both methods, PCS declined progressively across the lifespan, while MCS rose in young and mid-age women to a peak and subsequently declined in later life. Differences were apparent between correlated and uncorrelated scores, most notably for MCS in the older cohort, where correlated MCS reached 54.6 but still less than uncorrelated MCS, with a random effects model indicating 1.63 (95 % confidence intervals 1.58–1.67) units difference; it then declined to a score of 51.2 by the last survey and the difference widened to 3.44 (3.38–3.50) units compared with the uncorrelated MCS.

Conclusions PCS and MCS have distinct trajectories through life, with differences in results from correlated and uncorrelated component summary scores. The divergence is most notable with MCS, especially for older women, suggesting that correlated MCS and PCS should be used when examining change in health over time in this age group.

Keywords Physical component summary scores · Mental component summary scores · Uncorrelated scores · Oblique rotation · Varimax rotation · Australian women · Life course · SF-36

G. D. Mishra (✉) · R. Hockey · A. J. Dobson
School of Population Health, The University of Queensland, Herston Road, Herston, QLD 4006, Australia
e-mail: g.mishra@sph.uq.edu.au

Introduction

Health-related quality of life is often assessed via the short-form general health survey (SF-36), as this instrument provides a systematic approach for comparison between studies and populations [1]. SF-36 scores can be summarised using two widely accepted constructs labelled as 'physical component summary' (PCS) and 'mental component summary' (MCS) scores, which are based on exploratory factor analysis of the eight SF-36 subscales: physical functioning (PF), role limitation caused by physical health problems (RP), bodily pain (BP), general health perceptions (GH), vitality (VT) for energy levels and fatigue, social functioning (SF), role limitation because of emotional problems (RE), and mental health (MH) [2]. PCS and MCS scores are represented on a standardised scale (as a *T* score with mean 50 and standard deviation 10) and have better distributional properties (continuous and symmetrical) than individual SF-36 subscales. While studies often compute PCS and MCS scores for their sample based on population norms for weightings and some longitudinal studies have tracked individual subscales, such as MH and PF, the precise trajectory of component summary scores over the life course is less clear.

The use of PCS and MCS scores as a means of characterising health has been subject to considerable debate [3–8]. Much of this focuses on the method used in their calculation, specifically constraining factor analysis to have two uncorrelated axes (achieved via orthogonal rotation). The labelling of the resultant uncorrelated summary scores PCS_{uc} and MCS_{uc} implies that 'physical' and 'mental' characterisations of health occupy distinct dimensions. PCS_{uc} and MCS_{uc} comprise loadings from all subscales, with some subscales in PCS_{uc}, such as the mental health subscale, having a negative weighting. Thus, improved mental health (for instance reflected in higher MH and RE scores) can result in a counter-intuitive *decline* in PCS_{uc}. Similarly, PF and some other subscales have negative weightings on MCS_{uc}. Thus, concerns have been raised about the implications of these negative weightings when investigating changes in summary scores over time [5]. Some authors have suggested that due to links between physical and mental health, correlated summary scores (PCS_c and MCS_c) should be derived using factor analysis where correlation between axes is permitted, via the use of oblique rather than orthogonal rotation [3]. This approach tends to reduce rather than eliminate negative weightings entirely. Ware and Kosinski [9] have argued, however, that PCS_{uc} and MCS_{uc} should always be viewed alongside scores for individual subscales so that unusual results can be identified.

In spite of the extensive use of uncorrelated summary scores from SF-36 in many studies, differences between PCS and MCS derived from the two methods remains unclear; specifically how differences vary across life and consequent implications for interpreting changes in summary scores over long time periods. The issue being that the magnitude of difference in the summary scores as a measure of the rate of change in health of patients may be used, for instance, to assess the effectiveness or otherwise of a treatment programme. In this paper, we compare PCS and MCS scores across the life course via both orthogonal and oblique settings for factor analysis of SF-36 subscales using data from three cohorts of women in the Australian Longitudinal Study on Women's Health (ALSWH) [10].

Methods

Australian Longitudinal Study on Women's Health (ALSWH) is an ongoing, broad-ranging prospective study of factors affecting health and well-being of Australian women. In 1996, self-reported data on health including from SF-36, health service use, socio-demographics, and personal information were collected from women in three age cohorts: the young cohort, born in 1973–1978 (aged 18–23 years), the mid-age cohort, born in 1946–1951 (45–50 years), and the older cohort, born in 1921–1926 (70–75 years). To date all cohort members have been followed up in at least five waves (younger cohort: 1996, 2000, 2003, 2006, 2009; mid-age cohort: 1996, 1998, 2001, 2004, 2007, 2010; and older cohort: 1996, 1999, 2002, 2005, 2008). The study has been approved by Ethics Committees at the University of Queensland and University of Newcastle. Further details of recruitment and response rates have been published elsewhere [10]. A total of 40,394 women completed the baseline survey, of these 9.8 % died during the study, 9.5 % withdrew, 18 % did not complete any surveys, while 23 % only completed surveys intermittently. The present study focuses on the remaining 16,006 women who provided complete data for SF-36 at all surveys (younger cohort ($n = 5458$), mid-age cohort ($n = 7212$), and older cohort ($n = 3336$)).

Statistical method

Using SF-36 survey data from baseline and applying a very similar process to Ware and colleagues [2, 11], we obtained two versions of PCS and MCS scores for all women: uncorrelated (PCS_{uc} and MCS_{uc}) and correlated (PCS_c and MCS_c). First, scoring coefficients (factor loadings) for PCS_{uc} and MCS_{uc} were determined using factor analysis of SF-36 subscales with a two-factor orthogonal rotation. Second, all data for the eight SF-36 subscales for each women were rescaled by taking the difference between the subscale and the mean of the subscale at baseline and then dividing by the standard deviation at baseline. Note these results in SF-36 subscales that are z scores at baseline only (mean 0, standard deviation 1), while data for SF-36 subscales for subsequent surveys are standardised relative to baseline. Third, the two summary scores for the woman at every survey were computed by multiplying each of the standardised SF-36 subscales by its respective scoring coefficient (obtained in the first step) and summing the results. These two summary scores were then rescaled by multiplying by 10 and adding 50, so that at baseline the resultant PCS_{uc} and MCS_{uc} have mean 50 and standard deviation 10. Note that scores for subsequent surveys only differ by the standardised SF-36 subscales used in the calculation (scoring coefficients remain the same across surveys).

The correlated summary scores, PCS_c and MCS_c, were obtained by repeating steps above, but using two-factor oblique rotation in the factor analysis to determine the baseline scoring coefficients. Spearman correlation coefficients were also calculated between PF and MH subscales and component summary scores. The intra-class correlation (ICC) coefficients were calculated for each of the summary scores. Random effects models, a method for longitudinal regression for correlated data, [12] were used to investigate (1) if the summary scores produced by each method varied across surveys; (2) the difference in summary scores produced by the two methods was significantly different from zero; and (3) the extent that the magnitude of this difference varied across surveys. In these models, the intercept was considered as random effect and survey number (representing time) as fixed effect. All analyses were repeated for each cohort and performed in SAS version 9.1 [13].

Results

Table 1 shows the variation in individual subscales scores for each cohort through the study. Women in the younger cohort experienced a decline in PF, RP, and VT from survey 1 (aged 18–23) through to survey 5 (aged 31–36) but improved with respect to GH, SF, RE, and MH subscales. For the mid-age cohort PF, RP, BP, and GH declined from survey 1 (aged 45–50) to survey 6 (aged 59–64), but they also experienced a rise in RE and MH scores. Women in the older cohort, however, show a marked decline in all eight subscales of SF-36 from survey 1 (aged 70–75) to survey 5 (aged 85–90). For instance, PF for this cohort dropped from a mean score of 70.5 at baseline to 50.1 at S5 12 years later (Table 1).

For both the uncorrelated and correlated factors (Table 2), the factor 'physical component' had coefficients with high values (factor loadings of 0.2 or more) from PF, RP, BP, and GH, while 'mental component' had high coefficients for VT, SF, RE, and MH. Scoring coefficients with negative values, however, were considerably reduced for correlated factors compared with uncorrelated factors,

Table 1 Mean (standard deviation) of SF-36 subscales by age cohort

SF-36/cohort	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Survey 6
Physical functioning						
1973–1978	92.0 (12.6)	92.0 (13.6)	91.4 (14.5)	91.5 (14.2)	90.5 (15.4)	–
1946–1951	87.5 (15.8)	85.7 (16.8)	83.6 (17.8)	82.0 (18.8)	81.5 (19.4)	79.5 (20.5)
1921–1926	70.5 (21.5)	68.1 (21.9)	64.0 (24.6)	57.6 (26.3)	50.1 (27.2)	–
Role physical						
1973–1978	84.6 (29.2)	82.5 (30.4)	81.9 (31.4)	80.0 (33.6)	81.6 (32.7)	–
1946–1951	83.6 (32.2)	80.0 (33.8)	77.8 (35.0)	77.0 (35.8)	77.3 (36.3)	75.0 (37.4)
1921–1926	65.5 (40.5)	61.2 (40.6)	52.9 (41.3)	45.5 (41.6)	38.0 (40.6)	–
Bodily pain						
1973–1978	74.9 (20.9)	75.0 (21.1)	75.8 (20.6)	74.8 (20.6)	74.0 (20.9)	–
1946–1951	72.8 (22.3)	72.5 (22.9)	70.5 (22.9)	69.7 (22.8)	69.3 (23.0)	67.4 (23.2)
1921–1926	69.9 (24.3)	68.0 (25.0)	65.0 (25.3)	61.8 (26.0)	59.4 (26.6)	–
General health						
1973–1978	70.3 (19.9)	71.7 (19.5)	73.2 (19.0)	74.7 (18.8)	74.3 (18.6)	–
1946–1951	74.6 (18.8)	74.8 (19.1)	73.0 (19.5)	72.3 (19.8)	72.7 (19.7)	71.2 (20.0)
1921–1926	71.6 (18.7)	71.0 (18.4)	69.7 (18.6)	66.5 (19.5)	62.4 (20.2)	–
Vitality						
1973–1978	57.6 (19.6)	53.4 (19.6)	53.5 (19.8)	54.8 (19.7)	53.6 (20.0)	–
1946–1951	60.2 (20.1)	58.7 (21.5)	58.3 (21.4)	59.4 (21.2)	61.0 (20.8)	61.2 (20.8)
1921–1926	64.4 (18.8)	61.5 (19.1)	59.3 (19.4)	57.2 (20.4)	53.8 (20.5)	–
Social functioning						
1973–1978	78.4 (21.9)	79.6 (22.5)	81.6 (21.6)	82.2 (21.4)	83.5 (21.6)	–
1946–1951	84.9 (21.4)	84.0 (22.0)	84.1 (22.3)	84.3 (21.9)	84.9 (21.8)	84.3 (22.9)
1921–1926	87.6 (20.8)	87.2 (21.1)	83.2 (23.5)	79.6 (24.8)	75.2 (27.2)	–
Role emotion						
1973–1978	72.7 (36.3)	75.4 (36.4)	78.5 (34.9)	80.0 (33.9)	79.5 (34.8)	–
1946–1951	81.1 (33.1)	82.0 (33.0)	82.7 (32.4)	83.0 (32.3)	84.4 (31.5)	84.7 (31.4)
1921–1926	83.7 (31.7)	84.0 (30.7)	81.6 (32.7)	77.7 (35.5)	71.9 (38.7)	–
Mental health						
1973–1978	69.6 (17.4)	69.4 (17.0)	71.2 (16.7)	73.0 (16.2)	72.8 (16.4)	–
1946–1951	75.0 (16.6)	75.1 (17.5)	75.1 (17.2)	76.0 (17.0)	77.0 (16.7)	77.4 (16.8)
1921–1926	81.0 (14.8)	81.5 (14.0)	81.2 (14.1)	81.0 (14.6)	79.5 (15.7)	–

1973–1978 cohort: $n = 5{,}458$; 1946–1951 cohort: $n = 7{,}212$; 1921–1926 cohort: $n = 3{,}336$

for instance, MH reduced from −0.21 for PCS_{uc} to −0.1 for PCS_c; similarly for PF in MCS_{uc} compared with MCS_c. The percentage of variance explained by each of the two correlated factors (obtained using oblique rotation) was also higher than that for the uncorrelated factors (obtained using orthogonal factor rotation).

When tracking mean PCS obtained from both factor rotation methods, results from the three cohorts of women showed a declining trajectory over the life course (Fig. 1; Table 3), with PCS_{uc} dropping from 54.3 at baseline for younger women to 37.9 by the last survey for older women. PCS_c showed a similar decline, though not quite as steep overall: starting with slightly lower values in the younger cohort and then higher values than PCS_{uc} in the mid-age cohort that widens to an essentially parallel declining trajectory in the older cohort but at 1.5 units higher than PCS_{uc}. The higher value of the ICC for the correlated factors compared to the uncorrelated factors indicated a greater degree of agreement in measurements of the correlated factors across time. The ICC was lowest for the 1973–1978 cohort (Table 3).

Results from the random effects models shown in Table 4 confirm the difference in trajectories of the two measures for PCS: for the younger cohort PCS_{uc} are higher, but with the differences narrowing from 0.85 (0.81–0.88) at survey 1 to 0.33 (0.28–0.37) at survey 5. For the mid-age cohort, the difference is smaller with an overall estimate of PCS_{uc} being lower by 0.25 (0.22–0.28)—though the gap widens such that by survey 6 the difference is −0.61 (−0.65 to −0.57). For the older cohort, PCS_{uc} is consistently lower than PCS_c by an estimated 1.10 (1.07–1.13) across the five surveys.

A markedly different trajectory was indicated by MCS over the life course, rising from low levels in the young women and then at a slower rate in mid-life to peak at the baseline or second survey for the older women, before declining rapidly thereafter (Fig. 2). A noticeable divergence between MCS_{uc} and MCS_c was also evident across the three age cohorts. For the younger cohort, MCS_{uc} is consistently lower than MCS_c by an overall estimate of 0.90 (0.88–0.92) (Table 4). In the mid-age cohort, the two measures crossover with MCS_{uc} lower than MCS_c by 0.39 (0.35–0.43) at survey 1, with this difference reversing to 0.57 (0.53–0.61) by survey 6. By baseline in the older cohort, MCS_c reached 54.6, but was lower than MCS_{uc} by 1.63 (1.58–1.67) unit, and had declined to 51.2 by the last survey with the divergence widening to 3.44 (3.38–3.50) units lower than MCS_{uc}.

Discussion

As far as we are aware, this is the first study to use SF-36 summary component scores from a population-based

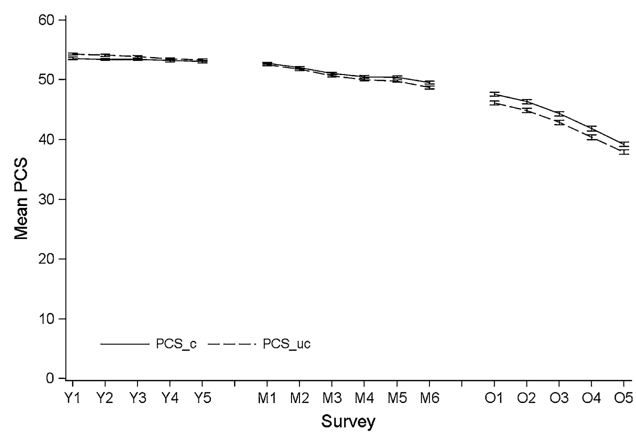

Fig. 1 Mean and 95 % confidence interval of physical component summary measures (PCS_uc derived from orthogonal rotation; PCS_c derived from oblique rotation) by survey wave and age cohort (Y1 to Y5 refers to survey 1–survey 5 for the 1973–1978 cohort, M1–M6 refers to survey 1–survey 6 for the 1946–1951 cohort, and O1–O5 refers to survey 1–survey 5 for the 1921–1926 cohort)

Table 2 Standardised scoring coefficients by the type of factor rotation using data at Survey 1 ($N = 39{,}241$)

	Physical component		Mental component	
	Orthogonal PCS_{uc}	Oblique PCS_c	Orthogonal MCS_{uc}	Oblique MCS_c
Physical functioning	0.44	0.38	−0.21	−0.10
Role physical	0.36	0.32	−0.10	−0.01
Bodily Pain Index	0.33	0.30	−0.08	0.00
General health perceptions	0.20	0.21	0.06	0.11
Vitality	0.02	0.08	0.24	0.24
Social functioning	−0.02	0.05	0.28	0.27
Role emotional	−0.13	−0.05	0.34	0.30
Mental health	−0.21	−0.10	0.42	0.36
Per cent variance explained by two factors (%)	35	44	33	41

Table 3 Mean (standard deviation) of SF-36 physical and mental health component summary scores by age cohort, derived with orthogonal and oblique rotation in factor analysis

SF-36 summary scores[a]/cohort	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Survey 6[b]	Standardised regression coefficient (95 % confidence interval) for time effect	Intra-class correlation (95 % confidence interval)
PCS_{uc}								
1973–1978	54.3 (6.7)	54.1 (7.0)	53.9 (7.2)	53.5 (7.5)	53.3 (7.6)	48.8 (9.6)	−0.20 (−0.24, −0.16)	0.34 (0.33, 0.35)
1946–1951	52.6 (7.9)	51.8 (8.4)	50.7 (8.8)	50.1 (9.1)	49.8 (9.3)		−0.74 (−0.77, −0.72)	0.57 (0.56, 0.58)
1921–1926	46.2 (9.9)	44.9 (10.1)	42.9 (10.6)	40.4 (11.0)	37.9 (11.0)		−1.98 (−2.03, −1.93)	0.60 (0.59, 0.61)
PCS_c								
1973–1978	53.6 (6.9)	53.4 (7.2)	53.4 (7.2)	53.2 (7.6)	53.1 (7.7)	49.5 (10.0)	−0.06 (−0.10, −0.03)	0.42 (0.41, 0.43)
1946–1951	52.8 (8.2)	52.1 (8.7)	51.1 (9.1)	50.6 (9.4)	50.5 (9.6)		−0.63 (−0.65, −0.60)	0.62 (0.62, 0.63)
1921–1926	47.6 (10.1)	46.4 (10.3)	44.4 (10.8)	41.9 (11.2)	39.2 (11.3)		−2.02 (−2.07, −1.96)	0.62 (0.61, 0.63)
MCS_{uc}								
1973–1978	47.5 (9.7)	47.4 (9.7)	48.5 (9.4)	49.4 (9.2)	49.3 (9.4)	53.2 (8.8)	0.59 (0.54, 0.64)	0.42 (0.41, 0.43)
1946–1951	51.2 (9.1)	51.3 (9.4)	51.6 (9.2)	52.2 (9.1)	52.9 (8.8)		0.44 (0.41, 0.47)	0.52 (0.51, 0.53)
1921–1926	55.7 (7.5)	55.8 (7.2)	55.5 (7.5)	55.2 (7.9)	54.2 (8.6)		−0.29 (−0.34, −0.24)	0.52 (0.51, 0.53)
MCS_c								
1973–1978	48.6 (9.5)	48.5 (9.6)	49.5 (9.3)	50.2 (9.1)	50.2 (9.2)	52.8 (9.3)	0.52 (0.48, 0.57)	0.46 (0.45, 0.47)
1946–1951	51.8 (9.2)	51.7 (9.6)	51.7 (9.5)	52.1 (9.4)	52.7 (9.2)		0.25 (0.22, 0.28)	0.58 (0.57, 0.59)
1921–1926	54.6 (8.0)	54.4 (7.8)	53.6 (8.1)	52.7 (8.6)	51.2 (9.2)		−0.77 (−0.82, −0.72)	0.58 (0.57, 0.59)

1973–1978 cohort: $n = 5{,}458$; 1946–1951 cohort: $n = 7{,}212$; 1921–1926 cohort: $n = 3{,}336$

[a] PCS_{uc} and MCS_{uc} were obtained from orthogonal rotation while PCS_c and MCS_c from oblique rotation

[b] Only 1946–1951 cohort had been followed up in six waves

Table 4 Estimates (means and 95 % confidence intervals) from random effects models for a difference in SF-36 physical and mental health component summary scores, derived with orthogonal and oblique rotation in factor analysis by age cohort

Difference scores[a]	Surveys[b]					
	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Survey 6
PCS_uc–PCS_c						
1973–1978	0.85	0.8	0.57	0.39	0.33	
	(0.81–0.88)	(0.76–0.85)	(0.53–0.62)	(0.34–0.43)	(0.28–0.37)	
1946–1951	−0.003	−0.1	−0.19	−0.32	−0.5	−0.61
	(−0.04 to 0.03)	(−0.14 to −0.06)	(−0.23 to −0.15)	(−0.36 to −0.28)	(−0.54 to −0.46)	(−0.65 to −0.57)
1921–1926	−1.05	−1.2	−1.11	−1.09	−1	
	(−1.09 to −1.02)	(−1.24 to −1.17)	(−1.16 to −1.07)	(−1.14 to −1.05)	(−1.05 to −0.96)	
MCS_uc–MCS_c						
1973–1978 Cohort	−0.99	−0.96	−0.89	−0.81	−0.74	
	(−1.02 to −0.96)	(−0.99 to −0.92)	(−0.93 to −0.86)	(−0.84 to −0.77)	(−0.78 to −0.70)	
1946–1951 Cohort	−0.39	−0.21	0.06	0.23	0.33	0.57
	(−0.43 to −0.35)	(−0.25 to −0.17)	(0.02–0.10)	(0.19–0.27)	(0.29–0.37)	(0.53–0.61)
1921–1926 Cohort	1.63	1.81	2.35	2.9	3.44	
	(1.58–1.67)	(1.76–1.86)	(2.30–2.41)	(2.85–2.96)	(3.38–3.50)	

[a] PCS_{uc}–PCS_c is the difference in physical component summary scores obtained from orthogonal rotation and oblique rotations for: MCS_{uc}–MCS_c is the difference in mental component summary scores obtained from orthogonal rotation and oblique rotations

[b] Least squares mean estimates (95 % confidence intervals) from random effects model for the difference in scores by survey year

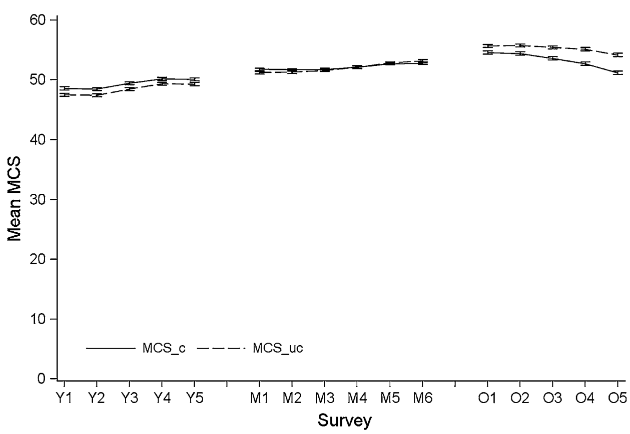

Fig. 2 Mean and 95 % confidence interval of mental component summary measures (MCS_uc derived from orthogonal rotation, MCS_c derived from oblique rotation) by survey wave and age cohort (Y1 to Y5 refers to survey 1–survey 5 for the 1973–1978 cohort, M1–M6 refers to survey 1–survey 6 for the 1946–1951 cohort, and O1–O5 refers to survey 1–survey 5 for the 1921–1926 cohort)

longitudinal study to present the trajectory of health-related quality of life for women over the life course. Using waves of SF-36, data from three cohorts of Australian women have also permitted a detailed comparison of results using uncorrelated and correlated PCS and MCS scores. The trajectory of PCS scores from the two methods produced similar trajectories that showed an accelerating decline across the lifespan, with correlated summary scores tracking below the uncorrelated scores from midlife onwards. In contrast, MCS scores increased in young women and through midlife to a peak in later life, before again declining in older women. There was also a notable divergence for MCS obtained from the two methods, with the trajectory of correlated MCS scores showing less of an increase in young and mid-age women leading a lower peak followed by a greater decline in older women than for the uncorrelated MCS score. This divergence occurs in part because by allowing for correlation between factors, or oblique rotation in the factor analysis, reduced the contribution of subscales with negative scoring coefficients.

The findings suggest that caution is needed to avoid misleading conclusions when using uncorrelated summary component scores to measure change in quality of life, especially when comparing changes in MCS among older women. The vicious circle of poor physical and mental health and poor functioning is well documented and this becomes more important with age [14, 15]. Correlated PCS and MCS scores can reflect such relationships in the underlying data, as was indicated by the correlation between the mental health subscale and the physical function and role physical subscales (Spearman correlation coefficient of 0.14 and 0.25, respectively). Mental and physical health may reflect common risk factors, however, and it may be difficult therefore to disentangle the precise relationship between the two [15].

In terms of previous longitudinal evidence on SF-36 subscales, the progressive decline in scores for the physical functioning subscale from this study is consistent with findings from the Whitehall II study [16]. Our results for the mental health subscale also show a similar trajectory to their findings, whereby they show an increase from midlife to a peak occurring when participants were aged around 70 years. The broader findings here regarding component summary scores also support previous evidence for the use of correlated PCS and MCS alongside uncorrelated PCS and PCS when measuring change in scores [8], though it should be noted that oblique factor rotation will generate health axes similar to the orthogonal setting if that provides an appropriate way of characterising the data.

The strength of this study lies in the population-based prospective study with three cohorts and multiple follow-up surveys of SF-36. One weakness is that using data only from women who responded to all surveys may have led to biased results. We would expect, however, that if other women had been included, such as those who missed surveys due to poor health, the divergence in results from the two methods would have increased.

In conclusion, our findings present contrasting trajectories for PCS and MCS until later life and that caution needs to be exercised when interpreting component summary scores as simple metrics of 'physical health' and 'mental health'. Specifically, divergence in the component summary scores obtained using correlated factors compared with uncorrelated factors, most notably where MCS declines for older women, suggests that correlated MCS and PCS provide a more reliable metric of change in health over time. Health professionals, including researchers, should include correlated PCS and MCS as part of their assessment of change in quality of life rather than relying on uncorrelated summary scores—especially for older women. Further research is needed to see if similar trajectories for summary component scores over the life course occur for men and for other populations.

Acknowledgments The Australian Longitudinal Study on Women's Health is funded by the Australian Department of Health and Ageing, and GDM is supported by the Australian National Health and Medical Research Council (APP1000986).

References

1. Contopoulos-Ioannidis, D. G., Karvouni, A., Kouri, I., & Ioannidis, J. P. (2009). Reporting and interpretation of SF-36 outcomes in randomised trials: Systematic review. *BMJ, 338*, a3006.
2. Ware, J. E., Kosinski, M., & Keller, S. D. (1994). *SF-36 Physical and Mental Health Summary Scales: A user's manual*. Boston, MA: The Health Institute, New England Medical Center.
3. Taft, C., Karlsson, J., & Sullivan, M. (2001). Do SF-36 summary component scores accurately summarize subscale scores? *Quality of Life Research, 10*(5), 395–404.
4. Anagnostopoulos, F., Niakas, D., & Tountas, Y. (2009). Comparison between exploratory factor-analytic and SEM-based approaches to constructing SF-36 summary scores. *Quality of Life Research, 18*(1), 53–63.
5. Simon, G. E., Revicki, D. A., Grothaus, L., & Vonkorff, M. (1998). SF-36 summary scores: Are physical and mental health truly distinct? *Medical Care, 36*(4), 567–572.
6. Hann, M., & Reeves, D. (2008). The SF-36 scales are not accurately summarised by independent physical and mental component scores. *Quality of Life Research, 17*(3), 413–423.
7. Wilson, D., Parsons, J., & Tucker, G. (2000). The SF-36 summary scales: Problems and solutions. *Sozial- und Praventivmedizin, 45*(6), 239–246.
8. Farivar, S. S., Cunningham, W. E., & Hays, R. D. (2007). Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.I. *Health Quality of Life Outcomes, 5*, 54.
9. Ware, J. E., & Kosinski, M. (2001). Interpreting SF-36 summary health measures: A response. *Quality of Life Research, 10*(5), 405–413.
10. Lee, C., Dobson, A. J., Brown, W. J., et al. (2005). Cohort profile: The Australian longitudinal study on women's health. *International Journal of Epidemiology, 34*(5), 987–991.
11. Mishra, G., & Schofield, M. J. (1998). Norms for the physical and mental health component summary scores of the SF-36 for young, middle-aged and older Australian women. *Quality of Life Research, 7*(3), 215–220.
12. Diggle, P. J., Heagerty, P., Liang, K.-Y., & Zeger, S. (2002). *Analysis of longitudinal data*. New York: Oxford University Press.
13. SAS Institute Inc. Base SAS® 9.3 Utilities: Reference. (2011). Cary, NC: SAS. Institute Inc.
14. Ruo, B., Baker, D. W., Thompson, J. A., Murray, P. K., Huber, G. M., & Sudano, J. J., Jr. (2008). Patients with worse mental health report more physical limitations after adjustment for physical performance. *Psychosomatic Medicine, 70*(4), 417–421.
15. Churchill, R. (2010). No health without mental health: A role for the Cochrane collaboration. *Cochrane Database Systematic Review, 8*, ED000012.
16. Jokela, M., Ferrie, J. E., Gimeno, D., et al. (2010). From midlife to early old age: Health trajectories associated with retirement. *Epidemiology, 21*(3), 284–290.