

Review

Physical activity level and health-related quality of life in the general adult population: A systematic review

Raphaël Bize^{a,b,*}, Jeffrey A. Johnson^{c,d}, Ronald C. Plotnikoff^{a,d,e}

^a Centre for Health Promotion Studies, School of Public Health, University of Alberta, Edmonton, Canada

^b Department of Ambulatory Care and Community Medicine, University of Lausanne, Switzerland

^c Institute of Health Economics, Edmonton, Canada

^d School of Public Health, University of Alberta, Edmonton, Canada

^e Faculty of Physical Education and Alberta Centre for Active Living, University of Alberta, Edmonton, Canada

Available online 21 July 2007

Abstract

Objective. Little is known regarding health-related quality of life and its relation with physical activity level in the general population. Our primary objective was to systematically review data examining this relationship.

Methods. We systematically searched MEDLINE, EMBASE, CINAHL, and PsycINFO for health-related quality of life and physical activity related keywords in titles, abstracts, or indexing fields.

Results. From 1426 retrieved references, 55 citations were judged to require further evaluation. Fourteen studies were retained for data extraction and analysis; seven were cross-sectional studies, two were cohort studies, four were randomized controlled trials and one used a combined cross sectional and longitudinal design. Thirteen different methods of physical activity assessment were used. Most health-related quality of life instruments related to the Medical Outcome Study SF-36 questionnaire. Cross-sectional studies showed a consistently positive association between self-reported physical activity and health-related quality of life. The largest cross-sectional study reported an adjusted odds ratio of “having 14 or more unhealthy days” during the previous month to be 0.40 (95% Confidence Interval 0.36–0.45) for those meeting recommended levels of physical activity compared to inactive subjects. Cohort studies and randomized controlled trials tended to show a positive effect of physical activity on health-related quality of life, but similar to the cross-sectional studies, had methodological limitations.

Conclusion. Cross-sectional data showed a consistently positive association between physical activity level and health-related quality of life. Limited evidence from randomized controlled trials and cohort studies precludes a definitive statement about the nature of this association.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Exercise; Quality of life; Health; Adult; Population

Contents

Introduction	402
Methods	402
Criteria for considering studies for this review	402
Search strategy for identification of studies	402
Data extraction	403
Results	403
Description of studies	403
Methodological quality	404
HRQL measurements	404
Physical activity measurements	404
Other methodological issues	404

* Corresponding author. Department of Ambulatory Care and Community Medicine, University of Lausanne, Bugnon 44, CH-1011 Lausanne, Switzerland. Fax: +41 21 314 48 93.

E-mail address: raphael.bize@hospvd.ch (R. Bize).

Summary of the evidence	405
Cross-sectional studies.	405
Cohort studies	409
Randomized controlled trials	409
Discussion	409
Conclusions	413
Acknowledgments	413
References	413

Introduction

Sedentary lifestyle increases the risk of developing many health conditions particularly cardiovascular diseases, hypertension, type 2 diabetes mellitus, colon and breast cancers, osteoporosis, depression, and backaches (U.S. Department of Health and Human Services, 1996; Morrow et al., 1999; Oja and Borms, 2004). Although many observational and experimental studies have demonstrated the benefit of increased physical activity (PA) in reducing chronic diseases in the general adult population (Hambrecht et al., 2004; Helmrach et al., 1994; Hoidrup et al., 2001; Hu et al., 2005; Laaksonen et al., 2005; McTiernan et al., 2003; Schnohr et al., 2003; Whelton et al., 2002), little is known however, regarding the association between health-related quality of life (HRQL), and physical activity (PA) level in this population.

Most conceptualizations of HRQL include physical, mental (including emotional dimensions), and social components (Revecki, 1989). HRQL encompasses the perceived, valued health attributes such as the sense of comfort or well-being, the ability to maintain good physical, emotional, and intellectual functions, and the ability to satisfactorily take part in social activities.

Research on HRQL in the PA domain has predominately been focused on elder populations with chronic conditions (e.g., cardiovascular diseases, arthritis, pulmonary diseases, and cancer). In their review on PA (including fitness level) and HRQL, Rejeski et al. (1996) report chronically diseased populations tend to improve their HRQL from increased PA. Courmeya and Friedenreich (1999) found physical exercise consistently had a positive effect on HRQL following cancer diagnosis in their review of 24 (mainly preliminary efficacy) studies. Although it appears no systematic review among diabetic patients has been conducted, some studies have shown a positive association between HRQL and PA level in this population (Chyun et al., 2006; Maddigan et al., 2005; Smith and McFall, 2005).

Older adults and those with chronic diseases present with specific HRQL profiles and have specific challenges and needs. For example, the elderly typically report poorer physical health than the younger population (and worse physical health than their mental health), and lack physical strength and balance (Hopman et al., 2000). Further, chronic disease groups, which consist of predominately older individuals in our populations, require tailored PA recommendations according to their disease status (Agency for Healthcare Research and Quality, 2002; Centers for Disease Control and Prevention, 2007). The relationship between PA and HRQL reported in these population

groups therefore may not be generalizable to younger, disease-free populations.

Moreover, the demonstration of a positive association between PA level and HRQL could provide healthy adults with a motivation to become more physically active, more so than the distal perspective of decreasing the risk for chronic diseases — especially when there is a general tendency for individuals to underestimate their health-related risks (Weinstein, 1989). In the field of health behavior change the need for finding alternative motivational strategies than those based on risk awareness is warranted. As advocated by Rose (1992), such public health approaches also represent a meaningful complement to strategies targeting high-risk populations.

Although the current literature examining the association between PA level and HRQL in the younger (<65 years of age), general adult population appears to be limited, a review of the studies on this population investigating this relationship is warranted. Our objective was to systematically review and analyze data examining the relationship between PA level (or aerobic fitness) and HRQL (or perceived health) among healthy subjects from the general adult population under 65 years of age.

Methods

Criteria for considering studies for this review

Eligible studies included those exploring HRQL in relation with PA among the general population (as opposed to clinical populations with specific chronic medical conditions). Eligible participants consisted of healthy adults (aged 15 or older) drawn from the general population who participated in studies related to PA. Studies specifically targeting older adults (≥ 65 years of age) were excluded. Studies examining aerobic fitness were also considered for inclusion. Aerobic fitness is a biological phenotype affected by both PA behavior and genetics, and therefore has to be interpreted as an indirect indicator of PA level. Cross-sectional studies, prospective cohort studies, and RCTs were considered for inclusion.

Evidence was appraised and synthesized according to the study design it originated from. For RCTs, eligible interventions comprised any PA promotion intervention in which the effect of PA was not confounded by the use of other interventions (co-intervention on diet for example). Eligible indicators of PA practice included self-report of PA, objective measures of PA measured through electronic devices (e.g., accelerometers, pedometers, heart rate monitors), and objective measures of physical fitness (e.g., treadmill exercise test). Eligible assessments of HRQL included any type of measurements capturing general (self-rated) health or its multidimensional components (physical, emotional, and social). Perceived health (e.g., as rated on an ordinal scale ranging from “poor” to “excellent”) was conceptualized as an acceptable surrogate for HRQL.

Search strategy for identification of studies

We accessed the following electronic databases in two discrete study phases: Phase 1 was conducted in December 2005 and examined MEDLINE (from 1966

to December 2005) and EMBASE (from 1988 to December 2005); Phase 2 was undertaken in May 2006 and investigated CINAHL (from 1982 to May 2006) and PsycINFO (from 1806 to May 2006). Search terms were based on terms used within other similar reviews (e.g., Courneya and Friedenreich, 1999; Rejeski et al., 1996) and on the 2007 MeSH Browser from the National Library of Medicine (MeSH Browser, 2007). Databases were searched for the following topic-related keywords in titles, abstracts, or indexing fields: [quality of life, or health status, or health profile\$, or HRQL, or HRQOL, or QOL]. These terms were combined with [physical activity, or exercise\$, or motor activity, or physical fitness], and with [health promotion, or primary prevention, or general population]. OVID was used to map search terms to keywords in each of the selected databases. Languages were limited to English, Spanish, French, German, and Italian. An additional search was conducted using manual references linkage, and electronic references linkage through the Web of Science Database.

Data extraction

All retrieved references were screened by two independent assessors on the basis of title and abstract for possible inclusion. References that met the inclusion criteria based on the abstract (or on the retrieved full-text article when the decision was not apparent from the abstract), were retained if they were deemed appropriate by at least one of the two assessors. References of the full text articles were used to further explore literature on this topic.

One of the authors (RB) initially extracted data (i.e., description of HRQL instruments and contexts of use, critical appraisal of validation methods, correlations between PA level and HRQL) from the studies meeting the inclusion criteria after a critical appraisal of the full-text articles. A second

author (RP) checked and edited all entries for accuracy and consistency. All disagreements concerning the information reported in the “table of included studies” were reconciled by the two assessors, in consultation with the third team member (JJ).

Results

Description of studies

Fig. 1 describes the progress through the stages of study selection. The electronic search strategy provided 1426 references. Based on titles and abstracts, 35 references were judged to require further evaluation (Andrijasevic et al., 2005; Annesi, 2002; Artazcoz et al., 2004; Ashley et al., 2001; Aurilio, 2000; Baiardi et al., 2005; Behmer et al., 1993; Berra, 2003; Daskapan et al., 2005; Fody-Urias et al., 2001; Fox, 1999; Hatziaandreu et al., 1988; Hiraoka et al., 1998; Hyland et al., 1999; Jones et al., 1998; Karvonen, 1996; Lahelma et al., 1997; Lee et al., 2006; Leonard, 2001; Malmberg et al., 2005; Marti, 1991; Marti and Martin, 2001; Miilunpalo, 2001; ODES Investigators, 1993; Penedo and Dahn, 2005; Periago, 2005; Rejeski et al., 1996; Riise et al., 2003; Rippe et al., 1998; Ross and Hayes, 1988; Shephard, 1997a,b; Vuillemin et al., 2005; Vuori, 1998; Wendel-Vos et al., 2004). A manual search and the

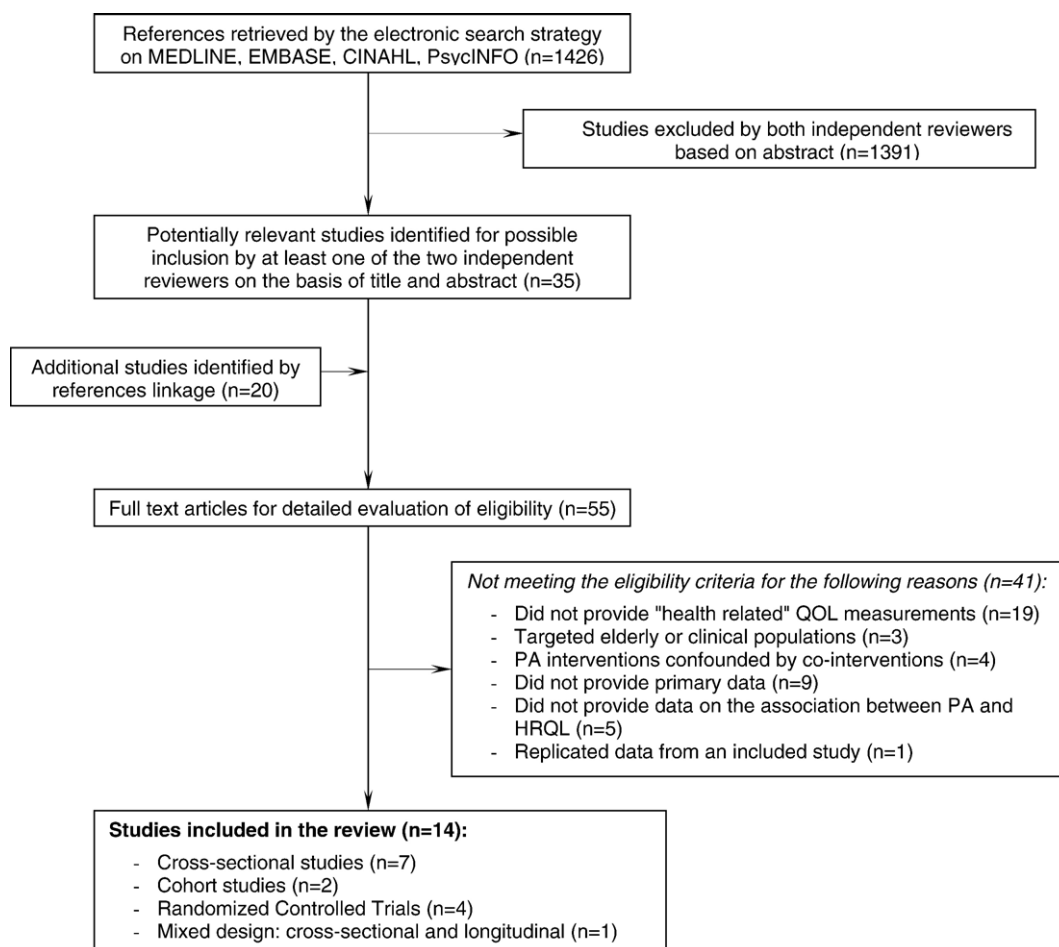


Fig. 1. Progress through the stages of study selection.

Web of Science database provided linkage with 20 additional relevant references (Atlantis et al., 2004; Berger and Motl, 2001; Brand et al., 2006; Brown et al., 1995, 2003, 2004; Hagberg and Lindholm, 2005; Herman et al., 2005; King et al., 1989; Kraemer et al., 2001; Laforge et al., 1999; Lindholm et al., 2003; Lustyk et al., 2004; Norris et al., 1990; Partonen et al., 1998; Sevic et al., 2000; Stewart et al., 1993; Tessier et al., 2007; Trine, 1999; Woodruff and Conway, 1992). The resulting 55 citations were therefore assessed as full-text articles. Of these, 41 were found not to meet eligibility criteria due to the following reasons: 19 did not provide “health related” QOL measurements (Andrijasevic et al., 2005; Annesi, 2002; Behmer et al., 1993; Berger and Motl, 2001; Brown et al., 1995; Fox, 1999; Hagberg and Lindholm, 2005; Hiraoka et al., 1998; King et al., 1989; Kraemer et al., 2001; Lustyk et al., 2004; Marti, 1991; Norris et al., 1990; Ross and Hayes, 1988; Sevic et al., 2000; Shephard, 1997a; Trine, 1999; Vuori, 1998; Woodruff and Conway, 1992); three targeted clinical or elderly populations (Fody-Urias et al., 2001; Lee et al., 2006; Penedo and Dahn, 2005); PA interventions were confounded by co-interventions on diet or other health behaviors in four (Atlantis et al., 2004; Berra, 2003; Herman et al., 2005; Rippe et al., 1998); nine did not provide primary data (Hatziaudreu et al., 1988; Jones et al., 1998; Karvonen, 1996; Leonard, 2001; Marti and Martin, 2001; Miilunpalo, 2001; Periago, 2005; Rejeski et al., 1996; Shephard, 1997b); five did not provide data on the association between physical activity and HRQL (Artazcoz et al., 2004; Baiardi et al., 2005; Hyland et al., 1999; Lahelma et al., 1997; ODES Investigators, 1993); one duplicated data from an included study (Brown et al., 2004).

The 14 remaining studies (Ashley et al., 2001; Aurilio, 2000; Brand et al., 2006; Brown et al., 2003; Daskapan et al., 2005; Laforge et al., 1999; Lindholm et al., 2003; Malmberg et al., 2005; Partonen et al., 1998; Riise et al., 2003; Stewart et al., 1993; Tessier et al., 2007; Vuillemin et al., 2005; Wendel-Vos et al., 2004) were included in our review. Seven studies were cross-sectional (Table 1a), collecting HRQL and PA data for discriminative purposes (Brown et al., 2003; Daskapan et al., 2005; Laforge et al., 1999; Lindholm et al., 2003; Riise et al., 2003; Stewart et al., 1993; Vuillemin et al., 2005), six were cohort studies or RCTs (Table 1b), exploring HRQL and PA changes over time (Ashley et al., 2001; Aurilio, 2000; Brand et al., 2006; Malmberg et al., 2005; Partonen et al., 1998; Tessier et al., 2007), and one used both a cross-sectional and a longitudinal approach (Wendel-Vos et al., 2004).

Six of the included papers reported *post hoc* data analysis from studies not primarily aimed at assessing HRQL and its relations with PA (Ashley et al., 2001; Brown et al., 2003; Malmberg et al., 2005; Tessier et al., 2007; Vuillemin et al., 2005; Wendel-Vos et al., 2004).

Most studies randomly recruited healthy adults from the general population through registries. Other recruitment settings included the workplace (Brand et al., 2006; Partonen et al., 1998), a university (Daskapan et al., 2005), and primary care practices (Ashley et al., 2001). The absence of chronic medical condition was rarely confirmed by objective data or a physician's report. The same was true for baseline PA or fitness

level, which was formally assessed during exercise tests in only two studies (Brand et al., 2006; Lindholm et al., 2003).

Methodological quality

As mentioned above, all studies faced the common methodological challenge associated with the self-reported nature of HRQL and PA level measurements. Only three studies provided contextual and psychometric arguments in favor of the validity of both their HRQL and PA instruments (Stewart et al., 1993; Tessier et al., 2007; Vuillemin et al., 2005).

HRQL measurements

Three studies did not report any aspect of the psychometric properties of their HRQL instruments (Daskapan et al., 2005; Riise et al., 2003; Wendel-Vos et al., 2004). Seven studies referred to previous validity assessments of their HRQL instruments in the literature but did not check for validity in their study context (Ashley et al., 2001; Aurilio, 2000; Brand et al., 2006; Brown et al., 2003; Lindholm et al., 2003; Malmberg et al., 2005; Partonen et al., 1998). Four studies provided internal consistency data (Cronbach's alpha coefficients) within their study sample (Laforge et al., 1999; Stewart et al., 1993; Tessier et al., 2007; Vuillemin et al., 2005). Two studies reported a ceiling effect on several scales of their HRQL instruments within their study sample (Ashley et al., 2001; Laforge et al., 1999). For more details, see Tables 1a and Tables 1b.

Physical activity measurements

Brand et al. (2006) and Lindholm et al. (2003) used objective measures of fitness level as an outcome. As mentioned above, physical fitness is an indirect indicator of PA level, as aerobic (or cardiorespiratory) fitness is a biological phenotype affected by both behavior (PA) and genetics. One randomized trial relied on the supervised nature of PA in its intervention group, but did not measure the amount of PA actually undertaken among participants (Partonen et al., 1998). All the other included studies relied on the self-reported assessment of PA level. Four studies did not provide any information about the validity of their PA instruments (Brown et al., 2003; Daskapan et al., 2005; Malmberg et al., 2005; Riise et al., 2003). One study referred to previous reliability assessment in the literature (Laforge et al., 1999). The remaining five studies used the following proxies as “gold standards” to validate their PA measurements: changes in sub-maximal aerobic capacity (Ashley et al., 2001); concurrent measurements with the “Behavioral Risk Factor Surveillance System” (Aurilio, 2000); activity logs combined with data from heart rate monitors, and activity sensors (Stewart et al., 1993); concurrent measurements with the original “Modifiable Activity Questionnaire” (Tessier et al., 2007; Vuillemin et al., 2005); or 3-day activity diaries (Wendel-Vos et al., 2004). For more details, see Tables 1a and Tables 1b.

Other methodological issues

Most of the cross-sectional and non-randomized studies provided adjusted estimates of association for potential confounding variables. For example Vuillemin et al. (2005)

considered an especially broad range of potential confounding variables including socio-demographic, lifestyle, and geographic variables. Results across these studies remained mainly unchanged after adjustment.

Based on the CONSORT statement criteria (Moher et al., 2001), many limitations in methodological quality were apparent in the four randomized controlled trials included in our review. Ashley et al. (2001) reported data from the trial conducted by Lamb et al. (2002). These investigators used an unorthodox randomization procedure with participants choosing a date among a randomly allocated series of dates, some of which corresponded to the intervention group while the others corresponded to the control group. Aurilio (2000) did not give details about the randomization procedure, nor about missing data. Brand et al. (2006) did not report concealment of allocation, had major loss to follow-up rates (41% in the intervention group, and 36% in the control group), and did not conduct an intention-to-treat analysis. Partonen et al. (1998) did not report blinding of assessors, or concealment of allocation.

Summary of the evidence

Cross-sectional data showed a consistently positive association between PA level and HRQL. Limited longitudinal evidence from RCTs and cohort studies precludes a definitive statement about the nature of this association.

Due to the high heterogeneity of measurements and study designs, we did not pool the data quantitatively. A narrative description of findings is presented instead, in the light of their methodological contexts.

Cross-sectional studies

Compared to inactive subjects (aged 45–64 years), Brown et al. (2003) found an adjusted odds ratio of “having 14 or more unhealthy days” during the previous month (physical or mental) to be 0.40 (95% CI: 0.36,0.45) for those practicing recommended PA levels (i.e., at least 30 min of moderate intensity PA per day on at least 5 days per week, or at least 20 min of vigorous intensity PA per day on at least 3 days per week). Compared to the latter category, HRQL scores were found to be poorer in subjects participating in extended periods (≥ 90 min per day) of physical activity. These findings thus suggested a bell-shaped dose–response relationship between HRQL and PA level. Analyses were conducted on data from 175,850 participants (BRFSS Survey).

Daskapan et al. (2005) reported energy expenditure was positively correlated with the physical functioning, vitality, social functioning, and mental health subscales of the SF-36 questionnaire, after adjusting for gender and body mass index. Pearson’s r ranged from 0.18 to 0.3 ($p < 0.05$). The other scales were uncorrelated.

Laforge et al. (1999) found stage of exercise adoption was associated with self-perceived HRQL, mostly in the domains of physical functioning, general health, and vitality. This study also reported an effect size, as the proportion of variance in the discriminant scores of the SF-36 that was attributable to

between-group differences in exercise stage. The effect size (ω^2) was 29% in the unadjusted model and 24% in the adjusted model. The researchers found a consistent pattern among scale scores showing an increase in SF-36 scores across the stages from ‘pre-contemplation’ to ‘maintenance’. The linear relationship was most direct for physical health, as measured by the physical functioning, bodily pain, and general health scales. A nonlinear relationship was found between exercise stage and the SF-36 scales that assess mental, emotional, and role functioning. With respect to mental health functioning, the strongest differences appeared for those who engaged in regular exercise behavior versus those who did not.

Lindholm et al. (2003) obtained a somewhat less demonstrative association between maximal exercise power and HRQL (SF-36), with a positive correlation limited to the physical functioning dimension. Maximal exercise power explained 25–35% of the variability of physical functioning scores. The researchers found a wide range of maximal exercise power values for those with high physical functioning scores, probably reflecting a ceiling effect in HRQL measurements in this population.

Riise et al. (2003) found that level of PA was positively correlated with scores on the Physical Composite Summary (PCS) and, to a slightly lesser degree, but still consistently, with Mental Composite Summary (MCS). For the PCS, the R^2 value was 0.03 for heavy and 0.01 for light PA.

In their post-test study design, Stewart et al. (1993) showed that level of participation to a one-year exercise program was positively correlated with a summary score of physical health ($p < 0.001$), but not with psychological well-being. This association was found to be independent of $\text{VO}_{2\text{max}}$.

Vuillemin et al. (2005) found a positive association between HRQL and leisure-time PA (LTPA) for both men and women, which remained statistically significant after adjustment for socio-demographic, lifestyle, and geographic variables. Regression coefficients for “vigorous LTPA” versus “inactivity” ranged from 4.83 to 8.58 (all p -values < 0.001) for the eight subscales of the SF-36. Differences in scores between inactive and vigorous exercisers were found to be at least four points in physical functioning, bodily pain, and mental health, approximately nine points in vitality and general health, and three points in physical composite score (a five points difference is generally considered as clinically meaningful (Norman et al., 2003)). Similar results were found between inactive and moderate LTPA groups, but with smaller differences in HRQL than those found between inactive and vigorous LTPA groups. HRQL scores showed a consistent positive gradient from the inactive group to the vigorous LTPA group, speaking for a dose–response relationship. The study sample showed higher HRQL and LTPA levels than the corresponding national population, which may compromise the external validity of these findings.

Wendel-Vos et al. (2004) found in both men and women, significant positive trends across quintiles of “PA of at least moderate intensity” for the following HRQL (RAND-36) scales: physical functioning, general health, and vitality dimensions, and physical health composite score (PHC). An additional significant trend was found for general mental health

Table 1a
Included studies and their characteristics: cross-sectional evidence

Study	Design	Participants	PA assessment/intervention	HRQL	Results
Brown et al. (2003) (RL)	Cross-sectional	• 175,850 Participants (aged 18 and older) from the 2001 Behavioral Risk Factor Surveillance System survey (BRFSS) • Invited to participate • Exercise level at BL: <20 min 2×/week • No medical condition	• Telephone based questionnaire • Six questions regarding practice of moderate or vigorous PA during a usual week • Three resulting PA categories: none (I), less than recommended (II), and recommended level or above (III) • No validation data reported	• Four measures of HRQL developed by the CDC (HRQOL-4): • 1) “Would you say in general your health is excellent, very good, good, fair, or poor?” • 2) Number of bad physical health in the previous 30 d • 3) Number of bad mental health in the previous 30 d • 4) Number of days without being able to conduct usual activities during the last 30 d • Last three items responses dichotomized in ≥ 14 unhealthy days or not • No validation data reported in the paper (“...construct and criterion validity have been reported elsewhere.”)	• Multivariate adjusted ORs of ≥ 14 unhealthy days (physical or mental) among level III vs level I: • 0.65 (95%CI 0.60–0.74) for those aged 18–44 years • 0.41 (95%CI 0.36–0.46) for those aged >64 years • 0.40 (95%CI 0.36–0.45) for those aged 45–64 years
Daskapan et al. (2005) (EMB)	Cross-sectional	• 180 out of 224 students (aged 17 to 25) participated in the study	• Paffenbarger physical activity questionnaire • “During the past month: how many flights of stairs climbed in each day; how many city blocks walked in each day; how many hours engaged in sports per week?” • Students’ expended energy per week extrapolated • No validation data reported	• Turkish version of the SF-36 questionnaire • No validation data reported	• After adjustment for gender and Body Mass Index, energy expenditure was positively correlated with PF, VT, SF, MH subscales of the SF-36 ($p < 0.05$) • Pearson’s r ranges from 0.18 to 0.3 • The other scales were uncorrelated
Laforge et al. (1999) (RL)	Cross-sectional	• 1387 Respondents (aged 18–75 years) to a random digit dial telephone survey of health behaviors • Response rate 75.1%	• “Do you consistently get regular exercise, that is, 3 times a week, for 20 min each time?” • Five answers corresponding each to a specific stage of exercise behavior: • 1) Yes, for ≥ 6 months (maintenance) • 2) Yes, but for <6 months (action) • 3) No, but I intend to within 1 month (preparation) • 4) No, but I intend to within 6 months (contemplation) • 5) No (precontemplation) • No validation data reported in the paper (“Validity assessed in several large population studies”)	• SF-36 • Employed here for discriminative purposes across stages of exercise adoption • All eight scales found to have good to excellent internal consistency (Cronbach’s α ranging from 0.74/GH to 0.92/PF) • Several scales (PF, RP, BP, RE, and SF) exhibited significant ceiling effect • Slightly higher scores than US population based norms (telephone effect?). Similar trends across age, but not across gender (women having rather better scores than men in this sample)	• Stages of exercise adoption are associated with self-perceived quality of life • The most strongly related areas are PF, GH, and VT • Effect size expressed as the proportion of the variance in the SF-36 scores explained by exercise stage (ω^2)=28.5% unadjusted, and 23.9% adjusted • Bonferroni correction applied with p-values having to be <0.006 (eight statistical testing, one for each scale)

Lindholm et al. (2003) (RL)	Cross-sectional	• 360 out of 430,000 randomly selected residents in Gotenburg (Sweden), aged 40–74 years • No incurable disseminated disease • Able to perform a maximal treadmill test • 345 eligible • 164 agreed to participate • 18% with significant disease (versus 40% in the 132 non-participants with available data)	• Maximal treadmill exercise test • Maximal mechanical power (W) at exercise was calculated as: • $W = \text{Body weight (kg)} \times \text{Speed (km/h)} \times [0.20 + 0.02 \times \text{Slope (\% grade)}]$	• Swedish standard version of the MOS SF-36 • “Normative data are available for the general Swedish population, allowing for matched comparisons with the study sample”	• Good concordance of sample health profiles with matched normative data • Slightly better scores however, mainly in PF, and RP • Maximal exercise power was only related to PF (explains 25–35% of the variability of PF scores) • There is however a wide range of Maximal exercise power values for those with high PF scores (probably due to a ceiling effect)
Riise et al. (2003) (EMB)	Cross-sectional	• 22,312 individuals aged 40–47 years who resided in Hordaland County (Norway), out of 33,136 invited	• Physical activity was registered as the number of hours of light and heavy leisure time physical activity (respectively) per week (self-report) • No validation data reported	• SF-12 Health Survey Questionnaire • No validation within the study population	• A high level of physical activity was correlated with high scores on the PCS and, to a slightly lesser degree, but still consistently, with MCS • For the PCS, the R^2 value was 0.03 for heavy physical activity, and 0.01 for light physical activity
Stewart et al. (1993) (RL)	Post-test design	• 194 healthy sedentary men and women aged 50–65 years participating in an RCT evaluating the effects of PA on health outcomes • Recruitment through a random digit dial telephone survey+media promotion among residents of Sunnyvale, CA	• No control group • Three intervention groups: • 1) Group-based training of high intensity (3×40 min/week $\times$ 1 year) • 2) Supervised (one initial face-to-face counseling session+telephone and mail instructions) home-based training of high intensity (3×40 min/week $\times$ 1 year) • 3) Supervised home-based training of moderate intensity (5×30 min/week $\times$ 1 year) • 4 groups defined in term of participation level across all arms of the study (validated by class attendance, activity logs, heart rate and activity monitors): 0–33%; 34–66%; 67–100%; >100%	• First 6 scales taken from MOS 20-item Short-Form. Items (6 for PF, 2 for RF, 1 for SF, 5 for GH, 1 for BP, 5 for MH) differ slightly in quality and quantity from those of the SF-36 • Last 2 scales=those developed for the longitudinal portion of the MOS (Energy/Fatigue, 5 items, and Sleep Problems Index, 6 items) • No BL HRQL data • Internal Consistency assessed in the study sample, ranging from 0.68 (PF) to 0.90 (E/F), with a tendency for lower reliability across scales than in the MOS • Oblique rotation used to summarize Physical Health (PH) and Psychological Well-being (PWB)	• No differences in HRQL according to group assignment • ANCOVA analysis showed that level of exercise participation during the year were positively correlated with PH scores at one year ($F=8.35$; $p<0.001$), but not with psychological well-being • This association was found to be independent of physiological parameters (VO max)
Vuillemin et al. (2005) (EMB+MED+PSY)	Cross-sectional	• 2333 men and 3321 women from the SUVIMAX cohort (RCT of vitamins and anti-oxydants supplements; 13017 eligible subjects aged 35–60 years [women], and 45–60 years [men], followed-up for 8 years [1994–2003]) • Only subjects with available data on HRQL and Modifiable Activity Questionnaire (MAQ) in 1998 were included	• Modifiable Activity Questionnaire (MAQ) • Initially designed to be interviewer-administered • Self-administered French version compared with original MAQ in a subsample of the cohort, with ICC of >0.80 • Assesses past 12-month PA during both leisure and work, and uses time spent daily watching TV as an indicator of sedentary behavior	• SF-36 validated in French • Internal consistency was ranging from Cronbach's α coefficients of 0.74 (RE) to 0.87 (BP) in the study sample	• When adjusting for all socio-demographic, lifestyle, and geographic variables, HRQL remained related to LTPA for men (except for RE) and women • Regression coefficients for “vigorous LTPA” versus “inactivity” ranged from 4.83 to 8.58 (all p-values < 0.001) for the eight subscales of the SF-36 • HRQL differences between inactive and vigorous were at least 4 points in PF, BP, and MH, around 9 points in VT, and GH, and 3 points in PCS • Higher proportion of subjects meeting PHR, as well as with high HRQL scores than national population

(continued on next page)

Table 1a (continued)

Study	Design	Participants	PA assessment/intervention	HRQL	Results
Vuillemin et al. (2005) (EMB+MED+PSY)			• Report of all activities ≥ 10 min/session and practiced = $10 \times$/year • 4 categories of LTPA derived: • 1) Inactivity: no LTPA • 2) Irregular activity: some LTPA but less than cat. 3 • 3) Moderate activity: ≥ 150 min/week of LTPA ≥ 3 mets but below cat. 4 • 4) Vigorous activity: ≥ 60 min/week of LTPA > 6 mets $\times = 20$ min/session		
Wendel-Vos et al. (2004) (EMB+MED+PSY)	Cross-sectional and Nested Cohort	• Subjects who participated in the Monitoring Project on Risk Factors for Chronic Diseases (MORGEN-project) • Annual random sample aged 20–59 years from municipal registries in the Netherlands • Data from year 1995 onward were used for consistency of instruments • In total 2129 (75%) participants completed measurements in 1995–1996 and 2000–2001	• Physical Activity Questionnaire designed for the “European prospective Investigation into Cancer and nutrition” (EPIC) • Questions on LTPA, Sports, and Work-related PA (type, frequency, duration) • Validated with a 3-day activity diary in a population of 126 men and women aged 20–70 years. Spearman correlation coefficients were between 0.28 and 0.81 • Quintiles of hours/week of at least moderately intense LTPA (4–6.5 METS)	• Dutch version of the standardized RAND-36 questionnaire • No validation within the study population	• Significant positive trends across quintiles of “PA of at least moderate intensity” for the following HRQL (RAND-36) scales: PF, GH, VT, and PCS scores • Differences in scores between the lowest and the highest quintiles ranged from 1.7 points (PCS in men) to 7.6 points (VT in women) • No major changes after adjustment for age, educational level, regression to the mean, and exclusion of extreme values

Abbreviation list: HRQL=health-related quality of life; PA=physical activity; CDC=centers for disease control and prevention; OR=odds ratio; RR=relative risk; BL=baseline; F-UP=follow-up; LTPA=leisure time physical activity; PF=physical functioning; RP=role-physical; BP=bodily pain; GH=general health; VT=vitality; SF=social functioning; RE=role-emotional; MH=mental health; PCS= physical composite summary; MCS=mental composite summary; MOS=medical outcomes study; PHR=public health recommendations; CIN=reference found in CINAHL; EMB=reference found in EMBASE; MED=reference found in MEDLINE; PSY=reference found in PsycINFO; RL=reference found by reference linkage.

in men, and for role limitations due to physical health problems and social functioning for women. Differences in scores between the lowest and the highest quintiles ranged from 1.7 (PHC in men) to 7.6 (vitality in women). The other scales were not correlated.

Among the five included cross-sectional studies that used the SF-36 or RAND-36 questionnaires for their HRQL assessment, five showed a statistically significant positive association of PA with the physical functioning dimension, four with the vitality dimension, three with the general health domain, and one with both the mental health and the role-physical domains. Estimations of effect size were given on the basis of different estimators: the proportion of the variance of the SF-36 scores explained by PA level ranged from 23.9% (SF-36 scale) (Laforge et al., 1999) to 35% (PF) (Lindholm et al., 2003); Pearson's correlations (r 's) ranged from 0.18 (SF) (Daskapan et al., 2005) to 0.3 (VT) (Daskapan et al., 2005); and, the differences in scores between the lowest and the highest PA categories ranged from 1.7 points (PCS in men) (Wendel-Vos et al., 2004) to 9 points (VT) (Vuillemin et al., 2005).

Cohort studies

Malmberg et al. (2005) revealed in their nested cohort study, among men only, significantly higher relative risks (1.49; 95% CI 1.12 to 1.98) of reporting poor perceived health (single-item questionnaire) after a 10-year follow-up among those subjects in the lowest LTPA categories at baseline (single-item self-assessment of LTPA).

Tessier et al. (2007) showed in their cohort study, after adjusting for socio-demographic, lifestyle and geographic variables, an increase of 1 h per week of LTPA was associated with 0.17 additional vitality points, and 0.15 mental health points for men. For women, an increase of 1 h per week of LTPA was associated with 0.39 additional vitality points, 0.28 mental health points, 0.40 social functioning points, and 0.23 points in the mental health composite score.

Wendel-Vos et al. (2004) reported in their nested cohort study, an increase in LTPA of 1 h per week over 5 years was significantly associated with an improvement in social functioning (beta coefficient of 0.38 for men, and 0.37 for women), even after adjustment for age, educational level, regression to the mean effect, and exclusion of extreme values. Using the same multiple linear regression model, no other scales were correlated to LTPA.

Randomized controlled trials

Ashley et al. (2001) in their secondary analysis of the RCT trial data from Lamb et al. (2002) found no evidence that people assigned to a walking group had HRQL improvement compared with the control group. To account for the poor compliance of subjects with the walking program, the authors further conducted a per protocol analysis, including only those participants that took part in at least one walk (33% of the initial intervention group). The lack of a relationship between walking and HRQL persisted in this secondary analysis.

Aurilio (2000) conducted a per protocol analysis of their RCT, defining adopters ($n=31$) as those who completed $\geq 80\%$ of prescribed PA (i.e., 2.4 days per week, or 72 min per week, or 3 miles per week). The author found higher vitality score increments among adopters than among non-adopters at 4 weeks (+10.48 versus +3.11, resp. $p=0.002$) and 8 weeks (+5.97 versus +0.95, resp. $p=0.029$).

Brand et al. (2006) report in their RCT that the mean score on the WHOQOL-BREF instrument for the intervention group increased by 13.7 points (versus 2.2 for the control group; $p=0.001$), after adjusting for baseline differences, and for gender. The authors stated that this difference on the WHOQOL-BREF can be interpreted as “moderate to large,” but did not state the scale range and standard deviation used as a reference.

In their RCT, Partonen et al. (1998) found few statistically significant differences in HRQL changes (RAND-36) between the active (8 weeks of supervised fitness training) and the placebo (supervised relaxation training) groups at the end of the intervention. The only dimension that improved statistically more in the active group was vitality (+13.7 versus +1.3, resp. $p=0.006$). No significant difference existed at the 4 months of follow up. Despite noticeable differences in baseline scores between groups, the authors did not mention any adjustment for a potential regression to the mean effect.

Discussion

To our knowledge, this systematic review is the first synthesis on the association between PA and HRQL in the general adult population. When reviewing cross-sectional evidence, higher PA levels were consistently associated with higher/better scores in various HRQL dimensions. The association between HRQL and PA however varies across HRQL dimensions, with higher physical functioning and vitality being more consistently associated with higher PA levels. This association remained after potential confounders were taken into account. RCTs and cohort studies were scarce and of insufficient quality to draw conclusions on the relationship between PA and HRQL changes over time in this population.

Cross-sectional studies showed moderate to strong associations between PA level and HRQL. The largest survey (Brown et al., 2003) was the only study to quantify the relationship in terms of a risk estimate, and found the odds of reporting at least 14 unhealthy days in the last month among the more active subjects was about half of the odds among those least active. A similar direction was reported in the two cross-sectional studies that validated both PA and HRQL instruments within their study samples (Stewart et al., 1993; Vuillemin et al., 2005). This constellation of findings constitutes good evidence for a positive association between PA level and HRQL.

The current review provides some support for a dose–response relationship between HRQL and PA level. Vuillemin et al. (2005) for example found a consistent positive gradient of HRQL scores across their four categories of LTPA level. Brown et al. (2003, 2004) described a bell-shaped dose–response curve with the best HRQL scores corresponding to a PA category (30–59 min of

Table 1b
Included studies and their characteristics: longitudinal evidence

Study	Design	Participants	PA assessment/intervention	HRQL	Results
Ashley et al. (2001) (CIN)	RCT	• A random sample of 2000 healthy sedentary people from a large GP practice (aged 40 to 70) were invited to participate • 260 enrolled in the study • 188 followed-up at 12 months	• Entry criteria: having undertaken less than 2 h of moderate intensity activity in the previous week • Participants randomized to receive either exercise advice or exercise advice plus the opportunity to attend a structured program of brisk walks to be completed in 1 h • Self-report of the absolute number of brisk walks completed • Validation: Spearman's correlation coefficient of 0.25 ($p=0.03$) between changes in self-reported activity and changes in sub-maximal aerobic capacity.	• SF-36 questionnaire • Substantial ceiling effects on some of the dimensions of the SF-36 in the study data (84% for "role function physical", 76% for "role function emotional", 76% for "social function"). • The trial was not powered to demonstrate a significant improvement in HRQL (secondary outcome)	• After 12 months, 61.5% of people in the "walks group" were still classified as sedentary, compared with 73.4% in the control arm ($p=0.05$) • Only 33% ($n=43$) of people in the "walks group" completed at least one walk • No significant evidence that the "walks group" showed an HRQL improvement over the control group (intention-to-treat analysis) • No significant evidence that those who actually walk among the "walks group" showed an HRQL improvement over the control group (per protocol analysis)
Aurilio (2000) (CIN)	RCT	• Convenience sample of 74 healthy sedentary women between 30 and 50 years of age • Data completed for 68 subjects (no information on lost to follow-up)	• Three arm study: I. Brochure only; II. Brochure+lecture; III. Brochure+lecture+supervised walking over a 12-week period • Self-report of PA using the "Leisure-Time Walking Physical Activity Diary" developed for this study • Self-reports diaries validated against the "Behavioral Risk Factor Surveillance System" (BRFSS) results among a random sample of subject ($n=30$) • Spearman's rho coefficients showed moderate correlation (0.4 to 0.6) between the diaries and the BRFSS measures using the following variables regarding walking: days/week; min/week; miles/week • ICC=0.77 for days walked per week	• SF-36 limited to general health, role-physical, role-emotional, vitality • No validation within the study population	• No significant differences between treatment groups for GH, RP, and RE at any of the three time periods (baseline, 4-week, and 12-week) • Significant changes in vitality scores between the baseline and 4-week scores ($p=0.0001$) • Women in group III were more likely to show this difference than others • Per protocol analysis for adopters ($n=31$), defined as those completing $\geq 80\%$ of prescribed PA (i.e., 2.4 days/week; 72 min/week; 3 miles/week): higher vitality score increments among adopters than among non-adopters at 4-week (+10.48 versus +3.11, resp. $p=0.002$) and 8-week (+5.97 versus +0.95, resp. $p=0.029$)
Brand et al. (2006) (RL)	RCT	• Employees of three business companies invited to participate by a marketing campaign (aged 36 to 45) • 250 employees assessed for eligibility • 202 randomized (reasons for not taking part unknown) • 177 agreed in testing for baseline measurements • Data completed for 109 subjects (no information on lost to follow-up)	• 26 sessions over 13 weeks provided by two medical fitness companies (endurance, strength, coordination) • PWC-150 bike ergometer test (starting from 50 Watts with 60 to 80 rounds per minute, raised by 25 Watts every 2 min, test stopped at reaching 150 heart beats per minute)	• WHOQOL-BREF • 26 questions based on four domains (physical health, psychological health, social relationships, and environment) • Responses made on a 5-point Lickert scale ranging from (1) very dissatisfied to (5) very satisfied • Cronbach's alpha ranging from 0.76 to 0.88 (report from another study, sample not specified)	• For WHOQOL-BREF global score, Group $\times$ Time interaction was significant (ANCOVA) • Mean quality of life increased for the intervention group by 13.70 score-points versus 2.2 for the control group; $p=0.001$ (adjusted for baseline differences in WHOQOL-BREF scores, and for gender)
Malmberg et al. (2005) (MED)	Nested Cohort Study	• Native cohort of 6787 residents aged 19–63 years (in 1979) in northeastern Finland (Study on health-related physical activity trends using a longitudinal design) • Sub-sample targeted to 2642 reporting good or rather good perceived health at BL (1980) • F-UP at 1, 5 and 10 years with response rates (for the native cohort) of 85.4% of the alive cohort at 10 years	• Focus on Leisure Time Physical Activity (LTPA) • Four different scales used: • 1) Single-item self-assessment of LTPA: 4 possible answers (vigorous $\geq 2 \times$/week; vigorous $1 \times$/week+some light; some/week; none); last two categories aggregated for analysis • 2) Compiled intensity–frequency of LTPA: 2 questions (frequency and duration) for each category of vigorous, moderate, and light PA (4 compiled cat.)	• "How do you rate your current state of health?": Good/Rather good/Average/Rather poor/Poor • Aggregated in 3 categories for data analysis (first two=good; average=average; last two=poor) • Test–Retest reliability (up to 42 days) found to be fair to good (weighted kappa 0.59–0.65) in a nationally representative sample and good (0.67) in a sample of middle-aged and elderly adults	• Outcome=Person-time Incidence Rate and Relative Risk of average or poor health according to LTPA level at BL • Significantly higher relative risks (RR=1.49; 95% CI 1.12 to 1.98) of reporting poor perceived health (single-item questionnaire) after a 10-year follow-up among those subjects in the lowest leisure time physical activity (LTPA) categories at baseline (single-item self-assessment of LTPA)

Malmberg et al. (2005) (MED)			• 3) LTPA energy expenditure index (kcal/week): 23 questions about specific LTPAs, then transformed in kcal/week, accounting for seasonality (3 cat.) • 4) Fitness activity (5 freq. cat.) + Sport (4 freq. cat.) + Commuting to and from work (5 freq. cat.)		
Partonen et al. (1998) (RL)	RCT	• 120 out of 15'000 indoor employees in Southern Finland (aged 22 to 57) • Healthy Subjects with or without subsyndromal seasonal affective disorder • Invited to participate by letter • No information on participation rate, reasons for non-participation, characteristics of non-participants • F-UP rate at 4 months: 78/103 • 17 early drop-outs	• Supervised Fitness training tailored to each individual according to a BL fitness evaluation, in order to achieve strenuous aerobic exercise 2–3 × 1 h/week × 8 week + Bright light (2500–4000 lx) = Group A • Idem with Normal light (400–600 lx) = Group B • Supervised relaxation training 1 × / week × 8 week as a placebo = Group C • PA assessed through compliance, checked once a week	• RAND 36-item Health Survey 1.0 • “Sensitive to changes in health among general populations” • “Translated into Finnish and back-translated into English”	• Only comparisons between group B and C are mentioned here, to avoid confounding by Bright light exposure • The only RAND 36-item score differences between BL and F-UP that were statistically favoring group B over group C at 8-week was vitality (+13.7 versus +1.3, resp. $p=0.006$). No significant difference persisted after 4 months of follow up (two independent samples t-test) • No mention of adjustment for BL differences in VT scores, or for regression to the mean • No mention of adjustment for multiple comparisons
Tessier et al. (2007) (RL)	Cohort Study	• 1644 men and 2247 women from the SUVIMAX cohort (RCT of vitamins and anti-oxydants supplements; 13017 eligible subjects aged 35–60 years [women], and 45–60 years [men], followed-up for 8 years [1994–2003]) • Only subjects free of cardiovascular disease and/or cancer before 2001, and with available data on HRQL and Modifiable Activity Questionnaire (MAQ) in 1998 and 2001 were included	• Modifiable Activity Questionnaire (MAQ) • Initially designed to be interviewer-administered • Self-administered French version compared with original MAQ in a subsample of the cohort, with ICC of >0.80 • Assesses past 12-month PA during both leisure and work, and uses time spent daily watching TV as an indicator of sedentary behavior • Report of all activities ≥ 10 min/session and practiced ≥ 10 x/year • 4 categories derived for LTPA: • 1) inactivity: no LTPA • 2) irregular activity: some LTPA but < than cat. 3 • 3) moderate activity: ≥ 150 min/week of LTPA ≥ 3 mets but below cat. 4 • 4) vigorous activity: ≥ 60 min/week of LTPA > 6 mets × = 20 min/session	• SF-36 validated in French • Internal consistency was ranging from Cronbach's α coefficients of 0.74 (RE) to 0.87 (BP) in the study sample	• When adjusting for all socio-demographic, lifestyle, and geographic variables, increased LTPA over 3 years was associated with improved scores in the following HRQL dimensions: PF, MH, VT for both sexes, and SF for women only • For men, an increase of 1 h/week of LTPA was associated in multiple variables analyses with 0.17 additional VT points, and 0.15 MH points • For women, an increase of 1 h/week of LTPA was associated in multiple variables analyses with 0.39 additional VT points, 0.28 MH points, 0.40 SF points, and 0.23 MCS points
Wendel-Vos et al. (2004) (EMB + MED + PSY)	Cross-sectional and Nested Cohort	• Subjects who participated in the Monitoring Project on Risk Factors for Chronic Diseases (MORGEN-project) • Annual random sample aged 20–59 years from municipal registries in the Netherlands. • Data from year 1995 onward were used for consistency of instruments • In total 2129 (75%) participants completed measurements in 1995–1996 and 2000–2001	• Physical Activity Questionnaire designed for the “European prospective Investigation into Cancer and nutrition” (EPIC) • Questions on LTPA, Sports, and Work-related PA (type, frequency, duration) • Validated with a 3-day activity diary in a population of 126 men and women aged 20–70 years. Spearman correlation coefficients were between 0.28 and 0.81 • Quintiles of hours/week of at least moderately intense LTPA (4–6.5 METS)	• Dutch version of the standardized RAND-36 questionnaire • No validation within the study population	• An increase in LTPA of 1 h per week over five years was significantly associated with an improvement in social functioning (beta coefficient of 0.38 for men, and 0.37 for women) • No major changes after adjustment for age, educational level, and regression to the mean

Abbreviation list: HRQL=health-related quality of life; PA=physical activity; CDC=centers for disease control and prevention; OR=odds ratio; RR=relative risk; BL=baseline; F-UP=follow-up; LTPA=leisure time physical activity; PF=physical functioning; RP=role-physical; BP=bodily pain; GH=general health; VT=vitality; SF=social functioning; RE=role-emotional; MH=mental health; PCS=physical composite summary; MCS=mental composite summary; MOS=medical outcomes study; PHR=public health recommendations; CIN=reference found in CINAHL; EMB=reference found in EMBASE; MED=reference found in MEDLINE; PSY=reference found in PsycINFO; RL=reference found by reference linkage.

moderate intensity PA per day) compatible with public health recommendations. Wendel-Vos et al. (2004) found significant positive trends across quintiles of “PA of at least moderate intensity” for the following HRQL (RAND-36) scales: physical functioning, general health, and vitality dimensions, and physical health composite score (PHC).

Support for the association between PA and HRQL, based on cohort studies and RCT data, is on the contrary, weaker. Three cohort studies (Malmberg et al., 2005; Tessier et al., 2007; Wendel-Vos et al., 2004), and three RCTs (Aurilio, 2000; Brand et al., 2006; Partonen et al., 1998) found differences in HRQL scores according to group allocation or to PA level. However, inadequate randomization procedures, lack of instrument validation, high rates of attrition, as well as inappropriate data analyses (not conducted on an intention-to-treat basis), call for caution in interpreting these positive findings.

Further, we cannot conclude whether the cross-sectional association between PA level and HRQL can be explained via a causal pathway (i.e., higher HRQL leading to higher level of PA, or vice versa, or mutual influence, or presence of an external factor leading to higher PA levels and HRQL scores). Other hypotheses could explain, at least in part, this association. For example, there may be cognitive constructs (e.g., self-efficacy) that may confound the association between physical activity level and HRQL. Some authors have argued that this type of construct would correspond to an intermediate (mediating) step on the causal pathway between PA and HRQL rather than a true confounding factor (Rejeski et al., 1996). Further, self-reported PA and HRQL may have conceptual overlap (especially in the physical functioning domains of HRQL), inflating the actual relationship between these two constructs. Moreover, the possibility of correlated misclassification cannot be excluded as both HRQL and PA measures rely on self-report.

Another concern is the ceiling effect in HRQL assessments which may have threatened the discriminative and evaluative properties of these instruments, especially among the relatively younger and healthier individuals from the included studies. Few of the selected studies assessed validity of their HRQL instrument in their study sample. Those who explored the distribution of their HRQL data tended to report a significant ceiling effect (Ashley et al., 2001; Laforge et al., 1999). As most HRQL instruments used in these studies derived from the MOS, they were initially conceived mainly for individuals with a chronic medical condition (Ware and Sherbourne, 1992; Ware et al., 1993). SF-36 and RAND-36 have since been widely used in the general population with relatively good psychometric properties (Hemingway et al., 1997). Scores of healthy adults tend however to aggregate in the highest categories. Together with this ceiling effect, high data variance may also require larger sample sizes to detect the effects of a physical activity intervention on HRQL, than for other outcomes on which trials are usually powered (e.g., an increase in fitness or PA level). This may explain why reviews focusing on clinical populations (Courneya and Friedenreich, 1999; Rejeski et al., 1996) have shown more consistent and stronger associations between physical activity and HRQL. If the benefit of regular physical activity among healthy adults is to preserve (rather than

improve) HRQL, data from long-term prospective cohort studies are needed to observe this phenomenon.

When regression methods ignore the presence of a ceiling effect or censoring in the HRQL measurements, the results tend to produce biased coefficient estimates (Grootendorst, 2000). Some authors therefore suggest employing the Tobit regression model (Austin, 2002; Austin et al., 2000) which includes a specific calculation of the maximum likelihood estimation. To perform better than ordinary least squares regression techniques, this method requires that the conditional distribution of the HRQL measure has uniform variance, which often does not occur.

Each included study in our review used a distinct method to assess PA level of which most relied on self-report. The validity of self-reported PA can be supported by comparison with data obtained in the same population while using an electro-mechanical device like a motion sensor, possibly combined with a heart rate monitor (Plasqui and Westerterp, 2005; Strath et al., 2001, 2002, 2003). Only one of the selected studies (Stewart et al., 1993) provided such type of validation for their PA self-report questionnaire. Two of the included studies (Brand et al., 2006; Lindholm et al., 2003) used physical fitness as an outcome. However, as noted earlier, aerobic (or cardio-respiratory) fitness has to be interpreted as an indirect indicator of PA level, as it is a biological phenotype affected by both PA behavior and genetics, and the subsequent interaction between the two. HRQL was assessed using a self-rated health item in two studies (Brown et al., 2003; Malmberg et al., 2005) e.g., “How do you rate your current state of health?” with response options of: *Good/Rather good/Average/Rather poor/Poor*. We included studies using this measure of HRQL, as it has been used to assess health-related quality of life in population surveys (U.S. Department of Health and Human Services, 2000). This construct has also demonstrated good reliability, and has shown ability to predict subsequent mortality (Scott et al., 1997; Wannamethee, 1991; Miiunpalo et al., 1997). Studies that did not measure general health or its multidimensional components (physical, emotional, and social) were excluded following consideration of the items in each of the measurement instruments.

Among the retrieved papers, only five were published during the 10-year period between 1993 and 2002, whereas nine were published in the following 5 years. This trend reflects the growing interest for physical activity related research, beyond the classical spectrum of chronic disease prevention and management. Replicating the methodology of this systematic review in older adults, and potentially in children/youth, would assist in providing a more comprehensive and perhaps contrasted picture of the relationships between PA and HRQL across our populations.

There is also debate on how to define meaningful differences on the SF-36 scores (especially when applied to healthy populations) as 3-, 5-, and 10-point changes have all been suggested as being clinically significant for clinical populations (Samsa et al., 1999). The observed changes in our reviewed studies ranged from +0.15 point per weekly hour of LTPA over 3 years for mental health in men (Tessier et al., 2007) to +13.7 in vitality (versus +1.3 in the control group) after an 8-week PA intervention (Partonen et al., 1998). Given this wide range, and

the limited studies, more work is also clearly needed to establish clinically significant PA-related HRQL cut-points for the general population.

Further studies are also needed to better understand the shape of the dose–response curve characterizing the relationship between PA and HRQL, and whether certain modes of PA show stronger positive association with HRQL than others. From a public health perspective, it would be particularly useful to know whether there are thresholds below and above certain PA levels that have no benefits on HRQL.

As some of our included studies were identified only in the manual reference linkage protocol, the electronic research terms that were used to characterize our target population may have decreased the sensitivity of our electronic search strategy. Despite the potential study limitations and the moderate outcomes, this review will provide the literature a relative comparison of the association between PA level and HRQL of the general population that can be examined against elderly and younger/clinical populations. It is hoped the study's results will guide the role and importance of PA promotion strategies in population-based contexts.

Conclusions

There appears to be a consistent association of higher HRQL scores with higher PA levels among apparently healthy adults in cross-sectional studies. These data will hopefully encourage researchers to improve the methods of this field, particularly by applying 'state of the art' PA measures and developing more responsive HRQL instruments. Further evidence regarding HRQL benefits of regular PA among the general population would support the role and importance of PA beyond chronic disease prevention and treatment contexts.

Acknowledgments

The authors would like to thank Cindy Forbes for her independent screening of the citations and her contribution to the selection of relevant studies, and Kylie Hugo, Erin Buhr and Nandini Karunamuni for their helpful comments on a previous draft of this manuscript. There are no known potential conflicts of interest. External sources of support are as follows: Raphaël Bize holds salary support from The Swiss National Science Foundation; Jeffrey Johnson holds a Canada Research Chair in Diabetes Health Outcomes and a Health Scholar Award from the Alberta Heritage Foundation for Medical Research; Ronald C. Plotnikoff holds a Health Scholar Award from the Alberta Heritage Foundation for Medical Research and a New Investigator Salary Award from the Canadian Institutes for Health Research.

References

Agency for Healthcare Research and Quality and the Centers for Disease Control, 2002. Physical Activity and Older Americans: Benefits and Strategies. URL <http://www.ahrq.gov/ppip/activity.htm>. (accessed May 7, 2007).

- Andrijasevic, M., Pausic, J., Bavcevic, T., Ciliga, D., 2005. Participation in leisure activities and self-perception of health in the students of the University of Split. *Kinesiology* 37 (1), 21–31.
- Annesi, J.J., 2002. Relation of rated fatigue and changes in energy after exercise and over 14 weeks in previously sedentary women exercisers. *Percept. Mot. Skills* 95 (3), 719–727.
- Artazcoz, L., Borrell, C., Benach, J., Cortes, I., Rohlfs, I., 2004. Women, family demands and health: the importance of employment status and socioeconomic position. *Soc. Sci. Med.* 59 (2), 263–274.
- Ashley, A., Lloyd, A., Lamb, S., Bartlett, H., 2001. Is health-related quality of life a suitable outcome measure for evaluating health promotion programmes? *NT Res.* 6 (3), 671–681.
- Atlantis, E., Chow, C.M., Kirby, A., Singh, M.F., 2004. An effective exercise-based intervention for improving mental health and quality of life measures: a randomized controlled trial. *Prev. Med.* 39 (2), 424–434.
- Aurilio, L.A.M., 2000. Promotion of adoption and adherence to regular leisure-time walking behavior in healthy mid-life women: a randomized controlled study. PhD Thesis, University of Pittsburgh, USA.
- Austin, P.C., 2002. A comparison of methods for analyzing health-related quality-of-life measures. *Value Health* 5 (4), 329–337.
- Austin, P.C., Escobar, M., Kopec, J.A., 2000. The use of the Tobit model for analyzing measures of health status. *Qual. Life Res.* 9 (8), 901–910.
- Baiardi, F., Puglia, M.G., Valpiani, G., et al., 2005. Relationship between body weight, quality of life and cardiovascular risk factors: a general population based Italian experience. *Eat. Weight Disord.* 10 (1), 19–24.
- Behmer, C.L., Hoffman, M.A., Freeman, V.S., 1993. A multifocal comparison of health status and physical readiness test results of Navy personnel. *Mil. Med.* 158 (4), 248–252.
- Berger, B.G., Motl, R., 2001. Physical activity and quality of life. In: Singer, R.N., Hausenblas, H.A., Janelle, C. (Eds.), *Handbook of Sport Psychology*. Wiley, New York, pp. 636–671.
- Berra, K., 2003. The effect of lifestyle interventions on quality of life and patient satisfaction with health and health care. *J. Cardiovasc. Nurs.* 18 (4), 319–325.
- Brand, R., Schlicht, W., Grossmann, K., Duhnsen, R., 2006. Effect of a physical exercise intervention on employees' perceptions of quality of life: a randomized controlled trial. *Soz.-Praventivmed.* 51, 14–23.
- Brown, D.R., Wang, Y., Ward, A., et al., 1995. Chronic psychological effects of exercise and exercise plus cognitive strategies. *Med. Sci. Sports Exerc.* 27 (5), 765–775.
- Brown, D.W., Balluz, L.S., Heath, G.W., et al., 2003. Associations between recommended levels of physical activity and health-related quality of life. Findings from the 2001 Behavioral Risk Factor Surveillance System (BRFSS) survey. *Prev. Med.* 37 (5), 520–528.
- Brown, D.W., Brown, D.R., Heath, G.W., et al., 2004. Associations between physical activity dose and health-related quality of life. *Med. Sci. Sports Exerc.* 36 (5), 890–896.
- Centers for Disease Control and Prevention, 2007. Promoting active lifestyles among older adults. Nutrition and Physical Activity. CDC, National Center for Chronic Disease Prevention and Health Promotion, Atlanta. URL: www.cdc.gov/nccdphp/dnpa/physical/recommendations/older_adults.htm (accessed May 7, 2007).
- Chyun, D.A., Melkus, G.D., Katten, D.M., et al., 2006. The association of psychological factors, physical activity, neuropathy, and quality of life in type 2 diabetes. *Biol. Res. Nurs.* 7 (4), 279–288.
- Courneya, K.S., Friedenreich, C.M., 1999. Physical exercise and quality of life following cancer diagnosis: a literature review. *Annals Behav. Med.* 21 (2), 171–179.
- Daskapan, A., Tuzun, E.H., Eker, L., 2005. Relationship between physical activity level and health related quality of life among university students. *Saudi Med. J.* 26 (6), 1026–1028.
- Fody-Urias, B.M., Fillit, H., Hill, J., 2001. The effect of a fitness program on health status and health care consumption in Medicare MCOs. *Manag. Care Interface* 14 (9), 58–64.
- Fox, K.R., 1999. The influence of physical activity on mental well-being. *Public Health Nutr.* 2 (3A), 411–418.
- Grootendorst, P., 2000. Censoring in statistical models of health status: what happens when one can do better than '1'. *Qual. Life Res.* 9 (8), 911–914.

- Hagberg, L.A., Lindholm, L., 2005. Is promotion of physical activity a wise use of societal resources? Issues of cost-effectiveness and equity in health. *Scand. J. Med. Sci. Sports* 15 (5), 304–312.
- Hambrecht, R., Walther, C., Mobius-Winkler, S., et al., 2004. Percutaneous coronary angioplasty compared with exercise training in patients with stable coronary artery disease: a randomized trial. *Circulation* 109 (11), 1371–1378.
- Hatzianandreu, E.I., Koplan, J.P., Weinstein, M.C., Caspersen, C.J., Warner, K.E., 1988. A cost-effectiveness analysis of exercise as a health promotion activity. *Am. J. Public Health* 78 (11), 1417–1421.
- Helmrich, S.P., Ragland, D.R., Paffenbarger Jr., R.S., 1994. Prevention of non-insulin-dependent diabetes mellitus with physical activity. *Med. Sci. Sports Exerc.* 26 (7), 824–830.
- Hemingway, H., Stafford, M., Stansfeld, S., Shipley, M., Marmot, M., 1997. Is the SF-36 a valid measure of change in population health? Results from the Whitehall II Study. *BMJ* 315 (7118), 1273–1279.
- Herman, W.H., Hoerger, T.J., Brandle, M., et al., 2005. The cost-effectiveness of lifestyle modification or metformin in preventing type 2 diabetes in adults with impaired glucose tolerance. *Ann. Intern. Med.* 142 (5), 323–332.
- Hiraoka, J., Ojima, T., Nakamura, Y., Yanagawa, H., 1998. A comparative epidemiological study of the effects of regular exercise on health level. *J. Epidemiol.* 8 (1), 15–23.
- Hoidrup, S., Sorensen, T.I., Stroger, U., Lauritzen, J.B., Schroll, M., Gronbaek, M., 2001. Leisure-time physical activity levels and changes in relation to risk of hip fracture in men and women. *Am. J. Epidemiol.* 154 (1), 60–68.
- Hopman, W.M., Towheed, T., Anastassiades, T., et al., 2000. Canadian normative data for the SF-36 health survey. Canadian Multicentre Osteoporosis Study Research Group. *Can. Med. Assoc. J.* 163 (3), 265–271.
- Hu, G., Jousilahti, P., Barengo, N.C., Qiao, Q., Lakka, T.A., Tuomilehto, J., 2005. Physical activity, cardiovascular risk factors, and mortality among Finnish adults with diabetes. *Diabetes Care* 28 (4), 799–805.
- Hyland, M.E., Sodergren, S.C., Singh, S.J., 1999. Variety of activity: relationship with health status, demographic variables and global quality of life. *Psychol. Health Med.* 4 (3), 241–254.
- Jones, W.K., Franks, A.L., Manson, J.E., Hoffman-Goetz, L., Otis, C.L., 1998. The health effects and benefits of physical activity in women. *Women's Health Issues* 8 (2), 74–97.
- Karvonen, M.J., 1996. Physical activity for a healthy life. *Res. Q. Exerc. Sport* 67 (2), 213–215.
- King, A.C., Taylor, C.B., Haskell, W.L., DeBusk, R.F., 1989. Influence of regular aerobic exercise on psychological health: a randomized, controlled trial of healthy middle-aged adults. *Health Psychol.* 8 (3), 305–324.
- Kraemer, W.J., Keuning, M., Ratamess, N.A., et al., 2001. Resistance training combined with bench-step aerobics enhances women's health profile. *Med. Sci. Sports Exerc.* 33 (2), 259–269.
- Laaksonen, D.E., Lindstrom, J., Lakka, T.A., et al., 2005. Physical activity in the prevention of type 2 diabetes: the Finnish diabetes prevention study. *Diabetes* 54 (1), 158–165.
- Laforge, R.G., Rossi, J.S., Prochaska, J.O., Velicer, W.F., Levesque, D.A., McHorney, C.A., 1999. Stage of regular exercise and health-related quality of life. *Prev. Med.* 28 (4), 349–360.
- Lahelma, E., Rahkonen, O., Berg, M., et al., 1997. Changes in health status and health behavior among Finnish adults 1978–1993. *Scand. J. Work, Environ. & Health* 23 (Suppl 3), 85–90.
- Lamb, S.E., Bartlett, H.P., Ashley, A., Bird, W., 2002. Can lay-led walking programmes increase physical activity in middle aged adults? A randomised controlled trial. *J. Epidemiol. Community Health* 56 (4), 246–252.
- Lee, P., Chang, W., Liou, T., Chang, P., 2006. Stage of exercise and health-related quality of life among overweight and obese adults. *J. Adv. Nurs.* 53 (3), 295–303.
- Leonard, W.R., 2001. Assessing the influence of physical activity on health and fitness. *Am. J. Human Biol.* 13 (2), 159–161.
- Lindholm, E., Brevinge, H., Bergh, C.H., Korner, U., Lundholm, K., 2003. Relationships between self-reported health related quality of life and measures of standardized exercise capacity and metabolic efficiency in a middle-aged and aged healthy population. *Qual. Life Res.* 12 (5), 575–582.
- Lustyk, M.K., Widman, L., Paschane, A.A., Olson, K.C., 2004. Physical activity and quality of life: assessing the influence of activity frequency, intensity, volume, and motives. *Behav. Med.* 30 (3), 124–131.
- Maddigan, S.L., Majumdar, S.R., Johnson, J.A., 2005. Understanding the complex associations between patient-provider relationships, self-care behaviours, and health-related quality of life in type 2 diabetes: a structural equation modeling approach. *Qual. Life Res.* 14 (6), 1489–1500.
- Malmberg, J., Miilunpalo, S., Pasanen, M., Vuori, I., Oja, P., 2005. Characteristics of leisure time physical activity associated with risk of decline in perceived health—a 10-year follow-up of middle-aged and elderly men and women. *Prev. Med.* 41 (1), 141–150.
- Marti, B., 1991. Health effects of recreational running in women. Some epidemiological and preventive aspects. *Sports Medicine*, vol. 11 (1), pp. 20–51.
- Marti, B., Martin, B.W., 2001. Athletic training or daily movement activities for optimal health and quality of life. *Ther. Umsch.* 58 (4), 189–195.
- MeSH Browser from the National Library of Medicine, 2007. URL: <http://www.nlm.nih.gov/mesh/2007/MBrowser.html> (accessed September 17, 2005).
- McTiernan, A., Kooperberg, C., White, E., et al., 2003. Recreational physical activity and the risk of breast cancer in postmenopausal women: the Women's Health Initiative Cohort Study. *JAMA* 290 (10), 1331–1336.
- Miilunpalo, S., 2001. Evidence and theory based promotion of health-enhancing physical activity. *Public Health Nutr.* 4 (2B), 725–728.
- Miilunpalo, S., Vuori, I., Oja, P., Pasanen, M., Urponen, H., 1997. Self-rated health status as a health measure: the predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *J. Clin. Epidemiol.* 50 (5), 517–528.
- Moher, D., Schulz, K.F., Altman, D.G., and the Consort Group, 2001. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Medical Research Methodology* 1 (2). The electronic version of this article can be found online at: <http://www.biomedcentral.com/1471-2288/1/2>.
- Morrow Jr., J.R., Jackson, A.W., Bazzarre, T.L., Milne, D., Blair, S.N., 1999. A one-year follow-up to physical activity and health. A report of the Surgeon General. *Am. J. Prev. Med.* 17 (1), 24–30.
- Norman, G.R., Sloan, J.A., Wyrrich, K.W., 2003. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med. Care* 41 (5), 582–592.
- Norris, R., Carroll, D., Cochrane, R., 1990. The effects of aerobic and anaerobic training on fitness, blood pressure, and psychological stress and well-being. *J. Psychosom. Res.* 34 (4), 367–375.
- ODES Investigators, 1993. The Oslo Diet and Exercise Study (ODES): design and objectives. *Control. Clin. Trials* 14 (3), 229–243.
- Health enhancing physical activity. In: Oja, P., Borms, J. (Eds.), *Perspectives — The Multidisciplinary Series of Physical Education and Sport Science*, vol. 6. Meyer & Meyer Sport (UK) Ltd, Oxford.
- Partonen, T., Leppamaki, S., Hurme, J., Lonqvist, J., 1998. Randomized trial of physical exercise alone or combined with bright light on mood and health-related quality of life. *Psychol. Med.* 28 (6), 1359–1364.
- Penedo, F.J., Dahn, J.R., 2005. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Curr. Opin. Psychiatry* 18 (2), 189–193.
- Periago, M.R., 2005. Longevity and the quality of life: a new challenge for public health in the Americas. *Pan Am. J. Public Health* 17 (5), 295–298.
- Plasqui, G., Westerterp, K.R., 2005. Accelerometry and heart rate as a measure of physical fitness: proof of concept. *Med. Sci. Sports Exerc.* 37 (5), 872–876.
- Rejeski, W.J., Brawley, L.R., Shumaker, S.A., 1996. Physical activity and health-related quality of life. *Exerc. Sport Sci. Rev.* 24, 71–108.
- Revicki, D.A., 1989. Health-related quality of life in the evaluation of medical therapy for chronic illness. *J. Fam. Pract.* 29 (4), 377–380.
- Riise, T., Moen, B.E., Nortvedt, M.W., 2003. Occupation, lifestyle factors and health-related quality of life: The Hordaland Health Study. *J. Occup. Environ. Med.* 45 (3), 324–332.
- Rippe, J.M., Price, J.M., Hess, S.A., et al., 1998. Improved psychological well-being, quality of life, and health practices in moderately overweight women participating in a 12-week structured weight loss program. *Obes. Res.* 6 (3), 208–218.
- Rose, G.A., 1992. *The Strategy of Preventive Medicine*. Oxford University Press, New York, USA.

- Ross, C.E., Hayes, D., 1988. Exercise and psychologic well-being in the community. *Am. J. Epidemiol.* 127 (4), 762–771.
- Samsa, G., Edelman, D., Rothman, M.L., Williams, G.R., Lipscomb, J., Matchar, D., 1999. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. *Pharmacoeconomics* 15, 141–155.
- Schnohr, P., Scharling, H., Jensen, J.S., 2003. Changes in leisure-time physical activity and risk of death: an observational study of 7000 men and women. *Am. J. Epidemiol.* 158 (7), 639–644.
- Scott, W.K., Macera, C.A., Cornman, C.B., Sharpe, P.A., 1997. Functional health status as a predictor of mortality in men and women over 65. *J. Clin. Epidemiol.* 50 (3), 291–296.
- Sevick, M.A., Dunn, A.L., Morrow, M.S., Marcus, B.H., Chen, G.J., Blair, S. N., 2000. Cost-effectiveness of lifestyle and structured exercise interventions in sedentary adults: results of project ACTIVE. *Am. J. Prev. Med.* 19 (1), 1–8.
- Shephard, R.J., 1997a. What is the optimal type of physical activity to enhance health? *Br. J. Sports Med.* 31 (4), 277–284.
- Shephard, R.J., 1997b. Exercise and relaxation in health promotion. *Sports Med.* 23 (4), 211–217.
- Smith, D.W., McFall, S.L., 2005. The relationship of diet and exercise for weight control and the quality of life gap associated with diabetes. *J. Psychosom. Res.* 59 (6), 385–392.
- Stewart, A.L., King, A.C., Haskell, W.L., 1993. Endurance exercise and health-related quality of life in 50–65-year-old adults. *Gerontologist* 33 (6), 782–789.
- Strath, S.J., Bassett Jr., D.R., Swartz, A.M., Thompson, D.L., 2001. Simultaneous heart rate-motion sensor technique to estimate energy expenditure. *Med. Sci. Sports Exerc.* 33 (12), 2118–2123.
- Strath, S.J., Bassett Jr., D.R., Thompson, D.L., Swartz, A.M., 2002. Validity of the simultaneous heart rate-motion sensor technique for measuring energy expenditure. *Med. Sci. Sports Exerc.* 34 (5), 888–894.
- Strath, S.J., Bassett Jr., D.R., Ham, S.A., Swartz, A.M., 2003. Assessment of physical activity by telephone interview versus objective monitoring. *Med. Sci. Sports Exerc.* 35 (12), 2112–2118.
- Tessier, S., Vuillemin, A., Bertrais, S., et al., 2007. Association between leisure-time physical activity and health-related quality of life changes over time. *Prev. Med.* 44, 202–208.
- Trine, M.R., 1999. Physical activity and quality of life. In: Rippe, J.M., Malden, M.A. (Eds.), *Lifestyle Medicine*. Blackwell Science Inc., London, pp. 989–997.
- U.S. Department of Health and Human Services, 1996. *Physical Activity and Health: A Report of the Surgeon General*. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta, USA.
- U.S. Department of Health and Human Services, 2000. *Healthy People 2010* (conference ed, in 2 vols). U.S. Department of Health and Human Services, Washington, DC. Available at <http://www.health.gov/healthypeople>.
- Vuillemin, A., Boini, S., Bertrais, S., et al., 2005. Leisure time physical activity and health-related quality of life. *Prev. Med.* 41 (2), 562–569.
- Vuori, I., 1998. Does physical activity enhance health? *Patient Educ. Couns.* 33 (Suppl 1), S95–S103.
- Wannamethee, G., Shaper, A.G., 1991. Self-assessment of health status and mortality in middle-aged British men. *Int. J. Epidemiol.* 20 (1), 239–245.
- Ware Jr., J.E., Sherbourne, C.D., 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med. Care* 30 (6), 473–483.
- Ware, J.E., Snow, K.K., Kosinski, M., Gandek, B. (Eds.), 1993. *SF-36 Health Survey: Manual and Interpretation Guide*. The Health Institute. New England Medical Center, Boston, USA.
- Weinstein, N.D., 1989. Optimistic biases about personal risks. *Science* 246 (4935), 1232–1233.
- Wendel-Vos, G.C., Schuit, A.J., Tijhuis, M.A., Kromhout, D., 2004. Leisure time physical activity and health-related quality of life: cross-sectional and longitudinal associations. *Qual. Life Res.* 13 (3), 667–677.
- Whelton, S.P., Chin, A., Xin, X., He, J., 2002. Effect of aerobic exercise on blood pressure: a meta-analysis of randomized, controlled trials. *Ann. Intern. Med.* 136 (7), 493–503.
- Woodruff, S.L., Conway, T.L., 1992. A longitudinal assessment of the impact of health/fitness status and health behavior on perceived quality of life. *Percept. Mot. Skills* 75 (1), 3–14.