World Journal of Preventive Medicine, 2015, Vol. 3, No. 2, 17-23
Available online at http://pubs.sciepub.com/jpm/3/2/1

DOI:10.12691/jpm-3-2-1

Reliability of the Short-Form Health Survey (SF-36) in Physical Activity Research Using Meta-Analysis

Peter D. Hart[1,*], Minsoo Kang[2]

[1]Health Promotion Program, Montana State University - Northern, & Health Demographics, Havre, MT 59501, USA
[2]Kinesmetrics Laboratory, Middle Tennessee State University, Murfreesboro, TN, USA
*Corresponding author: peter.hart@msun.edu

Received March 29, 2015; Revised April 04, 2015; Accepted April 07, 2015

Abstract Health-related quality of life (HRQOL) is a health outcome that has seen a growing interest in physical activity research. The Short-Form Health Survey (SF-36) is the most widely used HRQOL instrument in physical activity research. The purpose of this study was to perform a systematic review and meta-analysis of reliability coefficients on the SF-36 HRQOL assessment applied in physical activity research so as to assess the generalizability of HRQOL measurement reliability. A separate meta-analysis was performed for each HRQOL domain (physical and mental). A total of 87 effect sizes were analyzed in this study, 44 for the physical health domain and 43 for the mental health domain. The effect sizes were strong and significantly different from zero for both physical health (ES = .90 [95% CI: .88, .92], $p< .001$) and mental health (ES = .90 [95% CI: .89, .91], $p< .001$) domains. These findings show that both HRQOL scales of the SF-36 assessment are reliable across a wide variety of physical activity studies.

Keywords: *health-related quality of life, sf-36, reliability, generalizability, physical activity, measurement*

Cite This Article: Peter D. Hart, "Reliability of the Short-Form Health Survey (SF-36) in Physical Activity Research Using Meta-Analysis." *World Journal of Preventive Medicine*, vol. 3, no. 2 (2015): 17-23. doi: 10.12691/jpm-3-2-1.

1. Introduction

Health-related quality of life (HRQOL) is a health outcome that has seen a growing interest in physical activity research. The United States Department of Health and Human Services (USDHHS), in its publication entitled Healthy People 2020, includes the enhancement of quality of life as a major public health service goal [1]. HRQOL is a broad concept that includes both subjective and objective indicators of a people's lives that affect their physical and/or mental health status [2]. A more comprehensive description of HRQOL includes several dimensions including but not limited to physical functioning, psychological well-being, social functioning, role functioning, and health perceptions [3].

HRQOL has become a standard outcome measure in both intervention and observational studies [2]. A search on Pubmed.gov showed only 1,410 publication hits from the years 1980 to 1999 using the key word term "health-related quality of life." The same search resulted in 15,180 hits from the years 2000 to 2011. HRQOL is also being included in research studies alongside the more conventional and objective measures of health status [4]. Given the overwhelming interest in HRQOL as an outcome measure in physical activity research, there is a strong need for a better understanding of the measurement properties of HRQOL assessments commonly used in physical activity research.

The Short-Form Health Survey (SF-36) is the most widely used HRQOL instruments in physical activity research. The SF-36 was developed from the Medical Outcomes Study (MOS) conducted by RAND [5]. The SF-36 is a multi-dimensional scale consisting of 36 items, 8 health-related dimensions, and two domains. The dimensions include: 1) vitality, 2) physical functioning, 3) bodily pain, 4) general health, 5) physical role functioning, 6) emotional role functioning, 7) social role functioning, and 8) mental health. The physical domain consists of the physical functioning, bodily pain, general health, and physical role functioning dimensions and the mental domain consists of the vitality, emotional role functioning, social role functioning, and mental health dimensions [6]. The SF-36 is intended to measure HRQOL in adults and is easily self-administered in physical activity research. The SF-12 and SF-8 are both shorter versions of the original form with 12 and 8 items, respectively. Both shorter versions maintain the measurement of the 8 dimensions as well as the two domain-specific summary scores [7].

A Reliability index can be thought of as a percentage of variance in a set of scores that is true or from non-error factors [8]. Reliability also refers to the consistency or stability of test scores. This can be consistency of scores over time, consistency of scores across items (internal consistency) or consistency of scores across raters [9]. Of specific interest to this study are the internal consistency and test-retest reliability coefficients. Internal consistency refers to the extent to which items in an assessment tool are inter-correlated [10]. If such an inter-correlation exists, the items of the scale are said to measure a unidimensional construct. Test-retest reliability measures the stability of measurements over repeated trials [8]. Since reliability is not a property of an assessment tool itself but rather a property of the assessment scores, it should be common practice for researchers to report reliability of their scores at hand [11]. Despite the inability to infer an assessment's reliability from a single study to all future studies, it is still possible to strengthen the reliability generalization using meta-analysis [12].

To date, there are no studies which assess the reliability generalization of the SF-36 HRQOL assessment in physical activity research using meta-analysis. The purpose of this study was to perform a systematic review and meta-analysis of reliability coefficients on the SF-36 HRQOL assessment applied in physical activity research so as to assess the generalizability of HRQOL measurement reliability. A secondary purpose of the study was to examine potential moderators which may account for extra variance associated with the HRQOL measurements, such as age, gender, and even HRQOL assessment tool. This approach will provide valuable evidence as to the strengths and weaknesses of HRQOL scales in terms of measurement reliability.

2. Materials and Methods

2.1. Search Strategy

Using PubMed.gov, a systematic search was conducted to identify all physical activity studies using the SF-36 assessment or one of its variants (i.e., SF-12 or SF-8). The following search terms were used: ("physical activity" OR exercise) AND ("short-form 36" OR "sf-36" OR "MOS 36" OR "rand 36"). The same search strategy was used for identifying studies using the SF-12 and SF-8 forms. After all potential studies were identified; inspection of titles and abstracts ensued to identify exclusion criteria.

Studies were excluded if 1) they were published in a non-English language, 2) they were non-research based, 3) they were from a non-peer-reviewed journal, 4) HRQOL was not measured in the study using the SF-36 or one of its variants, 5) the study was a validation study where reliability measures were repeatedly computed due to changes made to the scale, 6) the HRQOL assessment tool was modified from its original form prior to its administration, 7) the study was not physical activity oriented, or 8) the study population involved youth under the age of 18 years.

By searching the remaining articles surviving exclusion criteria, reported measures of reliability were retrieved. Reliability measures were defined as Cronbach's alpha, Kuder-Richardson coefficient, Split-half reliability coefficient, or test-retest coefficients. When possible, reliability measures were extracted for each of the SF-36 domains (physical and mental). If a situation occurred where an article published reliability in the form of a range (i.e., .78 to .88), the midpoint of that range was used as the effect size. If a study reported a lower-bound value (i.e., > .80), the lower-bound was used as the effect size. Precaution was taken to not include redundant coefficients from multiple publications on the same study population, by matching author names and study characteristics.

2.2. Data Extraction and Coding

Data were extracted from each identified study using the following strategy. Reliability estimates along with the number of items for each assessment tool were recorded. If reliability estimates were reported separately for different populations (i.e., gender) then each population was considered a separate sample. Common demographic variables were included such as gender (male, female, both) and age (mean). Sample size was recorded for each study for its impact on reliability estimates. Other possible moderator data collected was disease status (diseased, non-diseased), reliability type (internal consistency, test-retest) and study design (randomized controlled trial [RCT], other). Also, the SF-36 form used for the HRQOL measure was recorded (SF-36, SF-12, SF-8).

2.3. Data analysis

A separate meta-analysis was performed for each HRQOL domain (physical and mental). Since internal consistency reliability coefficients represent a proportion of variance not accounted for by error, the square root of each coefficient was first taken. Fisher's r to Z transformation was then performed as recommended for correlation coefficients [13]. Meta-analytic mean effect sizes and 95% confidence intervals were computed for each reliability study using a random effects model. A mean effect size of .70 or greater was considered acceptable evidence of reliability [14]. A moderator analysis was performed to account for the extra variance by examining different study characteristics. The Q test of homogeneity was used to support the moderator analysis in determining if the variance in reliability coefficients was significantly different from zero. The Comprehensive Meta-Analysis Version 2.0 software was used for all meta-analyses [15].

3. Results

3.1. Search Results

A total of 1,358 articles were retrieved using the search strategy. After reviewing titles and/or abstracts, 1,080 were eliminated using exclusion criteria. After full review of the remaining 278 articles, 20 physical domain and 21 mental domain reliability coefficients were retrieved. After contacting study authors from the remaining articles, an additional 24 and 22 coefficients were retrieved for the physical and mental scales, respectively. Table 1 contains 87 effect sizes used in this study, 44 for the physical health domain and 43 for the mental health domain [16-62].

3.2. Physical Health Domain

Figure 1 shows the effect sizes and 95% confidence intervals for the physical health domain across each study. The weighted mean effect size from the physical health domain was computed using a random effects model. The effect size was strong and significantly different from zero, ES = .90 (95% CI: .88, .92), $p < .001$. Therefore, the physical health scales of the SF-36 assessment are reliable across a wide variety of physical activity studies. The test of homogeneity was significant, $Q = 2057.21$, $df = 43$, p

< .001, I2 = 97.91. This indicates that some effect sizes may come from different populations and a moderator analysis should be performed. Table 2 shows results of the moderator analysis for the physical health studies. Results showed that studies of various gender (female or both sexes) did not account for significant variance in effect sizes, Q = 1.77, df = 1, p = .184. Studies using different designs (RCT or other types) did not account for significant variance in effect sizes, Q = 0.02, df = 1, p = .884. Also, studies using participants of different health status (diseased or non-diseased) did not account for significant variance in effect sizes, Q = 0.77, df = 1, p = .379. However, studies using different forms of the SF-36 (SF-36, SF-12 or SF-8) did account for significant variance in effect sizes, Q = 12.82, df = 2, p = .002. As well, studies publishing different reliability coefficients (alpha or retest) accounted for significant variance in effect sizes, Q = 8.00, df = 1, p = .005. Finally, mean age of study participants was a significant predictor (b = -.003) of effect size, Q = 9.59, df = 1, p = .002. Figure 3 displays the linear relationship between age and the transformed effect sizes.

Table 1. Characteristics of the physical activity and SF-36 studies

Study	Np	Form	Type	Design	Disease	Age	Gender	ES_{phys}	ES_{ment}
Fisher 2004	582	SF-12	alpha	RCT	no	74	both	.911	.900
Li 2003	40	SF-12	alpha	RCT	no	72.6	both	.889	.883
Barnason 2009	55	SF-36	alpha	RCT	yes	71.6	both	.949	.949
Basen-Engquist 2006	60	SF-36	alpha	RCT	yes	55	female	.927	.927
Blacklock 2007	341	SF-36	alpha	other	no	55	both	NA	.917
Ciairano 2010	22	SF-36	alpha	RCT	no	80.6	both	.775	.812
Conroy 2007	497	SF-36	alpha	other	no	56.9	female	.943	.933
Cook 2011	539	SF-36	alpha	other	no	19.8	both	.900	.943
Coups 2009	175	SF-36	alpha	other	yes	68.7	both	.922	.922
Griffith 2009	126	SF-36	alpha	RCT	yes	60.2	both	.920	NA
Isaacs 2007	943	SF-36	retest	RCT	yes	57	both	.680	.800
Johansen 2001	38	SF-36	alpha	other	yes	52	both	.794	NA
Li 2009	599	SF-36	alpha	other	no	NG	both	.837	.837
Li 2010	187	SF-36	alpha	other	yes	59	both	.910	.910
McGrath 2011	143	SF-36	alpha	other	no	NG	both	.854	.854
Smith 2009	736	SF-36	alpha	other	yes	57	female	NA	.866
Tessier 2007	3891	SF-36	alpha	other	no	51.8	both	.906	.906
Tung 2010	70	SF-36	alpha	other	yes	69.6	both	.926	.896
Turner 2009	2995	SF-36	alpha	other	yes	55.3	both	.975	.927
Volkmann 2010	242	SF-36	alpha	other	yes	43	female	.959	NA
Yates 2003	64	SF-36	alpha	other	yes	NG	both	NA	.917
Zimmerman 2007	54	SF-36	alpha	RCT	yes	72.1	female	.967	.910
Tamari 2011	42	SF-36	alpha	other	no	75.7	both	.849	.900
Feldman 2009	50	SF-36	alpha	other	yes	51	both	.913	.931
Callaghan 2011	38	SF-12	retest	RCT	yes	53.5	female	.832	.832
Bennett 2007	56	SF-36	alpha	RCT	yes	58	both	.950	.961
Buessing 2009	388	SF-12	alpha	other	yes	60	female	.909	.922
Buys 2011	103	SF-36	alpha	other	yes	28.7	both	.905	.905
Vancampfort 2011	297	SF-12	alpha	other	yes	61.4	female	.831	.856
Rombaut 2010	64	SF-36	alpha	other	yes	38	female	.960	.920
Midtgaard 2006	55	SF-36	alpha	other	yes	42	both	.864	.879
Mueller 2009	57	SF-36	alpha	other	yes	NG	NG	.830	.869
Huisinga 2011	26	SF-36	alpha	other	yes	45.5	both	.922	.922
Aoyagi 2010	183	SF-36	alpha	other	no	73	both	.872	.922
Fassett 2009	120	SF-36	alpha	other	yes	60	both	.909	.884
Kerse 2008	193	SF-36	alpha	RCT	yes	NG	both	.949	.834
Logsdon 2009	37	SF-36	alpha	other	no	81.9	both	.854	.872
Ogilvie 2008	1322	SF-8	alpha	other	no	48	both	.930	.940
Van Uffelen 2007	152	SF-36	alpha	RCT	yes	75	both	.819	.793
Krousel-Wood 2008	76	SF-36	alpha	RCT	yes	56.6	both	.921	NA
Stroud 2009	121	SF-36	alpha	other	yes	50	both	.863	.871
Lund 2011	86	SF-36	alpha	RCT	yes	77	both	.906	.906
Lawton 2008	1089	SF-36	alpha	RCT	no	58.9	female	.869	.892
Poulin 2007	110	SF-36	alpha	other	yes	35.8	both	.900	.872
Park 2008	14	SF-36	alpha	other	no	72	both	.893	.815
Brandes 2011	53	SF-36	alpha	other	yes	65.8	both	.835	.870
MacMillan 2011	41	SF-36	alpha	RCT	no	NG	both	.815	.898

Note. NA indicates not applicable. NG indicates not given.

Table 2. Effect size by moderator for the SF-36 physical health studies

Moderator	*N*	*M ES*	*LL*	*UL*	*Q*	*df*	*p*
Gender					1.77	1	.184
Female	9	.92	.86	.98			
Both	34	.90	.86	.94			
Design					0.02	1	.884
RCT	16	.90	.86	.93			
Other	28	.90	.87	.93			
Disease					0.77	1	.379
No	14	.89	.87	.91			
Yes	29	.91	.87	.94			
Form					12.82	2	.002
SF-36	38	.90	.87	.93			
SF-12	5	.88	.84	.91			
SF-8	1	.93	.92	.94			
Reliability					8.00	1	.005
Alpha	41	.91	.88	.92			
Retest	2	.75	.55	.87			
Age	39	-.003	-.004	-.001	9.59	1	.002

Note. M ES (95% CI) for age is a regression coefficient.

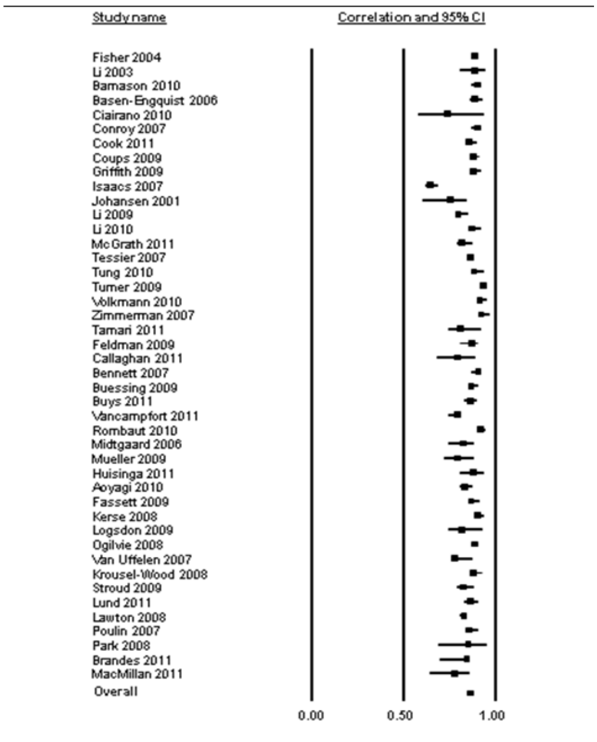


Figure 1. Forest plot of physical domain reliability effect sizes and 95% confidence intervals

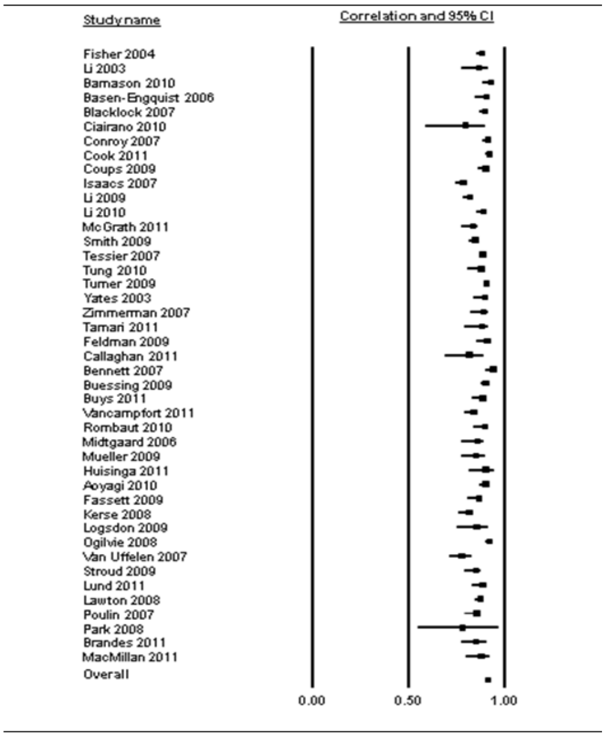
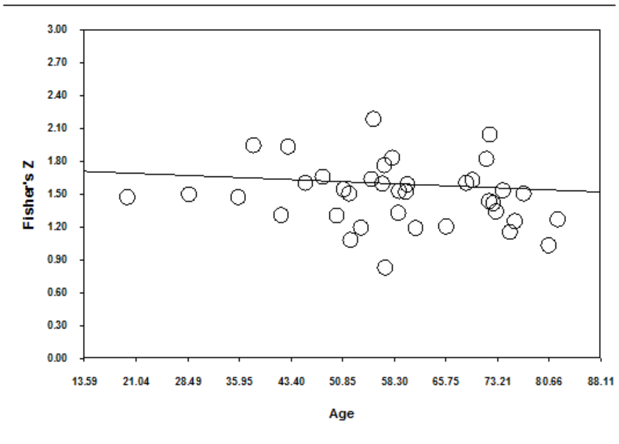


Figure 2. Forest plot of mental domain reliability effect sizes and 95% confidence intervals



Note. Age is a significant (p = .002) moderator of physical domain effect sizes. Effect sizes were transformed to Fisher's Z.

Figure 3. Weighted least squares regression of physical domain effect size regretted on age

3.3. Mental Health Domain

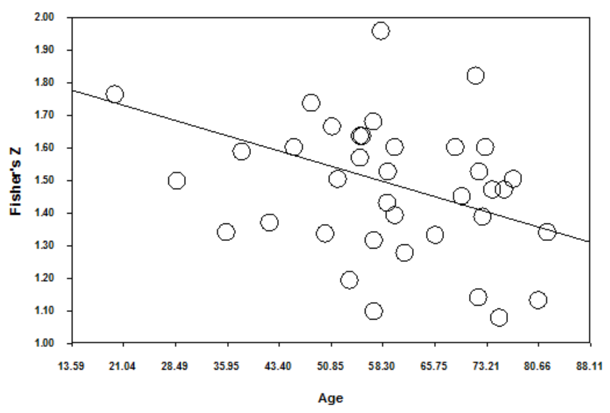
Figure 2 shows the effect sizes and 95% confidence intervals for the mental health domain across each study. The weighted mean effect size from the random effects model was strong and significantly different from zero, ES = .90 (95% CI: .89, .91), p < .001. Therefore, the mental health scales of the SF-36 assessment are also reliable across a various physical activity studies. The test of homogeneity was also significant for the mental health domain, Q = 529.90, df = 42, p < .001, I2 = 92.07. This also shows that a moderator analysis should be performed. Table 3 shows results of the moderator analysis for the mental health studies. Results showed that studies of various gender (female or both sexes) did not account for significant variance in effect sizes, Q = 0.03, df = 1, p

= .864. Studies using different designs (RCT or other types) did not account for significant variance in effect sizes, Q = 1.03, df = 1, p = .310. Also, studies using participants of different health status (diseased or non-diseased) did not account for significant variance in effect sizes, Q = 0.34, df = 1, p = .557. However, studies using different forms of the SF-36 (SF-36, SF-12 or SF-8) did account for significant variance in effect sizes, Q = 46.89, df = 2, p < .001. As well, studies publishing different reliability coefficients (alpha or retest) accounted for significant variance in effect sizes, Q = 77.39, df = 1, p < .001. Finally, mean age of study participants was also a significant predictor (b = -.006) of effect size, Q = 58.04, df = 1, p < .001. Figure 4 displays the relationship between age and the transformed effect sizes.

Table 3. Effect size by moderator for the SF-36 mental health studies

Moderator	*N*	*M ES*	*LL*	*UL*	*Q*	*df*	*p*
Gender					0.03	1	.864
Female	9	.90	.88	.92			
Both	33	.90	.88	.91			
Design					1.03	1	.310
RCT	14	.89	.86	.91			
Other	29	.90	.89	.91			
Disease					0.34	1	.557
No	15	.90	.89	.92			
Yes	27	.90	.88	.91			
Form					46.89	2	< .001
SF-36	37	.90	.88	.91			
SF-12	5	.89	.85	.92			
SF-8	1	.94	.93	.95			
Reliability					77.39	1	< .001
Alpha	41	.90	.89	.91			
Retest	2	.80	.78	.82			
Age	37	-.006	-.008	-.005	58.04	1	< .001

Note. M ES (95% CI) for age is a regression coefficient.



Note. Age is a significant (p < .001) moderator of mental domain effect sizes. Effect sizes were transformed to Fisher's Z.

Figure 4. Weighted least squares regression of mental domain effect size regressed on age

4. Discussion

The primary purpose of this study was to perform a systematic review and meta-analysis of reliability coefficients from the physical and mental health scales of the SF-36 HRQOL assessment applied to physical activity research. The results indicated that both SF-36 HRQOL domains are highly reliable in physical activity research. The secondary purpose of this study was to examine potential moderators which may account for extra variance associated with the HRQOL measurements. Results of the moderator analysis showed that the different forms, the different types of reliability, and mean age significantly contributed to effect size variability.

The SF-36 was more reliable when used in its compact SF-8 form compared to its SF-12 or SF-36 form. This result is surprising since, generally, scales longer in length provide greater reliability coefficients [8]. It is suggested that future research investigate the reliability comparison between the SF-8 and SF-12/SF-36 forms in physical activity research. The implications for finding the shorter form more reliable than the two longer forms are great when many physical activity studies are already burdened with many stages of intervention procedures. Caution should be taken; however, when interpreting the moderator analysis by form, since only one study was included using the SF-8 form.

More reliable scales were also seen when internal consistency reliability was the coefficient as compared to a test-retest coefficient. A lower reliability coefficient for test-retest situations compared to internal reliability seems appropriate since measures of stability are heavily affected by participant's carry-over effects such as memory, mood, or actual changes in the traits as well as length of time between administrations [8]. Finally, study mean age was indirectly related to reliability for both scales, with the relationship stronger for the mental scales than for the physical scales. This observation could be explained by the fact that older research participants have more problems interpreting the SF-36 items as well as more problems responding and completing all items of the assessment [63]. The reasoning behind the stronger age effect in the mental health scales; however, is not as clear. It is possible that perceptions of mental health is not as clear to older individuals as is their perceptions of physical health.

Results of the moderator analyses showed that studies with females only did not have significantly different reliability of SF-36 scales from studies using both males and females. Results showed that studies utilizing RCTs did not have significantly different reliability from studies using other designs. Also, studies using the SF-36 to assess subjects with chronic disease did not have significantly different reliability from studies using non-diseased subjects. The non-significant moderators provide evidence of the robustness of the SF-36 reliability across gender, design, and disease.

This study has some limitations. First, the results of this study may have been affected by publication bias. The fact that the reliability coefficients were published in peer-reviewed journals makes it possible that studies reporting lower reliabilities for the SF-36 were not accepted by reviewers and therefore biased the results toward higher reliability estimates. It is also possible that researchers using the SF-36 in physical activity research, who computed low reliability measures, removed the SF-36 as an outcome variable from their study. Another limitation in this study is the fact that there were no male only studies used in the moderator analysis. Therefore, we are uncertain whether the reliability of the SF-36 is different for males compared to females or different from both male and female studies. The reason reliability coefficients from male only studies were not analyzed in this study was not because there were no physical activity articles using the SF-36 assessment on males only; but rather, there were no studies of this kind reporting the reliability. This bias should be recognized and future research is needed to assess the specific reliability generalization to the male population in physical activity research.

A strength of this study is that it was limited to the collection of reliability coefficients coming from publications on physical activity research. This type of focus allows for a more precise generalization of reliability. To know that the SF-36 assessment was on average reliable across a broad area of research topics would indeed be valuable; it is perhaps more valuable for the physical activity researchers to know that it performs reliably across a large span of physical activity research areas. Another strength of this study was its ability to separately analyze reliability in the physical and mental health domains. Many physical activity researchers use both physical and mental health components in their research [17], while others use only the physical component [25] or the mental component [36]. From a measurement perspective, it is more valuable to research the reliability of a set of scores in the way in which they are used in research. Finding that the SF-36 on average produces scores that are reliable for both domains in physical activity research provides more psychometric information than if reliability was only generalized to the SF-36 as a whole.

5. Conclusion

In conclusion, the evidence from this study overwhelmingly supports the use of the SF-36 HRQOL assessment in physical activity research. Both physical and mental health domains maintain very strong reliability across studies of different gender, different research design, and different diseases states. The SF-8 may provide slightly greater reliability compared to its longer counterparts. Measures of internal consistency provide greater reliability coefficients as compared to measures of stability. Finally, reliability may slightly decrease among the aging population.

References

[1] United States Department of Health & Human Services. (2011). *Healthy people 2020* from: http://www.healthypeople.gov/2020/topicsobjectives2020/default.aspx

[2] Centers for Disease Control and Prevention. *Measuring healthy days: Population assessment of health-related quality of life*. Centers for Disease Control and Prevention, Atlanta, Georgia 2000.

[3] Hennessy, C. H., Moriarty, D. G., Zack, M. M., Scherr, P. A., & Brackbill, R. (1994). Measuring health-related quality of life for public health surveillance. *Public Health Reports, 109* (5), 665-672.

[4] Dominick, K. L., Ahern, F. M., Gold, C. H., &Heller, D. A. (2002). Relationship of health-related quality to health care utilization and mortality among older adults. *Aging Clinical and Experimental Research, 14* (6), 499-508.

[5] Ware, J.E., Sherbourne, C.D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care, 30* (6), 473.

[6] Ware, J.E. (2004). SF-36 Health Survey Update. Retrieved August 15, 2011, from http://www.sf-36.org/announcements/Updated_SF36_bookChapter_Sept04.pdf

[7] QualityMetric. (2011). SF Health Surveys. Retrieved August 15, 2011, from: http://www.qualitymetric.com/WhatWeDo/GenericHealthSurveys/tabid/184/Default.aspx

[8] Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, Calif.: Brooks/Cole Pub. Co.

[9] Kline, T. (2005). Psychological testing : a practical approach to design and evaluation. Thousand Oaks, Calif.: Sage Publications.

[10] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16* (3), 297-334.

[11] Ragan, B. G., & Kang, M. (2005). Reliability: Current Issues and Concerns. *Athletic Therapy Today, 10* (6), 30-33.

[12] Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement, 58* (1), 6-20.

[13] Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, Calif.: Sage Publications.

[14] Gliner, J. A., Morgan, G. A., & Harmon, R. J. (2001). Measurement reliability. *Journal of the American Academy of Child and Adolescent Psychiatry, 40* (4), 486-488.

[15] Comprehensive Meta-Analysis. [Computer software]. Englewood, NJ: Biostat; 2006.

[16] Fisher, K. J., & Li, F. Z. (2004). A community-based walking trial to improve neighborhood quality of life in older adults: A multilevel analysis. *Annals of Behavioral Medicine, 28* (3), 186-194.

[17] Li, F. Z., Harmer, P., Wilson, N. L., & Fisher, K. J. (2003). Health benefits of cobblestone-mat walking: Preliminary findings. *Journal of Aging and Physical Activity, 11* (4), 487-501.

[18] Barnason, S., Zimmerman, L., Schulz, P., & Tu, C. (2009). Influence of an early recovery telehealth intervention on physical activity and functioning after coronary artery bypass surgery among older adults with high disease burden. *Heart & Lung, 38* (6), 459-468.

[19] Basen-Engquist, K., Taylor, C. L. C., Rosenblum, C., Smith, M. A., Shinn, E. H., Greisinger, A., Rivera, E. (2006). Randomized pilot test of a lifestyle physical activity intervention for breast cancer survivors. *Patient Education and Counseling, 64* (1-3), 225-234.

[20] Blacklock, R. E., Rhodes, R. E., & Brown, S. G. (2007). Relationship between regular walking, physical activity, and health-related quality of life. *Journal of physical activity & health, 4* (2), 138-152.

[21] Ciairano, S., Liubicich, M. E., & Rabaglietti, E. (2010). The effects of a physical activity programme on the psychological wellbeing of older people in a residential care facility: an experimental study. *Ageing & Society, 30*, 609-626.

[22] Conroy, M. B., Simkin-Silverman, L. R., Pettee, K. K., Hess, R., Kuller, L. H., & Kriska, A. M. (2007). Lapses and psychosocial factors related to physical activity in early postmenopause. *Medicine and Science in Sports and Exercise, 39* (10), 1858-1866.

[23] Cook, B., Hausenblas, H., Tuccitto, D., & Giacobbi, P. R., Jr. (2011). Eating Disorders and Exercise: A Structural Equation Modelling Analysis of a Conceptual Model. *European Eating Disorders Review, 19* (3), 216-225.

[24] Coups, E. J., Park, B. J., Feinstein, M. B., Steingart, R. M., Egleston, B. L., Wilson, D. J., & Ostroff, J. S. (2009). Physical activity among lung cancer survivors: changes across the cancer trajectory and associations with quality of life. *Cancer Epidemiology Biomarkers &Prevention, 18* (2), 664-672.

[25] Griffith, K., Wenzel, J., Shang, J., Thompson, C., Stewart, K., & Mock, V. (2009). Impact of a Walking Intervention on Cardiorespiratory Fitness, Self-Reported Physical Function, and Pain in Patients Undergoing Treatment for Solid Tumors. *Cancer, 115* (20), 4874-4884.

[26] Isaacs, A. J., Critchley, J. A., Tai, S. S., Buckingham, K., Westley, D., Harridge, S. D. R., Gottlieb, J. M. (2007). Exercise Evaluation Randomised Trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. *Health Technology Assessment, 11* (10), 1-165.

[27] Johansen, K. L., Painter, P., Kent-Braun, J. A., Ng, A. V., Carey, S., Da Silva, M., & Chertow, G. M. (2001). Validation of questionnaires to estimate physical activity and functioning in end-stage renal disease. *Kidney International, 59* (3), 1121-1127.

[28] Li, G. S. F., Lu, F. J. H., & Wang, A. H. H. (2009). Exploring the relationships of physical activity, emotional intelligence and health in Taiwan college students. *Journal of Exercise Science & Fitness, 7* (1), 55-63.

[29] Li, M., Li, L., & Fan, X. (2010). Patients having haemodialysis: physical activity and associated factors. *Journal of Advanced Nursing, 66* (6), 1338-1345.

[30] McGrath, J. A., O'Malley, M., & Hendrix, T. J. (2011). Group exercise mode and health-related quality of life among healthy adults. *Journal of Advanced Nursing, 67* (3), 491-500.

[31] Smith, A. W., Alfano, C. M., Reeve, B. B., Irwin, M. L., Bernstein, L., Baumgartner, K., Ballard-Barbash, R. (2009). Race/ethnicity, physical activity, and quality of life in breast cancer survivors. *Cancer Epidemiology & Biomarkers Prevention, 18* (2), 656-663.

[32] Tessier, S., Vuillemin, A., Bertrais, S., Boini, S., Le Bihan, E., Oppert, J. M., Briancon, S. (2007). Association between leisure-time physical activity and health-related quality of life changes over time. *Preventive Medicine, 44* (3), 202-208.

[33] Tung, H.-H., Chen, Y.-C., Wei, J., Liu, C.-Y., Chang, C.-Y., & Wang, T.-J. (2010). Leisure physical activity and quality of life after coronary artery bypass graft surgery for patients with metabolic syndrome in Taiwan. *Heart & Lung, 39* (5), 410-420.

[34] Turner, A. P., Kivlahan, D. R., & Haselkorn, J. K. (2009). Exercise and Quality of Life Among People With Multiple Sclerosis: Looking Beyond Physical Functioning to Mental Health and Participation in Life. *Archives of Physical Medicine and Rehabilitation, 90* (3), 420-428.

[35] Volkmann, E. R., Grossman, J. M., Sahakian, L. J., Skaggs, B. J., FitzGerald, J., Ragavendra, N., McMahon, M. (2010). Low Physical Activity Is Associated With Proinflammatory High-Density Lipoprotein and Increased Subclinical Atherosclerosis in Women With Systemic Lupus Erythematosus. *Arthritis Care & Research, 62* (2), 258-265.

[36] Yates, B. C., Price-Fowlkes, T., & Agrawal, S. (2003). Barriers and facilitators of self-reported physical activity in cardiac patients. *Research in Nursing & Health, 26* (6), 459-469.

[37] Zimmerman, L., Barnason, S., Schulz, P., Nieveen, J., Miller, C., Hertzog, M., Tu, C. (2007). The effects of a symptom management intervention on symptom evaluation, physical functioning, and physical activity for women after coronary artery bypass surgery. *Journal of Cardiovascular Nursing, 22* (6), 493-500.

[38] Tamari, K. (2011). Self-reported Home Exercise and Younger Age Predict Improved Health-related Quality of Life among Community-Dwelling Elderly Participants in a Three-month Group Exercise Class. *Journal of Physiological Anthropology, 30* (2), 77-85.

[39] Feldman, L. S., Kaneva, P., Demyttenaere, S., Carli, F., Fried, G. M., & Mayo, N. E. (2009). Validation of a physical activity questionnaire (CHAMPS) as an indicator of postoperative recovery after laparoscopic cholecystectomy. *Surgery, 146* (1), 31-39.

[40] Callaghan, P., Khalil, E., Morres, I., & Carter, T. (2011). Pragmatic randomised controlled trial of preferred intensity exercise in women living with depression. *BMC Public Health, 12* (11), 465.

[41] Bennett, J. A., Lyons, K. S., Winters-Stone, K., Nail, L. M., & Scherer, J. (2007). Motivational interviewing to increase physical activity in long-term cancer survivors - A randomized controlled trial. *Nursing Research, 56* (1), 18-27.

[42] Buessing, A., Fischer, J., Fouladbakhsh, J. M., Ostermann, T., & Matthiessen, P. F. (2009). Perceived and expected effects of physical activity in female cancer survivors. *European Journal of Integrative Medicine, 1* (1), 9-17.

[43] Buys, R., Budts, W., Delecluse, C., & Vanhees, L. (2011). Exercise Capacity, Physical Activity, and Obesity in Adults With Repaired Aortic Coarctation. *Journal of Cardiovascular Nursing*. In Print.

[44] Vancampfort, D., Probst, M., Scheewe, T., Maurissen, K., Sweers, K., Knapen, J., & De Heft, M. (2011). Lack of physical activity during leisure time contributes to an impaired health related quality of life in patients with schizophrenia. *Schizophrenia Research, 129* (2-3), 122-127.

[45] Rombaut, L., Malfait, F., Cools, A., De Paepe, A., & Calders, P. (2010). Musculoskeletal complaints, physical activity and health-related quality of life among patients with the Ehlers-Danlos syndrome hypermobility type. *Disability and Rehabilitation, 32* (16), 1339-1345.

[46] Midtgaard, J., Rorth, M., Stelter, R., & Adamsen, L. (2006). The group matters: an explorative study of group cohesion and quality of life in cancer patients participating in physical exercise intervention during treatment. *European Journal of Cancer Care, 15* (1), 25-33.

[47] Mueller, J., Christov, F., Schreiber, C., Hess, J., & Hager, A. (2009). Exercise capacity, quality of life, and daily activity in the long-term follow-up of patients with univentricular heart and total cavopulmonary connection. *European Heart Journal, 30* (23), 2915-2920.

[48] Huisinga, J. M., Filipi, M. L., & Stergiou, N. (2011). Elliptical exercise improves fatigue ratings and quality of life in patients with multiple sclerosis. *Journal of Rehabilitation Research and Development, 48* (7), 881-889.

[49] Aoyagi, Y., Park, H., Park, S., & Shephard, R. J. (2010). Habitual physical activity and health-related quality of life in older adults: interactions between the amount and intensity of activity (the Nakanojo Study). *Quality of Life Research, 19* (3), 333-338.

[50] Fassett, R. G., Robertson, I. K., Geraghty, D. P., Ball, M. J., Burton, N. W., & Coombes, J. S. (2009). Physical Activity Levels in Patients with Chronic Kidney Disease Entering the LORD Trial. *Medicine and Science in Sports and Exercise, 41* (5), 985-991.

[51] Kerse, N., Falloon, K., Moyes, S. A., Hayman, K. J., Dowell, T., Kolt, G. S., &Arroll, B. (2008). DeLLITE depression in late life: an intervention trial of exercise. Design and recruitment of a randomised controlled trial. *BMC Geriatrics, 24* (8), 12-18.

[52] Logsdon, R. G., McCurry, S. M., Pike, K. C., & Teri, L. (2009). Making Physical Activity Accessible to Older Adults With Memory Loss: A Feasibility Study. *Gerontologist, 49*.

[53] Ogilvie, D., Mitchell, R., Mutrie, N., Petticrew, M., & Platt, S. (2008). Personal and environmental correlates of active travel and physical activity in a deprived urban population. *International Journal of Behavioral Nutrition and Physical Activity, 5*.

[54] van Uffelen, J. G. Z., Paw, M. J. M. C. A., Hopman-Rock, M., & van Mechelen, W. (2007). The effect of walking and vitamin B supplementation on quality of life in community-dwelling adults with mild cognitive impairment: a randomized, controlled trial. *Quality of Life Research, 16* (7), 1137-1146.

[55] Krousel-Wood, M. A., Berger, L., Jiang, X., Blonde, L., Myers, L., & Webber, L. (2008). Does home-based exercise improve body mass index in patients with type 2 diabetes? Results of a feasibility trial. *Diabetes Research and Clinical Practice, 79* (2), 230-236.

[56] Stroud, N. M., & Minahan, C. L. (2009). The impact of regular physical activity on fatigue, depression and quality of life in persons with multiple sclerosis. *Health and Quality of Life Outcomes, 7.*

[57] Lund, A., Michelet, M., Sandvik, L., Wyller, T. B., & Sveen, U. (2011). A lifestyle intervention as supplement to a physical activity programme in rehabilitation after stroke: a randomized controlled trial. *Clinical Rehabilitation*.

[58] Lawton, B. A., Rose, S. B., Elley, C. R., Dowell, A. C., Fenton, A., & Moyes, S. A. (2008). Exercise on prescription for women aged 40-74 recruited through primary care: two year randomised controlled trial. *British Medical Journal, 337*.

[59] Poulin, M.-J., Chaput, J.-P., Simard, V., Vincent, P., Bernier, J., Gauthier, Y., Tremblay, A. (2007). Management of anti psychotic-induced weight gain: a prospective naturalistic study of the effectiveness of a supervised exercise program. *Australian and New Zealand Journal of Psychiatry, 41* (12), 980-989.

[60] Park, S. A., Shoemaker, C., & Haub, M. (2008). Can Older Gardeners Meet the Physical Activity Recommendation through Gardening? *HortTechnology, 18* (4), 639-643.

[61] Brandes, M., Ringling, M., Winter, C., Hillmann, A., & Rosenbaum, D. (2011). Changes in Physical Activity and Health-Related Quality of Life During the First Year After Total Knee Arthroplasty. *Arthritis Care & Research, 63* (3), 328-334.

[62] MacMillan, F., Fitzsimons, C., Black, K., Granat, M. H., Grant, M. P., Grealy, M., Mutrie, N. (2011). West End Walkers 65+: A randomised controlled trial of a primary care-based walking intervention for older adults: Study rationale and design. *BMC Public Health, 11*.

[63] Sullivan, M., Karlsson, J., & Ware, J. E., Jr. (1995). The Swedish SF-36 Health Survey--I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. *Social Science Medicine, 41* (10), 1349-1358.