Acta Clin Croat 2016; 55:440-452 Review

doi: 10.20471/acc.2016.55.03.13

MEASURING PHYSICAL ACTIVITY IN PREGNANCY USING QUESTIONNAIRES: A META-ANALYSIS

Snježana Schuster[1], Iva Šklempe Kokić[2] and Joško Sindik[3]

[1]University of Applied Health Sciences, Zagreb; [2]Lavoslav Ružička College of Applied Sciences, Vukovar; [3]Institute for Anthropological Research, Zagreb, Croatia

SUMMARY – Physical activity (PA) during normal pregnancy has various positive effects on pregnant women's health. Determination of the relationship between PA and health outcomes requires accurate measurement of PA in pregnant women. The purpose of this review is to provide a summary of valid and reliable PA questionnaires for pregnant women. During 2013, Pubmed, OvidSP and Web of Science databases were searched for trials on measurement properties of PA questionnaires for pregnant population. Six studies and four questionnaires met the inclusion criteria: Pregnancy Physical Activity Questionnaire, Modified Kaiser Physical Activity Survey, Short Pregnancy Leisure Time Physical Activity Questionnaire and Third Pregnancy Infection and Nutrition Study Physical Activity Questionnaire. Assessment of validity and reliability was performed using correlations of the scores in these questionnaires with objective measures and subjective measures (self-report) of PA, as well as test-retest reliability coefficients. Sample sizes included in analysis varied from 45 to 177 subjects. The best validity and reliability characteristics (together with effect sizes) were identified for the Modified Kaiser Physical Activity Survey and Pregnancy Physical Activity Questionnaire (French, Vietnamese, standard). In conclusion, assessment of PA during pregnancy remains a challenging and complex task. Questionnaires are a simple and effective, yet limited tool for assessing PA.

Key words: *Motor activity; Exercise; Pregnancy, Surveys and questionnaires; Self-report*

Introduction

Physical activity (PA) in pregnancy reduces the risk of gestational diabetes mellitus, excessive weight gain, hypertension, and preterm delivery[1-5]. However, the amount of activity required for favorable pregnancy outcomes remains to be determined[6]. Determination of the relationship between PA levels in pregnancy and health outcomes demands accurate assessment of PA levels in pregnant population. Currently, there is no gold standard for assessing PA during pregnancy.

Correspondence to: *Snježana Schuster, PhD*, Department of Physiotherapy, University of Applied Health Sciences, Mlinarska cesta 38, HR-10000 Zagreb, Croatia
E-mail: snjezana.schuster@zvu.hr
Received June 26, 2015, accepted June 8, 2016

Valid and reliable measures of physical activity are required to document the frequency, duration and distribution of PA in defined populations; evaluate the prevalence of individuals meeting health recommendations; examine the effect of various intensities of PA on specific health parameters; make cross-cultural comparisons and evaluate the effects of interventions[7]. Measurement of physical activity can be divided into self-report methods (questionnaires and diaries) and objective assessment (accelerometers, pedometers and heart rate monitors). Self-reported questionnaires can be self- or interviewer-administered. They can collect the mode or type of activity, frequency, duration and perceived exertion. Mode and perceived exertion of PA cannot be collected by objective measures. However, physical activity questionnaires (PAQs) are prone to measurement error and bias[8,9].

Objective measurements are not subject to self-report error, but they have other limitations, like inability to measure various types of PA, inaccurate measurement of upper body movement, stationary exercise, weight lifting, and water activities. Also, cut-off points necessary to translate accelerometer data into intensity categories have not been developed for pregnant women[10]. Pedometers cannot differentiate intensity level of PA and they are only able to assess walking and running. Heart rate monitors have even more limitations for use in pregnancy.

Despite frequent use of objective assessment methods to measure PA, PAQs provide a practical method for PA assessment in surveillance studies, for risk stratification, and when examining the etiology of disease in large observational studies[11]. They are simple to use and are cost-effective. Most PAQs report the type, time, frequency, intensity and context of the activity. They can be designed for general or specific population and a few of them are designed specifically for pregnant women. Results from studies aimed at evaluating the validity of PAQs assessed in one population cannot be systematically extrapolated to other populations, and consequently, a great variety of PAQs have been developed and tested for reliability and validity in recent years[11]. The majority of currently available PAQs have been developed and validated in men and emphasize moderate and vigorous intensity PA. Also, they do not include household and childcare activities, which comprise a substantial portion of pregnancy PA[12]. The growing body of literature recommends that questionnaires should assess PA in the full range of physical activities related to sports and recreational activities, but should also include work, transportation, childcare and household activities[13,14].

Only several PA questionnaires have been developed and evaluated for validity and reliability in pregnant population. Still, they are the most common method used in surveillance and epidemiological studies of the impact of PA during pregnancy on health outcomes because of their cost and time-effectiveness[10,15,16]. Results of these studies impact PA guidelines for pregnant women and, since their findings are based on PAQs, it is very important that these PAQs are valid and reliable.

The purpose of this systematic review and respective meta-analysis is to provide an overview and summary of self-reported PAQs for pregnant population with evidence of validity and reliability. It contains description and measurement properties of PAQs for pregnant women, but also provides the information about effect sizes and advantages and disadvantages of certain PAQs for pregnant women.

Materials and Methods

A comprehensive database search was performed during 2013 within PubMed, OvidSP and Web of Science databases with the purpose of finding trials on measurement properties of PAQs for pregnant population. We included only trials in English, published within the last ten years and reporting on PAQs specifically tested for both validity and reliability in pregnant women. The search was concluded on August 2014. The key terms and their combinations used for the search were: "physical activity", "pregnancy", "questionnaire", "self-report", "validity", "reliability", "measurement" and "assessment". The procedure was concluded by using references found in all relevant papers. We excluded trials on non-pregnant or postpartum female populations and PAQs that did not take into account at least two determinants of PA (mode, frequency, duration and perceived intensity). Articles without report of agreement statistics between questionnaire and criterion measure were excluded from the review. Trials without reported reliability were also excluded.

For the purpose of meta-analysis, the correlation coefficient effect size (r) is used, designed for contrasting two continuous variables. Fisher's Z-r transformation and variance are useful for meta-analysis when given the correlation and sample size. Thus, values of correlations were transformed in standardized correlations (Z-r), with belonging confidence intervals (Table 1)[17]:

$$z = \frac{1}{2}\ln\left(\frac{1+r}{1r}\right)$$

Then, variances for each correlation coefficient were calculated[17]

$$v = \frac{1}{n-3}$$

After this step, the correlation coefficient (r) was recalculated again, as well as weight (w)[17].

$$r = \frac{e^{2n2}-1}{e^{2n2}+1} \quad w = n-3$$

Table 1. Validity and reliability of physical activity questionnaires – an insight in effect sizes (variances and transformed correlations, with sample sizes)

Reference	Sample size	Validity – objective measures	Validity – self-report measures	Reliability test-retest
Chasan-Taber *et al.*, 2004	54	SCC-Freedson 0.08 (Fisher Z_r=0.08; v=0.02; 95% C.I.=-0.19±0.36) SCC-Swartz 0.32 (Fisher Z_r=0.33; v=0.02; 95% C.I.=0.06±0.61) SCC-Hendelman 0.43 (Fisher Z_r=0.46; v=0.02; 95% C.I.=0.19±0.73) Average SCC 0.27 (Fisher Z_r=0.28; v=0.02; 95% C.I.=0.00±0.55)	N/A	0.79 sedentary PA (Fisher Z_r=1.07; v=0.02; 95% C.I.=0.80±1.35) 0.78 light PA (Fisher Z_r=1.05; v=0.02; 95% C.I.=0.77±1.32) 0.82 moderate PA (Fisher Z_r=1.16; v=0.02; 95% C.I.=-0.71±0.89) 0.81 vigorous PA (Fisher Z_r=1.13; v=0.02; 95% C.I.=-0.85±1.40) Total PA - ICC 0.78 (Fisher Z_r=1.05; v=0.02; 95% C.I.=-0.77±1.32)
Schmidt *et al.*, 2006	54	SCC-Freedson, 0.59 (Fisher Z_r=0.68; v=0.02; 95% C.I.=0.40±0.95) SCC-Swartz, 0.49 (Fisher Z_r=0.54; v=0.02; 95% C.I.=0.26±0.81) SCC-Hendelman 0.55 (Fisher Z_r=0.62; v=0.02; 95% C.I. =0.34±0.89) Average SCC 0.52 (Fisher Z_r=0.58; v=0.02; 95% C.I. =0.30±0.85)	SCC-total PA 0.37 (Fisher Z_r=0.39; v=0.02; 95% C.I. =0.11±0.66) SCC-total PA (weighted) 0.51 (Fisher Z_r=0.56; v=0.02; 95% C.I. =0.29±0.84)	ICC-total PA 0.84 (Fisher Z_r=1.22; v=0.02; 95% C.I.=0.95±1.50)
Ota *et al.*, 2008	59	PCC 0.29 (Fisher Z_r=0.30; v=0.02; 95% C.I.=0.04±0.56)	N/A	ICC total PA 0.88 (Fisher Z_r=1.38; v=0.02; 95% C.I.=1.11±1.64) 0.94 sedentary (Fisher Z_r=1.74; v=0.02; 95% C.I.=1.47±2.00) 0.88 light (Fisher Z_r=1.38; v=0.02; 95% C.I.=1.11±1.64) 0.90 moderate (Fisher Z_r=1.47; v=0.02; 95% C.I.=1.21±1.73) 0.87 vigorous PA (Fisher Z_r=1.33; v=0.02; 95% C.I.=1.07±1.60)

Table 1. Continue

Reference	Sample size	Validity – objective measures	Validity – self-report measures	Reliability test-retest
Aittasalo *et al.*, 2010	45-accelero-meter 47-logbook	SCC-frequency 0.16 (Fisher Z_r=0.16; v=0.02; 95% C.I.=-0.14±0.46) SCC-duration -0.18 (Fisher Z_r=-0.18; v=0.02; 95% C.I.=-0.48±0.18)	SCC-frequency 0.27 (Fisher Z_r=1.33; v=0.02; 95% C.I.=1.07±1.60) SCC-duration 0.47 (Fisher Z_r=0.51; v=0.02; 95% C.I.=0.21±0.80)	N/A
Evenson & Wen, 2010	177	MET-hours/week accelerometer SCC-moderate 0.04 (Fisher Z_r=0.04; v=0.01; 95% C.I.=-0.11±0.19) SCC-MVPA 0.18-0.21; (Fisher Z_r=0.20; v=0.01; 95% C.I.=0.05±0.35) SCC-vigorous 0.29-0.31 (Fisher Z_r=0.31; v=0.01; 95% C.I.=0.16±0.46) Hours/week per accelerometer SCC-moderate 0.25-0.33 (Fisher Z_r=0.30; v=0.01; 95% C.I.=-0.15±0.45) SCC- MVPA 0.28-0.34 (Fisher Z_r=0.32; v=0.01; 95% C.I.=-0.17±0.47) SCC-vigorous 0.26-0.32 (Fisher Z_r=0.30; v=0.01; 95% C.I.=-0.15±0.45)	SCC (MET-hours/ week): 0.61 moderate (Fisher Z_r=0.71; v=0.01; 95% C.I.=0.56±0.86) 0.69 MVPA (Fisher Z_r=0.85; v=0.01; 95% C.I.=0.70±1.00) 0.47 vigorous (Fisher Z_r=0.51; v=0.01; 95% C.I.=0.36±0.66) SCC (hours/week): 0.63 moderate (Fisher Z_r=0.74; v=0.01; 95% C.I.=0.59±0.89) 0.66 MVPA (Fisher Z_r=0.79; v=0.01; 95% C.I.=0.64±0.94) 0.68 vigorous (Fisher Z_r=0.83; v=0.01; 95% C.I.=0.68±0.98)	ICC (MET-hours/week) 0.72 moderate (Fisher Z_r=0.91; v=0.01; 95% C.I.=0.76±1.06) 0.74 MVPA (Fisher Z_r=0.95; v=0.01; 95% C.I.=0.80±1.10) 0.73 vigorous (Fisher Z_r=0.93; v=0.01; 95% C.I.=0.78±1.08) ICC total hours/week 0.84 (Fisher Z_r=1.22; v=0.01; 95% C.I.=1.07±1.37)
Chandonnet *et al.*, 2012	49	SCC-total activity 0.50 (Fisher Z_r=0.55; v=0.01; 95% C.I.=0.26±0.84) SCC-vigorous 0.25 (Fisher Z_r=0.26; v=0.01; 95% C.I.=-0.03±0.50) SCC-moderate 0.40 (Fisher Z_r=0.42; v=0.01; 95% C.I.=0.13±0.71)	N/A	ICC 0.90 total activity (Fisher Z_r=1.47; v=0.01; 95% C.I.=1.18±1.76) 0.86 light and moderate (Fisher Z_r=1.29; v=0.01; 95% C.I.=1.00±1.58) 0.81 vigorous (Fisher Z_r=1.13; v=0.01; 95% C.I.=0.84±1.42) 0.59 transportation (Fisher Z_r=0.68; v=0.01; 95% C.I.=0.39±0.97) 0.89 household and caregiving (Fisher Z_r=1.42; v=0.01; 95% C.I.=1.13±1.97)

SCC = Spearman's correlation coefficient; MET = metabolic equivalent; MVPA = moderate to vigorous physical activity; PA = physical activity; N/A = not available

Table 2 shows belonging effect sizes and overall (average) effect sizes for subsets of correlations (calculated in the step described above) using the formula[17]:

$$E = \frac{\sum_{i=1}^{n} w_i r_i}{\sum_{i=1}^{n} w_i}$$

Finally, in Table 3, average values of overall transformed validity and reliability indicators with belonging effect sizes are presented. It should be mentioned that average values of effect sizes are calculated in terms of the means of specific measures (Table 2), as well as overall validity and reliability indicators and belonging effect sizes (Table 3). In other words, for example, overall physical activity indicators are calculated as the mean of sedentary, moderate and vigorous physical activity, not the indicator of total physical activity.

Results

Four questionnaires in six studies met the inclusion criteria: Pregnancy Physical Activity Questionnaire (PPAQ)[6,18,19], Modified Kaiser Physical Activity Survey (MKPAS)[20], Short Pregnancy Leisure Time Physical Activity Questionnaire (LTPA)[21], and Third Pregnancy Infection and Nutrition Study (PIN3) Physical Activity Questionnaire[15] (Table 4). Another six questionnaires were excluded because reliability data were not reported in trials[22-27]. Two questionnaires were self-administered (PPAQ and LTPA) and two interviewer-administered (MKPAS and PIN3 PAQ). Trials were conducted mostly in the North American geographical region, i.e. United States[6,15,20] and Canada[19]. One trial was conducted in Vietnam[18] and one in Finland[21]. Recall period for PA assessment was current trimester of pregnancy (PPAQ, MKPAS), the last two weeks for LTPA, and the last week for PIN3 PAQ. All questionnaires assessed the mode, frequency and duration of PA, and perceived intensity of PA was assessed in PIN3 PAQ and LTPA.

Table 1 shows confidence intervals, standardized correlations (Zr) and variances for each correlation coefficient, while Table 2 shows belonging effect sizes and overall (average) effect sizes for subsets of correlations[17].

Assessment of validity was done using an objective measure of PA in all trials (Table 5). In four trials, it was done by accelerometer[6,15,19,20], and in two by pedometer[18,21]. Validity analysis sample varied from 45 to 177 pregnant women. Results of the comparison ranged from slight to moderate agreement and were reported as Spearman's correlation coefficient (SCC) or Pearson's correlation coefficient (PCC). In three trials, an additional subjective measure was used for calculating validity, i.e. Schmidt *et al.*[20] compared MKPAS and PPAQ results, Aittasalo *et al.*[21] used leisure activity logbook, and Evenson and Wen[15] used PIN3 structured diary. The period of objective criterion measurement was seven days in all trials, except for the trial conducted by Ota *et al.*[18], where it was ten days. Agreement between subjective measures varied from fair to strong and was mostly calculated as SCC and PCC. Overall effect sizes (objective measures) varied from 0.08 to 0.29, i.e. they were mainly small. Overall effect sizes (self-report measures) were higher and varied from 0.22 to 0.42, i.e. they were mostly medium (Table 2).

Sample size for test-retest reliability varied from 49 to 109 participants (Table 5). Time between test and retest ranged from one-two days[15] to two weeks[18,21], being mostly seven days[6,19,20]. The same time period was recalled in most reliability assessments. Test-retest reliability estimates were mostly shown as intraclass correlation coefficients (ICC) with substantial to strong values. Overall effect sizes were higher and varied from 0.48 to 0.63, i.e. they were medium to large (Table 2).

However, the clearest insight into the validity and reliability of measuring instruments for PA based on questionnaires is given in Table 3. The most desirable characteristics in terms of their validity based on objective measures were identified for the Modified Kaiser Physical Activity Survey, and then the Pregnancy Physical Activity Questionnaire (French, Vietnamese, standard); they showed highest correlations with objective measures, as well as small but the best effect sizes (among all questionnaires included in this meta-analysis) (Table 3).

The most desirable characteristics in terms of their validity based on subjective (self-report) measures were recorded in the Third Pregnancy Infection and Nutrition Study (PIN3) Physical Activity Questionnaire, and then the Modified Kaiser Physical Activity

Table 2. Validity and reliability of physical activity questionnaires – overall effect sizes

Reference	Name of the questionnaire	Sample size	Validity – objective measures Effect sizes (total)	Validity – self-report measures Effect sizes (total)	Reliability test-retest Effect sizes (total)
Chasan--Taber *et al.*, 2004	Pregnancy Physical Activity Questionnaire	54	SCC-Freedson 0.04 SCC-Swartz 0.16 SCC-Hendelman 0.21 Average SCC 0.14 Overall 0.14	N/A	sedentary PA 0.49 light PA 0.48 moderate PA 0.52 vigorous PA 0.51 Total PA - ICC 0.48 Overall 0.49
Schmidt *et al.*, 2006	Modified Kaiser Physical Activity Survey	54	SCC-Freedson 0.33 SCC-Swartz, 0.26 SCC-Hendelman 0.27 Average SCC 0.28 Overall 0.29	SCC-total PA 0.19 SCC-total PA (weighted) 0.25 Overall 0.22	ICC-total PA 0.54 Overall 0.54
Ota *et al.*, 2008	Pregnancy Physical Activity Questionnaire – Vietnamese version	59	PCC 0.15	N/A	sedentary 0.70 light 0.60 moderate 0.63 vigorous PA 0.58 ICC total PA 0.60 Overall 0.63
Aittasalo *et al.*, 2010	Short Pregnancy Leisure Time Physical Activity Questionnaire	45-accelero-meter 47-logbook	SCC-frequency 0.08 SCC-duration 0.09 Overall 0.08	SCC-frequency 0.58 SCC-duration 0.25 Overall 0.42	N/A
Evenson & Wen, 2010	Third Pregnancy Infection and Nutrition Study (PIN3) Physical Activity Questionnaire	177	MET-hours/week accelerometer SCC-moderate 0.02 SCC-MVPA 0.10 SCC-vigorous 0.15 Overall MET 0.09 Hours/week per accelerometer SCC-moderate 0.15 SCC- MVPA 0.16 SCC-vigorous 0.15 Overall Hours/week 0.15	SCC (MET-hours/ week): SCC-moderate 0.34 SCC- MVPA 0.40 SCC-vigorous 0.25 Overall 0.33 SCC (hours/week): SCC-moderate 0.35 SCC- MVPA 0.38 SCC-vigorous 0.39 Overall 0.37	ICC (MET-hours/ week) moderate 0.43 MVPA 0.44 vigorous 0.43 Overall 0.43 ICC total hours/ week 0.54
Chandonnet *et al.*, 2012	Pregnancy Physical Activity Questionnaire –French version	49	SCC-moderate 0.21 SCC-vigorous 0.13 SCC-total activity 0.27 Overall: 0.17	N/A	0.57 light and moderate 0.51 vigorous 0.33 transportation 0.61 household and caregiving Overall 0.48 ICC - total activity 0.63

ICC = intraclass correlation coefficient; SCC = Spearman's correlation coefficient; MET = metabolic equivalent; PA = physical activity; MVPA = moderate to vigorous physical activity; N/A = not available

Table 3. Physical activity questionnaires – validity, reliability and overall effect sizes

Reference	Overall validity – objective measures (Effect sizes -total)	Overall validity – self-report measures (Effect sizes - total)	Overall reliability (Effect sizes - total)
Pregnancy Physical Activity Questionnaire (French, Vietnamese, standard)	range 0.08-0.50 average 0.36 effect sizes (0.14-0.17; average 0.16)	N/A	range 0.79-0.89 average 0.83 effect sizes (0.48-0.63; average 0.53)
Modified Kaiser Physical Activity Survey	average 0.36 average effect size (0.29)	average 0.44 average effect size (0.22)	average 0.84 average effect size (0.54)
Short Pregnancy Leisure Time Physical Activity Questionnaire	average 0.01 average effect size (0.08)	average 0.37 average effect size (0.42)	N/A
Third Pregnancy Infection and Nutrition Study (PIN3) Physical Activity Questionnaire	average 0.24 average effect size (0.12)	average 0.57 average effect size (0.35)	average 0.73 average effect size (0.43)

N/A = not available

Table 4. Descriptive characteristics of physical activity questionnaires

Reference	Name of the questionnaire	Administration delivery	Recall period	Country	Population		Mode/ frequency/ duration/ perceived intensity of physical activity	Domain of activity
					Size	Age		
Chasan--Taber *et al.*, 2004	Pregnancy Physical Activity Questionnaire	Self--administered	Current trimester	USA	54	N/A	Yes/yes/ yes/no	Household activities, sport, transportation, work, inactivity
Schmidt *et al.*, 2006	Modified Kaiser Physical Activity Survey	Interviewer--administered	Current trimester	USA	54	18-47 (mean age: 29.5)	Yes/yes/ yes/no	Active living, household activities, sport, work
Ota *et al.*, 2008	Pregnancy Physical Activity Questionnairee – Vietnamese version	Self--administered	Current trimester	Vietnam	59	26.8±5.0	Yes/yes/ yes/no	Household activities, sport, transportation, work, inactivity
Aittasalo *et al.*, 2010	Short Pregnancy Leisure Time Physical Activity Questionnaire	Self--administered	Past 2 weeks	Finland	92	28.7±5.8	Yes/yes/ yes/yes	Household activities, recreation
Evenson & Wen, 2010	Third Pregnancy Infection and Nutrition Study (PIN3) Physical Activity Questionnaire	Interviewer--administered	Past week	USA	177	4% <20 yrs, 43% 20-29 yrs, 53.1% >30 yrs	Yes/yes/ yes/yes	Any activity that was at least fairly light in intensity
Chandonnet *et al.*, 2012	Pregnancy Physical Activity Questionnairee – French version	Self--administered	Current trimester	Canada	49	29.8±4.2	Yes/yes/ yes/no	Household activities, sport, transportation, work, inactivity

N/A = not available

Table 5. Validity and reliability of physical activity questionnaires

Reference	Validity analysis sample	Validity						Reliability analysis sample	Reliability		
		Criterion method – objective measures			Criterion method – self-report measures						
		Measure	Recall exact same period of time	Results	Measure	Recall exact same period of time	Results		Time between test an retest	Recall same period of time	Results
Chasan-Taber *et al.*, 2004	54	Actigraph 7164 accelerometer, 7 days	Yes	SCC (MVPA accelerometer to total PA from the questionnaire) 0.08 Freedson, 0.32 Swartz, 0.43 Hendelman; SCC (average counts/minute accelerometer to questionnaire) 0.27	None	N/A	N/A	54	7 days	Yes, if still in the same trimester 1 week later	ICC 0.78 total PA, 0.79 sedentary, 0.78 light, 0.82 moderate, 0.81 vigorous PA
Schmidt *et al.*, 2006	54	ActiGraph 7164 accelerometer, 7 days	Yes	SCC (MVPA accelerometer to total PA from questionnaire) 0.59 Freedson, 0.49 Swartz, 0.55 Hendelman; SCC (average counts/minute accelerometer to questionnaire) 0.52	Pregnancy Physical Activity Questionnaire (Chasan-Taber *et al.*, 2004)	Yes	SCC 0.37 total PA; SCC 0.51 total PA with weighted domains	54	7 days	Yes, if still in the same trimester 1 week later	ICC 0.84 total PA
Ota *et al.*, 2008	59	Digiwalker SW-200 Yamax pedometer 10-day averages	No	PCC (overall): 0.29	None	N/A	N/A	60	2 weeks	Yes, if still in the same trimester 2 weeks later	ICC 0.88 total PA, 0.94 sedentary, 0.88 light, 0.90 moderate, 0.87 vigorous PA

Aittasalo *et al.*, 2010	45-accelero-meter; 47-logbook	Omron HJ-113 pedometer, 7 days	No	SCC (pedometer to average weekly leisure activity on questionnaire) 0.16 frequency, -0.18 duration	Leisure activity logbook – duration and intensity, 1 week	No	SCC (logbook to average weekly leisure activity on questionnaire) 0.27 frequency, 0.47 duration	49	2 weeks	No	Weekly leisure activity: change in mean frequency of sessions -0.7; change in geometric mean duration 0.9
Evenson & Wen, 2010	177	ActiGraph 7164 accelerometer, 7 days	Yes	SCC (questionnaire in MET-hours/ week to accelerometer using 3 cutpoints): 0.04 moderate, 0.29-0.31 vigorous, 0.18-0.21 MVPA; SCC (questionnaire in hours/week to accelerometers using 3 cutpoints): 0-25-0.33 moderate, 0.26-0.32 vigorous, 0.28-0.34 MVPA	Third Pregnancy Infection and Nutrition Study (PIN3) structured diary	Yes	SCC (questionnaire to diary, in MET-hours/week): 0.61 moderate, 0.47 vigorous, 0.69 MVPA; SCC (questionnaire to diary, hours/ week): 0.63 moderate, 0.68 vigorous, 0.66 MVPA	109	1-2 days	Yes	ICC (MET-hours/ week) 0.72 moderate, 0.73 vigorous, 0.74 MVPA; ICC total hours/ week using perceived intensity 0.84
Chandonnet *et al.*, 2012	49	Actigraph GT1M accelerometer, 7 days	Yes	SCC (questionnaire to accelerometer, Matthews' cut point): 0.50 total activity, 0.25 vigorous, 0.40 moderate	None	N/A	N/A		7 days	Yes, if still in the same trimester 1 week later	ICC 0.90 total activity, 0.86 light and moderate, 0.81 vigorous; 0.59 transportation, 0.89 household and caregiving

SCC = Spearman's correlation coefficient; PCC = Pearson's correlation coefficient; ICC = intraclass correlation coefficient; PA = physical activity; MVPA = moderate to vigorous physical activity; MET = metabolic equivalent; N/A = not available

Survey; they showed highest correlations with subjective measures, as well as small to moderate but the best effect sizes (among all questionnaires included in this meta-analysis). However, even the Short Pregnancy Leisure Time Physical Activity Questionnaire also had satisfactory validity based on subjective measures, yet something lower size than the abovementioned questionnaires, but with moderate effect size (Table 3).

Finally, the most desirable characteristics in terms of their reliability based on test-retest measures were identified for the Modified Kaiser Physical Activity Survey, and then the Pregnancy Physical Activity Questionnaire (French, Vietnamese, standard); they showed highest correlations with subjective measures, as well as moderate effect sizes (among all questionnaires included in this meta-analysis). However, even the Third Pregnancy Infection and Nutrition Study (PIN3) Physical Activity Questionnaire also had satisfactory validity based on subjective measures, but something lower size than the abovementioned questionnaires, also with moderate effect size (Table 3).

Overall, among all questionnaires in this short meta-analysis, the best validity and reliability characteristics (together with effect sizes) were recorded for the Modified Kaiser Physical Activity Survey and Pregnancy Physical Activity Questionnaire (French, Vietnamese, standard). Very close to them was the Third Pregnancy Infection and Nutrition Study (PIN3) Physical Activity Questionnaire, while the worst characteristics were found for the Short Pregnancy Leisure Time Physical Activity Questionnaire.

Discussion

In non-pregnant population, the gold standard of energy expenditure is the use of doubly-labeled water (DLW)[28]. However, this method is expensive, time-consuming and not feasible for general use in assessing PA in large-scale studies. Objective measures of PA in pregnancy use accelerometers, pedometers and heart rate monitors, but they do have limitations.

Accelerometry is validated against DLW in non-pregnant population[28] and more feasible to use in pregnant women. However, accelerometers are not able to measure precisely upper body movements, weight-bearing activities, cycling and swimming[10,29]. Cut-off points from accelerometer-calibration studies needed to categorize counts into levels of intensity are not validated for pregnant women. Pedometers estimate total steps and distance and can be used as a valid and reliable method for PA measurement in non-pregnant populations. However, their role in pregnancy remains unclear[30]. Heart rate monitors can be uncomfortable and limited in assessing most determinants of PA. Heart rate is variable during pregnancy and could be affected by numerous factors other than PA. Also, objective activity monitors heighten awareness of PA among pregnant women and potentially cause bias.

This review identified four questionnaires with proven validity compared with objective measures and test-retest reliability. While reliability was substantial to strong, validity was moderate at best. Pearson's and Spearman's correlations may not be the most appropriate statistical methods for reporting validity of PAQs[11]. The challenge for questionnaires used to assess PA during pregnancy is to rank pregnant women in categories from sedentary to most active within a narrower range of PA than in non-pregnant samples[6]. Questionnaires for pregnant population have to take into account the potentially different metabolic cost of PA in pregnancy in comparison to general population. The potential misclassification can bias studies of the relationship between pregnancy PA and maternal and fetal health, limiting their ability to detect important associations with disease[31-33].

The Pregnancy Physical Activity Questionnaire is a semi-quantitative questionnaire that reports on the time spent on 32 activities, including household, caregiving, occupational activities, sports and exercise, transportation, and inactivity. There is also an open-ended section for adding activities not already listed. It has been developed and validated only for pregnant population. PIN3 PAQ is also specifically designed for pregnant women, as well as to capture moderate and vigorous PA. It reports on the frequency, duration and perceived exertion of participation in recreation, occupational activities, transportation, childcare and household activities.

The Leisure Time Physical Activity Questionnaire is based on the International Physical Activity Questionnaire (IPAQ)[34] with several modifications for better distinguishing structured and unstructured features of PA[21]. It includes two PA domains: leisure-time PA and household PA. It does not report on occupational

PA. It reports on the mode, frequency, duration and perceived intensity of PA. KPAS is adapted from the Baecke Physical Activity Survey[35] and designed specifically to assess PA of women[36]. It includes assessment of multiple domains of PA: household activities and family care, occupational activities, active living habits and participation in sports and exercise. It reports on the mode, frequency and duration of PA. It has also been validated in non-pregnant women[36].

Self-reported measurements of PA are the most common method in epidemiological studies of pregnant women[10], which makes their validity and reliability critical for making conclusions and developing future guidelines. Neither self-reporting nor objective measures are perfect for assessment of PA in pregnancy. Accuracy and precision of PAQs are still a methodological problem which can be partially prevented by choosing valid and reliable PAQ tested and adjusted for a specific population, and by using both self-reporting and objective measures of PA in a single trial. When deciding which PAQ to use, researchers should take into account the measurement properties of a specific PAQ and the determinants of PA they wish to measure, according to a specific research question, because not all PAQs measure every aspect of PA. If they plan to measure and compare PA in different populations or pre-, post- and during pregnancy, they should use PAQs validated for both pregnant and non-pregnant populations, i.e. for all participants.

In conclusion, there is no gold standard for the assessment of PA during pregnancy. It remains a challenging and complex task. Questionnaires are the most common, simple and effective, yet imperfect assessment method used in surveillance and epidemiological studies. They should be meticulously tested for validity and reliability before being used in trials and before drawing conclusions from the results of these trials. Also, questionnaires should contain information on all aspects of PA (mode, frequency, duration and perceived intensity) for best determination of the causative relationship between various aspects of PA, especially dose-response and health-related outcomes. In this short meta-analysis, we determined the validity and reliability characteristics of four questionnaires, revealing which could be more desirable for future research in this field. Therefore, continuous research is necessary to improve PA assessments for pregnant women, both self-reported and objective.

References

1. Evenson KR, Siega-Riz AM, Savitz DA, Leiferman JA, Thorp JM. Vigorous leisure activity and pregnancy outcome. Epidemiology. 2002 Nov;13(6):653-9. PubMed PMID: 12410006. doi: 10.1097/01.EDE.0000021463.45041.95
2. Hatch MC, Shu X, McLean DE, Levin B, Begg M, Reuss L, *et al*. Maternal exercise during pregnancy, physical fitness, and fetal growth. Am J Epidemiol. 1993 May 15;137(10):1105-14. PubMed PMID: 8317440
3. Solomon CG, Willett WC, Carey VJ, Rich-Edwards J, Hunter DJ, Colditz GA, *et al*. A prospective study of pregravid determinants of gestational diabetes mellitus. JAMA. 1997 Oct 1;278(13):1078-83. PubMed PMID: 9315766
4. Radić B, Radić P, Duraković D. Sports and Health: equivalence or contrariety. Acta Clin Croat 2014; 53:430-6.
5. Vuković Bobić M, Habek D, Čerkez Habek J. Perinatal epidemiological risk factors for preeclampsia. Acta Clin Croat 2015; 54:9-13.
6. Chasan-Taber L, Schmidt MD, Roberts DE, Hosmer D, Markenson G, Freedson PS. Development and validation of a Pregnancy Physical Activity Questionnaire. Med Sci Sports Exerc. 2004 Oct;36(10):1750-60. PubMed PMID: 15595297
7. Warehan NJ, Rennie KL. The assessment of physical activity in individuals and populations: why try to be more precise about how physical activity is assessed? Int J Obes Relat Metab Disord. 1998 Aug;22 Suppl 2:S30-8. PubMed PMID: 9778094
8. Jobe JB, Mingay DJ. Cognitive research improves questionnaires. Am J Public Health. 1989 Aug;79(8):1053-5. PubMed PMID: 2751028
9. Durante R, Ainsworth BE. The recall of physical activity: using a cognitive model of the question-answering process. Med Sci Sports Exerc. 1996 Oct;28(10):1282-91. PubMed PMID: 8897386
10. Evenson KR, Chasan-Taber L, Symons Downs D, Pearce EE. Review of self-reported physical activity assessments for pregnancy: summary of the evidence for validity and reliability. Paediatr Perinat Epidemiol. 2012 Sep;26(5):479-94. doi: 10.1111/j.1365-3016.2012.01311.x. Epub 2012 Jul 5
11. Helmerhorst HJ, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. Int J Behav Nutr Phys Act. 2012 Aug 31;9:103. doi: 10.1186/1479-5868-9-103
12. Schmidt MD, Erickson JB, Freedson PS, Markenson G, Chasan-Taber L. Physical activity patterns during pregnancy in a low income racially diverse population. Am J Epidemiol. 2002;155:S103
13. Jacobs DR, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. Med Sci Sports Exerc. 1993 Jan;25(1):81-91. PubMed PMID: 8423759
14. Chasan-Taber L, Evenson KR, Sternfeld B, Kengeri S. Assessment of recreational physical activity during pregnancy in epi-

demiologic studies of birthweight and length of gestation: methodologic aspects. Women Health. 2007;45(4):85-107. PubMed PMID: 18032169. doi: 10.1300/J013v45n04_05

15. Evenson KR, Wen F. National trends in self-reported physical activity and sedentary behaviors among pregnant women: NHANES 1999-2006. Prev Med. 2010 Mar;50(3):123-8. doi: 10.1016/j.ypmed.2009.12.015. Epub 2010 Jan 4
16. Bovbjerg ML, Siega-Riz AM. Exercise during pregnancy and cesarean delivery: North Carolina PRAMS, 2004-2005. Birth. 2009 Sep;36(3):200-7. doi: 10.1111/j.1523-536X.2009.00324.x
17. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale NJ: Lawrence Earlbaum Associates; 1988.
18. Ota E, Haruna M, Yanai H, Suzuki M, Anh DD, Matsuzaki M, *et al*. Reliability and validity of the Vietnamese version of the Pregnancy Physical Activity Questionnaire (PPAQ). Southeast Asian J Trop Med Public Health. 2008 May;39 (3):562-70. PubMed PMID: 18564699
19. Chandonnet N, Saey D, Alméras N, Marc I. French Pregnancy Physical Activity Questionnaire compared with an accelerometer cut point to classify physical activity among pregnant obese women. PLoS One. 2012;7(6):e38818 doi:10.1371/ journal. pone.0038818. Epub 2012 Jun 11
20. Schmidt MD, Freedson PS, Pekow P, Roberts D, Sternfeld B, Chasan-Taber L. Validation of the Kaiser Physical Activity Survey in pregnant women. Med Sci Sports Exerc. 2006 Jan;38(1):42-50. PubMed PMID: 16394952
21. Aittasalo M, Pasanen M, Fogelholm M, Ojala K. Validity and repeatability of a short pregnancy leisure time physical activity questionnaire. J Phys Act Health. 2010 Jan;7(1):109-18. PubMed PMID: 20231762
22. Poudevigne MS, O'Connor PJ. Physical activity and mood during pregnancy. Med Sci Sports Exerc. 2005 Aug;37(8):1374-80. PubMed PMID: 16118585
23. Rousham EK, Clarke PE, Gross H. Significant changes in physical activity among pregnant women in the UK as assessed by accelerometry and self-reported activity. Eur J Clin Nutr. 2006 Mar;60(3):393-400. PubMed PMID: 16306930. doi: 10.1038/sj.ejcn.1602329
24. Symons Downs D, LeMasurier G, DiNallo J. Baby steps: pedometer determined and self-reported leisure-time exercise behaviours of pregnant women. J Phys Act Health. 2009 Jan;6(1):63-72. PubMed PMID: 19211959
25. Brantsaeter AL, Owe KM, Haugen M, Alexander J, Meltzer HM, Longnecker MP. Validation of self-reported recreational exercise in pregnant women in the Norwegian Mother and Child Cohort Study. Scand J Med Sci Sports. 2009; 20(1):e48-55. doi: 10.1111/j.1600-0838.2009.00896.x. Epub 2009 Mar 31
26. Haakstad LA, Gundersen I, Bø K. Self-reporting compared to motion monitor in the measurement of physical activity during pregnancy. Acta Obstet Gynecol Scand. 2010 Jun;89(6): 749-56. doi: 10.3109/00016349.2010.484482. PubMed PMID: 20450260
27. Harrison CL, Thompson RG, Teede HJ, Lombard CB. Measuring physical activity during pregnancy. Int J Behav Nutr Phys Act. 2011 Mar 21;8:19. doi: 10.1186/1479-5868-8-19
28. Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. Obesity (Silver Spring). 2007 Oct;15(10):2371-9. PubMed PMID: 17925461. doi: 10.1038/oby.2007.281
29. Chinapaw MJ, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, *et al.* Reliability and validity of the Activity Questionnaire for Adults and Adolescents (AQuAA). BMC Med Res Methodol. 2009 Aug 10;9:58. doi: 10.1186/ 1471-2288-9-58
30. Linseth G, Vari P. Measuring physical activity during pregnancy. West J Nurs Res. 2005 Oct;27(6):722-34. PubMed PMID: 16157944. doi: 10.1177/0193945905276523
31. Ainsworth BE. Issues in the assessment of physical activity in women. Res Q Exerc Sport. 2000 Jun;71(2 Suppl):S37-42. PubMed PMID: 10925823
32. Kriska AM, Caspersen CJ. Introduction to a collection of physical activity questionnaires. Med Sci Sports Exerc. 1997;29 (Suppl 6):S5-9.
33. Masse LC, Ainsworth BE, Tortolero S, Levin S, Fulton JE, Henderson KA, *et al*. Measuring physical activity in midlife, older, and minority women: issues from an expert panel. J Womens Health. 1998 Feb;7(1):57-67. PubMed PMID: 9511133.
34. Craig CL, Marshal AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, *et al*. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc. 2003 Aug;35(8):1381-95. PubMed PMID: 12900694. doi: 10.1249/01.MSS.0000078924.61453.FB
35. Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr. 1982 Nov;36(5):936-42. PubMed PMID: 7137077
36. Ainsworth BE, Sternfeld B, Richardson MT, Jackson K. Evaluation of the Kaiser Physical Activity Survey in women. Med Sci Sports Exerc. 2000 Jul;32(7):1327-38. PubMed PMID: 10912901

Sažetak

UPITNICI ZA PROCJENU TJELESNE AKTIVNOSTI U TRUDNOĆI

S. Schuster, I. Šklempe Kokić i *J. Sindik*

Tjelesna aktivnost (TA) tijekom normalne trudnoće ima niz pozitivnih učinaka na zdravlje trudnice. Upitnicima je moguće prikupljati informacije o načinu, frekvenciji, trajanju i percipiranom intenzitetu TA. Svrha ovoga članka je dati pregled valjanih i pouzdanih upitnika za primjenu u trudnoći. Za potrebe izrade ovoga članka, s ciljem pronalaženja studija o mjernim svojstvima upitnika za procjenu TA u trudnoći, tijekom 2013. godine pretražene su baze Pubmed, OvidSP i Web of Science. Kriterije za uključivanje ispunilo je 6 studija i 4 upitnika: *Pregnancy Physical Activity Questionnaire, Modified Kaiser Physical Activity Survey, Short Pregnancy Leisure Time Physical Activity Questionnaire* i *Third Pregnancy Infection and Nutrition Study Physical Activity Questionnaire*. Uzorak za analizu valjanosti varirao je od 45 do 177 ispitanica. Rezultati su se kretali od slabog do umjerenog slaganja, a iskazani su kao Spearmanov ili Pearsonov koeficijent korelacije. Uzorak za analizu pouzdanosti kretao se od 49 do 109 ispitanica te su procjene pouzdanosti iskazane kao koeficijenti intraklasne korelacije s utvrđenom značajnom do većom povezanošću. Upitnici za procjenu TA su praktični i financijski pristupačni, ali skloni greškama u mjerenju i pristranosti. Ovo se djelomice može spriječiti primjenom valjanih i pouzdanih upitnika za procjenu TA. Zaključno, upitnici su jednostavni, učinkoviti, ali još uvijek alat s ograničenjima za procjenu TA.

Ključne riječi: *Motorika; Vježbanje; Trudnoća; Ankete i upitnici; Samoprocjena*