Comparison of total and activity energy expenditure estimates from physical activity questionnaires and doubly labeled water: a systematic review and meta-analysis

Mohammad Sharifzadeh[1], Minoo Bagheri[2], John R. Speakman[3,4], Kurosh Djafarian[5]

[1]Department of Clinical Nutrition, School of Nutritional Sciences and Dietetics, Tehran University of Medical Sciences (TUMS), Tehran, Iran

[2]Department of Medicine, Division of Cardiovascular Medicine, Vanderbilt University Medical Center, US

[3]State Key Laboratory of Molecular Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China

[4]Institute of Biological and Environmental Sciences, University of Aberdeen, Scotland, UK

[5]Department of Clinical Nutrition, School of Nutritional Sciences and Dietetics, Tehran University of Medical Sciences (TUMS), Tehran, Iran, No 44, Hojjat-dost Alley, Naderi St., Keshavarz Blvd, Tehran 1416-643931, Iran

Corresponding Author:

Department of Clinical Nutrition, School of Nutritional Sciences and Dietetics, Tehran University of Medical Sciences (TUMS), Tehran, Iran, No 44, Hojjat-dost Alley, Naderi St., Keshavarz Blvd, Tehran 1416-643931, Iran. Email: kdjafarian@tums.ac.ir
Telephone number: 0098 (0) 21 6465405
Fax Number: 0098 (0) 21 88974462

This peer-reviewed article has been accepted for publication but not yet copyedited or typeset, and so may be subject to change during the production process. The article is considered published and may be cited using its DOI
10.1017/S0007114520003049
The British Journal of Nutrition is published by Cambridge University Press on behalf of The Nutrition Society

Physical activity questionnaires (PAQs) could be suitable tools in free-living people for measures of physical activity, total and activity energy expenditure (TEE and AEE). This meta-analysis was performed to determine valid PAQs for estimating TEE and AEE using doubly labeled water (DLW). We identified data from relevant studies by searching Google Scholar, PubMed and Scopus databases. This revealed 38 studies that had validated PAQs with DLW, and reported the mean differences between PAQs and DLW measures of TEE (TEE_{DLW} -TEE_{PAQ}) and AEE (AEE_{DLW} -AEE_{PAQ}). We assessed 78 PAQs consisting of 59 PAQs that assessed TEE and 35 PAQs that examined AEE. There was no significant difference between TEE_{PAQs} and TEE_{DLW} with a weighted mean difference of -243.3 and a range of -841.4 to 354.6 kJ/day, and a significant weighted mean difference of AEE_{DLW} – AEE_{PAQs} 414.6 and a range of 78.7 to 750.5 . To determine whether any PAQ was a valid tool for estimating TEE and AEE, we carried out a subgroup analysis by type of PAQ. Only Active-Q, administered in two seasons, and Three-day PA diaries were correlated with TEE by DLW at the population level, however, these two PAQs did not demostrate an acceptable limit of agreement at individual level. For AEE, no PAQ was correlated with DLW either at the population or at the individual levels. Active-Q and Three-day PA diaries were identified as the only valid PAQs for TEE estimation. Further well designed studies are needed to verify this result and identify additional valid PAQs

Introduction

Total Energy Expenditure (TEE) consists of three components: Basal metabolic rate (BMR or BEE) ≈ 60-75% of TEE, activity energy expenditure (AEE) ≈15-30% of TEE and dietary thermogenesis ≈10% of TEE(1, 2). TEE, BEE and AEE change during the life course and are different between the sexes, with males usually higher than females, and older individuals lower than younger ones(3). TEE and AEE may also be affected by different disease states (4). BEE as a part of TEE decreases with age and this age-related reduction is affected by sex and body composition (5, 6). TEE is balanced by energy intake. When this balance is disrupted individuals become obese(7)

One of the most important means of decreasing risk of diabetes and cardiovascular diseases is to increase physical activity (8, 9). Also, previous research demonstrated that TEE changes in some diseases, including advance pancreatic cancer, sepsis (10, 11) and resistant training (12). Therefore, measuring TEE and PA are essential to set up efficient strategies for prevention and treatment of these disorders. The gold standard method for assessing TEE (and AEE by difference between TEE and BEE) is the doubly-labeled water method (13). DLW can also be used to estimate food intake rates as individuals are generally in energy balance during measurements. However this technique is relatively expensive (currently around 500-800US$ per subject) and hence is unsuitable for large scale survey work. As an alternative self-report questionnaires are often used in epidemiological studies to assess physical activity levels and food intake, and these may be extended to estimate AEE. In addition, since AEE is the most variable part of the TEE, they are also often used to evaluate TEE (14-16). Questionnaires are advantageous because they are inexpensive, relatively easy to administer and generally well tolerated by participants (17-19). However, self-report questionnaires for food intake have come under considerable criticism recently, because people are unreliable monitors of their own behavior and have poor recall of detailed past events. ***Research demonstrated that self report questionnaires were not reliable measures of not only food intake*** (20), ***but also physical activity*** (21)**.** Previous comparisons of physical activity questionnaires (PAQ_s) and DLW have shown that misreporting of energy expenditure by PAQs is also common (21).

Physical activity questionnaires are being developed continuously and hence it is necessary to validate which PAQs provide valid estimates of TEE and AEE (22) by comparison to the gold standard DLW methodology. Systematic reviews conducted a decade ago by Nielsen et al. [1] and Prince et al. (23), examined the correlation between self-report (PAQ) and direct measures of adult physical activity. The latter study focused on the ineffectiveness of self-

report assessment tools of physical activity. At present , the validity and reliability of many recently developed PAQs has not been established. Furthermore, it is unknown if these questionnaires are valid to evaluate TEE and AEE in either clinical settings or epidemiological studies (1). Some PAQs may be useful in epidemiological studies, and some in individual studies like clinical research. To find PAQs suitable for these two kinds of studies, we need to follow two criteria; first, at the population level, suitable PAQs must have a mean difference of <10% in differences with a gold standard method like DLW and a Spearman correlation of >0.6 [1]. At the individual level PAQs must have a acceptable limit of agreement which can be defined by the Bland-Altman method(21). Therefore, the purpose of the present work was to perform a meta-analysis of studies exploring the validity of existing PAQs to estimate TEE and/or AEE, across all age groups.

Methods

Search strategy

The following databases were searched to identify studies published up to 2 october 2019: Google Scholar, PubMed and Scopus database using the following lists and terms:

List A: "Doubly labeled water" OR "doubly-labeled water" OR "isotope labeled water" OR "doubly labelled water"

List B: "Activity monitor*" OR "physical Activity*" OR "Motor Activity*" OR “physical activity level" OR "Activity energy expenditure"

List C: "Energy expenditure" OR “TEE”

List D: "Resting metabolic rate"

List E:”Questionnaire*” OR “Survey” OR “Record” OR “Recall”

List F: valid*

Key search terms in Lists A, B, C, D, E and F were combined together.

Three independent reviewers screened the studies and extracted relevant research. When duplicate reports were removed, the full-texts of studies were further assessed to extract the required data for the current study.

We included studies that A) validated PAQs with DLW based on measurements of TEE and/or AEE and B) included PAQs that calculated TEE or AEE. Our search was limited to studies written in English, with no constraint on publication year, and with no restriction on subject age, disease status , sex and gestation and lactation status.

Data extraction

We extracted the following information from each study: Publication year, country, sample size, sex, mean (± standard deviation [SD]) age, weight, body mass index (BMI; in kg/m^2),

Body fat percent (BF %), (Table 1), TEE (in kJ/d), (Table 2), AEE (in kJ/d) measured by both DLW and PAQ (Table 3).

Quality assessment

The quality of each eligible study was assessed using Newcastle-Ottawa Scale adapted for cross-sectional studies (24). This quality assessment was performed based on seven questions in three main domains including selection, comparability and outcome (Supplementary Table 1).

Statistical analysis

In our meta-analysis, the means and SDs of the differences in TEEs or AEE measured by PAQ and DLW (the study outcome) were pooled using the weighted averages of the mean differences. Between-study heterogeneity was assessed using Cochran's Q test and I squared. According to previous research, we considered I^2 values of 25%, 50% and 75% as low, moderate and high heterogeneity, respectively (25). Random-effects models (Der- Simonian-Laird approach) were administered if heterogeneity was significant (26). To explore potential sources of heterogeneity, we performed subgroup analysis with the following covariates: sex, age, BMI, disease and body fat. Age was categorized as < 13, ≥13 and < 24, ≥ 24 and < 44, ≥ 44 and < 64 and ≥ 65 years. Subgroup analysis according to type of diseases was also conducted by classifying studies based on the health status of the study population: healthy or having either chronic kidney disease (CKD) or spinal cord injury. BMI was classified as BMI< 18.5, 18.5 ≤ BMI < 25, 25 ≤ BMI < 30 and 30 ≤ BMI< 35 and body fat percent diveded to the follwoing groups 15≤ body fat <25, 25≤ body fat <35 and body fat ≥35. All statistical tests for this meta-analysis were performed using STATA software (version 14.0; Stata Corporation, College Station, TX, USA).

Results

We identified 1780 studies of which 69 were identified in PubMed and 1711 in Scopus and Google Scholar. A total of 113 studies remained after a preliminary title and abstract review, 75 records were excluded from our analysis since they didn't report TEE or AEE (n=15) or didn't validate self-report measures with DLW (n=31) or didn't use PAQs (n=13) or reported AEE in an inappropriate way like PA score or MET (Metabolic Equivalent) category (n=16). In the end, 38 articles met the inclusion criteria of our study and were considered for further assessment (Figure 1).

Study characteristics

The 38 studies included 5997 individuals. There were seven studies performed in Sweden (27-33), 1 in Australia (34), 1 in France (35), 17 in the United States (36-50), 1 in Canada

(51), 1 in New Zealand (52,), 1 in Brazil (53), 3 in the United Kingdom (54-56), 1 in China (57), 1 in India (58), 2 in the Netherlands (59, 60), 1 in Japan (61), and 1 in Finland (62). For studies that included more than one PAQs, each of these PAQs was entered separately into our meta-analysis. Therefore, the total number of PAQs extracted for the analysis was 78. Of these, 59 of the PAQs reported TEE and 35 of them reported AEE. Forty different PAQs were identified. Thirty-one PAQs included women only, 25 included men only and the remaining 22 included both sexes. The mean age of the study population that was reported in 64 studies using PAQs ranged from 8.2 to 73.4 years. The mean BMI that was recorded in 57 studies using PAQs ranged from 16 to 34 kg/m2. The Mean BF that was recorded in 42 studies ranged from 14 to 44 (%).

Main analysis

Forest plots of the mean differences between the estimates of DLW and PAQ measures of TEE are shown in **Figure 2**. The weighted mean difference was not significant between TEE_{DLW} -TEE_{PAQ} (WMD : -243, 95% CI (-841.4 to 354.6), I^2=97.9%, p<0.0001). The mean differences between the estimates of AEE_{DLW} and AEE_{PAQs} are shown in **Figure 3**. A significant difference was found between AEEs examined by various indirect measures and the direct measures derived from DLW (WMD: 414.6, 95%CI(78.7 to 750.5), I^2= 92%, p<0.001) in which AEE assessed by DLW was higher than that of measured by PAQ.

Subgroup analysis

Since we observed significant between-study heterogeneity for both TEE and AEE, we examined possible sources of heterogeneity within the included studies using subgroup analyses. We conducted subgroup analysis to explore the effect of PAQ types on the mean difference between the estimates of TEE and AEE measured by DLW and PAQ (Table 4, 5). In 13 studies that reported information at the individual level, agreement, only 2 of them showed good agreement. In the study that was conducted by Conway et al (63) on 24 subjects, for 10 subjects the difference between TEE_{DLW} and TEE7-dPArecord was <10%, and also in the study conducted by Sridharan et al, (2015) A (64) RPAQ had a narrow limit of agreement with a mean bias of 451 kj/day(6%). In the group level, our findings indicated that heterogeneity disappeared in five subgroups of TEE_{PAQ} types including Physical activity questionnaire for adolescents (PAQA), Active-Q, 7d physical activity record (7-dPArecord), the Sedentary Time and Activity Reporting Questionnaire (STAR-Q) and Three-day PA diaries. Weighted mean differences of TEE were significant for PAQA, 7-dPArecord, STAR-Q and non-significant for Active Q(0.403) and Three-day PA diaries (0.341). Active Q and Three-

day PA diaries were the only PAQ where their estimated report of TEE was within the prespecified minimum difference with TEE_{DLW}.

Also heterogeneity disappeared in one of the AEE_{PAQ} types (STAR-Q) but the weighted mean differences of AEE were significant for this questionnaire. Also, for AEE only 8 studies reported information at the individual level and non of them showed acceptable agreement.

Additional subgroup analyses were also performed by comparing results grouped by sex, age, BMI, disease and body fat (Tables 6, 7). Results showed that mean differences between PAQ and DLW to estimate TEE may be different based on age groups. Differences was significant only in those who were in the range of 13<age<24. Although BMI was not source of heterogeneity, there was significant difference between PAQ and DLW for estimating TEE in those who were overweight.

Subgroup analysis was performed to find potential sources of heterogeneity for the mean differences between physical activity questionnaire (PAQ) and doubly labeled water (DLW) estimates of Activity energy expenditure (AEE). Results showed that all the predefined criteria were potential sources of heterogeneity except for sex. According to the subgroup analysis the greatest differences were observed in women, aged more that 44 years old, all categories of BMI except those who were overweight, healthy people, and body fat percent between 25< body fat <35.

Discussion

In this meta-analysis, we identified Active-Q and Three-day PA diaries as indirect tools that had acceptable mean differences and heterogeneity for measuring TEE at the population level. Subgroup analyses showed that the weighted mean difference in TEE measured by PAQs and DLW was influenced by age and disease status, but not by sex and the percentage of body fat (BF%). Moreover, except for sex, all of other pre-defined criteria including age, disease status, BMI and percentage of body fat were potential sources of heterogeneity.

According to previous studies, a PAQ was considered useful for estimating TEE at population level for epidemiological study if the percentage difference in means between TEE_{DLW} and TEE_{PAQ} $[(TEE_{DLW}\_TEE_{PAQ})/ TEE\_DLW] \times 100\%$ was <10% and correlations between these two estimations were > 0.60 (1). More precisely, there are some criteria that explain how good a PAQ is at the individual level and illustrate whether the questionnare is good for clinical purposes. To compare two measurements methods, a Bland-Altman plot or 'difference plot' might be used. A wide limit of agreement in this method represents PAQs are not suitable for the clinical and individual purpose. Acceptable limit of agreement is defined as a 10% of mean difference for example in the study by Bonnefoy et al, (2001) (65),

QAPSE questionnaire underestimated TEE by 358 kcal/d with limit of agreement -1075 to 1625 which means that QAPSE have wide limit of agreement for this purpose (1). In the small number of questionnaires validated against DLW, few studies have demonstrated Spearman correlation coefficients above 0.60 (RPAQ (r=0.67) (66), MARCA (r=0.7) (67), SAPAC (r=0.6) (68), MNLTPA (r=0.73) (69), 3-day activity registration (r=0.98) and JALSPAQ (r=0.742) (70)).

To estimate AEE, we did not find any PAQ as a suitable measure. Moreover, none of the questionaires estimating AEE showed acceptable correlation with DLW. Subgroup analyses showed that, in the AEE_{PAQ} group, the weighted mean difference was influenced by age, disease status, BMI and percentage of body fat.

All the studies included in the review by Nielson et al. (1) were evaluated based on the two methods of finding a good PAQ for TEE and AEE estimation: correlation coefficient and mean difference. Also these studies were divided into two groups; the first group included AEE and DLW and the second group was composed of TEE and DLW. The emphasis in the review by Nielson et al. (1) was on the first group. Furthermore, in another study by Prince et al. (23), only AEE was compared with DLW. In our study, the difference between $TEE_{DLW} - TEE_{PAQ}$ and $AEE_{DLW} - AEE_{PAQ}$ were both evaluated and the included PAQs were further assessed using a classification based on their types. Previous reviews were limited by small sample sizes (1), sex (they included studies conducted exclusively on women) and age (1, 23). In our study, however, we did not have any limitation regarding these parameters.

Studies used both predicted and measured (assessed by indirect calorimetry) resting metabolic rate (RMR) for estimating TEE and AEE but as PAQs are considered as feasible approaches to be used in epidemiological studies, it is more sensible to use predicted RMR (RMRp) rather than measured RMR (RMRm) (71). To reduce the level of over and underestimation of TEE and AEE that are blinded to the use of PAQs in different population with diverse specifications, the best PAQ with the lowest mean differences with DLW should be identified and utilized in epidemiological studies.

There are several causes for over and underestimation of TEE and AEE that are measured with PAQs. First, most equations used to measure RMRp, overestimated the BMR compared with the indirect calorimetry, including Schofield et al (72), Henry et al (73), WHO (74), Schofield BW (body weight) and ht(height) (72), WHO BW and ht (74) (in these equations age is an essential parameter and some of them need height or weight for calculating RMR), On the other hand, Molnar's equation (75) yielded a lower RMR compared with the indirect

calorimetry. In fact, use of this equation is one of the important factors leading to an underestimation in TEE (23). Of the 46 PAQs types which were assessed in our study, 25 underestimated and 21 overestimated TEE. Therefore, both underreporting and overreporting of activities were observed with respect to mean difference of(TEE_{DLW} -TEE_{PAQ})and (AEE_{DLW}- AEE_{PAQ}). This pattern is inconsistent with self reported food intake questionnaires in which underreporting is far more common. Second, consistent with our findings, Neilson et al. (1) revealed that lower body weight was associated with smaller mean differences between AEE_{PAQ} and TEE_{DLW}. Likewise, the study by Walsh et al. (42) demonstrated that the order of total energy expenditure (TEE) overestimation (large mean differences between TEE_{PAQ} and TEE_{DLW}) in premenopausal women from highest to lowest was observed in overweight black, overweight white, lean white, and lean black women. In fact, for overweight women the TEE was overestimated 49% more than normal weight control subjects (42). After weight loss, the TEE overestimation in white women was reduced by 48% whereas it did not significantly change in black women (42). Therefore, PAQ may not be a suitable tool for estimating TEE in black women. Another study conducted in obese women reported a TEE overestimation but following a 12-week weight-reducing diet, the participants underestimated TEE (the mean difference decreased from 205 kJ/day to 50 kJ/day). Third, all of the included articles used MET values for calculating TEE except for the studies by Barnard et al. (76) and Bonnefoy et al. (35) (that used the physical activity level) and Walsh et al.(42) (that used the instructions described in the study by Montoya et al) (77). In most physical activity questionnaires, the use of MET values for estimating the energy expenditure of a particular activity is considered a limitation (42). When the MET value is administered for a specific activity, the same energy cost per kilogram of body weight is calculated for all participants, regardless of differences in metabolic rate and this might be the reason attributed to the decrease in TEE overestimation in obese women after weight loss (42).

For TEE, we observed that only two PAQs had the least mean difference with DLW and none of the PAQs showed good measure of AEE. This is because the magnitude of difference between PAQs and DLW estimates of TEE and AEE depends on some factors including the type of PAQs, the sex of the population on which the questionnaire was used and the number of activities measured by the PAQs. For instance, when the 7D-PAR was used, mean daily EE was overestimated in women while it was underestimated in men (1). Also, for the questionnaires Tecumseh Occupational (past year) and Minnesota Leisure Time (past month)

which measured sleep and general activities, when watching TV, reading, and childcare activities were ignored from EE calculated by these questionnaires, an excellent agreement with DLW measure of TEE was obtained (36). As some PAQs do not estimate all physical activity especially in low-intensity level, an underreporting of AEE is anticipated (23). However, some PAQs like IPAQ and PAQA can capture low to high-intensity level physical activities and the underreporting of TEE in these questionnaires is compensated by over reporting of vigorous physical activity (78).

In conclusion, our meta-analysis identified PAQ (Active-Q) and Three-day PA diaries that had sufficient validity for measuring TEE based on the mean correspondence in group level. However, as each of these questionnaires were used only in one study we may conclude that this finding might be due to a chance and requires further verification. This study provides evidence highlighting that the majority of PAQs compared to DLW might not be qualified tools for estimating TEE or AEE. Therefore, it is recommended that until further research is performed to investigate the agreement between direct and indirect measures of TEE and AEE, the use of either Active-Q and Three-day PA diaries or direct measurement methods in epidemiological studies might yield more reliable findings.

Abbreviations: Total Energy Expenditure (TEE), Basal metabolic rate (BMR), Activity energy expenditure (AEE), Doubly labeled water (DLW), standard deviation (SD), body mass index (BMI), Body fat percent (BF %), Chronic kidney disease (CKD), measured RMR (RMRm), predicted RMR (RMRp), BW (body weight), ht(height), Metabolic Equivalent (MET), Physical activity questionnaire for adolescents(PAQA), Modifiable activity questionnaire(MAQ), recent physical activity questionnaire(RPAQ), 7-day physical activity recall questionnaire(7d-PAR), Questionnaire d'Activité physique saint-etienne(QAPSE), (TEC = Tecumseh occupational activity questionnaire) + (MNLTPA = Minnesota leisure time physical activity questionnaire) + (EE SLEEP = EE from sleep) (TEC+MNLTPA+EESLEEP), 7-day physical activity record questionnaire(7-dPArecord), Sedentary time and activity reporting questionnaire(STAR-Q), Multimedia activity recall for children and adolescents(MARCA), 24-h Physical activity diaries (PAD), Self-Administered physical activity checklist(SAPAC), Physical activity recall assessment for people with spinal cord injury(PARA-SCI), Physical activity scale for individuals with physical disabilities(PASIPD), Five City Project Questionnaire(FCQ), modified Yale PhysicalActivity Survey(Modified YAPS), the Japan Arteriosclerosis Longitudinal Study Physical Activity Questionnaire(JALSPAQ), Cross-Cultural Activity Participation Study (CAPS)

Declarations

Ethics Approval

Meta-analysis collects relevant data from published articles, and thus, no ethics committee approval was need for this meta-analysis.

Consent to Participate

Not applicable

Conflict of Interest

They had no conflict of interests to disclose

Authors' contributions

These authors contributed equally to this work

Acknowledgements

The authors would like to thank all research staff involved in the study, and tehran university of medical science financial support.

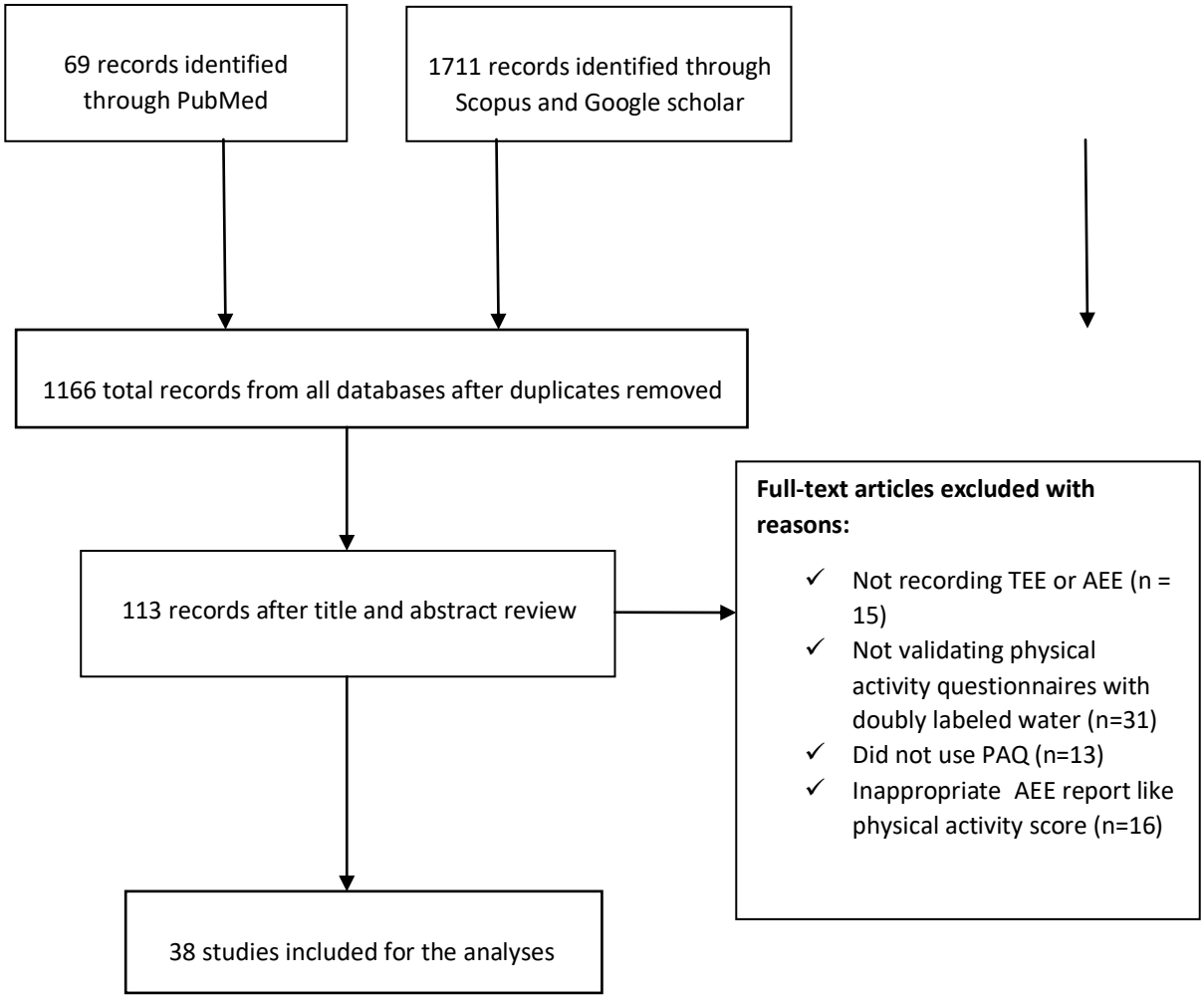
Funding

This article was financially supported by Tehran university of medical science

Availability of data and materials

Not Applicable

figure 1- Study selection process.



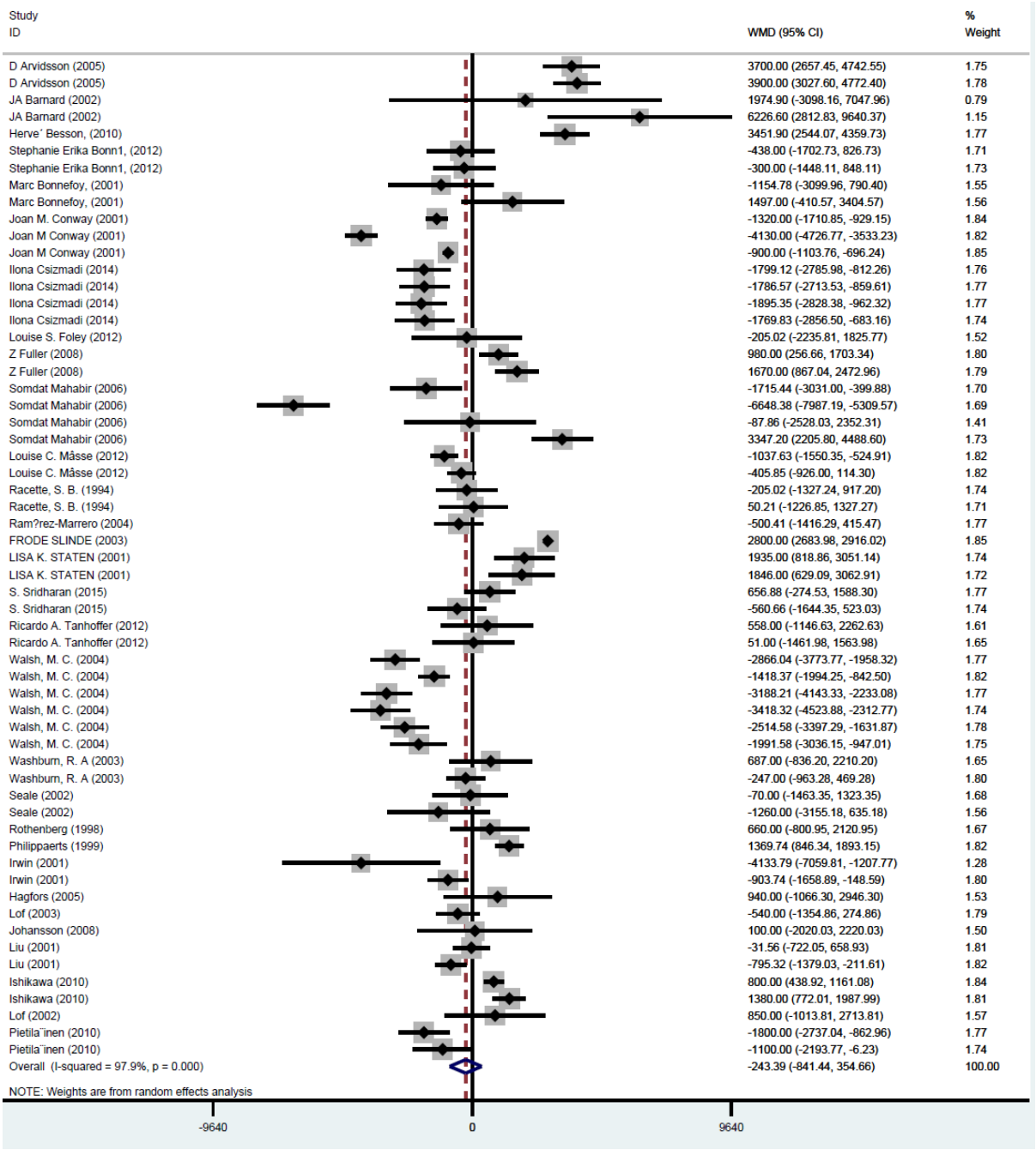


Figure 2- Forest plots of mean difference of TEE_{DLW} and TEE_{PAQ}

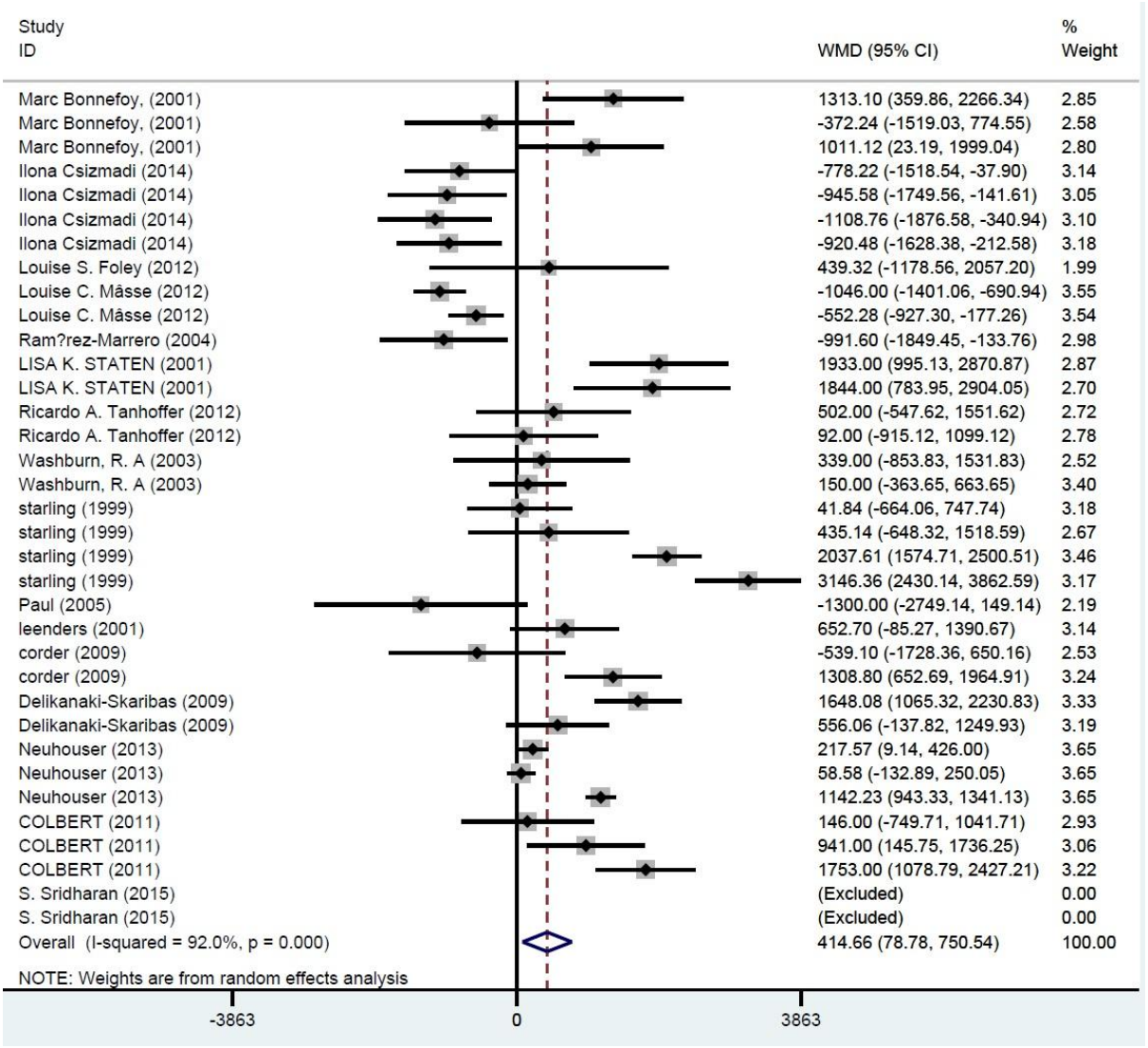


Figure 3- Forest plots of mean difference of AEE_{DLW} and AEE_{PAQ}

Table 1- Characteristics* of the studies included into the meta-analysis

Study	Sample size	Sex	Heath status of the participants	Age	BMI	Weight	Body fat (%)
Arvidsson et al, (2005) A (78)	17	Boy	Healthy	15.8	21±2.6	64.1±9	16.4±4.7
Arvidsson et al, (2005)B (78)	16	Girl	Healthy	15.7	21±2.7	56.4±9.4	27.5±5.2
Barnard et al, (2002) A (76)	8	Men	Healthy	35.4	25.9±3.9	Not reported	21.9±6.8
Barnard et al, (2002)B (76)	7	Women	Healthy	37.1	23.8±5.3	Not reported	38.4±9
Besson et al, (2010) A (66)	50	Men(50%) Women(50%)	Healthy	34.3	25.1±3.1	Not reported	22±7.9
Erika Bonn1 et al, (2012) A (65)	37	Men (19%) Women(81%)	Healthy	20-65	Not reported	Not reported	Not reported
Erika Bonn1 et al, (2012) B (65)	37	Men (19%) Women (81%)	Healthy	21-65	Not reported	Not reported	Not reported
Bonnefoy et al, (2001) A, B, C, D, E (65)	19	Men	Healthy	73.4	Not reported	74.3±9.7	Not reported
Conway et al, (2002) A (36)	24	Men	Healthy	42	25.6±.6	81.5±2.1	21.1±6.8
Conway et al, (2002) B,C (63)	24	Men	Healthy	41.2	25.1±.5	79.5±1.8	Not reported
Csizmadi et al, (2014) A, B, C, D (79)	102	Men (86%) Women (14%)	Healthy	48	24±.3	Not reported	Not reported
Foley et al, (2012) (67)	32	Men (56%) Women (44%)	Healthy	14.3	20.3±3.3	57±16	17.3±7
Fuller et al, (2008) A (80)	59	Men (51%) Women (49%)	Healthy	42.7	22.3±2.25	77.1±9.6	27.2±2.9
Fuller et al, (2008) B (80)	59	Men (51%) Women (49%)	Healthy	42.7	22.3±2.25	77.1±9.6	27.2±2.9
Mahabir et al, (2006) A, B, C, D (37)	65	Women	Postmenopausal	59.9	27.7±5.6	Not reported	41.2±8.6
Mâsse et al, (2012) A, B (81)	130	Women	Healthy	49.2	30±6.3	76.9±17.3	Not reported
Racette et al, (1994) A (39)	14	Women	Healthy	40	34±8.8	91.2±.06	44.8±2.9
Racette et al, (1994) B (39)	14	Women	Healthy	40	30.2±4.48	81±4.48	Not reported
Marrero et al, (2004) (68)	12	Men (43%) Women (57%)	Healthy	8.18	16.7±9.5	27.6±5.45	Not reported
SLINDE et al, (2003) (69)	2400	Boys (48%) Girls (52%)	Health	15	20.8±2.6	60.4±9.6	Not reported
Staten et al, (2001) A, B (82)	35	Women	Healthy	43.8	28±8.1	73±20.4	Not reported
Sridharan et al, (2015) A, B (64)	40	Men (55%) Women (45%)	Chronic kidney disease (stages 1–5)	54	26.8±4.2	77.1±12.2	Not reported

Tanhoffer et al, (2012) A, B (83)	14	Men (93%) Women (7%)	Spinal cord injury	40	25±3	79±15	33±9
Walsh et al, (2004) A (42)	21	Women	Healthy	36.5	29.1±1.7	73±20.4	42.6±3.6
Walsh et al, (2004) B (42)	21	Women	Healthy	36.5	23.9±1.1	78.7±5.3	32.5±4.7
Walsh et al, (2004) C (42)	20	Women	Healthy	36	28.6±1.8	65.2±4.5	42.1±3.7
Walsh et al, (2004) D (42)	20	Women	Healthy	36	24±.9	78±9.2	33.1±4.5
Walsh et al, (2004) E (42)	20	Women	Healthy	31.8	23.1±1	65.5±7.9	32.4±4
Walsh et al, (2004) F (42)	14	Women	Healthy	31.8	23±1.6	62.3±4.7	31.5±5.3
Washburn et al, (2003) A (84)	17	Men	Healthy	23.9	29.8±2.7	62.3±4.7	28.2±4.7
Washburn et al, (2003) B (84)	29	Women	Healthy	23.3	29.4±2.8	95.1±11.9	36.6±4.2
Starling et al, (1999) A, B (85)	35	Women	Healthy	67	24.8±3.9	63.9±10.2	35±8
Starling et al, (1999) C, D (85)	32	Men	Healthy	66	25.7±4.5	79.5±14.5	21±7
Seale et al, (2002) A (86)	13	Women	Healthy	73.5	27.6±3.2	69.8±9.5	Not reported
Seale et al, (2002) B (86)	14	Men	Healthy	74.1	28.2±2.4	83.6±7.9	Not reported
Rothenberg et al, (1998) (30)	12	Men (40%) Women (60%)	Healthy	73	24.3	62	Not reported
Philippaerts et al, (1999) (87)	90	Men	Healthy	40	24.6±2.8	78±8	20.3
Paul et al, (2005) (47)	12	Men	Healthy	39	24.1±1.4	79.9±8.3	18.1
Leenders et al, (2001) (88)	13	Women	Healthy	25.8	23.5±.6	65.5±2	26.3
Irwin et al, (2001) A, B (89)	24	Men	Healthy	41.2	25.1±2.7	79.5±9	21.1
Hagfors et al, (2005) (90)	9	Men (60%) Women (40%)	Healthy	58.8	28.1±4.4	77.8±14.1	Not reported
Lof et al, (2003) (91)	34	Women	Healthy	30	24±4	67±10	34±8
Corder et al, (2010) A (92)	13	Men	Healthy	15.9	17.4±2.6	46.1±7.1	14.3±10
Corder et al, (2010) B (92)	15	Women	Healthy	15.7	20.8±4.2	49.4±12.5	29.8±8.7
Skaribas et al, (2009) A, B (93)	20	Men	Healthy	72.9	Not reported	77.4±9.5	24.2±7.9
Johansson et al, (2008) (94)	9	Men (34%) Women (66%)	Healthy	60	27.4±4.5	Not reported	Not reported
Liu et al, (2001) A (95)	18	Women	Renal, cancer, healthy	64-84	Not reported	Not reported	Not reported
Liu et al, (2001) B (95)	13	Men	Renal, cancer, healthy	64-84	Not reported	Not reported	Not reported
Neuhouser et al, (2013) A (96)	450	Women	Healthy	50-80	Not reported	Not reported	Not reported
Neuhouser et al, (2013) B (96)	444	Women	Healthy	50-81	Not reported	Not reported	Not reported
Neuhouser et al, (2013) C	426	Women	Healthy	50-82	Not	Not	Not reported

(96)					reported	reported	
Ishikawa et al, (2010) A(70)	118	Women	Healthy	50.4	22.3±2.5	52.7±7.3	Not reported
Ishikawa et al, (2010) B (70)	108	men	Healthy	50.4	23±3	66.6±10.9	Not reported
Colbert et al, (2011) A (97)	56	Women(79%) Men(21%)	Healthy	74.7	25.8±4.2	69.2±14.5	Not reported
Colbert et al, (2011) B (97)	56	Women(79%) Men(21%)	Healthy	74.7	25.8±4.2	69.2±14.5	Not reported
Colbert et al, (2011) C (97)	56	Women(79%) Men(21%)	Healthy	74.7	25.8±4.2	69.2±14.5	Not reported
Lof et al, (2002) (98)	24	women	Healthy	30	24±4	67±10	Not reported
Pietila¨inen et al, (2010) A (99)	7	men	Healthy	25.5	30±0.5	88±2.3	38.3±1.8
Pietila¨inen et al, (2010) B (99)	7	men	Healthy	25.5	25±0.5	73±2.3	29.4±2.3

* n (%), mean or mean ± standard deviation (SD)

Table 2- Summary of results for the difference in total energy expenditure (TEE) means between physical activity questionnaires (PAQs) and doubly labeled water (DLW). All data in KJ/day.

Study	PAQ type	TEE_{DLW}	TEE_{PAQ}
Arvidsson et al, (2005) A (78)	PAQA	11300± 1500	7600±1600
Arvidsson et al, (2005)B (78)	PAQA	9100±1400	5200±1100
Barnard et al, (2002) A (76)	MAQ	29409± 6857.9	15243.6±2562.3
Barnard et al, (2002)B (76)	MAQ	15189.4±4531.7	8962.8±836.4
Besson et al, (2010) A (66)	RPAQ	11967.9±2574.1	8516±2025.1
Bonn1 et al, (2012) A (65)	Active-Q	11229±2256	11667±3212
Bonn1 et al, (2012) B (65)	Active-Q	11229±2256	11529±2758
Bonnefoy et al, (2001) B (65)	7d-PAQR	11181±1647	12335.78±1658.4
Bonnefoy et al, (2001) D (65)	QAPSE	11181±1647	9684±856.017
Conway et al, (2002) A (36)	(TEC+MNLTPA+EESLEEP+EEGEN)	13550±380	14870±900
Conway et al, (2002) B (63)	7-dPAR	13270±350	17400±1450
Conway et al, (2002) C (63)	7-dPArecord	13270± 350	14170±370
Csizmadi et al, (2014) A (79)	STAR-Q	11748.67±3213.31	13547.79±3941.33
Csizmadi et al, (2014) B (79)	STAR-Q	11748.67±3213.31	13535.24±3338.83
Csizmadi et al, (2014) C (79)	STAR-Q	11748.67±3213.31	13644.02±3414.14
Csizmadi et al, (2014) D (79)	7d-PAQR	11748.67±3213.31	13518.50±4619.14
Foley et al, (2012) (67)	MARCA	13346.96±3778.15	13551.98±4481.064
Fuller et al, (2008) A (80)	24-h PAD	11030±2190	10050±1800
Fuller et al, (2008) B (80)	7-dPAR	11040±2200	9370±2250
Mahabir et al, (2006) A (37)	Five City Project questionnaire	10711.04±2602.45	12426.48±4744.656
Mahabir et al, (2006) B (37)	Harvard Alumni questionnaire	10711.04±2602.45	17359.42±4853.44
Mahabir et al, (2006) C (37)	CAPS Study Four Week Activity Recall	10711.04±2602.45	10798.9±9694.328
Mahabir et al, (2006) D (37)	CAPS Study Typical Week Activity Recall	10711.04±2602.45	7363.84±3907.86
Mâsse et al, (2012) A (81)	The Checklist questionnaire	9552.072±1824.22	10589.7±2359.78
Mâsse et al, (2012) B (81)	Global Questionnaire	9552.072±1824.22	9957.92±2414.17
Racette et al, (1994) A (39)	7-dPAR	10945.34±1765.65	11150.36±1213.36
Racette et al, (1994) B (39)	7-dPAR	10259.17±1840.96	10208.96±1598.29
Marrero et al, (2004) (68)	SAPAC	7004.016±999.1392	7504.4224±1273.6096
SLINDE et al, (2003) (69)	MNLTPA	11400±2100	8600±2000
Staten et al, (2001) A (82)	The Arizona Activity FrequencyQuestionnaire28 day	9847±2555	7912±2196
Staten et al, (2001) B (82)	The Arizona Activity FrequencyQuestionnaire7 day	9847±2555	8001±2639
Sridharan et al, (2015) A (64)	RPAQ	10380.5±1991.58	9723.616±2250.99
Sridharan et al, (2015) B (64)	7-dPAR	10380.5±1991.58	10941.16±2874.41
Tanhoffer et al, (2012) A (83)	PARA-SCI	9817±2491	9259±2094
Tanhoffer et al, (2012) B (83)	PASIPD	9817±2491	9766±1462
Walsh et al, (2004) A (42)	TEC+MNLTPA	9347.056±1656.86	12213.1±1326.33
Walsh et al, (2004) B (42)	TEC+MNLTPA	8711.088±1071.1	10129.46±815.88
Walsh et al, (2004) C (42)	TEC+MNLTPA	8861.712±1435.11	12049.92±1640.13
Walsh et al, (2004) D (42)	TEC+MNLTPA	7819.896±1669.42	11238.22±1891.17
Walsh et al, (2004) E (42)	TEC+MNLTPA	8439.128±991.608	10953.71±1753.1
Walsh et al, (2004) F (42)	TEC+MNLTPA	8334.528±1422.56	10326.11±1397.46
Washburn et al, (2003) A (84)	7-dPAR	13885±2754	13198±1638
Washburn et al, (2003) B (84)	7-dPAR	10771±1457	11018±1323
Seale et al, (2002) A (86)	7-dPAR	9440± 900	9510± 2400
Seale et al, (2002) B (86)	7-dPAR	12430± 1630	13690± 3230
Rothenberg et al, (1998) (30)	Activity diary in four day	9900± 1430	9240± 2150
Philippaerts et al, (1999) (87)	FCQ 7 day index	13400± 1800	12030.26± 1782.8
Irwin et al, (2001) A (89)	7-dPAR	13259.10± 1719.62	17392.89± 7108.62
Irwin et al, (2001) B (89)	7-dPArecord	13259.10± 1719.62	14162.84± 778.22
Hagfors et al, (2005) (90)	3-day activity registration	10760± 2590	9820± 1650
Lof et al, (2003) (91)	2-wk Recall	10670± 1370	11210± 2000
Johansson et al, (2008) (94)	Two-question questionnaire on physical activity	10900±2700	10800±1800
Liu et al, (2001) A (95)	Modified YPAS	8121.80	8153.36±1118.38
Liu et al, (2001) B (95)	Modified YPAS	1017.20	10967.52±585.7

Ishikawa et al, (2010) A(70)	JALSPAQ	8420±1400	7620±1430
Ishikawa et al, (2010) B (70)	JALSPAQ	11210±3000	9830±1180
Lof et al, (2002) (98)	LOF questionnare	11420	10570
Pietila¨inen et al, (2010) A (99)	Three-day PA diaries.	12400±400	14200
Pietila¨inen et al, (2010) B (99)	Three-day PA diaries.	11500±700	12600

Abbreviations: PAQA = Physical activity questionnaire for adolescents; MAQ = Modifiable activity questionnaire; RPAQ = recent physical activity questionnaire; 7d-PAR = 7-day physical activity recall questionnaire; QAPSE = Questionnaire d'Activité physique saint-etienne; TEC+MNLTPA+EESLEEP = (TEC = Tecumseh occupational activity questionnaire) + (MNLTPA = Minnesota leisure time physical activity questionnaire) + (EE SLEEP = EE from sleep); 7-dPArecord = 7-day physical activity record questionnaire; STAR-Q = Sedentary time and activity reporting questionnaire; MARCA = Multimedia activity recall for children and adolescents; PAD = 24-h Physical activity diaries; SAPAC = Self-Administered physical activity checklist; PARA-SCI = Physical activity recall assessment for people with spinal cord injury; PASIPD = Physical activity scale for individuals with physical disabilities; FCQ= Five City Project Questionnaire; Modified YAPS = modified Yale PhysicalActivity Survey; JALSPAQ= the Japan Arteriosclerosis Longitudinal Study Physical Activity Questionnaire; CAPS = Cross-Cultural Activity Participation Study

Table 3- Summary of results from difference in Activity energy expenditure (AEE) means between physical activity questionnaires (PAQs) and doubly labeled water (DLW)

Study	PAQ type	AEE_{DLW}	AEE_{PAQ}
Bonnefoy et al, (2001) A (65)	MNLTPA	3367± 1940	2053.900±854.790
Bonnefoy et al, (2001) C (65)	YPAS	3367± 1940	3739.241±1655.609
Bonnefoy et al, (2001) E (65)	College Alumni Questionnaire	3367± 1940	2355.885±1031.356
Csizmadi et al, (2014) A (79)	STAR-Q	4250.944± 2765.620	5029.168±2627.550
Csizmadi et al, (2014) B (79)	STAR-Q	4250.944± 2765.620	5196.528±2916.250
Csizmadi et al, (2014) C (79)	STAR-Q	4250.944± 2765.620	5359.704±2690.310
Csizmadi et al, (2014) D(79)	7-dPAR	4250.944± 2765.620	5171.424±2405.800
Foley et al, (2012) (67)	MARCA	6895.232±3234.230	6455.912±3368.120
Mâsse et al, (2012) A (81)	Checklist questionnaire	2882.780±1292.860	3928.780±2359.78
Mâsse et al, (2012) B (81)	Global Questionnaire	2882.780±1292.860	3435.060±1757.280
Marrero et al, (2004) (68)	SAPAC	1271.936 ±778.224	2263.540 ±1301.220
Staten et al, (2001) A (82)	The Arizona Activity FrequencyQuestionnaire28 day	5578±2084	3645±1916
Staten et al, (2001) B (82)	The Arizona Activity FrequencyQuestionnaire7 day	5578±2084	3734±2428
Sridharan et al, (2015) A (64)	RPAQ	2627.550	9723.616±2250.99
Sridharan et al, (2015) B (64)	7-dPAR	2627.550	10941.16±2874.41
Tanhoffer et al, (2012) A (83)	PARA-SCI	2841±1626	2339±1171
Tanhoffer et al, (2012) B (83)	PASIPD	2841±1626	2749±1026
Washburn et al, (2003) A (84)	7-dPAR	3989 ±2461	3650 ±490
Washburn et al, (2003) B (84)	7-dPAR	3223 ±1360	3073 ±377
Starling et al, (1999) A (85)	YPAS	3652.630± 1020.9	3610.790± 1870.25
Starling et al, (1999) B (85)	YPAS	5066.824± 1794.94	4631.688± 2560.61
Starling et al, (1999) C (85)	MNLTPA	3652.630± 1020.9	1615.020± 953.952
Starling et al, (1999) D (85)	MNLTPA	5066.824± 1794.94	1920.460± 1204.99
Paul et al, (2005) (47)	7-dPArecord	10500± 1600	11800± 2000
Leenders et al, (2001) (88)	7-dPAR	3338.830± 1251.02	2686.13± 527.184
Corder et al, (2010) A (92)	Youth Physical Activity Questionnaire recall in past week	2349.2± 1187.7	2888.3±1837.3
Corder et al, (2010) B (92)	Youth Physical Activity Questionnaire recall in past week	1990.5± 1185	681.7±526
Skaribas et al, (2009) A (93)	YPAS	2179.446± 1297.04	531.368±292.88
Skaribas et al, (2009) B (93)	PASE	2179.446± 1297.04	1623.39±907.928
Neuhouser et al, (2013) A (96)	Arizona Activity Frequency Questionnaire28 day	3075.240	2857.670
Neuhouser et al, (2013) B (96)	7-dPAR	3075.240	3016.660
Neuhouser et al, (2013) C (96)	PHQ	3075.240	1933.010
Colbert et al, (2011) A (97)	YPAS	2845±1138	2699
Colbert et al, (2011) B (97)	modPASE	2845±1138	1904
Colbert et al, (2011) C (97)	CHAMPS	2845±1138	1092

Abbreviations: MNLTPA = Minnesota leisure time physical activity questionnaire; Modified YAPS = modified Yale Physical Activity Survey; STAR-Q = Sedentary time and activity reporting questionnaire; 7-dPAR = 7-day physical activity recall questionnaire; MARCA = Multimedia activity recall for children and adolescents; SAPAC = Self-Administered physical activity checklist; PARA-SCI = Physical activity recall assessment for people with spinal cord injury; PASIPD = Physical activity scale for individuals with physical disabilities; 7-dPArecord = 7-day physical activity record questionnaire; PASE=Physical Activity Scale for the Elderly; PHQ= Personal Habits Questionnaire; modPASE= modified Physical Activity Scale for theElderly; CHAMPS= Community HealthActivities Model Program for Seniors

Table 4- Agreement between physical activity questionnaire (PAQ) and doubly labeled water (DLW) estimates of total energy expenditure (TEE) stratified by PAQ type

Type of physical activity questionnaire	No. of studies	Mean difference (95% confidence interval (CI)) [kJ/day]	*P*[*]	Test of heterogeneity[†]	
				P	***I*^{2} (%)**
PAQA (78)	2	3817.631 [3148.5 4486.6]	< 0.001	0.773	0.0
MAQ (76)	2	4531.851 [451.834 8611.868]	0.029	0 .173	46.2
RPAQ (66), [27	2	2056.412[-682.65 4795.4]	0.141	< 0.001	94.4
Active-Q (65)	2	-362.345 [-1.2e+03 487.737]	0.403	0 .874	0.0
MNLTPAQ (71)	1	2800.000 [2683.978 2916.022]	< 0.001	-	-
7-dPAQ (65), (63), (80), (37), (39), (64), (84), (79), (86), (89)	12	-857.43.766[-2.1e+03 394.454]	0 .179	< 0.001	93.5
QAPSE (65)	1	1497 [-410.57 3404.56]	0 .124	-	-
(TEC+MNLTPA+EESLEEP) (36)	1	-1.3e+03 [-1.7e+03 -929.152]	< 0.001	-	-
7-dPArecord (89), (63)	2	-900.254 [-1.1e+03 -703.526]	< 0.001	.993	0.0
STAR-Q (79)	3	-1.8e+03 [-2.4e+03 -1.3e+03]	< 0.001	0.985	0.0
MARCA (67)	1	-205.020 [-2.2e+03 1825.765]	0.843	-	-
24-PAD (80)	1	980 [256.656 1703.344]	0.008	-	-
Five City Project questionnaire (37)	1	-1.7e+03 [-3.0e+03 -399.881]	0.011	-	-
Harvard Alumni questionnaire (37)	1	-6.6e+03 [-8.0e+03 -5.3e+03]	< 0.001	-	-
CAPS Four Week Activity Recall (37)	1	-87.860[-2.5e+03 2352.309]	0.944	-	-
CAPS Typical Week Activity Recall (37)	1	3347.2[2205.8 4488.6]	< 0.001		
The Checklist questionnaire (81)	1	-1.0e+03 [-1.6e+03 -524.906]	< 0.001	-	-
Global Questionnaire (81)	1	-405.848 [-925.999 114.303]	0 .126	-	-
SAPAC (68)	1	-500.406[-1.4e+03 415.472]	0 .284	-	-
Arizona Activity Frequency Questionnaire28 day (82)	1	1935[818.855 3051.145]	0 .001	-	-
Arizona Activity Frequency Questionnaire7 day (82)	1	1846[629.092 3062.908]	0.003	-	-
PARA-SCI (83)	1	558.000 [-1.1e+03 2262.631]	0 .521	-	-
PASIPD (83)	1	51.000 [-1.5e+03 1563.979]	0.947	-	-
TEC+MNLTPA (42)	6	-2.5e+03 [-3.2e+03 -1.8e+03]	< 0.001	.003	72.7

activity diary in four day (30)	1	660.000[-800.951 2120.951]	0.376	-	-
FCQ 7 day index (87)	1	1369.745[846.338 1893.152]	< 0.001	-	-
3-day activity registration (90)	1	940.000 [-1.1e+03 2946.303]	0.358	-	-
2-wk Recall (91)	1	-540.000 [-1.4e+03 274.860]	0.194	-	-
Two-question questionnaire on physical activity (94)	1	100.000 [-2.0e+03 2220.025]	0.926	-	-
Modified YPAS (95)	2	-436.627 [-1.2e+03 310.461]	0.252	0.098	63.5
JALSPAQ (70)	2	1036.305 [477.743 1594.867]	< 0.001	0.108	61.3
Lof questionnaire (98)	1	850.000 [-1.0e+03 2713.807]	0.371	-	-
Three-day PA diaries	2	-1.5e+03 [-2.2e+03 -792.095]	< 0.001	0.341	0.0

* *P* for the meta-analysis. $P < 0.05$ indicates a lack of agreement between PAQ and DLW estimates of TEE by using a random-effects model.

† P-heterogeneity: heterogeneity was evaluated using Cochran's test, and $P < 0.5$ indicates significant heterogeneity across studies.

Abbreviations: PAQA = Physical activity questionnaire for adolescents; MAQ = Modifiable activity questionnaire ; RPAQ = Recent physical activity questionnaire ; 7-dPAQ = 7-day physical activity recall questionnaire; QAPSE = Questionnaire d'Activité physique saint-etienne; TEC+MNLTPA+EESLEEP = (TEC = Tecumseh occupational activity questionnaire) + (MNLTPA = Minnesota leisure time physical activity questionnaire)+ (EE SLEEP = EE from sleep); STAR-Q = Sedentary time and activity reporting questionnaire; MARCA = Multimedia activity recall for children and adolescents; 24-PAD = 24-h Physical activity diaries; SAPAC = Self-administered physical activity checklist; PARA-SCI = Physical activity recall assessment for people with spinal cord injury; PASIPD = Physical activity scale for individuals with physical disabilities; CAPS= Cross-Cultural Activity Participation Study; JALSPAQ= the Japan Arteriosclerosis Longitudinal Study Physical Activity Questionnaire

Table 5- Agreement between physical activity questionnaire (PAQ) and doubly labeled water (DLW) estimates of Activity energy expenditure (AEE) stratified by PAQ type

Type of physical activity questionnaire	No. of studies	Mean difference (95% confidence interval (CI)) [kJ/day]	*P*[*]	Test of heterogeneity[†]	
				P	***I*^{2} (%)**
YPAS (65), (85), (97), (93)	5	433.077 [-376.955 1243.109]	0.330	0.001	78.4
College Alumni Questionnaire (65)	1	1011.115 [23.192 1999.038]	0.045	-	-
STAR-Q (79)	3	-939.945 [-1.4e+03 -495.738]	<0.001	0.831	0.0
7-dPAR (79), [55], (88), (96), (64)	6	33.070 [-369.996 436.137]	0.872	0.038	60.6
MARCA (67)	1	439.320 [-1.2e+03 2057.198]	0.595	-	-
Checklist questionnaire (81)	1	-1.0e+03 [-1.4e+03 -690.940]	<0.001	-	-
Global Questionnaire (81)	1	-552.280 [-927.303 -177.257]	0.004	-	-
SAPAC (68)	1	-991.604 [-1.8e+03 -133.759]	0.023	-	-
MNLTPA (69)	3	2198.583 [1282.793 3114.374]	<0.001	0.005	81
The Arizona ActivityFrequencyQuestionnaire28 day (82), (96)	2	1011.841 [-664.644 2688.326]	0.237	<0.001	91.8
The Arizona ActivityFrequencyQuestionnaire7 day (82)	1	1844.000 [783.949 2904.051]	0.001	-	-
PARA-SCI (83)	1	502.000 [-547.623 1551.623]	0.349	-	-
PASIPD (83)	1	92.000 [-915.123 1099.123]	0.858	-	-
7-dPArecord (47)	1	-1.3e+03 [-2.7e+03 149.137]	0.079	-	-
Youth Physical Activity Questionnaire recall in past week (92)	2	454.150 [-1.4e+03 2259.958]	0.622	0.008	85.9
PASE (93)	1	556.056 [-137.817 1249.928]	0.116	-	-
PHQ (96)	1	1142.230 [1009.320 1275.141]	<0.001	-	-
CHAMPS (97)	1	1753.000 [1078.787 2427.213]	<0.001	-	-
modPASE (97)	1	1753.000 [1078.787 2427.213]	0.020	-	-

Abbreviations: MNLTPA = Minnesota leisure time physical activity questionnaire; Modified YAPS = modified Yale Physical Activity Survey; STAR-Q = Sedentary time and activity reporting questionnaire; 7-dPAR = 7-day physical activity recall questionnaire; MARCA = Multimedia activity recall for children and adolescents; SAPAC = Self-Administered physical activity checklist; PARA-SCI = Physical activity recall assessment for people with spinal cord injury; PASIPD = Physical activity scale for individuals with physical disabilities; 7-dPArecord = 7-day physical activity record questionnaire; PASE=Physical Activity Scale for the Elderly; PHQ= Personal Habits Questionnaire; CHAMPS= Community HealthActivities Model Program for Seniors; ; modPASE= modified Physical Activity Scale for theElderly

Table 6-Subgroup analysis of mean differences between physical activity questionnaire (PAQ) and doubly labeled water (DLW) estimates of total energy expenditure (TEE) stratified by identified study characteristics

Variables	No. of studies	Mean difference (95% confidence interval (CI)) [kJ/day]	*P**	Test of heterogeneity†	
				P	I^2 (%)
Sex					
men	16	-467.036 [-1.3e+03 363.780]	0.271	< 0.001	95.5
women	24	-432.043 [-1.2e+03 344.451]	0.275	< 0.001	94.9
Men and women	19	144.580 [-856.523 1145.682]	0.777	< 0.001	96.4
Age					
Age<13	1	-500.406[-1.4e+03 415.472]	0.284	.	.
13≤age<24	6	1879.012 [541.481 3216.543]	0.006	< 0.001	94.6
24≤age<44	27	-533.133 [-1.2e+03 122.301]	0.111	< 0.001	94.8
44≤age<64	18	-596.864 [-1.4e+03 177.626]	0.131	< 0.001	93.1
Age≥64	7	-234.563 [-819.655 350.529]	0.432	0.117	41.1
BMI					
BMI< 18.5	1	-500.406[-1.4e 415.472]	0.284	.	.
18.5 ≤ BMI < 25	22	387.865 [-515.405 1291.135]	0.400	< 0.001	97.8
25 ≤ BMI < 30	25	-754.668 [-1.4e+03 -72.568]	0.030	< 0.001	93.3
30 ≤ BMI< 35	5	-742.724 [-1.3e+03 -183.225]	0.009	0.038	60.5
Disease					
Healthy	55	-244.285 [-941.282 452.712]	0.545	< 0.001	98.1
Chronic kidney disease	2	80.917[-1.1e+03 1272.354]	0.894	0.095	64.1
spinal cord injury	2	274.408[-857.147 1607.082]	0.635	0.663	0.0
Body fat (%)					
15≤ body fat <25	9	-574.335 [-1.8e+03 642.891]	0.355	< 0.001	97.1
25≤ body fat <35	13	25.160 [-1.2e+03 1249.203]	0.968	< 0.001	95.2
body fat ≥35	11	-1.0e+03 [-2.3e+03 181.036]	0.095	< 0.001	94.8

* P for the meta-analysis. $P < 0.05$ indicates a lack of agreement between PAQ and DLW estimates of TEE by using a random-effects model.

† P-heterogeneity: heterogeneity was evaluated using Cochran's test, and $P < 0.5$ indicates significant heterogeneity across studies.

Table 7-Subgroup analysis of mean differences between physical activity questionnaire (PAQ) and doubly labeled water (DLW) estimates of Activity energy expenditure (AEE) stratified by identified study characteristics

Variables	No. of studies	Mean difference (95% confidence interval (CI)) [kJ/day]	*P**	Test of heterogeneity†	
				P	I^2 (%)
Sex					
men	10	702.976 [-79.624 1485.576]	0.078	< 0.001	86
women	12	591.859 [105.076 1078.641]	0.017	< 0.001	94.9
Men and women	13	-97.471 [-732.735 537.793]	0.764	< 0.001	83.6
Age					
Age<13	1	-991.604 [-1.8e+03 -133.759]	0.023	.	.
13≤age<24	5	404.631 [-260.130 1069.393]	0.223	0.032	62.2
24≤age<44	6	694.203 [-123.296 1511.703]	0.096	0.001	74.7
44≤age<64	8	-851.553 [-1.1e+03 -638.864]	< 0.001	0.527	0.0
Age≥64	15	958.987 [529.831 1388.144]	< 0.001	< 0.001	92.6
BMI					
BMI< 18.5	2	-836.739 [-1.5e+03 -141.006]	0.018	0.545	0.0
18.5 ≤ BMI < 25	10	-30.264 [-871.242 810.714]	0.944	< 0.001	91.9
25 ≤ BMI < 30	13	1044.680 [389.432 1699.928]	0.002	< 0.001	84.7
30 ≤ BMI< 35	2	-802.982 [-1.3e+03 -319.204]	0.001	0.061	71.5
Disease					
Healthy	31	421.428 [72.707 770.14]9	0.018	< 0.001	92.1
spinal cord injury	2	288.532 [-438.172 1015.235]	0.436	0.581	0.0
Body fat (%)					
15≤ body fat <25	7	712.941 [-351.025 1776.907]	0.189	< 0.001	89.3
25≤ body fat <35	5	701.396 [253.319 1149.474]	0.002	0.271	22.5
body fat ≥35	5	121.714 [-972.305 1215.733]	0.827	< 0.001	96.5

* P for the meta-analysis. $P < 0.05$ indicates a lack of agreement between PAQ and DLW estimates of TEE by using a random-effects model.

† P-heterogeneity: heterogeneity was evaluated using Cochran's test, and $P < 0.5$ indicates significant heterogeneity across studies.

Supplementary Table 1- Quality assessment of the included studies in this meta-analysis

Study	Study year	Selection	Comparability	Outcome	Overall score	Grade*
Arvidsson et al. (78)	2005	****	**	***	9	high
Barnard et al. (76)	2002	**	**	***	7	high
Besson et al. (66)	2010	*****	**	***	10	high
Erika Bonn1 et al. (100)	2012	**	*	**	5	median
Bonnefoy et al. (35)	2001	****	**	***	9	high
Conway et al. (36)	2002	**	**	***	7	high
Conway et al. (63)	2002	**	**	***	7	high
Csizmadi et al. (79)	2014	*****	*	**	8	high
Foley et al. (67)	2012	*****	**	***	10	high
Fuller et al. (80)	2008	*****	**	***	10	high
Mahabir et al. (37)	2006	**	**	***	7	high
Mâsse et al. (81)	2012	***	**	***	8	High
Racette et al. (39)	1994	***	**	***	8	high
Marrero et al. (68)	2004	****	**	***	9	high
SLINDE et al. (69)	2003	*****	**	***	10	high
STATEN et al. (82)	2001	****	**	***	9	high
Sridharan et al. (64)	2015	*****	**	***	10	high
Tanhoffer et al. (83)	2012	****	**	***	9	high
Walsh et al. (42)	2004	***	**	***	8	high
Washburn et al. (84)	2003	***	**	***	8	high
Corder et al. (92)	2010	**	**	***	7	high
Skaribas et al. (93)	2009	**	**	***	7	high
Johansson et al. (94)	2008	*	**	**	5	median
Liu et al. (95)	2001	**	*	*	4	median
Lof et al. (91)	2003	***	**	***	8	high
Neuhouser et al. (96)	2013	*****	**	*	8	high
Starling et al. (85)	1999	****	**	***	9	high
Hagfors et al. (90)	2005	*	**	***	6	median
Irwin et al. (89)	2001	**	**	**	6	median
Leenders et al. (88)	2001	**	**	***	7	high

Paul et al. (47)	2005	**	**	***	7	high
Philippaerts et al. (87)	1999	****	**	***	9	high
Rothenberg et al. (30)	1998	**	*	**	5	median
Seale et al. (86)	2002	***	**	***	8	high
Starling et al. (85)	1999	*****	**	***	10	high
Ishikawa et al, (70)	2010	*****	**	***	10	high
Colbert et al, (97)	2011	***	*	**	6	median
Lof et al, (2002) (98)	2002	***	**	**	7	high
Pietila¨inen et al, (99)	2010	**	*	**	5	median
*Grade was categorized as low, median and high when overall quality score ranged from 1 to 3, 4 to 6, and 7 to 10, respectively.						

References

1. Neilson HK, Robson PJ, Friedenreich CM, Csizmadi I. Estimating activity energy expenditure: how valid are physical activity questionnaires? The American journal of clinical nutrition. 2008;87(2):279-91.
2. Donahoo WT, Levine JA, Melanson EL. Variability in energy expenditure and its components. Current Opinion in Clinical Nutrition & Metabolic Care. 2004;7(6):599-605.
3. Speakman JR, Westerterp KR. Associations between energy demands, physical activity, and body composition in adult humans between 18 and 96 y of age–. The American journal of clinical nutrition. 2010;92(4):826-34.
4. Li J, Siegrist J. Physical activity and risk of cardiovascular disease—a meta-analysis of prospective cohort studies. International journal of environmental research and public health. 2012;9(2):391-407.
5. Siervo M, Oggioni C, Lara J, Celis-Morales C, Mathers JC, Battezzati A, et al. Age-related changes in resting energy expenditure in normal weight, overweight and obese men and women. Maturitas. 2015;80(4):406-13.
6. Milanović Z, Pantelić S, Trajković N, Sporiš G, Kostić R, James N. Age-related decrease in physical activity and functional fitness among elderly men and women. Clinical interventions in aging. 2013;8:549.
7. Hall KD, Heymsfield SB, Kemnitz JW, Klein S, Schoeller DA, Speakman JR. Energy balance and its components: implications for body weight regulation. The American Journal of Clinical Nutrition. 2012;95(4):989-94.
8. Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. Archives of internal medicine. 2001;161(12):1542-8.
9. Bassuk SS, Manson JE. Epidemiological evidence for the role of physical activity in reducing risk of type 2 diabetes and cardiovascular disease. Journal of applied physiology. 2005;99(3):1193-204.
10. Moses A, Slater C, Preston T, Barber M, Fearon K. Reduced total energy expenditure and physical activity in cachectic patients with pancreatic cancer can be modulated by an energy and protein dense oral supplement enriched with n-3 fatty acids. British journal of cancer. 2004;90(5):996.
11. Uehara M, Plank LD, Hill GL. Components of energy expenditure in patients with severe sepsis and major trauma: a basis for clinical care. Critical care medicine. 1999;27(7):1295-302.
12. Hunter GR, Wetzstein CJ, Fields DA, Brown A, Bamman MM. Resistance training increases total energy expenditure and free-living physical activity in older adults. Journal of Applied Physiology. 2000;89(3):977-84.
13. Speakman J. Doubly labelled water: theory and practice: Springer Science & Business Media; 1997.
14. Westerterp KR. Physical activity as determinant of daily energy expenditure. Physiology & Behavior. 2008;93(4-5):1039-43.
15. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Medicine & science in sports & exercise. 2003;35(8):1381-95.
16. Fujii H, Yamamoto S, Takeda-Imai F, Inoue M, Tsugane S, Kadowaki T, et al. Validity and applicability of a simple questionnaire for the estimation of total and domain-specific physical activity. Diabetology international. 2011;2(2):47-54.
17. Health USDo, Services H. Physical activity and health: a report of the Surgeon General: diane Publishing; 1996.

18. KATZMARZYK PT, Montoye HJ, Kemper HC, Saris WH, Washburn RA. Measuring Physical Activity and Energy Expenditure. JSTOR; 1996.
19. Wendel-Vos GW, Schuit AJ, Saris WH, Kromhout D. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. Journal of clinical epidemiology. 2003;56(12):1163-9.
20. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TI, et al. Energy balance measurement: when something is not better than nothing. International journal of obesity. 2015;39(7):1109.
21. Neilson HK, Robson PJ, Friedenreich CM, Csizmadi I. Estimating activity energy expenditure: how valid are physical activity questionnaires?–. The American journal of clinical nutrition. 2008;87(2):279-91.
22. Melanson Jr EL, Freedson PS, Blair S. Physical activity assessment: a review of methods. Critical Reviews in Food Science & Nutrition. 1996;36(5):385-96.
23. Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber SC, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. International Journal of Behavioral Nutrition and Physical Activity. 2008;5(1):56.
24. Herzog R, Álvarez-Pasquin MJ, Díaz C, Del Barrio JL, Estrada JM, Gil Á. Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review. BMC public health. 2013;13(1):154.
25. Higgins J, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses [journal article as teaching resource, deposited by John Flynn]. British medical journal. 2003;327:557-60.
26. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. 1959.
27. Arvidsson D, Slinde F, Hulthen L. Physical activity questionnaire for adolescents validated against doubly labelled water. Eur J Clin Nutr. 2005;59(3):376-83.
28. Bonn SE, Trolle Lagerros Y, Christensen SE, Moller E, Wright A, Sjolander A, et al. Active-Q: validation of the web-based physical activity questionnaire using doubly labeled water. J Med Internet Res. 2012;14(1):e29.
29. Slinde F, Arvidsson D, Sjoberg A, Rossander-Hulthen L. Minnesota leisure time activity questionnaire and doubly labeled water in adolescents. Med Sci Sports Exerc. 2003;35(11):1923-8.
30. Rothenberg E, Bosaeus I, Lernfelt B, Landahl S, Steen B. Energy intake and expenditure: validation of a diet history by heart rate monitoring, activity diary and doubly labeled water. European Journal of Clinical Nutrition. 1998;52(11):832-8.
31. Löf M, Hannestad U, Forsum E. Assessing physical activity of women of childbearing age. Ongoing work to develop and evaluate simple methods. Food and nutrition bulletin. 2002;23(3_suppl1):30-3.
32. Lof M, Hannestad U, Forsum E. Comparison of commonly used procedures, including the doubly-labelled water technique, in the estimation of total energy expenditure of women with special reference to the significance of body fatness. Br J Nutr. 2003;90(5):961-8.
33. Hagfors L, Westerterp K, Skoldstam L, Johansson G. Validity of reported energy expenditure and reported intake of energy, protein, sodium and potassium in rheumatoid arthritis patients in a dietary intervention study. Eur J Clin Nutr. 2005;59(2):238-45.
34. Barnard J, Tapsell LC, Davies P, Brenninger V, Storlien L. Relationship of high energy expenditure and variation in dietary intake with reporting accuracy on 7 day food records and diet histories in a group of healthy adult volunteers. European journal of clinical nutrition. 2002;56(4):358-67.
35. Bonnefoy M, Normand S, Pachiaudi C, Lacour JR, Laville M, Kostka T. Simultaneous validation of ten physical activity questionnaires in older men: a doubly labeled water study. Journal of the American Geriatrics Society. 2001;49(1):28-35.

36. Conway JM, Irwin ML, Ainsworth BE. Estimating energy expenditure from the Minnesota Leisure Time Physical Activity and Tecumseh Occupational Activity questionnaires–a doubly labeled water validation. Journal of clinical epidemiology. 2002;55(4):392-9.
37. Mahabir S, Baer DJ, Giffen C, Clevidence BA, Campbell WS, Taylor PR, et al. Comparison of energy expenditure estimates from 4 physical activity questionnaires with doubly labeled water estimates in postmenopausal women. The American journal of clinical nutrition. 2006;84(1):230-6.
38. Mâsse LC, Fulton JE, Watson KB, Tortolero S, Kohl HW, Meyers MC, et al. Comparing the validity of 2 physical activity questionnaire formats in African-American and Hispanic women. Journal of Physical Activity and Health. 2012;9(2):237-48.
39. Racette SB, Schoeller DA, Kushner RF. Comparison of heart rate and physical activity recall with doubly labeled water in obese women. Medicine and science in sports and exercise. 1995;27(1):126-33.
40. Ramirez-Marrero FA, Smith BA, Sherman WM, Kirby TE. Comparison of methods to estimate physical activity and energy expenditure in African American children. Int J Sports Med. 2005;26(5):363-71.
41. STATEN LK, Taren DL, Howell WH, TOBAR M, POEHLMAN ET, HILL A, et al. Validation of the Arizona Activity Frequency Questionnaire using doubly labeled water. Medicine & Science in Sports & Exercise. 2001;33(11):1959-67.
42. Walsh MC, Hunter GR, Sirikul B, Gower BA. Comparison of self-reported with objectively assessed energy expenditure in black and white women before and after weight loss. The American journal of clinical nutrition. 2004;79(6):1013-9.
43. Washburn RA, Jacobsen DJ, Sonko BJ, Hill JO, Donnelly JE. The validity of the Stanford Seven-Day Physical Activity Recall in young adults. Med Sci Sports Exerc. 2003;35(8):1374-80.
44. Starling RD, Matthews DE, Ades PA, Poehlman ET. Assessment of physical activity in older individuals: a doubly labeled water study. Journal of Applied Physiology. 1999;86(6):2090-6.
45. Seale JL, Klein G, Friedmann J, Jensen GL, Mitchell DC, Smiciklas-Wright H. Energy expenditure measured by doubly labeled water, activity recall, and diet records in the rural elderly. Nutrition. 2002;18(7-8):568-73.
46. Leenders NY, Sherman WM, Nagaraja H, Kien CL. Evaluation of methods to assess physical activity in free-living conditions. Medicine & Science in Sports & Exercise. 2001;33(7):1233-40.
47. Paul DR, Rhodes DG, Kramer M, Baer DJ, Rumpler WV. Validation of a food frequency questionnaire by direct measurement of habitual ad libitum food intake. American journal of epidemiology. 2005;162(8):806-14.
48. Irwin ML, Ainsworth BE, Conway JM. Estimation of energy expenditure from physical activity measures: determinants of accuracy. Obesity Research. 2001;9(9):517-25.
49. Delikanaki-Skaribas E, Trail M, Wong WW, Lai EC. Daily energy expenditure, physical activity, and weight loss in Parkinson's disease patients. Mov Disord. 2009;24(5):667-71.
50. Neuhouser ML, Di C, Tinker LF, Thomson C, Sternfeld B, Mossavar-Rahmani Y, et al. Physical activity assessment: biomarkers and self-report of activity-related energy expenditure in the WHI. Am J Epidemiol. 2013;177(6):576-85.
51. Csizmadi I, Neilson HK, Kopciuk KA, Khandwala F, Liu A, Friedenreich CM, et al. The Sedentary Time and Activity Reporting Questionnaire (STAR-Q): reliability and validity against doubly labeled water and 7-day activity diaries. Am J Epidemiol. 2014;180(4):424-35.
52. Foley LS, Maddison R, Rush E, Olds TS, Ridley K, Jiang Y. Doubly labeled water validation of a computerized use-of-time recall in active young people. Metabolism. 2013;62(1):163-9.
53. Tanhoffer RA, Tanhoffer AI, Raymond J, Hills AP, Davis GM. Comparison of methods to assess energy expenditure and physical activity in people with spinal cord injury. J Spinal Cord Med. 2012;35(1):35-45.
54. Besson H, Brage S, Jakes RW, Ekelund U, Wareham NJ. Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. Am J Clin Nutr. 2010;91(1):106-14.

55. Fuller Z, Horgan G, O'Reilly LM, Ritz P, Milne E, Stubbs RJ. Comparing different measures of energy expenditure in human subjects resident in a metabolic facility. Eur J Clin Nutr. 2008;62(4):560-9.
56. Sridharan S, Wong J, Vilar E, Farrington K. Comparison of energy estimates in chronic kidney disease using doubly-labelled water. J Hum Nutr Diet. 2016;29(1):59-66.
57. Liu B, Woo J, Tang N, Ng K, Ip R, Yu A. Assessment of total energy expenditure in a Chinese population by a physical activity questionnaire: examination of validity. Int J Food Sci Nutr. 2001;52(3):269-82.
58. Corder K, Brage S, Wright A, Ramachandran A, Snehalatha C, Yamuna A, et al. Physical activity energy expenditure of adolescents in India. Obesity (Silver Spring). 2010;18(11):2212-9.
59. Philippaerts R, Westerterp K, Lefevre J. Doubly labeled water validation of three physical activity questionnaires. International journal of sports medicine. 1999;20(05):284-9.
60. Johansson G, Westerterp KR. Assessment of the physical activity level with two questions: validation with doubly labeled water. Int J Obes (Lond). 2008;32(6):1031-3.
61. Ishikawa-Takata K, Naito Y, Tanaka S, Ebine N, Tabata I. Use of doubly labeled water to validate a physical activity questionnaire developed for the Japanese population. J Epidemiol. 2011;21(2):114-21.
62. Pietilainen KH, Korkeila M, Bogl LH, Westerterp KR, Yki-Jarvinen H, Kaprio J, et al. Inaccuracies in food and physical activity diaries of obese subjects: complementary evidence from doubly labeled water and co-twin assessments. Int J Obes (Lond). 2010;34(3):437-45.
63. Conway JM, Seale JL, Jacobs DR, Irwin ML, Ainsworth BE. Comparison of energy expenditure estimates from doubly labeled water, a physical activity questionnaire, and physical activity records. The American journal of clinical nutrition. 2002;75(3):519-25.
64. Sridharan S, Wong J, Vilar E, Farrington K. Comparison of energy estimates in chronic kidney disease using doubly-labelled water. Journal of Human Nutrition and Dietetics. 2016;29(1):59-66.
65. Bonn SE, Lagerros YT, Christensen SE, Möller E, Wright A, Sjölander A, et al. Active-Q: validation of the web-based physical activity questionnaire using doubly labeled water. Journal of medical Internet research. 2012;14(1):e29.
66. Besson H, Brage S, Jakes RW, Ekelund U, Wareham NJ. Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. The American journal of clinical nutrition. 2010;91(1):106-14.
67. Foley LS, Maddison R, Rush E, Olds TS, Ridley K, Jiang Y. Doubly labeled water validation of a computerized use-of-time recall in active young people. Metabolism: clinical and experimental. 2013;62(1):163-9.
68. Ramirez-Marrero F, Smith B, Sherman W, Kirby T. Comparison of methods to estimate physical activity and energy expenditure in African American children. International journal of sports medicine. 2005;26(05):363-71.
69. Slinde F, Arvidsson D, Sjoberg A, Rossander-Hulthén L. Minnesota leisure time activity questionnaire and doubly labeled water in adolescents. Medicine and science in sports and exercise. 2003;35(11):1923-8.
70. Ishikawa-Takata K, Naito Y, Tanaka S, Ebine N, Tabata I. Use of doubly labeled water to validate a physical activity questionnaire developed for the Japanese population. Journal of epidemiology. 2011;21(2):114-21.
71. Slinde F, Arvidsson D, Sjöberg A, Rossander-hulthén L. Minnesota leisure time activity questionnaire and doubly labeled water in adolescents. Medicine & Science in Sports & Exercise. 2003;35(11):1923-8.
72. Schofield W. Predicting basal metabolic rate, new standards and review of previous work. Human nutrition Clinical nutrition. 1985;39:5-41.
73. Henry C, Dyer S, Ghusain-Choueiri A. New equations to estimate basal metabolic rate in children aged 10-15 years. European journal of clinical nutrition. 1999;53(2):134-42.

74. Organization WH. Energy and protein requirements: report of a joint FAO/WHO/UNU expert consultation (WHO Technical Report Series, no 724). Geneva: World Health Organization. 1985.
75. Molnár D, Jeges S, Erhardt E, Schutz Y. Measured and predicted resting metabolic rate in obese and nonobese adolescents. The Journal of pediatrics. 1995;127(4):571-7.
76. Barnard J, Tapsell LC, Davies P, Brenninger V, Storlien L. Relationship of high energy expenditure and variation in dietary intake with reporting accuracy on 7 day food records and diet histories in a group of healthy adult volunteers. European journal of clinical nutrition. 2002;56(4):358-68.
77. Montoye HJ, Kemper HC, Saris WH, Washburn RA. Measuring physical activity and energy expenditure: Human Kinetics Champaign, IL; 1996.
78. Arvidsson D, Slinde F, Hulthén L. Physical activity questionnaire for adolescents validated against doubly labelled water. European journal of clinical nutrition. 2005;59(3):376-83.
79. Csizmadi I, Neilson HK, Kopciuk KA, Khandwala F, Liu A, Friedenreich CM, et al. The Sedentary Time and Activity Reporting Questionnaire (STAR-Q): reliability and validity against doubly labeled water and 7-day activity diaries. American journal of epidemiology. 2014;180(4):424-35.
80. Fuller Z, Horgan G, O'reilly L, Ritz P, Milne E, Stubbs R. Comparing different measures of energy expenditure in human subjects resident in a metabolic facility. European journal of clinical nutrition. 2008;62(4):560-9.
81. Mâsse LC, Fulton JE, Watson KB, Tortolero S, Kohl III HW, Meyers MC, et al. Comparing the validity of 2 physical activity questionnaire formats in African-American and Hispanic women. Journal of Physical Activity and Health. 2012;9(2):237-48.
82. Staten LK, Taren DL, Howell WH, Tobar M, Poehlman ET, Hill A, et al. Validation of the Arizona Activity Frequency Questionnaire using doubly labeled water. Medicine and science in sports and exercise. 2001;33(11):1959-67.
83. Tanhoffer RA, Tanhoffer AI, Raymond J, Hills AP, Davis GM. Comparison of methods to assess energy expenditure and physical activity in people with spinal cord injury. The journal of spinal cord medicine. 2012;35(1):35-45.
84. Washburn RA, Jacobsen DJ, Sonko BJ, Hill JO, Donnelly JE. The validity of the Stanford Seven-Day Physical Activity Recall in young adults. Medicine and science in sports and exercise. 2003;35(8):1374-80.
85. Starling RD, Matthews DE, Ades PA, Poehlman ET. Assessment of physical activity in older individuals: a doubly labeled water study. Journal of applied physiology (Bethesda, Md : 1985). 1999;86(6):2090-6.
86. Seale JL, Klein G, Friedmann J, Jensen GL, Mitchell DC, Smiciklas-Wright H. Energy expenditure measured by doubly labeled water, activity recall, and diet records in the rural elderly. Nutrition. 2002;18(7):568-73.
87. Philippaerts RM, Westerterp KR, Lefevre J. Doubly labelled water validation of three physical activity questionnaires. International journal of sports medicine. 1999;20(5):284-9.
88. Leenders N, Sherman WM, Nagaraja H, Kien CL. Evaluation of methods to assess physical activity in free-living conditions. Medicine and science in sports and exercise. 2001;33(7):1233-40.
89. Irwin ML, Ainsworth BE, Conway JM. Estimation of energy expenditure from physical activity measures: determinants of accuracy. Obesity. 2001;9(9):517-25.
90. Hagfors L, Westerterp K, Sköldstam L, Johansson G. Validity of reported energy expenditure and reported intake of energy, protein, sodium and potassium in rheumatoid arthritis patients in a dietary intervention study. European journal of clinical nutrition. 2005;59(2):238-45.
91. Lof M, Hannestad U, Forsum E. Comparison of commonly used procedures, including the doubly-labelled water technique, in the estimation of total energy expenditure of women with special reference to the significance of body fatness. British journal of nutrition. 2003;90(5):961-8.
92. Corder K, Brage S, Wright A, Ramachandran A, Snehalatha C, Yamuna A, et al. Physical activity energy expenditure of adolescents in India. Obesity. 2010;18(11):2212-9.

93. Delikanaki-Skaribas E, Trail M, Wong WWL, Lai EC. Daily energy expenditure, physical activity, and weight loss in Parkinson's disease patients. Movement Disorders. 2009;24(5):667-71.
94. Johansson G, Westerterp K. Assessment of the physical activity level with two questions: validation with doubly labeled water. International Journal of Obesity. 2008;32(6):1031.
95. Liu B, Woo J, Tang N, Ng K, Ip R, Yu A. Assessment of total energy expenditure in a Chinese population by a physical activity questionnaire: examination of validity. International journal of food sciences and nutrition. 2001;52(3):269-82.
96. Neuhouser ML, Di C, Tinker LF, Thomson C, Sternfeld B, Mossavar-Rahmani Y, et al. Physical activity assessment: biomarkers and self-report of activity-related energy expenditure in the WHI. American journal of epidemiology. 2013;177(6):576-85.
97. Colbert LH, Matthews CE, Havighurst TC, Kim K, Schoeller DA. Comparative validity of physical activity measures in older adults. Medicine and science in sports and exercise. 2011;43(5):867.
98. Lof M, Hannestad U, Forsum E. Assessing physical activity of women of childbearing age. Ongoing work to develop and evaluate simple methods. FOOD AND NUTRITION BULLETIN-UNITED NATIONS UNIVERSITY-. 2002;23(3; SUPP):30-3.
99. Pietiläinen K, Korkeila M, Bogl L, Westerterp K, Yki-Järvinen H, Kaprio J, et al. Inaccuracies in food and physical activity diaries of obese subjects: complementary evidence from doubly labeled water and co-twin assessments. International Journal of Obesity. 2010;34(3):437.
100. Bonn SE, Lagerros YT, Christensen SE, Möller E, Wright A, Sjölander A, et al. Active-Q: validation of the web-based physical activity questionnaire using doubly labeled water. Journal of medical Internet research. 2012;14(1).