British Journal of Nutrition (2021), 125, 983–997 doi:10.1017/S0007114520003049

British Journal of Nutrition

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

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

[1]*Department of Clinical Nutrition, School of Nutritional Sciences and Dietetics, Tehran University of Medical Sciences (TUMS), Tehran 1416-643931, Iran*

[2]*Department of Medicine, Division of Cardiovascular Medicine, Vanderbilt University Medical Center, Nashville, TN 37232, USA*

[3]*Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, People's Republic of China*

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

(Submitted 12 February 2020 – Final revision received 6 July 2020 – Accepted 17 July 2020 – First published online 28 July 2020)

Abstract

Physical activity questionnaires (PAQ) 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 PAQ for estimating TEE and AEE using doubly labelled water (DLW). We identified data from relevant studies by searching Google Scholar, PubMed and Scopus databases. This revealed thirty-eight studies that had validated PAQ with DLW and reported the mean differences between PAQ and DLW measures of TEE ($TEE_{DLW} - TEE_{PAQ}$) and AEE ($AEE_{DLW} - AEE_{PAQ}$). We assessed seventy-eight PAQ consisting of fifty-nine PAQ that assessed TEE and thirty-five PAQ that examined AEE. There was no significant difference between TEE_{PAQ} and TEE_{DLW} with a weighted mean difference of –243·3 and a range of –841·4 to 354·6 kJ/d, and a significant weighted mean difference of $AEE_{DLW} - AEE_{PAQ}$ 414·6 and a range of 78·7–750·5. 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 3-d PA diaries were correlated with TEE by DLW at the population level; however, these two PAQ did not demonstrate 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 3-d PA diaries were identified as the only valid PAQ for TEE estimation. Further well-designed studies are needed to verify this result and identify additional valid PAQ.

Key words: Physical activity: Total energy expenditure: Activity energy expenditure: Doubly labelled water method

Total energy expenditure (TEE) consists of three components: BMR (or basal energy expenditure; 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 CVD 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 resistance training[12]. Therefore, measuring TEE and PA is 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 labelled water (DLW) 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

Abbreviations: AEE, activity energy expenditure; DLW, doubly labelled water; PAQ, physical activity questionnaire; TEE, total energy expenditure; WMD, weighted mean difference.

* **Corresponding author:** Kurosh Djafarian, fax +98 21 88974462, email kdjafarian@tums.ac.ir

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 behaviour 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) and DLW have shown that misreporting of energy expenditure by PAQ is also common[21].

PAQ are being developed continuously and hence it is necessary to validate which PAQ provide valid estimates of TEE and AEE[22] by comparison to the 'gold standard' DLW methodology. Systematic reviews conducted a decade ago by Neilson *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 PAQ have 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 PAQ may be useful in epidemiological studies, and some in individual studies like clinical research. To find PAQ suitable for these two kinds of studies, we need to follow two criteria: first, at the population level, suitable PAQ 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, PAQ must have an 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 PAQ 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 present study.

We included studies that (A) validated PAQ with DLW based on measurements of TEE and/or AEE and (B) included PAQ 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 values and standard deviations, age, weight, BMI (kg/m^2), body fat percentage (BF %) (Table 1), TEE (kJ/d) (Table 2) and AEE (kJ/d) measured by both DLW and PAQ (Table 3).

Quality assessment

The quality of each eligible study was assessed using the 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 (online Supplementary Table S1).

Statistical analysis

In our meta-analysis, the means and standard deviations of the differences in TEE 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^2. 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 (DerSimonian–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 categorised 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 or spinal cord injury. BMI (kg/m^2) was classified as BMI < 18·5, 18·5 ≤ BMI < 25, 25 ≤ BMI < 30 and 30 ≤ BMI < 35 and BF % divided into the following 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).

Results

We identified 1780 studies of which sixty-nine were identified in PubMed and 1711 in Scopus and Google Scholar. A total of 113 studies remained after a preliminary title and abstract review, seventy-five records were excluded from our analysis since they did not report TEE or AEE (*n* 15) or did not validate self-report measures with DLW (*n* 31) or did not use PAQ (*n* 13) or reported AEE in an inappropriate way like PA score or metabolic equivalent category (*n* 16). In the end, thirty-eight articles met the inclusion criteria of our study and were considered for further assessment (Fig. 1).

https://doi.org/10.1017/S0007114520003049 Published online by Cambridge University Press

Table 1. Characteristics of the studies included into the meta-analysis
(Numbers and percentages; mean values and standard deviations)

Study	Sample size	Sex	Heath status of the participants	Age (years)	BMI (kg/m^2)		Weight (kg)		Body fat (%)	
					Mean	SD	Mean	SD	Mean	SD
Arvidsson *et al.* A[78]	17	Boy	Healthy	8	2·6	21	9	1	4·7	4
Arvidsson *et al.* B[78]	16	Girl	Healthy	7	2·7	21	9·4	4	5·2	5
Barnard *et al.* A[76]	8	Men	Healthy	4	3·9	9		Not reported	6·8	9
Barnard *et al.* B[76]	7	Women	Healthy	1	5·3	8		Not reported	9	4
Besson *et al.* A[66]	50	Men (50 %) Women (50 %)	Healthy	3	3·1	1		Not reported	7·9	22
Bonn *et al.* A[65]	37	Men (19 %) Women (81 %)	Healthy	20–65		Not reported		Not reported		Not reported
Bonn *et al.* B[65]	37	Men (19 %) Women (81 %)	Healthy	21–65		Not reported		Not reported		Not reported
Bonnefoy *et al.* A, B, C, D, E[35]	19	Men	Healthy	4		Not reported	9·7	3		Not reported
Conway *et al.* A[36]	24	Men	Healthy	42	0·6	6	2·1	5	6·8	1
Conway *et al.* B, C[63]	24	Men	Healthy	2	0·5	1	1·8	5		Not reported
Csizmadi *et al.* A, B, C, D[79]	102	Men (86 %) Women (14 %)	Healthy	48	0·3	24		Not reported		Not reported
Foley *et al.*[67]	32	Men (56 %) Women (44 %)	Healthy	3	3·3	20·3	16	57	7	3
Fuller *et al.* A[80]	59	Men (51 %) Women (49 %)	Healthy	7	2·25	3	9·6	1	2·9	2
Fuller *et al.* B[80]	59	Men (51 %) Women (49 %)	Healthy	7	2·25	3	9·6	1	2·9	2
Mahabir *et al.* A, B, C, D[37]	65	Women	Postmenopausal	9	5·6	7		Not reported	8·6	2
Mâsse *et al.* A, B[81]	130	Women	Healthy	2	6·3	30	17·3	76·9		Not reported
Racette *et al.* A[39]	14	Women	Healthy	40	8·8	34	0·06	2	2·9	8
Racette *et al.* B[39]	14	Women	Healthy	40	4·48	30·2	4·48	81		Not reported
Ramírez-Marrero *et al.*[68]	12	Men (43 %) Women (57 %)	Healthy	18	9·5	7	5·45	6		Not reported
Slinde *et al.*[69]	2400	Boys (48 %) Girls (52 %)	Health	15	2·6	20·8	9·6	60·4		Not reported
Staten *et al.* A, B[82]	35	Women	Healthy	8	8·1	28	20·4	73		Not reported
Sridharan *et al.* A, B[64]	40	Men (55 %) Women (45 %)	Chronic kidney disease (stages 1–5)	54	4·2	8	12·2	1		Not reported
Tanhoffer *et al.* A, B[83]	14	Men (93 %) Women (7 %)	Spinal cord injury	40	3	25	15	79	9	33
Walsh *et al.* A [42]	21	Women	Healthy	5	1·7	1	20·4	73	3·6	6
Walsh *et al.* B[42]	21	Women	Healthy	5	1·1	9	5·3	7	4·7	5
Walsh *et al.* C[42]	20	Women	Healthy	36	1·8	6	4·5	2	3·7	1
Walsh *et al.* D[42]	20	Women	Healthy	36	0·9	24	9·2	78	4·5	1
Walsh *et al.* E[42]	20	Women	Healthy	8	1	1	7·9	5	4	4
Walsh *et al.* F[42]	14	Women	Healthy	8	1·6	23	4·7	3	5·3	5
Washburn *et al.* A[84]	17	Men	Healthy	9	2·7	8	4·7	3	4·7	2
Washburn *et al.* B[84]	29	Women	Healthy	3	2·8	4	11·9	1	4·2	6
Starling *et al.* A, B[85]	35	Women	Healthy	67	3·9	8	10·2	9	8	35
Starling *et al.* C, D[85]	32	Men	Healthy	66	4·5	7	14·5	5	7	21
Seale *et al.* A[86]	13	Women	Healthy	5	3·2	6	9·5	8		Not reported
Seale *et al.* B[86]	14	Men	Healthy	1	2·4	2	7·9	6		Not reported

https://doi.org/10.1017/S0007114520003049 Published online by Cambridge University Press

Table 1. (*Continued*)

Study	Sample size	Sex	Heath status of the participants	Age (years)	BMI (kg/m^2)		Weight (kg)		Body fat (%)	
					Mean	SD	Mean	SD	Mean	SD
Rothenberg *et al.*[30]	12	Men (40 %) Women (60 %)	Healthy	73		24·3		62		Not reported
Philippaerts *et al.*[87]	90	Men	Healthy	40	2·8	24·6	8	78		20·3
Paul *et al.*[47]	12	Men	Healthy	39	1·4	24·1	8·3	79·9		18·1
Leenders *et al.*[88]	13	Women	Healthy	25·8	0·6	23·5	2	65·5		26·3
Irwin *et al.* A, B[89]	24	Men	Healthy	41·2	2·7	25·1	9	79·5		21·1
Hagfors *et al.*[90]	9	Men (60 %) Women (40 %)	Healthy	8	4·4	28·1	14·1	77·8		Not reported
Lof *et al.*[91]	34	Women	Healthy	30	4	24	10	67	8	34
Corder *et al.* A[92]	13	Men	Healthy	15·9	2·6	17·4	7·1	46·1	10	14·3
Corder *et al.* B[92]	15	Women	Healthy	15·7	4·2	20·8	12·5	49·4	8·7	29·8
Skaribas *et al.* A, B[93]	20	Men	Healthy	72·9		Not reported	9·5	77·4	7·9	24·2
Johansson *et al.*[94]	9	Men (34 %) Women (66 %)	Healthy	60	4·5	4		Not reported		Not reported
Liu *et al.* A[95]	18	Women	Renal, cancer, healthy	64–84		Not reported		Not reported		Not reported
Liu *et al.* B[95]	13	Men	Renal, cancer, healthy	64–84		Not reported		Not reported		Not reported
Neuhouser *et al.* A[96]	450	Women	Healthy	50–80		Not reported		Not reported		Not reported
Neuhouser *et al.* B[96]	444	Women	Healthy	50–81		Not reported		Not reported		Not reported
Neuhouser *et al.* C[96]	426	Women	Healthy	50–82		Not reported		Not reported		Not reported
Ishikawa *et al.* A[70]	118	Women	Healthy	50·4	2·5	3	7·3	7		Not reported
Ishikawa *et al.* B[70]	108	Men	Healthy	50·4	3	23	10·9	6		Not reported
Colbert *et al.* A[97]	56	Women (79 %) Men (21 %)	Healthy	74·7	4·2	8	14·5	2		Not reported
Colbert *et al.* B[97]	56	Women (79 %) Men (21 %)	Healthy	74·7	4·2	8	14·5	2		Not reported
Colbert *et al.* C[97]	56	Women (79 %) Men (21 %)	Healthy	74·7	4·2	8	14·5	2		Not reported
Lof *et al.*[98]	24	Women	Healthy	30	4	24	10	67		Not reported
Pietiläinen *et al.* A[99]	7	Men	Healthy	25·5	0·5	30	2·3	88	1·8	3
Pietiläinen *et al.* B[99]	7	Men	Healthy	25·5	0·5	25	2·3	73	2·3	4

https://doi.org/10.1017/S0007114520003049 Published online by Cambridge University Press

Table 2. Summary of results for the difference in total energy expenditure (TEE) means between physical activity questionnaires (PAQ) and doubly labelled water (DLW)*
(Mean values and standard deviations)

		TEE_{DLW}		TEE_{PAQ}	
Study	PAQ type	Mean	SD	Mean	SD
Arvidsson *et al.* A[78]	PAQA	11 300	1500	7600	1600
Arvidsson *et al.* B[78]	PAQA	9100	1400	5200	1100
Barnard *et al.* A[76]	MAQ	29 409	6857·9	6	2562·3
Barnard *et al.* B[76]	MAQ	4	4531·7	8	836·4
Besson *et al.* A[66]	RPAQ	9	2574·1	8516	2025·1
Bonn *et al.* A[65]	Active-Q	11 229	2256	11 667	3212
Bonn *et al.* B[65]	Active-Q	11 229	2256	11 529	2758
Bonnefoy *et al.* B[35]	7 d-PAQR	11 181	1647	12 335·78	1658·4
Bonnefoy *et al.* D[35]	QAPSE	11 181	1647	9684	856·017
Conway *et al.* A[36]	(TEC + MNLTPA + EESLEEP + EEGEN)	13 550	380	14 870	900
Conway *et al.* B[63]	7-dPAR	13 270	350	17 400	1450
Conway *et al.* C[63]	7-dPArecord	13 270	350	14 170	370
Csizmadi *et al.* A[79]	Star-Q	67	3213·31	79	3941·33
Csizmadi *et al.* B[79]	Star-Q	67	3213·31	24	3338·83
Csizmadi *et al.* C[79]	Star-Q	67	3213·31	2	3414·14
Csizmadi *et al.* D[79]	7 d-PAQR	67	3213·31	50	4619·14
Foley *et al.*[67]	MARCA	96	3778·15	98	4481·064
Fuller *et al.* A[80]	24-h PAD	11 030	2190	10 050	1800
Fuller *et al.* B[80]	7-dPAR	11 040	2200	9370	2250
Mahabir *et al.* A[37]	Five city project questionnaire	10 711·04	2602·45	48	4744·656
Mahabir *et al.* B[37]	Harvard Alumni questionnaire	10 711·04	2602·45	42	4853·44
Mahabir *et al.* C[37]	CAPS study 4 week activity recall	10 711·04	2602·45	10 798·9	9694·328
Mahabir *et al.* D[37]	CAPS study typical week activity recall	10 711·04	2602·45	84	3907·86
Mâsse *et al.* A[81]	The checklist questionnaire	72	1824·22	10 589·7	2359·78
Mâsse *et al.* B[81]	Global questionnaire	72	1824·22	92	2414·17
Racette *et al.* A[39]	7-dPAR	10 945·34	1765·65	11 150·36	1213·36
Racette *et al.* B[39]	7-dPAR	10 259·17	1840·96	10 208·96	1598·29
Ramírez-Marrero *et al.*[68]	SAPAC	7004·016	999·1392	7504·4224	1273·6096
Slinde *et al.*[69]	MNLTPA	11 400	2100	8600	2000
Staten *et al.* A[82]	The Arizona activity Frequency questionnaire 28 d	9847	2555	7912	2196
Staten *et al.* B[82]	The Arizona activity Frequency questionnaire 7 d	9847	2555	8001	2639
Sridharan *et al.* A[64]	RPAQ	10 380·5	1991·58	616	2250·99
Sridharan *et al.* B[64]	7-dPAR	10 380·5	1991·58	10 941·16	2874·41
Tanhoffer *et al.* A[83]	Para-Sci	9817	2491	9259	2094
Tanhoffer *et al.* B[83]	PASIPD	9817	2491	9766	1462
Walsh *et al.* A[42]	TEC + MNLTPA	56	1656·86	1	1326·33
Walsh *et al.* B[42]	TEC + MNLTPA	88	1071·1	10 129·46	815·88
Walsh *et al.* C[42]	TEC + MNLTPA	712	1435·11	12 049·92	1640·13
Walsh *et al.* D[42]	TEC + MNLTPA	896	1669·42	22	1891·17
Walsh *et al.* E[42]	TEC + MNLTPA	128	991·608	10 953·71	1753·1
Walsh *et al.* F[42]	TEC + MNLTPA	528	1422·56	10 326·11	1397·46
Washburn *et al.* A[84]	7-dPAR	13 885	2754	13 198	1638
Washburn *et al.* B[84]	7-dPAR	10 771	1457	11 018	1323
Seale *et al.* A[86]	7-dPAR	9440	900	9510	2400
Seale *et al.* B[86]	7-dPAR	12 430	1630	13 690	3230
Rothenberg *et al.*[30]	Activity diary in 4 d	9900	1430	9240	2150
Philippaerts *et al.*[87]	FCQ 7 d index	13 400	1800	12 030·26	1782·8
Irwin *et al.* A[89]	7-dPAR	10	1719·62	89	7108·62
Irwin *et al.* B[89]	7-dPArecord	10	1719·62	84	778·22
Hagfors *et al.*[90]	3-d activity registration	10 760	2590	9820	1650
Lof *et al.*[91]	2-week recall	10 670	1370	11 210	2000
Johansson *et al.*[94]	Two-question questionnaire on physical activity	10 900	2700	10 800	1800
Liu *et al.* A[95]	Modified YPAS	80		36	1118·38
Liu *et al.* B[95]	Modified YPAS	1017·20		10 967·52	585·7
Ishikawa *et al.* A[70]	JALSPAQ	8420	1400	7620	1430
Ishikawa *et al.* B[70]	JALSPAQ	11 210	3000	9830	1180
Lof *et al.*[98]	LOF questionnaire	11 420		10 570	
Pietiläinen *et al.* A[99]	3-d PA diaries	12 400	400	14 200	
Pietiläinen *et al.* B[99]	3-d PA diaries	11 500	700	12 600	

PAQA, Physical Activity Questionnaire for Adolescents; MAQ, Modifiable Activity Questionnaire; RPAQ, Recent Physical Activity Questionnaire; 7 d-PAR, 7-d 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-d 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 Physical Activity Survey; JALSPAQ, the Japan Arteriosclerosis Longitudinal Study Physical Activity Questionnaire; CAPS, Cross-Cultural Activity Participation Study.
* All data in kJ/d.

British Journal of Nutrition

British Journal of Nutrition

Table 3. Summary of results from difference in activity energy expenditure (AEE) means between physical activity questionnaires (PAQ) and doubly labelled water (DLW)*
(Mean values and standard deviations)

		AEE_{DLW}		AEE_{PAQ}	
Study	PAQ type	Mean	SD	Mean	SD
Bonnefoy *et al.* A[35]	MNLTPA	3367	1940	2053·900	854·790
Bonnefoy *et al.* C[35]	YPAS	3367	1940	241	1655·609
Bonnefoy *et al.* E[35]	College Alumni questionnaire	3367	1940	885	1031·356
Csizmadi *et al.* A[79]	Star-Q	4250·944	2765·620	5029·168	2627·550
Csizmadi *et al.* B[79]	Star-Q	4250·944	2765·620	528	2916·250
Csizmadi *et al.* C[79]	Star-Q	4250·944	2765·620	704	2690·310
Csizmadi *et al.* D[79]	7-dPAR	4250·944	2765·620	424	2405·800
Foley *et al.*[67]	MARCA	232	3234·230	912	3368·120
Mâsse *et al.* A[81]	Checklist questionnaire	780	1292·860	780	2359·78
Mâsse *et al.* B[81]	Global questionnaire	780	1292·860	60	1757·280
Ramírez-Marrero *et al.*[68]	SAPAC	778·224	1271·936	1301·220	2263·540
Staten *et al.* A[82]	The Arizona activity Frequency questionnaire 28 d	5578	2084	3645	1916
Staten *et al.* B[82]	The Arizona activity Frequency questionnaire 7 d	5578	2084	3734	2428
Sridharan *et al.* A[64]	RPAQ	550		616	2250·99
Sridharan *et al.* B[64]	7-dPAR	550		10 941·16	2874·41
Tanhoffer *et al.* A[83]	Para-Sci	2841	1626	2339	1171
Tanhoffer *et al.* B[83]	PASIPD	2841	1626	2749	1026
Washburn *et al.* A[84]	7-dPAR	3989	2461	3650	490
Washburn *et al.* B[84]	7-dPAR	3223	1360	3073	377
Starling *et al.* A[85]	YPAS	630	1020·9	3610·790	1870·25
Starling *et al.* B[85]	YPAS	5066·824	1794·94	688	2560·61
Starling *et al.* C[85]	MNLTPA	630	1020·9	20	953·952
Starling *et al.* D[85]	MNLTPA	5066·824	1794·94	1920·460	1204·99
Paul *et al.*[47]	7-dPArecord	10 500	1600	11 800	2000
Leenders *et al.*[88]	7-dPAR	830	1251·02	13	527·184
Corder *et al.* A[92]	Youth physical activity questionnaire recall in past week	2	1187·7	3	1837·3
Corder *et al.* B[92]	Youth physical activity questionnaire recall in past week	1990·5	1185	7	526
Skaribas *et al.* A[93]	YPAS	446	1297·04	368	292·88
Skaribas *et al.* B[93]	PASE	446	1297·04	39	907·928
Neuhouser *et al.* A[96]	Arizona activity FFQ 28 d	3075·240		670	
Neuhouser *et al.* B[96]	7-dPAR	3075·240		3016·660	
Neuhouser *et al.* C[96]	PHQ	3075·240		10	
Colbert *et al.* A[97]	YPAS	2845	1138	2699	
Colbert *et al.* B[97]	modPASE	2845	1138	1904	
Colbert *et al.* C[97]	Champs	2845	1138	1092	

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-d 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-d physical activity record questionnaire; PASE, Physical Activity Scale for the Elderly; PHQ, Personal Habits Questionnaire; modPASE, modified Physical Activity Scale for the Elderly; CHAMPS, Community Health Activities Model Program for Seniors.
* All data in kJ/d.

Study characteristics

The thirty-eight studies included 5997 individuals. There were seven studies performed in Sweden[27–33], one in Australia[34], one in France[35], seventeen in the USA[36–50], one in Canada[51], one in New Zealand[52], one in Brazil[53], three in the UK[54–56], one in China[57], one in India[58], two in the Netherlands[59,60], one in Japan[61] and one in Finland[62]. For studies that included more than one PAQ, each of these PAQ was entered separately into our meta-analysis. Therefore, the total number of PAQ extracted for the analysis was seventy-eight. Of these, fifty-nine of the PAQ reported TEE and thirty-five of them reported AEE. Forty different PAQ were identified. Thirty-one PAQ included women only, twenty-five included men only and the remaining twenty-two included both sexes. The mean age of the study population that was reported in sixty-four studies using PAQ ranged from 8·2 to 73·4 years. The mean BMI that was recorded in fifty-seven studies using PAQ ranged from 16 to 34 kg/m^2. The mean body fat that was recorded in forty-two 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 Fig. 2. The weighted mean difference (WMD) was not significant between TEE_{DLW} − TEE_{PAQ} (WMD –243, 95 % CI –841·4, 354·6), $I^2 = 97·9\,\%$, $P < 0·0001$). The mean differences between the estimates of AEE_{DLW} and AEE_{PAQ} are shown in Fig. 3. A significant difference was found between AEE examined by various indirect measures and the direct measures derived from

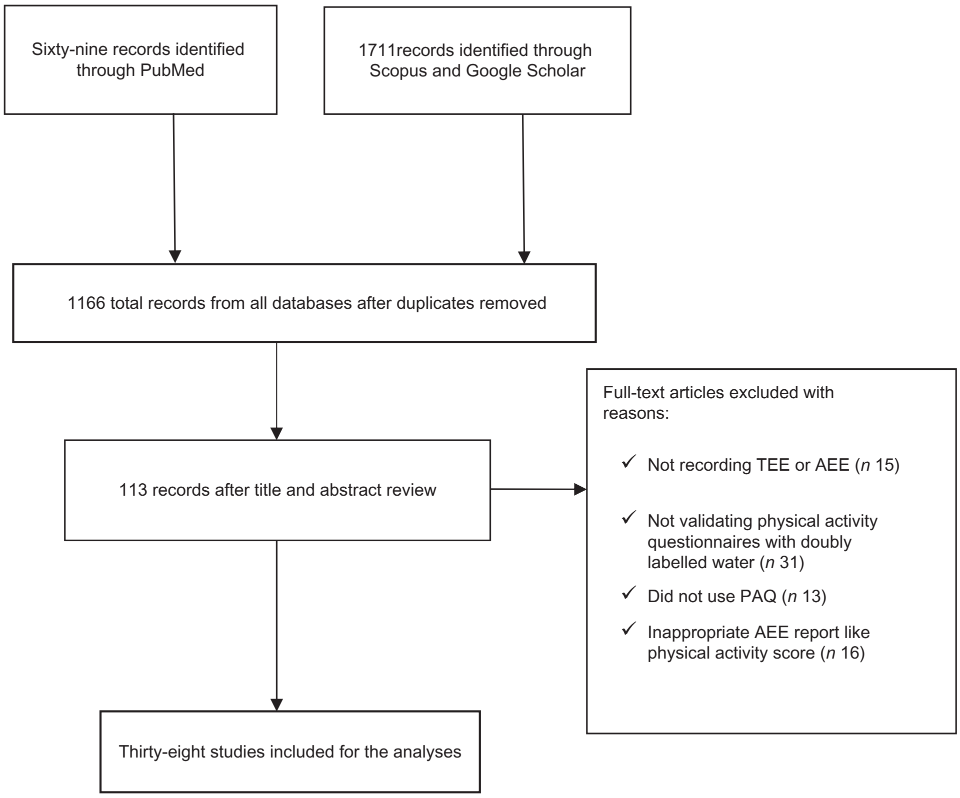

Fig. 1. Study selection process. TEE, total energy expenditure; AEE, activity energy expenditure; PAQ, physical activity questionnaire.

DLW (WMD 414·6, 95 % CI 78·7, 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 (Tables 4 and 5). In thirteen studies that reported information at the individual level, agreement, only two of them showed good agreement. In the study that was conducted by Conway *et al.*[63] on twenty-four subjects, as well as in the study conducted by Sridharan *et al.*[64], for ten subjects, the difference between TEE_{DLW} and TEE 7-d physical activity record was <10 %. A Recent Physical Activity Questionnaire had a narrow limit of agreement with a mean bias of 451 kJ/d (6 %). At the group level, our findings indicated that heterogeneity disappeared in five subgroups of TEE_{PAQ} types including Physical Activity Questionnaire for Adolescents, Active-Q, 7 d physical activity record, the Sedentary Time and Activity Reporting Questionnaire and 3-d PA diaries. Weighted mean differences of TEE were significant for Physical Activity Questionnaire for Adolescents, 7 d physical activity record, Sedentary Time and Activity Reporting Questionnaire and non-significant for Active-Q (0·403) and 3-d PA diaries (0·341). Active Q and 3-d 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 (Sedentary Time and Activity Reporting Questionnaire) but the WMD of AEE were significant for this questionnaire. Also, for AEE only eight studies reported information at the individual level and none 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 and 7). Results showed that mean differences between PAQ and DLW to estimate TEE may be different based on age groups. Differences were significant only in those who were in the range of $13 < \text{age} < 24$ years. 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 PAQ and DLW estimates of 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 than 44 years old, all categories of BMI except those who were overweight, healthy people and BF % between $25 < \text{body fat} < 35$.

Discussion

In this meta-analysis, we identified Active-Q and 3-d PA diaries as indirect tools that had acceptable mean differences and

https://doi.org/10.1017/S0007114520003049 Published online by Cambridge University Press

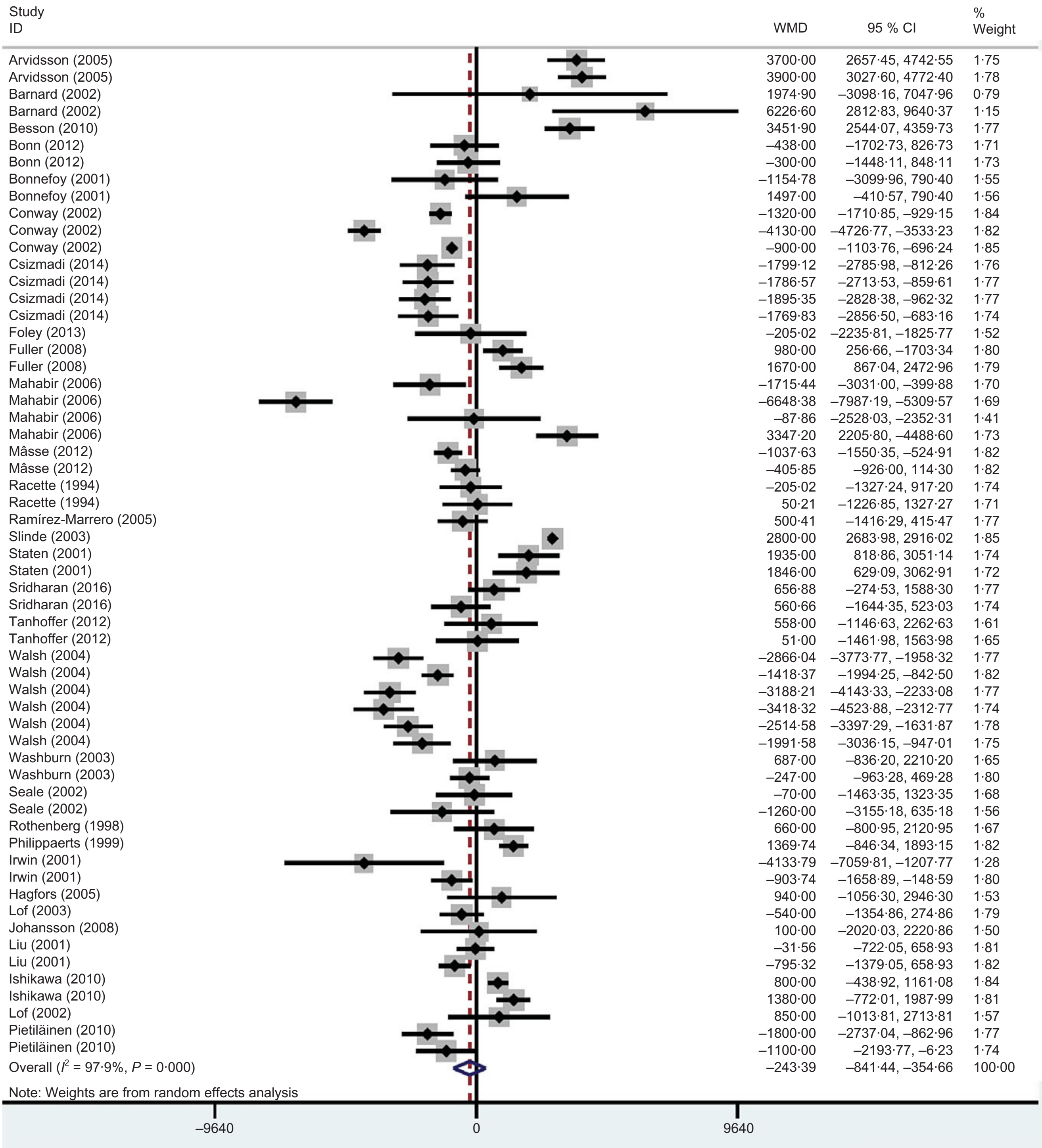

Fig. 2. Forest plot of mean differences of total energy expenditure (TEE) measured by the doubly labelled water method and TEE measured using physical activity questionnaires. WMD, weighted mean difference.

heterogeneity for measuring TEE at the population level. Subgroup analyses showed that the WMD in TEE measured by PAQ and DLW was influenced by age and disease status, but not by sex and the BF %. Moreover, except for sex, all of other predefined criteria including age, disease status, BMI and BF % 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 $>0{\cdot}60$[1]. More precisely, there are some criteria that explain how good

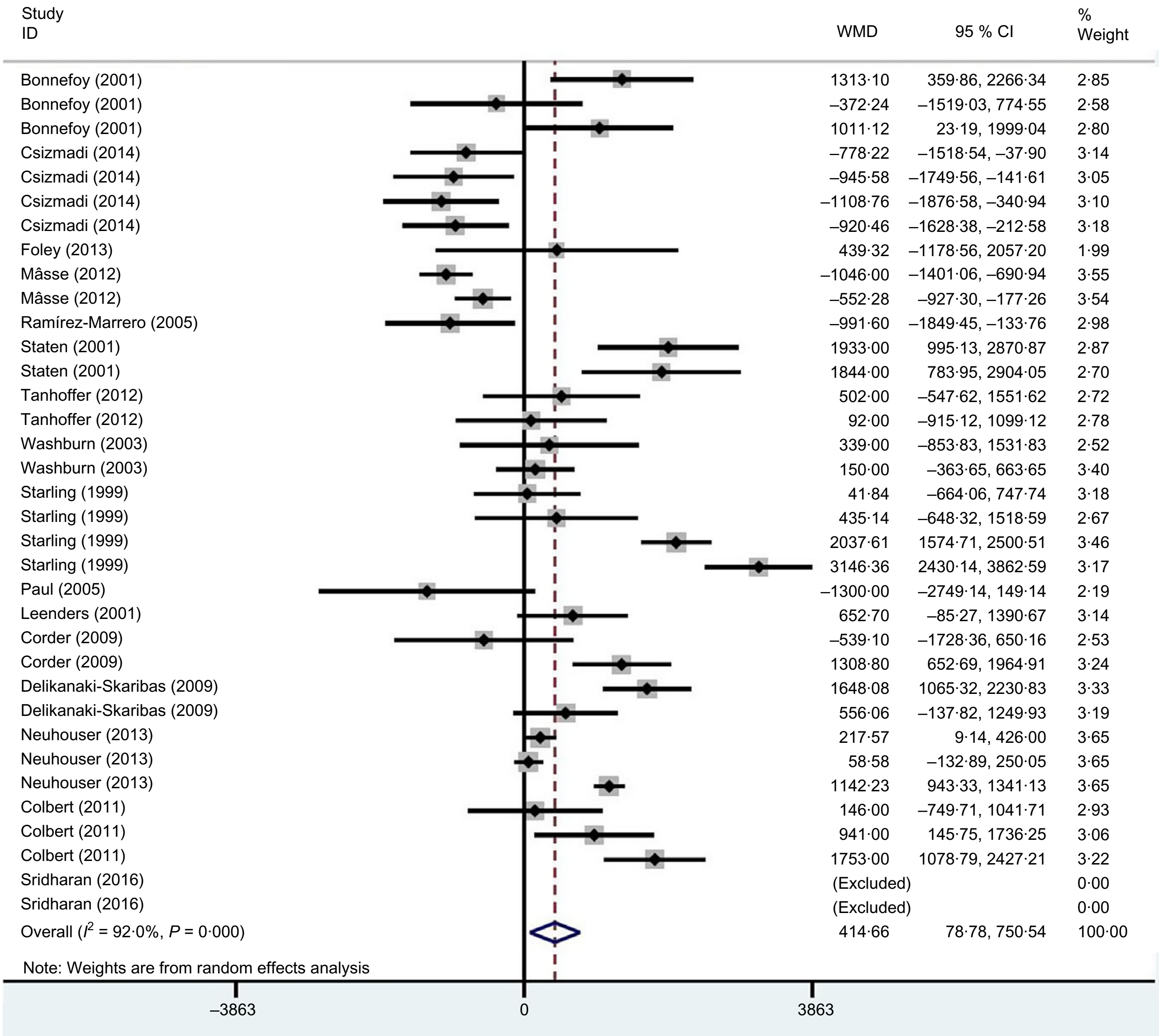

Fig. 3. Forest plot of mean differences of activity energy expenditure (AEE) measured by the doubly labelled water method and AEE measured using physical activity questionnaires. WMD, weighted mean difference.

a PAQ is at the individual level and illustrate whether the questionnaire 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 PAQ 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 Bonn *et al.*[65], the Questionnaire d'Activité Physique Saint-Etienne questionnaire underestimated TEE by 1498 kJ/d (358 kcal/d) with limit of agreement –1075 to 1625 which means that the Questionnaire d'Activité Physique Saint-Etienne has 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 (Recent Physical Activity Questionnaire (*r* 0·67)[66], Multimedia Activity Recall for Children and Adolescents (*r* 0·7)[67], Self-Administered Physical Activity Checklist (*r* 0·6)[68], Minnesota Leisure Time Physical Activity Questionnaire (*r* 0·73)[69], 3-d activity registration (*r* 0·98) and Japan Arteriosclerosis Longitudinal Study Physical Activity Questionnaire (*r* 0·742)[70]).

To estimate AEE, we did not find any PAQ as a suitable measure. Moreover, none of the questionnaires estimating AEE showed acceptable correlation with DLW. Subgroup analyses showed that, in the AEE_{PAQ} group, the WMD was influenced by age, disease status, BMI and BF %.

All the studies included in the review by Neilson *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

Table 4. Agreement between physical activity questionnaire (PAQ) and doubly labelled water (DLW) estimates of total energy expenditure (TEE) stratified by PAQ type
(Mean values and 95 % confidence intervals)

Type of physical activity questionnaire	No. of studies	Mean difference (kJ/d)	95 % CI	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	464·2
RPAQ[27,66]	2	2056·412	–682·65, 4795·4	0·141	<0·001	94·4
Active-Q[65]	2	–362·345	$-1{\cdot}2\times10^3$, 487·737	0·403	0·874	0·0
MNLTPAQ[71]	1	2800·000	2683·978, 2916·022	<0·001	–	–
7-dPAQ[37,39,63–65,79,80,84,86,89]	12	–857·43·766	$-2{\cdot}1\times10^3$, 394·454	0·179	<0·001	93·5
QAPSE[65]	1	1497	–410·57, 3404·56	0·124	–	–
(TEC + MNLTPA + EESLEEP)[36]	1	$-1{\cdot}3\times10^3$	$-1{\cdot}7\times10^3$, –929·152	<0·001	–	–
7-dPArecord[63,89]	2	–900·254	$-1{\cdot}1\times10^3$, –703·526	<0·001	0·993	0·0
STAR-Q[79]	3	$-1{\cdot}8\times10^3$	$-2{\cdot}4\times10^3$, $-1{\cdot}3\times10^3$	<0·001	0·985	0·0
MARCA[67]	1	–205·020	$-2{\cdot}2\times10^3$, 1825·765	0·843	–	–
24-PAD[80]	1	980	256·656, 1703·344	0·008	–	–
Five City Project questionnaire[37]	1	$-1{\cdot}7\times10^3$	$-3{\cdot}0\times10^3$, –399·881	0·011	–	–
Harvard Alumni questionnaire[37]	1	$-6{\cdot}6\times10^3$	$-8{\cdot}0\times10^3$, $-5{\cdot}3\times10^3$	<0·001	–	–
CAPS Four Week Activity Recall[37]	1	–87·860	$-2{\cdot}5\times10^3$, 2352·309	0·944	–	–
CAPS Typical Week Activity Recall[37]	1	3347·2	2205·8, 4488·6	<0·001	–	–
The Checklist questionnaire[81]	1	$-1{\cdot}0\times10^3$	$-1{\cdot}6\times10^3$, –524·906	<0·001	–	–
Global Questionnaire[81]	1	–405·848	–925·999, 114·303	0·126	–	–
SAPAC[68]	1	–500·406	$-1{\cdot}4\times10^3$, 415·472	0·284	–	–
Arizona Activity FFQ 28 d[82]	1	1935	818·855, 3051·145	0·001	–	–
Arizona Activity FFQ 7 d[82]	1	1846	629·092, 3062·908	0·003	–	–
PARA-SCI[83]	1	558·000	$-1{\cdot}1\times10^3$, 2262·631	0·521	–	–
PASIPD[83]	1	51·000	$-1{\cdot}5\times10^3$, 1563·979	0·947	–	–
TEC + MNLTPA[42]	6	$-2{\cdot}5\times10^3$	$-3{\cdot}2\times10^3$, $-1{\cdot}8\times10^3$	<0·001	0·003	72·7
Activity diary in 4 d[30]	1	660·000	–800·951, 2120·951	0·376	–	–
FCQ 7 d index[87]	1	1369·745	846·338, 1893·152	<0·001	–	–
3-d activity registration[90]	1	940·000	$-1{\cdot}1\times10^3$, 2946·303	0·358	–	–
2-week recall[91]	1	–540·000	$-1{\cdot}4\times10^3$, 274·860	0·194	–	–
Two-question questionnaire on physical activity[94]	1	100·000	$-2{\cdot}0\times10^3$, 2220·025	0·926	–	–
Modified YPAS[95]	2	–436·627	$-1{\cdot}2\times10^3$, 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	$-1{\cdot}0\times10^3$, 2713·807	0·371	–	–
3-d PA diaries	2	$-1{\cdot}5\times10^3$	$-2{\cdot}2\times10^3$, –792·095	<0·001	0·341	0·0

PAQA, Physical activity questionnaire for adolescents; MAQ, Modifiable Activity Questionnaire ; RPAQ, Recent Physical Activity Questionnaire ; 7-dPAQ, 7-d 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.
* P for the meta-analysis. $P < 0{\cdot}05$ 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 $P < 0{\cdot}5$ indicates significant heterogeneity across studies.

composed of TEE and DLW. The emphasis in the review by Neilson *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 $AEE_{DLW} - AEE_{PAQ}$ was both evaluated and the included PAQ 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) RMR for estimating TEE and AEE, but as PAQ are considered as feasible approaches to be used in epidemiological studies, it is more sensible to use predicted RMR (RMRp) rather than measured RMR[71]. To reduce the level of over and underestimation of TEE and AEE that are blinded to the use of PAQ in different population with diverse specifications, the best PAQ with the lowest mean differences with DLW should be identified and utilised in epidemiological studies.

There are several causes for over and underestimation of TEE and AEE that are measured with PAQ. First, most equations used to measure predicted RMR, overestimated the BMR compared with the indirect calorimetry, including Schofield[72], Henry *et al.*[73], WHO[74], Schofield BW (body weight) and ht (height)[72] and 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 forty-six PAQ types which were assessed in our study, twenty-five underestimated and twenty-one overestimated TEE. Therefore, both underreporting and overreporting of activities were observed with

British Journal of Nutrition

Table 5. Agreement between physical activity questionnaire (PAQ) and doubly labelled water (DLW) estimates of activity energy expenditure (AEE) stratified by PAQ type
(Mean values and 95 % confidence intervals)

Type of physical activity questionnaire	No. of studies	Mean difference (kJ/d)	95 % CI	P	Test of heterogeneity P	Test of heterogeneity I^2 (%)
YPAS[65,85,93,97]	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	$-1·4 \times 10^3$, –495·738	<0·001	0·831	0·0
7-dPAR[55,64,79,88,96]	6	33·070	–369·996, 436·137	0·872	0·038	60·6
MARCA[67]	1	439·320	$-1·2 \times 10^3$, 2057·198	0·595	–	–
Checklist questionnaire[81]	1	$-1·0 \times 10^3$	$-1·4 \times 10^3$, –690·940	<0·001	–	–
Global Questionnaire[81]	1	–552·280	–927·303, –177·257	0·004	–	–
SAPAC[68]	1	–991·604	$-1·8 \times 10^3$, –133·759	0·023	–	–
MNLTPA[69]	3	2198·583	1282·793, 3114·374	<0·001	0·005	81
The Arizona Activity Frequency Questionnaire 28 d[82,96]	2	1011·841	–664·644, 2688·326	0·237	<0·001	91·8
The Arizona Activity Frequency Questionnaire 7 d[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	$-1·3 \times 10^3$	$-2·7 \times 10^3$, 149·137	0·079	–	–
Youth Physical Activity Questionnaire recall in past week[92]	2	454·150	$-1·4 \times 10^3$, 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	–	–

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-d 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-d physical activity record questionnaire; PASE, Physical Activity Scale for the Elderly; PHQ, Personal Habits Questionnaire; CHAMPS, Community Health Activities Model Program for Seniors; modPASE, modified Physical Activity Scale for the Elderly.

Table 6. Subgroup analysis of mean differences between physical activity questionnaire (PAQ) and doubly labelled water (DLW) estimates of total energy expenditure (TEE) stratified by identified study characteristics
(Mean values and 95 % confidence intervals)

Variables	No. of studies	Mean difference (kJ/d)	95 % CI	P*	Test of heterogeneity† P	Test of heterogeneity† I^2 (%)
Sex						
Men	16	–467·036	$-1·3 \times 10^3$, 363·780	0·271	<0·001	95·5
Women	24	–432·043	$-1·2 \times 10^3$, 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 (years)						
Age < 13	1	–500·406	$-1·4 \times 10^3$, 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	$-1·2 \times 10^3$, 122·301	0·111	<0·001	94·8
44 ≤ age < 64	18	–596·864	$-1·4 \times 10^3$, 177·626	0·131	<0·001	93·1
Age ≥ 64	7	–234·563	–819·655, 350·529	0·432	0·117	41·1
BMI (kg/m^2)						
BMI < 18·5	1	–500·406	$-1·4 \times 10^3$, 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	$-1·4 \times 10^3$, –72·568	0·030	<0·001	93·3
30 ≤ BMI < 35	5	–742·724	$-1·3 \times 10^3$, –183·225	0·009	0·038	60·5
Disease						
Healthy individuals	55	–244·285	–941·282, 452·712	0·545	<0·001	98·1
Chronic kidney disease	2	80·917	$-1·1 \times 10^3$, 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	$-1·8 \times 10^3$, 642·891	0·355	<0·001	97·1
25 ≤ body fat <35	13	25·160	$-1·2 \times 10^3$, 1249·203	0·968	<0·001	95·2
Body fat ≥ 35	11	$-1·0 \times 10^3$	$-2·3 \times 10^3$, 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.

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}.

Table 7. Subgroup analysis of mean differences between physical activity questionnaire (PAQ) and doubly labelled water (DLW) estimates of Activity energy expenditure (AEE) stratified by identified study characteristics
(Mean values and 95 % confidence intervals)

Variables	No. of studies	Mean difference (kJ/d)	95 % CI	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 (years)						
Age < 13	1	–991·604	$-1·8 \times 10^3$, –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	$-1·1 \times 10^3$, –638·864	<0·001	0·527	0·0
Age ≥ 64	15	958·987	529·831, 1388·144	<0·001	<0·001	92·6
BMI (kg/m²)						
BMI < 18·5	2	–836·739	$-1·5 \times 10^3$, –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	$-1·3 \times 10^3$, –319·204	0·001	0·061	71·5
Disease						
Healthy individuals	31	421·428	72·707, 770·14	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 ≥ 5	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.

Likewise, the study by Walsh *et al.*[42] demonstrated that the order of 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/d to 50 kJ/d). Third, all of the included articles used metabolic equivalent 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 Montoye *et al.*)[77]. In most PAQ, the use of metabolic equivalent values for estimating the energy expenditure of a particular activity is considered a limitation[42]. When the metabolic equivalent value is administered for a specific activity, the same energy cost per kg 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 PAQ had the least mean difference with DLW and none of the PAQ showed good measure of AEE. This is because the magnitude of difference between PAQ and DLW estimates of TEE and AEE depends on some factors including the type of PAQ, the sex of the population on which the questionnaire was used and the number of activities measured by the PAQ. 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 television, 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 PAQ do not estimate all physical activity especially in low-intensity level, an underreporting of AEE is anticipated[23]. However, some PAQ like IPAQ and Physical Activity Questionnaire for Adolescents can capture low- to high-intensity level physical activities and the underreporting of TEE in these questionnaires is compensated by overreporting of vigorous physical activity[78].

In conclusion, our meta-analysis identified PAQ (Active-Q) and 3-d PA diaries that had sufficient validity for measuring TEE based on the mean correspondence in group level. However, as each of these questionnaires was used only in one study, we may conclude that this finding might be due to a chance and requires further verification. The present study provides evidence highlighting that the majority of PAQ compared with 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 3-d PA diaries or direct measurement methods in epidemiological studies might yield more reliable findings.

British Journal of Nutrition

Acknowledgements

The authors would like to thank all research staff involved in the study, and Tehran University of Medical Science for financial support.

This article was financially supported by Tehran University of Medical Science.

The authors contributed equally to this work.

The authors have no conflicts of interest to disclose.

Supplementary material

For supplementary material referred to in this article, please visit https://doi.org/10.1017/S0007114520003049

References

1. Neilson HK, Robson PJ, Friedenreich CM, *et al.* (2008) Estimating activity energy expenditure: how valid are physical activity questionnaires? *Am J Clin Nutr* **87**, 279–291.
2. Donahoo WT, Levine JA & Melanson EL (2004) Variability in energy expenditure and its components. *Curr Opin Clin Nutr Metab Care* **7**, 599–605.
3. Speakman JR & Westerterp KR (2010) Associations between energy demands, physical activity, and body composition in adult humans between 18 and 96 y of age. *Am J Clin Nutr* **92**, 826–834.
4. Li J & Siegrist J (2012) Physical activity and risk of cardiovascular disease—a meta-analysis of prospective cohort studies. *Int J Environ Res Public Health* **9**, 391–407.
5. Siervo M, Oggioni C, Lara J, *et al.* (2015) Age-related changes in resting energy expenditure in normal weight, overweight and obese men and women. *Maturitas* **80**, 406–413.
6. Milanović Z, Pantelić S, Trajković N, *et al.* (2013) Age-related decrease in physical activity and functional fitness among elderly men and women. *Clin Interv Aging* **8**, 549.
7. Hall KD, Heymsfield SB, Kemnitz JW, *et al.* (2012) Energy balance and its components: implications for body weight regulation. *Am J Clin Nutr* **95**, 989–994.
8. Hu FB, Leitzmann MF, Stampfer MJ, *et al.* (2001) Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med* **161**, 1542–1548.
9. Bassuk SS & Manson JE (2005) Epidemiological evidence for the role of physical activity in reducing risk of type 2 diabetes and cardiovascular disease. *J Appl Physiol (1985)* **99**, 1193–1204.
10. Moses A, Slater C, Preston T, *et al.* (2004) 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. *Br J Cancer* **90**, 996.
11. Uehara M, Plank LD & Hill GL (1999) Components of energy expenditure in patients with severe sepsis and major trauma: a basis for clinical care. *Crit Care Med* **27**, 1295–1302.
12. Hunter GR, Wetzstein CJ, Fields DA, *et al.* (2000) Resistance training increases total energy expenditure and free-living physical activity in older adults. *J Appl Physiol (1985)* **89**, 977–984.
13. Speakman J (1997) *Doubly Labelled Water: Theory and Practice*. Berlin: Springer Science & Business Media.
14. Westerterp KR (2008) Physical activity as determinant of daily energy expenditure. *Physiol Behav* **93**, 1039–1043.
15. Craig CL, Marshall AL, Sjöström M, *et al.* (2003) International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* **35**, 1381–1395.
16. Fujii H, Yamamoto S, Takeda-Imai F, *et al.* (2011) Validity and applicability of a simple questionnaire for the estimation of total and domain-specific physical activity. *Diabetol Int* **2**, 47–54.
17. Department of Health and Human Services (1996) *Physical Activity and Health: A Report of the Surgeon General*. Atlanta, GA: USDHSS, Centers for Disease Control and Prevention.
18. Montoye HJ, Kemper HCG, Saris WHM, *et al.* (1996) *Measuring Physical Activity and Energy Expenditure*. Champaign, IL: Human Kinetics.
19. Wendel-Vos GW, Schuit AJ, Saris WH, *et al.* (2003) Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. *J Clin Epidemiol* **56**, 1163–1169.
20. Dhurandhar NV, Schoeller D, Brown AW, *et al.* (2015) Energy balance measurement: when something is not better than nothing. *Int J Obes* **39**, 1109.
21. Neilson HK, Robson PJ, Friedenreich CM, *et al.* (2008) Estimating activity energy expenditure: how valid are physical activity questionnaires? *Am J Clin Nutr* **87**, 279–291.
22. Melanson Jr EL, Freedson PS & Blair S (1996) Physical activity assessment: a review of methods. *Crit Rev Food Sci Nutr* **36**, 385–396.
23. Prince SA, Adamo KB, Hamel ME, *et al.* (2008) A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act* **5**, 56.
24. Herzog R, Álvarez-Pasquin MJ, Díaz C, *et al.* (2013) Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review. *BMC Public Health* **13**, 154.
25. Higgins JPT, Thompson SG, Deeks JJ, *et al.* (2003) Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560.
26. Mantel N & Haenszel W (1959) Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* **22**, 719–748.
27. Arvidsson D, Slinde F & Hulthen L (2005) Physical activity questionnaire for adolescents validated against doubly labelled water. *Eur J Clin Nutr* **59**, 376–383.
28. Bonn SE, Trolle Lagerros Y, Christensen SE, *et al.* (2012) Active-Q: validation of the web-based physical activity questionnaire using doubly labeled water. *J Med Internet Res* **14**, e29.
29. Slinde F, Arvidsson D, Sjoberg A, *et al.* (2003) Minnesota Leisure Time Activity Questionnaire and doubly labeled water in adolescents. *Med Sci Sports Exerc* **35**, 1923–1928.
30. Rothenberg E, Bosaeus I, Lernfelt B, *et al.* (1998) Energy intake and expenditure: validation of a diet history by heart rate monitoring, activity diary and doubly labeled water. *Eur J Clin Nutr* **52**, 832–838.
31. Löf M, Hannestad U & Forsum E (2002) Assessing physical activity of women of childbearing age. Ongoing work to develop and evaluate simple methods. *Food Nutr Bull* **23**, Suppl. 3, 30–33.
32. Lof M, Hannestad U & Forsum E (2003) 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* **90**, 961–968.
33. Hagfors L, Westerterp K, Skoldstam L, *et al.* (2005) 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* **59**, 238–245.
34. Barnard J, Tapsell LC, Davies P, *et al.* (2002) 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. *Eur J Clin Nutr* **56**, 358–367.
35. Bonnefoy M, Normand S, Pachiaudi C, *et al.* (2001) Simultaneous validation of ten physical activity questionnaires in older men: a doubly labeled water study. *J Am Geriatr Soc* **49**, 28–35.
36. Conway JM, Irwin ML & Ainsworth BE (2002) Estimating energy expenditure from the Minnesota Leisure Time Physical Activity and Tecumseh Occupational Activity questionnaires – a doubly labeled water validation. *J Clin Epidemiol* **55**, 392–399.
37. Mahabir S, Baer DJ, Giffen C, *et al.* (2006) Comparison of energy expenditure estimates from 4 physical activity questionnaires with doubly labeled water estimates in postmenopausal women. *Am J Clin Nutr* **84**, 230–236.
38. Mâsse LC, Fulton JE, Watson KB, *et al.* (2012) Comparing the validity of 2 physical activity questionnaire formats in African-American and Hispanic women. *J Phys Act Health* **9**, 237–248.
39. Racette SB, Schoeller DA & Kushner RF (1995) Comparison of heart rate and physical activity recall with doubly labeled water in obese women. *Med Sci Sports Exerc* **27**, 126–133.
40. Ramírez-Marrero FA, Smith BA, Sherman WM, *et al.* (2005) Comparison of methods to estimate physical activity and energy expenditure in African American children. *Int J Sports Med* **26**, 363–371.
41. Staten LK, Taren DL, Howell WH, *et al.* (2001) Validation of the Arizona Activity Frequency Questionnaire using doubly labeled water. *Med Sci Sports Exerc* **33**, 1959–1967.
42. Walsh MC, Hunter GR, Sirikul B, *et al.* (2004) Comparison of self-reported with objectively assessed energy expenditure in black and white women before and after weight loss. *Am J Clin Nutr* **79**, 1013–1039.
43. Washburn RA, Jacobsen DJ, Sonko BJ, *et al.* (2003) The validity of the Stanford Seven-Day Physical Activity Recall in young adults. *Med Sci Sports Exerc* **35**, 1374–1380.
44. Starling RD, Matthews DE, Ades PA, *et al.* (1999) Assessment of physical activity in older individuals: a doubly labeled water study. *J Appl Physiol (1985)* **86**, 2090–2096.
45. Seale JL, Klein G, Friedmann J, *et al.* (2002) Energy expenditure measured by doubly labeled water, activity recall, and diet records in the rural elderly. *Nutrition* **18**, 568–573.
46. Leenders NY, Sherman WM, Nagaraja H, *et al.* (2001) Evaluation of methods to assess physical activity in free-living conditions. *Med Sci Sports Exerc* **33**, 1233–1240.
47. Paul DR, Rhodes DG, Kramer M, *et al.* (2005) Validation of a food frequency questionnaire by direct measurement of habitual ad libitum food intake. *Am J Epidemiol* **162**, 806–814.
48. Irwin ML, Ainsworth BE & Conway JM (2001) Estimation of energy expenditure from physical activity measures: determinants of accuracy. *Obes Res* **9**, 517–525.
49. Delikanaki-Skaribas E, Trail M, Wong WW, *et al.* (2009) Daily energy expenditure, physical activity, and weight loss in Parkinson's disease patients. *Mov Disord* **24**, 667–671.
50. Neuhouser ML, Di C, Tinker LF, *et al.* (2013) Physical activity assessment: biomarkers and self-report of activity-related energy expenditure in the WHI. *Am J Epidemiol* **177**, 576–585.
51. Csizmadi I, Neilson HK, Kopciuk KA, *et al.* (2014) The Sedentary Time and Activity Reporting Questionnaire (STAR-Q): reliability and validity against doubly labeled water and 7-day activity diaries. *Am J Epidemiol* **180**, 424–435.
52. Foley LS, Maddison R, Rush E, *et al.* (2013) Doubly labeled water validation of a computerized use-of-time recall in active young people. *Metabolism* **62**, 163–169.
53. Tanhoffer RA, Tanhoffer AI, Raymond J, *et al.* (2012) Comparison of methods to assess energy expenditure and physical activity in people with spinal cord injury. *J Spinal Cord Med* **35**, 35–45.
54. Besson H, Brage S, Jakes RW, *et al.* (2010) Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. *Am J Clin Nutr* **91**, 106–114.
55. Fuller Z, Horgan G, O'Reilly LM, *et al.* (2008) Comparing different measures of energy expenditure in human subjects resident in a metabolic facility. *Eur J Clin Nutr* **62**, 560–569.
56. Sridharan S, Wong J, Vilar E, *et al.* (2016) Comparison of energy estimates in chronic kidney disease using doubly-labelled water. *J Hum Nutr Diet* **29**, 59–66.
57. Liu B, Woo J, Tang N, *et al.* (2001) Assessment of total energy expenditure in a Chinese population by a physical activity questionnaire: examination of validity. *Int J Food Sci Nutr* **52**, 269–282.
58. Corder K, Brage S, Wright A, *et al.* (2010) Physical activity energy expenditure of adolescents in India. *Obesity* **18**, 2212–2219.
59. Philippaerts R, Westerterp K & Lefevre J (1999) Doubly labeled water validation of three physical activity questionnaires. *Int J Sports Med* **20**, 284–289.
60. Johansson G & Westerterp KR (2008) Assessment of the physical activity level with two questions: validation with doubly labeled water. *Int J Obes* **32**, 1031–1033.
61. Ishikawa-Takata K, Naito Y, Tanaka S, *et al.* (2011) Use of doubly labeled water to validate a physical activity questionnaire developed for the Japanese population. *J Epidemiol* **21**, 114–121.
62. Pietilainen KH, Korkeila M, Bogl LH, *et al.* (2010) Inaccuracies in food and physical activity diaries of obese subjects: complementary evidence from doubly labeled water and co-twin assessments. *Int J Obes* **34**, 437–445.
63. Conway JM, Seale JL, Jacobs DR, *et al.* (2002) Comparison of energy expenditure estimates from doubly labeled water, a physical activity questionnaire, and physical activity records. *Am J Clin Nutr* **75**, 519–525.
64. Sridharan S, Wong J, Vilar E, *et al.* (2016) Comparison of energy estimates in chronic kidney disease using doubly-labelled water. *J Hum Nutr Diet* **29**, 59–66.
65. Bonn SE, Lagerros YT, Christensen SE, *et al.* (2012) Active-Q: validation of the web-based physical activity questionnaire using doubly labeled water. *J Med Internet Res* **14**, e29.
66. Besson H, Brage S, Jakes RW, *et al.* (2010) Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. *Am J Clin Nutr* **91**, 106–114.
67. Foley LS, Maddison R, Rush E, *et al.* (2013) Doubly labeled water validation of a computerized use-of-time recall in active young people. *Metab Clin Exp* **62**, 163–169.
68. Ramírez-Marrero F, Smith B, Sherman W, *et al.* (2005) Comparison of methods to estimate physical activity and energy expenditure in African American children. *Int J Sports Med* **26**, 363–371.
69. Slinde F, Arvidsson D, Sjoberg A, *et al.* (2003) Minnesota leisure time activity questionnaire and doubly labeled water in adolescents. *Med Sci Sports Exerc* **35**, 1923–1928.
70. Ishikawa-Takata K, Naito Y, Tanaka S, *et al.* (2011) Use of doubly labeled water to validate a physical activity questionnaire developed for the Japanese population. *J Epidemiol* **21**, 114–121.

British Journal of Nutrition

71. Slinde F, Arvidsson D & Sjöberg A, *et al.* (2003) Minnesota leisure time activity questionnaire and doubly labeled water in adolescents. *Med Sci Sports Exerc* **35**, 1923–1928.
72. Schofield W (1985) Predicting basal metabolic rate, new standards and review of previous work. *Hum Nutr Clin Nutr* **39**, 5–41.
73. Henry C, Dyer S & Ghusain-Choueiri A (1999) New equations to estimate basal metabolic rate in children aged 10–15 years. *Eur J Clin Nutr* **53**, 134–142.
74. World Health Organization (1985) *Energy and Protein Requirements: Report of a Joint FAO/WHO/UNU Expert Consultation (WHO Technical Report Series, no 724)*. Geneva: World Health Organization.
75. Molnár D, Jeges S, Erhardt E, *et al.* (1995) Measured and predicted resting metabolic rate in obese and nonobese adolescents. *J Pediatr* **127**, 571–577.
76. Barnard J, Tapsell LC, Davies P, *et al.* (2002) 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. *Eur J Clin Nutr* **56**, 358–368.
77. Montoye HJ, Kemper HC, Saris WH, *et al.* (1996) *Measuring Physical Activity and Energy Expenditure*. Champaign, IL: Human Kinetics.
78. Arvidsson D, Slinde F & Hulthén L (2005) Physical activity questionnaire for adolescents validated against doubly labelled water. *Eur J Clin Nutr* **59**, 376–383.
79. Csizmadi I, Neilson HK, Kopciuk KA, *et al.* (2014) The Sedentary Time and Activity Reporting Questionnaire (STAR-Q): reliability and validity against doubly labeled water and 7-day activity diaries. *Am J Epidemiol* **180**, 424–435.
80. Fuller Z, Horgan G, O'Reilly L, *et al.* (2008) Comparing different measures of energy expenditure in human subjects resident in a metabolic facility. *Eur J Clin Nutr* **62**, 560–569.
81. Mâsse LC, Fulton JE, Watson KB, *et al.* (2012) Comparing the validity of 2 physical activity questionnaire formats in African-American and Hispanic women. *J Phys Act Health* **9**, 237–248.
82. Staten LK, Taren DL, Howell WH, *et al.* (2001) Validation of the Arizona Activity Frequency Questionnaire using doubly labeled water. *Med Sci Sports Exerc* **33**, 1959–1967.
83. Tanhoffer RA, Tanhoffer AI, Raymond J, *et al.* (2012) Comparison of methods to assess energy expenditure and physical activity in people with spinal cord injury. *J Spinal Cord Med* **35**, 35–45.
84. Washburn RA, Jacobsen DJ, Sonko BJ, *et al.* (2003) The validity of the Stanford Seven-Day Physical Activity Recall in young adults. *Med Sci Sports Exerc* **35**, 1374–1380.
85. Starling RD, Matthews DE, Ades PA, *et al.* (1999) Assessment of physical activity in older individuals: a doubly labeled water study. *J Appl Physiol (1985)* **86**, 2090–2096.
86. Seale JL, Klein G, Friedmann J, *et al.* (2002) Energy expenditure measured by doubly labeled water, activity recall, and diet records in the rural elderly. *Nutrition* **18**, 568–573.
87. Philippaerts RM, Westerterp KR & Lefevre J (1999) Doubly labelled water validation of three physical activity questionnaires. *Int J Sports Med* **20**, 284–289.
88. Leenders N, Sherman WM, Nagaraja H, *et al.* (2001) Evaluation of methods to assess physical activity in free-living conditions. *Med Sci Sports Exerc* **33**, 1233–1240.
89. Irwin ML, Ainsworth BE & Conway JM (2001) Estimation of energy expenditure from physical activity measures: determinants of accuracy. *Obesity* **9**, 517–525.
90. Hagfors L, Westerterp K, Sköldstam L, *et al.* (2005) 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* **59**, 238–245.
91. Lof M, Hannestad U & Forsum E (2003) 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* **90**, 961–968.
92. Corder K, Brage S, Wright A, *et al.* (2010) Physical activity energy expenditure of adolescents in India. *Obesity* **18**, 2212–2219.
93. Delikanaki-Skaribas E, Trail M, Wong WWL, *et al.* (2009) Daily energy expenditure, physical activity, and weight loss in Parkinson's disease patients. *Mov Disord* **24**, 667–671.
94. Johansson G & Westerterp K (2008) Assessment of the physical activity level with two questions: validation with doubly labeled water. *Int J Obes* **32**, 1031.
95. Liu B, Woo J, Tang N, *et al.* (2001) Assessment of total energy expenditure in a Chinese population by a physical activity questionnaire: examination of validity. *Int J Food Sci Nutr* **52**, 269–282.
96. Neuhouser ML, Di C, Tinker LF, *et al.* (2013) Physical activity assessment: biomarkers and self-report of activity-related energy expenditure in the WHI. *Am J Epidemiol* **177**, 576–585.
97. Colbert LH, Matthews CE, Havighurst TC, *et al.* (2011) Comparative validity of physical activity measures in older adults. *Med Sci Sports Exerc* **43**, 867.
98. Lof M, Hannestad U & Forsum E (2002) Assessing physical activity of women of childbearing age. Ongoing work to develop and evaluate simple methods. *Food Nutr Bull* **23**, Suppl. 3, 30–33.
99. Pietiläinen K, Korkeila M, Bogl L, *et al.* (2010) Inaccuracies in food and physical activity diaries of obese subjects: complementary evidence from doubly labeled water and co-twin assessments. *Int J Obes* **34**, 437.
100. Bonn SE, Lagerros YT, Christensen SE, *et al.* (2012) Active-Q: validation of the web-based physical activity questionnaire using doubly labeled water. *J Med Int Res* **14**, e29.

British Journal of Nutrition