International Journal of Obesity (2011) 35, S79–S87

www.nature.com/ijo

ORIGINAL ARTICLE

The IDEFICS validation study on field methods for assessing physical activity and body composition in children: design and data collection

K Bammann[1], I Sioen[2], I Huybrechts[2], JA Casajús[3], G Vicente-Rodríguez[3], R Cuthill[4], K Konstabel[5], B Tubić[6], N Wawro[1], M Rayson[7], K Westerterp[8], S Mårild[6], YP Pitsiladis[4], JJ Reilly[9], LA Moreno[3] and S De Henauw[2], on behalf of the IDEFICS Consortium

[1]*Department of Biometry, Bremen Institute for Prevention Research and Social Medicine, University of Bremen, Bremen, Germany;* [2]*Department of Public Health, Ghent University, Ghent, Belgium;* [3]*Growth, Exercise, Nutrition and Development Research Group, University of Zaragoza, Zaragoza, Spain;* [4]*Faculty of Biomedical and Life Sciences, Integrative and Systems Biology, University of Glasgow, Glasgow, UK;* [5]*Department of Chronic Diseases, National Institute for Health Development, Tallinn, Estonia;* [6]*Department of Paediatrics, Institute for Clinical Sciences, The Queen Silvia Children's Hospital, The Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden;* [7]*BioTel Ltd, Bristol, UK;* [8]*Department of Human Biology, Maastricht University Medical Centre, Maastricht, The Netherlands and* [9]*Physical Activity for Health Research Group, School of Psychological Sciences and Health, University of Strathclyde, Glasgow, UK*

Objective: To describe the design, measurements and fieldwork of the IDEFICS (Identification and prevention of dietary- and lifestyle-induced health effects in children and infants) physical activity and body composition validation study, and to determine the potential and limitations of the data obtained.
Design: Multicentre validation study.
Subjects: A total of 98 children from four different European countries (age: 4–10 years).
Methods: An 8-day measurement protocol was carried out in all children using a collaborative protocol. Reference methods were the doubly labelled water method for physical activity, and a three- and a four-compartment model for body composition. Investigated field methods were accelerometers, a physical activity questionnaire and various anthropometric measurements.
Results: For the validation of physical activity field methods, it was possible to gather data from 83 to 89 children, laying the basis for age- and sex-specific results. The validation of body composition field methods is possible in 64–80 children and allows sex-specific analyses but has only limited statistical power in the youngest age group (<6 years). The amount of activity energy expenditure (AEE) varied between centres, sexes and age groups, with boys and older children having higher estimates of AEE. After normalisation of AEE by body weight, most group-specific differences diminished, except for country-specific differences.
Conclusion: The IDEFICS validation study will allow age- and sex-specific investigation of questions pertaining to the validity of several field methods of body composition and physical activity, using established reference methods in four different European countries. From the participant analyses it can be concluded that the compliance for the investigated field methods was higher than that for the reference methods used in this validation study.
International Journal of Obesity (2011) **35**, S79–S87; doi:10.1038/ijo.2011.38

Keywords: validation study; body composition; physical activity; accelerometry; doubly labelled water

Introduction

For studying childhood obesity, the identification of children with excess body fat is crucial. For this purpose, valid methods for assessing body composition in children are necessary. The same applies to physical activity as one important determinant of obesity. Within the IDEFICS (Identification and prevention of dietary- and lifestyle-induced health effects in children and infants) surveys, physical activity and body composition were assessed in 16 224 children 2–9 years of age in eight different European countries.[1,2] Application of laboratory methods is not feasible in such a large-scale epidemiological study, as they are rather expensive and too time consuming for fieldwork. All potential field methods, however, do not measure energy

Correspondence: Dr K Bammann, Bremen Institute for Prevention Research and Social Medicine, University of Bremen, Achterstrasse 30, 28359 Bremen, Germany.
E-mail: bammann@bips.uni-bremen.de

expenditure or body composition directly but rather estimate these using other parameters. Depending on the measurements involved, the various field methods produce different systematic errors regarding the estimation of energy expenditure as well as body composition, and this error is dependent on other variables such as age or sex. For a review of field methods for the assessment of body composition and levels of physical activity, see Goran,[3] Reilly *et al*.[4] For IDEFICS surveys, it was decided to use different field methods for assessing physical activity and body composition in parallel. Previous validation studies have found ambiguous results pertaining to the validity of accelerometer[5–9] and body composition measurements[10–13] in children. These findings might be partly due to different devices and measurement procedures, and also due to insufficient sample size, especially in the younger IDEFICS children. Therefore, a validation study was carried out within the framework of the IDEFICS study.

This paper describes the design and data collection of the IDEFICS validation study. Results from the recruitment phase are presented and discussed in light of the young age group and the multicentre nature of the validation study. The potential and limitations for investigations using these data are critically discussed.

Materials and methods

Study design

The fieldwork of the validation study was carried out from October 2008 to July 2009 in convenience samples of healthy children aged 4–10 years in four different centres, namely the universities of Ghent, Belgium; Glasgow, United Kingdom; Gothenburg, Sweden and Zaragoza, Spain. The decision for taking convenience samples was made because the burden of taking measurements for participating children and their parents was deemed too high for a random sample. Sample size calculations showed that 31 children in each stratum are sufficient for detecting mean differences of 5% between reference and field methods at a significance level of 0.05, with a statistical power of 80% for body composition and physical activity assessment methods. It was decided *a priori* that stratification by two age groups (4 to <6 years, $\geqslant 6$ years) should be carried out for all analyses. For cost reasons, stratification by sex was restricted to the upper age group. Thus, the targeted sample size was 93 children.

To obtain the required age and sex distribution in the sample, recruitment was monitored throughout the data collection phase and recruitment efforts were reinforced in underrepresented cells of the recruitment scheme. Ways of approaching the study subjects comprised recruitment through schools and newspapers, asking colleagues at the university or friends of the researchers. Except for the Spanish centre, where all subjects were recruited through schools, all other centres reported having problems recruiting study subjects because of the high burden the study protocol placed on the subjects and their parents. Gothenburg was different from the other validation study centres as they recruited children who were being treated in obesity clinics and were free of concurrent diseases. This approach was chosen to increase *post hoc* the proportion of obese children in the validation study sample, as other validation study centres reported difficulties recruiting overweight and obese children from the general population. All measurements were usually taken within 8 days with a fixed schedule that is depicted in Figure 1.

Physical activity and body composition assessment methods

Three field methods for assessing physical activity and two field methods for assessing body composition were included in the IDEFICS validation study to be compared with reference methods. All methods are presented in Table 1 and described in the following sections.

Reference method for assessing physical activity. The most widespread reference method for human energy expenditure is the doubly labelled water (DLW) method. The principle behind this method is well described by Prentice *et al*.[14] Stable isotopes of oxygen (^{18}O) and hydrogen (^{2}H) are given in a single oral dose to the participant and they equilibrate in body water. The hydrogen isotope is eliminated as water, and the oxygen isotope as water and CO_2. Urine samples are

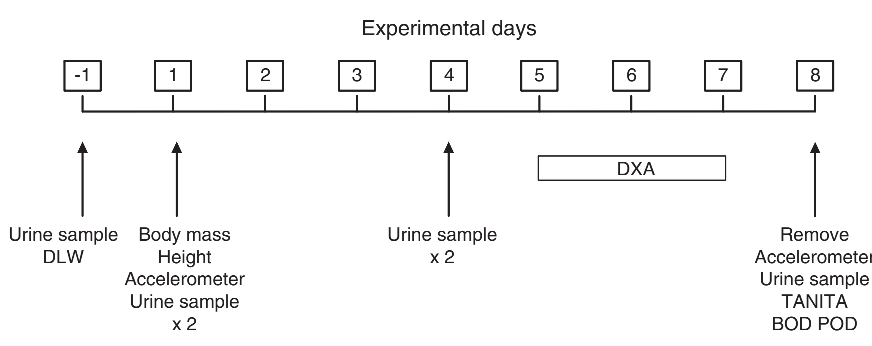

Figure 1 Measurement schedule of the IDEFICS validation study.

Table 1 Methods included in the IDEFICS validation study

Outcome	*Method*	*Measurement*	*Used in the IDEFICS survey*
Physical activity	Doubly labelled water (reference method)	Total energy expenditure by doubly labelled water, estimated resting energy expenditure, estimated diet-induced energy expenditure, activity energy expenditure calculated from these	No
	Uniaxial (1D) accelerometer	Movement during waking hours in one plane	Yes
	Triaxial (3D) accelerometer	Movement during waking hours in three planes	No[a]
	Outdoor playtime checklist	Outdoor playtime assessed by self-administered parental questionnaire	Yes
Body composition	Three-component model (reference method A)	Body mass by TANITA BC 420 scale, actual body volume by BOD POD, total body water by DLW, fat mass calculated from these	No
	Four-component model (reference method B)	Body mass by TANITA BC 420 scale, actual body volume by BOD POD, total body water by DLW, bone mineral mass by DXA, fat mass calculated from these	No
	Skinfold thickness: two sites	Subcutaneous fat at defined sites: triceps, subscapular	Yes
	Skinfold thickness: four sites	Subcutaneous fat at defined sites: biceps, triceps, subscapular, suprailliac	Yes
	Skinfold thickness: six sites	Subcutaneous fat at defined sites: biceps, triceps, subscapular, suprailliac, thigh, calf	Yes
	Circumferences: four sites	Limb girths at defined sites: waist, hip, mid-upper arm, neck	Yes
	Prototype TANITA BC 420 scale	Leg-to-leg bioelectrical impedance	Yes

Abbreviation: DXA, dual-energy X-ray absorptiometry; IDEFICS, Identification and prevention of dietary- and lifestyle-induced health effects in children and infants.
[a]The triaxial accelerometer was initially planned to be included in the IDEFICS survey, but had to be discarded because of cost reasons. Nevertheless, it was decided to leave it in the validation study.

collected at fixed time points and the relative abundance of the isotopes is measured by isotope ratio mass spectrometry. The difference between both isotopes in the urine samples is then used to estimate the expired CO_2 or total energy expenditure (TEE). For obtaining an estimate of activity energy expenditure (AEE), the resting energy expenditure (REE) and diet-induced energy expenditure (thermic effect of food) have to be subtracted from the TEE.

In our study, the Schofield equations specific to 3- to 10-year-old girls and boys based on body mass were used to estimate REE,[15] because accurate measurement of REE is impractical in young children. Diet-induced energy expenditure was set to 10%, assuming the children are in energy balance.[16]

Field methods for assessing physical activity. Three field methods for assessing levels of physical activity were included in the IDEFICS validation study, in two accelerometers and in one questionnaire. The uniaxial accelerometer, Actitrainer (ActiGraph, Pensacola, FL, USA), measures acceleration in a single plane. The Actitrainer has a sampling range of 0.25–2.5 Hz and a sampling frequency of 30 Hz. In addition, a triaxial accelerometer that was not used in the IDEFICS survey was included in the validation study. The triaxial accelerometer, 3DNX (BioTel Ltd, Bristol, UK), measures acceleration in three movement planes. It has a sampling range of 0.2–10 Hz and a sampling frequency of 100 Hz. Movement was recorded at 15 s epochs. Parents were asked to attach both accelerometer units to the right hip of the child during their waking day for seven consecutive days and to fill in a diary for assessing times and reasons when the accelerometers were not worn. Accelerometer measurements were considered to be valid if at least 3-day measurements with a minimum of 6-h daily wearing time were available.

As a third method, indicators of physical activity were assessed using the outdoor playtime checklist, a self-administered questionnaire that was answered by parents. The questionnaire was significantly correlated with objective measures of physical activity in preschool children in a previous study.[17] The questionnaire can be found in Appendix.

Reference methods for assessing body composition. The most interesting compartment of the human body in the context of obesity, fat mass, cannot be directly measured in living individuals.[18] Therefore, models are used to derive fat mass from the measurements of other compartments. Both three- and four-compartment (3C, 4C) models are considered to be valid reference methods for estimating fat mass in children.[19] These model the human body as

$$\text{Body mass} = \text{fat mass} + \text{fat} - \text{free dry mass} + \text{total body water} \quad \text{(3C)}$$

$$\text{Body mass} = \text{fat mass} + \text{lean dry mass} + \text{bone dry mass} + \text{total body water} \quad \text{(4C)}$$

For deriving estimates of fat mass, all other compartments of the respective model, including body mass, have to be estimated. In the IDEFICS validation study, we measured total body water using DLW, fat-free dry mass using air displacement plethysmography, bone dry mass using dual-energy X-ray absorptiometry (DXA) and body mass using a TANITA BC 420 SMA digital weighing scale (TANITA, Tokyo, Japan).

The measurement of total body water by the DLW method is quite straightforward. Isotope intake and the concentration of isotopes in the urine samples are used to estimate the volume of the pool in which the isotope equilibrates; that is, body water.[14]

Body volume is measured by air displacement plethysmography using a BOD POD device (Life Measurement, Inc., Concord, CA, USA) and corrected for measured thoracic gas volume and estimated surface area artefact. Whole-body density can then be calculated as body mass divided by body volume. Fat-free dry mass is estimated assuming a fixed density and hydration for fat mass and age- and sex-specific densities for fat-free mass (cf. for example, Wells *et al*.[20]).

Bone mineral content is measured by DXA. DXA uses two X-ray beams to distinguish between fat and lean tissues on the one hand and bone and soft tissues on the other, on the basis of the extent to which the pairs of tissues attenuate the two X-rays to different degrees. A whole-body scan is taken and fat mass, fat-free mass and bone mass can be calculated from the measurement results using a computer algorithm provided by the manufacturer. As it was not possible to acquire new DXA devices for the study, DXA devices that were available at the study centres had to be used: universities of Ghent, Belgium and Zaragoza, Spain used Hologic QDR 4500 devices (Hologic, Inc., Bedford, MA, USA), whereas University of Gothenburg, Sweden and University of Glasgow, United Kingdom used Lunar Prodigy (GE Healthcare UK Ltd, Chalford, UK). Although results are known to vary by manufacturer, this especially applies to the differentiation of soft tissue into fat and fat-free mass.[21] For the 4C model, only bone mass is used, which is measured much more accurately by DXA than soft tissue mass. To evaluate the bias induced by using different DXA devices, fat mass derived by the 4C model was compared with the 3C model that does not use the DXA measurements.

Field methods for assessing body composition. The field methods for assessing body composition comprise anthropometric and bioelectrical impedance measurements. Skinfold thickness was measured after previous landmarking using Holtain Tanner/Whitehouse skinfold calipers (Holtain Ltd., Crosswell, UK) from six sites (triceps, biceps, subscapular, suprailliac, thigh, calf) according to the International standards for anthropometric assessment.[22] Fat mass was estimated using single skinfolds as well as different generalised and population-specific empirical equations involving two to six sites (for an overview see Gibson[18]). In addition, limb girths were measured after previous landmarking in four sites (waist, hip, neck, mid-upper arm) using a Seca 200 tape (Seca GmbH & KG, Hamburg, Germany) and standing height using a Seca 225 stadiometer (Seca GmbH & KG) according to International standards for anthropometric assessment.

Bioelectrical impedance and body mass were assessed using a prototype leg-to-leg device that is based on the TANITA BC 420 SMA digital scale (TANITA). The prototype was developed by TANITA Europe (TANITA Europe GmbH, Sindelfingen, Germany) specifically for the IDEFICS surveys to assess leg-to-leg bioelectrical impedance in children whose feet are too small for the currently produced devices.

Statistical procedures

The comparison between fat mass estimated by the 4C and 3C models was carried out using the method of Bland and Altman.[23] In addition, linear regression was performed with the difference of both models as the dependent variable and the average of both methods as the independent variable; variance ratio between the 4C and 3C models was calculated for each of the devices.

AEE/day was estimated by subtracting REE/day and diet-induced energy expenditure/day from TEE/day. As described above, TEE/day was estimated using the DLW procedure, REE/day was estimated using Schofield's equation[15] and diet-induced energy expenditure/day was set at 10% of TEE/day.

$$\text{Time spent outdoors/day was calculated by } 5/7 \times (\text{minutes spent outdoors on weekdays}) + 2/7 \times (\text{minutes spent outdoors on weekend days}).$$

AEE/day per kg was estimated by dividing AEE/day by body mass in kilograms measured on day 1 on the TANITA scale.

Body mass index was calculated by dividing body mass in kilograms measured on day 1 on the TANITA scale by squared body height in metres. Body mass index categories were interpolated for continuous age as proposed by Cole *et al*.[24,25] For this interpolation, cubic splines were used. All statistical analyses were performed with SAS 9.2 (SAS Institute, Cary, NC, USA).

Results

From the 98 children initially enrolled in the study, six withdrew after drinking the initial dose of DLW (two from university of Ghent, Belgium, one from University of Glasgow, United Kingdom and three from University of Zaragoza, Spain). Because of the high cost of the DLW it was not feasible to replace these children, leaving 92 children in the validation study. Comparing the included children with the sampling scheme of the study protocol it can be observed that for older girls (ages 6–8) the target number was not reached (see Table 2). This was partly compensated for by including four 9- to 10-year-old girls, an age group not foreseen in the study protocol.

Table 3 shows the numbers of completed measurements and the data available for the planned comparison of the field methods with the respective reference methods by centre, sex and age group. For all 92 children, an estimate for AEE by the chosen reference method is available; therefore, all 83 children (90.2%) with 1D accelerometer measurements, 89 children (96.7%) with 3D accelerometer measurements and 88 children (95.7%) with a completed outdoor playtime checklist can be included in the comparison. Sample size is sufficient to perform analyses stratified by sex, as well as for the Spanish sample and the older age group. The younger age group does not have sufficient 1D accelerometer measurements; comparisons in this group are restricted to 3D accelerometer measurements and to the outdoor playtime checklist.

It was not possible to perform the BOD POD measurements in the full sample of children in all centres; hence, fat mass estimates by reference method are only available in 87.0% ($N=80$; 3C model) and 86.0% ($N=79$; 4C model) of the children. Field methods were applied without problems in most centres in the full sample. The only exception is Glasgow, where measurements of more than two skinfolds are available in only some children. Because of the reduced number of valid observations for the reference method, comparisons of field methods with reference models in single strata are possible only in the older age group and in girls. For boys, comparisons are sufficiently statistically powered for all field methods except for the six-skinfold model.

Table 2 Comparison of reached numbers of study subjects with foreseen sampling scheme of the study protocol

Age (in years)	*Foreseen in study protocol*		*Included in validation study*[a]	
	Boys	*Girls*	*Boys*	*Girls*
4	~31/4	~31/4	4	10
5	~31/4	~31/4	8	11
4 to <6	31		33	
	Boys	*Girls*	*Boys*	*Girls*
6	~31/3	~31/3	5	13
7	~31/3	~31/3	12	6
8	~31/3	~31/3	11	5
9	0	0	2	3
10	0	0	1	1
⩾6	31	31	31	28

[a]Only participants with complete doubly labelled water/urine collection are considered.

Overall, the Bland and Altman plot of the 4C versus 3C models shows a good agreement between both models, with a mean difference of −0.17 kg and limits of agreement of 0.39 kg and −0.74 kg (see Figure 2). However, the two devices yield quite distinct patterns. The regression slope for the Lunar Prodigy device was almost flat, with $b=-0.0003$ ($P=0.9669$), and contrasted with that of the Hologic QDR 4500 ($b=0.0344$; $P=0.0003$), suggesting that here the difference of both models increased with increasing fat mass. The mean difference between the 4C and 3C models was −0.39 kg (s.d. = 0.26) for the Lunar Prodigy and −0.06 kg (s.d. = 0.23 kg) for the Hologic QDR 4500. Variance between fat mass derived by the 4C and 3C models did not differ substantially in both devices (variance ratio of 4C/3C: 0.999 for Lunar Prodigy; 1.071 for Hologic QDR 4500).

The basic characteristics of physical activity level and body composition of the included children are given in Table 4. The amount of AEE varied between centres, between sexes (boys: 1.79 MJ/day; girls: 1.46 MJ/day) and between age groups

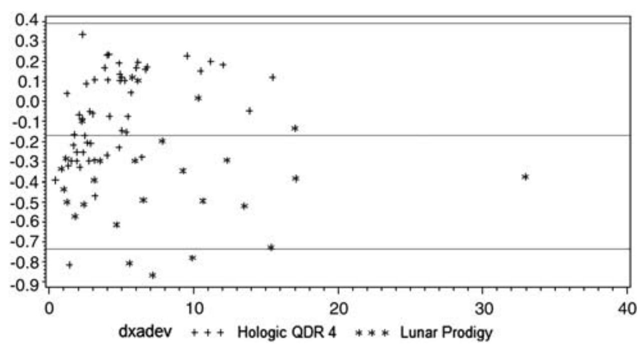

Figure 2 Bland–Altman plot of the agreement between the 3C and 4C models for estimating fat mass in 79 children.

Table 3 Available data for statistical testing of investigated field methods[a]

Method	*Study centre*				*Sex*		*Age group*		*All*
	UGENT	*UGLW*	*UGOT*	*UZAZ*	*Boys*	*Girls*	*4–5 years*	*6+ years*	
	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Physical activity									
Reference method: Doubly labelled water	**32 (100)**	19 (100)	10 (100)	**31 (100)**	**43 (100)**	**49 (100)**	**33 (100)**	**59 (100)**	**92 (100)**
1D accelerometer	29 (90.6)	16 (84.2)	7 (70.0)	**31 (100)**	**39 (90.7)**	**44 (89.8)**	30 (90.9)	**53 (89.8)**	**83 (90.2)**
3D accelerometer	**31 (96.9)**	17 (89.5)	10 (100)	**31 (100)**	**40 (93.0)**	**49 (100)**	**31 (96.8)**	**58 (98.3)**	**89 (96.7)**
Outdoor playtime checklist	29 (90.6)	19 (100)	10 (100)	**30 (96.8)**	**42 (97.7)**	**46 (93.9)**	**31 (93.9)**	**57 (96.6)**	**88 (95.7)**
Body composition[b]									
Reference method A: 3-compartment model	23 (71.9)	18 (94.7)	10 (100)	29 (93.5)	**37 (86.0)**	**43 (87.8)**	24 (72.7)	**56 (94.9)**	**80 (87.0)**
Reference method B: 4-compartment model	23 (71.9)	18 (94.7)	10 (100)	28 (90.3)	**36 (83.7)**	**43 (87.8)**	23 (69.7)	**56 (94.9)**	**79 (86.0)**
Skinfolds: two sites	23 (71.9)	18 (94.7)	10 (100)	29 (93.5)	**37 (86.0)**	**43 (87.8)**	24 (72.7)	**56 (94.9)**	**80 (87.0)**
Skinfolds: four sites	23 (71.9)	6 (31.6)	10 (100)	29 (93.5)	**31 (72.1)**	**37 (75.5)**	24 (72.7)	**44 (74.6)**	**68 (73.9)**
Skinfolds: six sites	23 (71.9)	2 (10.5)	10 (100)	29 (93.5)	28 (65.1)	**36 (73.5)**	23 (69.7)	**41 (69.5)**	**64 (69.6)**
Circumferences	23 (71.9)	18 (94.7)	10 (100)	28 (90.3)	**36 (83.7)**	**43 (87.8)**	23 (69.7)	**56 (94.9)**	**79 (86.0)**
Prototype TANITA BC 420 scale	23 (71.9)	18 (94.7)	10 (100)	29 (93.5)	**37 (86.0)**	**43 (87.8)**	24 (72.7)	**56 (94.9)**	**80 (87.0)**

Abbreviations: UGENT, University of Ghent; UGLW, University of Glasgow; UGOT, University of Gothenburg; UZAZ, University of Zaragoza. [a]Bold numbers indicate that sample sizes are sufficient for statistical testing within the respective stratum, that is, $N \geqslant 31$. [b]Numbers for field methods are given for complete cases, for which a valid measurement by reference method is available, only.

Table 4 Basic descriptions of included children

Level of physical activity	*Study centre*				*Sex*		*Age group (in years)*		***All***
	UGENT	*UGLW*	*UGOT*	*UZAZ*	*Boys*	*Girls*	*4–5*	*6+*	
	Mean (s.d.)	*Mean (s.d.)*	*Mean (s.d.)*	*Mean (s.d.)*	*Mean (s.d.)*	*Mean (s.d.)*	*Mean (s.d.)*	*Mean (s.d.)*	*Mean (s.d.)*
AEE[a] (MJ per day)	1.57 (0.56)	1.87 (0.70)	1.71 (0.71)	1.46 (0.63)	1.79 (0.64)	1.46 (0.62)	1.35 (0.63)	1.77 (0.61)	1.62 (0.64)
Time spent outdoors (min per day)	93 (46)	141 (77)	282 (118)	148 (79)	155 (99)	133 (85)	140 (88)	145 (95)	143 (92)
AEE[b] (MJ/(day × kg))	0.071 (0.019)	0.069 (0.025)	0.040 (0.015)	0.061 (0.028)	0.068 (0.026)	0.061 (0.023)	0.063 (0.028)	0.065 (0.023)	0.064 (0.025)
BMI category	N *(%)*	N *(%)*	N *(%)*	N *(%)*	N *(%)*	N *(%)*	N *(%)*	N *(%)*	N *(%)*
Thinness grade II	2 (6.3)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	2 (4.1)	1 (3.0)	1 (1.7)	2 (2.2)
Thinness grade I	2 (6.3)	0(0.0)	0(0.0)	2 (6.5)	3 (7.0)	1 (2.0)	1 (3.0)	3 (5.1)	4 (4.4)
Normal weight	23 (71.9)	13 (68.4)	0(0.0)	22 (71.0)	30 (69.8)	28 (57.1)	22 (66.7)	36 (61.0)	58 (63.0)
Overweight	5 (15.6)	6 (31.6)	0(0.0)	4 (12.9)	3 (7.0)	12 (24.5)	5 (15.2)	10 (17.0)	15 (16.3)
Obesity	0 (0.0)	0(0.0)	10 (100)	3 (9.7)	7 (16.3)	6 (12.2)	4 (12.1)	9 (15.3)	13 (14.1)

Abbreviations: AEE, activity energy expenditure; BMI, body mass index; UGENT, University of Ghent; UGLW, University of Glasgow; UGOT, University of Gothenburg; UZAZ, University of Zaragoza. [a]Activity energy expenditure calculated by doubly labelled water method given in mega joule (MJ) per day. [b]Activity energy expenditure calculated by doubly labelled water method given in mega joule (MJ) per day per kg body weight.

(4–5 years: 1.35 MJ/day; 6+ years: 1.77 MJ/days), with boys and the older age group having higher estimates of AEE. Similarly, the minutes spent outdoors varied between the different groups and the pattern did not coincide with that for mean AEE. For example, the Swedish children (University of Gothenburg, Sweden), who were recruited from a clinical setting and were receiving obesity therapy, spent on average 282 min per day outside, nearly twice as much as the average of the total group (143 min per day) without having a particularly high AEE mean (1.71 MJ/day as compared with 1.62 MJ/day in all children). After normalisation of AEE by body weight as proposed by Ekelund *et al.*,[26] most group-specific differences of AEE diminished, except for centre-specific differences. It was especially observed that Swedish children had a markedly lower AEE per kg compared with the other centres.

The distribution of body mass index categories of the children from the three non-clinical settings (thinness grades I/II: 7.3%; normal weight: 70.7%; overweight/obesity: 22.0%) is similar to the distribution in the general population. However, the inclusion of the 28 overweight and obese children in the Swedish sample ensured that these categories were well represented in the validation study (overweight: 16.3%; obesity: 14.1%).

Discussion

The IDEFICS validation study was planned and conducted to compare the field methods used to assess physical activity and body composition in the IDEFICS survey with established reference methods and to derive sex- and age-sensitive analysis strategies for the IDEFICS survey data. For physical activity, this will be possible, as data from 83 to 89 children are available for this comparison, depending on the considered field method. For body composition, age-specific analyses can only be carried out in the oldest age group; sex-specific analyses are possible for all field methods except for the six-site skinfold model.

Urine samples were collected as foreseen from 93.9% of the 98 children who were initially included in the validation study, resulting in 92 children for whom TEE and total body water were assessed by the DLW method. The most limiting factor for assessing fat mass by the reference method was the BOD POD measurement. Nine observations were lost in Ghent, where the BOD POD device was located in another city and parents and children had to travel to this destination. In Glasgow, one child refused the measurement. In Zaragoza, two 4-year-old children could not be measured, because it was too difficult to keep them quiet during the measurement procedure. Summarising, a valid Bod Pod measurement was only taken in 80 of 92 children (87.0%). From these 80 children, a DXA measurement was possible in all but one 4-year-old boy, who would not rest long enough to be measured. Not surprisingly, assessment of reference methods for body composition turned out to be more difficult in the youngest age group (4 and 5 years) compared with older children.

The investigated field methods were largely assessed without any problems for both the measurements taken in the study centres and for the accelerometer measurements that required collaboration and compliance of the parents. The accelerometer criterion was chosen using the IDEFICS main survey criterion (3 days/6 h); however, for most children in the validation study, more data are available. For 86 children (89.6%), accelerometer measurements of 6 or more days are available; the mean wear time is 10.9 h/day. It could be speculated that this high compliance is partly due to the highly motivated group of parents of the children in the validation study, which might not be expected in larger studies using random samples. In a Canadian random sample, wear times of children did not differ from that of adults, and in the youngest age group (6–11 years) 86% of boys and 88% of girls showed ⩾4 days out of 7, with a minimum of 10 h wear time, which is slightly above the overall mean of the sample (84%).[27] Furthermore, in this study, we did not observe differences in compliance of accelerometer measurements in the two different age groups.

Because of the study size, the IDEFICS validation study can give insight into the sex- and age-specific validity of physical activity and body composition assessment in small children. To date, paediatric studies in which body composition methods have been validated against 3C and 4C models have been scarce and usually very small, with typically less than 20 subjects per age and sex group.[28] Similarly, accelerometers have not been validated against DLW to a satisfying extent,[29] particularly not in preschool children.[30] Validation studies on physical activity are often not conducted after stratifying by sex, which may be inappropriate.[31] Although in this study with 98 children a slightly larger sample (107%) than needed based on sample size calculations was initially approached, this was not enough to compensate for all non-compliant study participants: for example, 115% would have been necessary for the 4C model and 111% for the 1-D accelerometer measurements. Therefore, similar losses should be considered when planning such a study.

With the inclusion of the Swedish group, artificial oversampling was carried out for the group of obese children. As a result, at least exploratory analyses will be possible for overweight and obese children. In theory, this approach might have been possible for highly physically active children as well. However, it might be difficult to define and identify such highly active groups in this particular age group. Moreover, the usefulness of this information might be quite limited for the general population.

The fieldwork of the validation study was conducted in four centres from different regions in Europe. This clearly introduces heterogeneity in data, both from potential measurement error and potential country-by-country differences. One important difference between centres was the use of different DXA devices. A systematic difference between both devices has been observed and has to be accounted for in future analyses.[32,33] Similarly, centre-specific measurement errors will have to be controlled for by appropriate statistical methods.[34,35] Country-by-country differences will have to be analysed carefully. They could already be observed for the association between AEE and minutes spent outdoors, in which it could be speculated that cultural aspects and/or different weather conditions between centres might be responsible for the observed discrepancies between centres. The multicentred approach might be seen as a drawback of the study, but on the other hand it provides valuable information on whether the use of unified analysis approaches for data on physical activity and body composition are justified not only between age groups and sexes but also in different European countries. However, it is of utmost importance to use a common study protocol, the same devices wherever possible and same measurement procedures and to implement quality control measures when conducting a multicentre validation study.

Conflict of interest

MR has declared equity ownership/stock options with BioTel Ltd and Optimal Performance Ltd. The remaining authors declare no conflict of interest.

Acknowledgements

We thank all members of the study teams and especially the children and their parents for their enthusiastic participation in the study. This study was conducted as part of the IDEFICS study (http://www.idefics.eu). We gratefully acknowledge the financial support of the European Community within the Sixth RTD Framework Programme under Contract No. 016181 (FOOD).

The information in this document reflects the author's view and is provided as it is.

Statement of ethics

We certify that all applicable institutional and governmental regulations pertaining to the ethical use of human volunteers were followed during this research. Approval by the appropriate ethics committees was obtained by each of the eight centres carrying out the fieldwork. Study children did not undergo any procedure before both they and their parents gave consent for examinations, collection of samples, subsequent analysis and storage of personal data and collected samples. Participating children and their parents could consent to single components of the study while abstaining from others.

References

1 Bammann K, Peplies J, Sjöström M, Lissner L, De Henauw S, Galli C, on behalf of the IDEFICS Consortium *et al*. Assessment of diet, physical activity, biological, social and environmental factors in a multi-centre European project on diet- and lifestyle-related disorders in children (IDEFICS). *J Public Health* 2006; **14**: 279–289.
2 Ahrens W, Bammann K, Siani A, Buchecker K, De Henauw S, Iacoviello L *et al*. The IDEFICS cohort: design, characteristics and participation in the baseline survey. *Int J Obes* 2011; **35** (Suppl 1): S3–S15.
3 Goran MI. Measurement issues related to studies of childhood obesity: assessment of body composition, body fat distribution, physical activity, and food intake. *Pediatrics* 1998; **101**: 505–518.
4 Reilly JJ, Penpraze V, Hislop J, Davies G, Grant S, Paton JY. Objective measurement of physical activity and sedentary behaviour: review with new data. *Arch Dis Child* 2008; **93**: 614–619.
5 Butte NF, Wong WW, Adolph AL, Puyau MR, Vohra FA, Zakeri IF. Validation of cross-sectional time series and multivariate adaptive regression splines models for the prediction of energy expenditure in children and adolescents using doubly labeled water. *J Nutr* 2010; **140**: 1516–1523.
6 Eisenmann JC, Strath SJ, Shadrick D, Rigsby P, Hirsch N, Jacobson L. Validity of uniaxial accelerometry during activities of daily living in children. *Eur J Appl Physiol* 2004; **91**: 259–263.

7 Ekelund U, Sjöström M, Yngve A, Poortvliet E, Nilsson A, Froberg K *et al.* Physical activity assessed by activity monitor and doubly labeled water in children. *Med Sci Sports Exerc* 2001; **33**: 275–281.
8 Krishnaveni GV, Veena SR, Kuriyan R, Kishore RP, Wills AK, Nalinakshi M *et al.* Relationship between physical activity measured using accelerometers and energy expenditure measured using doubly labelled water in Indian children. *Eur J Clin Nutr* 2009; **63**: 1313–1319.
9 Reilly JJ, Kelly LA, Montgomery C, Jackson DM, Slater C, Grant S *et al.* Validation of actigraph accelerometer estimates of total energy expenditure in young children. *Int J Pediatr Obes* 2006; **1**: 161–167.
10 Bray GA, DeLany JP, Harsha DW, Volaufova J, Champagne CC. Evaluation of body fat in fatter and leaner 10-y-old African American and White children: the Baton Rouge Children's Study. *Am J Clin Nutr* 2001; **73**: 687–702.
11 Eisenmann JC, Heelan KA, Welk GJ. Assessing body composition among 3- to 8-year-old children: anthropometry, BIA, and DXA. *Obes Res* 2004; **12**: 1633–1640.
12 Elberg J, McDuffie JR, Sebring NG, Salaita C, Keil M, Robotham D *et al.* Comparison of methods to assess change in children's body composition. *Am J Clin Nutr* 2004; **80**: 64–69.
13 Parker L, Reilly JJ, Slater C, Wells JC, Pitsiladis Y. Validity of six field and laboratory methods for measurement of body composition in boys. *Obes Res* 2003; **11**: 852–858.
14 Prentice AM (ed). *The doubly-labelled water method for measuring energy expenditure: A consensus report by the IDECG Working Group, NAHRES-4*. International Atomic Energy Agency: Vienna, 1990.
15 Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. *Hum Nutr Clin Nutr Suppl* 1985; **39/1**: 5–41.
16 Westerterp KR. Diet induced thermogenesis. *Nutr Metab* 2004; **1**: 5.
17 Burdette HL, Whitaker RC, Daniels SR. Parental report of outdoor playtime as a measure of physical activity in preschool-aged children. *Arch Pediatr Adolesc Med* 2004; **158**: 353–357.
18 Gibson RS. *Principles of Nutritional Assessment*. 2nd edn. Oxford University Press: Oxford, New York, 2005.
19 Wells JC, Fuller NJ, Dewit O, Fewtrell MS, Elia M, Cole TJ. Four-component model of body composition in children: density and hydration of fat-free mass and comparison with simpler models. *Am J Clin Nutr* 1999; **69**: 904–912.
20 Wells JC, Williams JE, Chomtho S, Darch T, Grijalva-Eternod C, Kennedy K *et al.* Pediatric reference data for lean tissue properties: density and hydration from age 5 to 20 y. *Am J Clin Nutr* 2010; **91**: 610–618.
21 Tothill P, Hannan WJ, Wilkinson S. Comparisons between a pencil beam and two fan beam dual energy X-ray absorptiometers used for measuring total body bone and soft tissue. *Brit J Radiol* 2001; **74**: 166–167.
22 Marfell-Jones M, Olds T, Stewart A, Carter L. *International Standards for Anthropometric Assessment*. ISAK: Potchefstroom, 2006.
23 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **i**: 307–310.
24 Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000; **320**: 1–6.
25 Cole TJ, Flegal KM, Nicholls D, Jackson AA. Body mass index cut offs to define thinness in children and adolescents: international survey. *BMJ* 2007; **335**: 18.
26 Ekelund U, Yngve A, Brage S, Westerterp K, Sjöström M. Body movement and physical activity related energy expenditure in children and adolescents: How to adjust for differences in body size and age? *Am J Clin Nutr* 2004; **79**: 851–856.
27 Colley R, Gorber SC, Tremblay MS. Quality control and data reduction procedures for accelerometry-derived measures of physical activity. *Health Rep* 2010; **21**: 63–69.
28 Reilly JJ, Gerasimidis K, Paparacleous N, Sherriff A, Carmichael A, Ness AR *et al.* Validation of dual-energy x-ray absorptiometry and foot-foot impedance against deuterium dilution measures of fatness in children. *Int J Ped Obes* 2010; **5**: 111–115.
29 Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. *Obes* 2007; **15**: 2371–2379.
30 Cliff DP, Reilly JJ, Okely AD. Methodological considerations in using accelerometers to assess habitual physical activity in children aged 0-5 years. *J Sci Med Sport* 2009; **12**: 557–567.
31 Rennie KL, Wareham NJ. The validation of physical activity instruments for measuring energy expenditure: problems and pitfalls. *Public Health Nutr* 1998; **1**: 265–271.
32 Hui SL, Gao S, Zhou XH, Johnston Jr CC, Lu Y, Gluer CC *et al.* Universal standardization of bone density measurements: a method with optimal properties for calibration among several instruments. *J Bone Miner Res* 1997; **12**: 1463–1470.
33 Fan B, Lu Y, Genant H, Fuerst T, Shepherd J. Does standardized BMD still remove differences between Hologic and GE-Lunar state-of-the-art DXA systems? *Osteoporos Int* 2010; **21**: 1227–1236.
34 Guolo A. Robust techniques for measurement error correction: a review. *Stat Method Med Res* 2008; **17**: 555.
35 Stukal TA, Demidenko E, Dykes J, Karagas MR. Two-stage methods for the analysis of pooled data. *Stat Med* 2001; **20**: 2115–2130.

Appendix: Outdoor playtime checklist (Burdette *et al.*, 2004)

How much time does your child usually spend per day playing in the yard or street around your house (or the house of a friend, neighbour or relative)?
Please indicate for every time frame.

	0 minutes	*1-15 minutes*	*16-30 minutes*	*31-60 minutes*	*Over 60 minutes*
Wake-up time until noon	❍$_1$	❍$_2$	❍$_3$	❍$_4$	❍$_5$
Noon until 6 PM	❍$_1$	❍$_2$	❍$_3$	❍$_4$	❍$_5$
6 PM until bedtime	❍$_1$	❍$_2$	❍$_3$	❍$_4$	❍$_5$

How much time does your child usually spend per day at a park, playground or outdoor recreation area (for example, swimming pool, zoo or amusement park)?
Please indicate for every time frame. Include times that the child is at daycare, kindergarten, preschool or school.

	0 minutes	***1-15 minutes***	***16-30 minutes***	***31-60 minutes***	***Over 60 minutes***
Wake-up time until noon	❍1	❍2	❍3	❍4	❍5
Noon until 6 PM	❍1	❍2	❍3	❍4	❍5
6 PM until bedtime	❍1	❍2	❍3	❍4	❍5

Think for a moment about a typical weekday for your child in the last month. How much time would you say your child spends playing outdoors on a typical weekday?

|___|___| hours |___|___| minutes

Now think about a typical weekend day for your child in the last month. How much time would you say your child spends playing outdoors on a typical weekend day?

|___|___| hours |___|___| minutes