National Diet and Nutrition Survey Rolling Programme

Comparison Study

Measuring physical activity and energy expenditure

January 2009

Dr Jennifer Mindell[1], Dr Emannuel Stamatakis[1], Dr Soren Brage[2]

[1] Department of Epidemiology & Public Health, UCL

[2] MRC Epidemiology Unit, Cambridge

Contents

Acknowledgments

We are grateful to the members of the Physical Activity Working Group, and in particular to Professor Nick Wareham, Dr Ulf Ekelund, Dr Herve Besson, and Dr Kirsten Corder of the MRC Epidemiology Unit, Cambridge, for their help, advice and support with this work.

Executive summary

Background

The principal purpose of estimating energy expenditure in a dietary survey is so that energy intake can be analysed together with energy expenditure, to assess the degree of under-reporting of food and drink intake. Physical activity is measured in the National Diet and Nutrition Survey (NDNS) primarily to estimate energy expenditure and not to measure the population's activity levels per se. Resting metabolism and physical activity account for 90% of energy use on average. Despite their many limitations, questionnaires have been used widely in the past to assess energy expenditure as this is an inexpensive and convenient method. Newer objective measures, by monitoring heart rate or movement, promise greater accuracy but at greater cost.

Method

The NDNS Comparison Study carried out in 2007 (see Comparison Study for the NDNS RP) included a doubly-labelled water (DLW) sub-study of a quota sample of 160 participants (eight per age-group/sex/dietary method cell) to assess energy expenditure and measurement methods. The DLW method involves drinking a known, weighed dose of water labelled containing two non-radioactive and natural 'stable' isotopes (oxygen-18 and deuterium). Data are obtained from one pre-dose urine sample and ten spot urine samples from ten consecutive days. This is considered the gold standard for measurement of energy expenditure during free-living.

Questionnaires were developed to assess physical activity in sufficient detail to enable estimation of energy expenditure. These were based on existing questionnaires where suitable, with new questions developed where required to provide additional detail which was considered missing. Separate questionnaires were prepared for children aged 11-15yrs and participants aged 16+. No questionnaire can record the repeated, intermittent, short bursts of activity of children under 11 accurately enough to estimate energy expenditure.

An Actigraph (a uniaxial accelerometer) was worn around the waist over one week by each participant in the DLW sub-study and by all children aged 4-10yrs. As it is not feasible for children under four to wear Actigraphs and no questionnaire was used for them either, energy expenditure is not being estimated in the youngest children. Actigraphs do not record all activities equally well, e.g. cycling, and are removed for some sports (e.g. swimming), so participants or their parents were asked to keep a logbook as well. Actigraphs were also worn by all children aged 11-15yrs to provide sufficient numbers to validate the questionnaire. Actigraph data were flagged for non-wear time by the zero-string-length method and translated to physical activity energy expenditure (PAEE) by four previously published equations, using two different approaches for dealing with missing data (non-wear time).

Findings

Actigraph data underestimate mean energy expenditure levels compared with DLW data but give results within acceptable limits, with acceptable variance across individuals enabling categorisation.

The new questionnaire for children aged 11- 15yrs underestimates the mean energy expenditure less than the Actigraphs but with extremely wide variance (twice that of Actigraph estimates), such that data are not useful except at the population level.

The new questionnaire for those aged 16+ underestimated mean energy expenditure with fairly wide variance. It would be possible to use the results to divide participants into categories of physical activity. However, the questionnaire was very long (15-20minutes) and repetitive. The recently validated but much shorter self-completion RPAQ (Recent Physical Activity Questionnaire) performed in that validation study at least as well against DLW measurements as the NDNS questionnaire did in this study.

Decisions for measuring physical activity in NDNS Years 2-4 (April 2009 – Mar 2012)

- No physical activity measurement aged 18mths to three years (as currently).
- Actigraph for all children aged 4-15 years, but a simple record sheet to be used without the detailed log book. There will be no questionnaire for children, except to ask about time spent sleeping.
- RPAQ to be administered at the diary pick-up visit for all participants aged 16+. Existing physical activity questions will be omitted, except for time spent sleeping.

1 Background

1.1 Need for measurement of energy expenditure

Data on energy expenditure are collected for a number of purposes, which determine the level of accuracy required.

1.1.1 Assessment of energy expenditure at the population level

This may be desired to assess broad differences in energy expenditure by age-group, sex, or other socio-demographic categories, or to measure changes over time, for example, the extent to which rising obesity can be attributed to increases in energy intake or to decreases in energy expenditure. For these purposes, the important features of the method chosen are comparability over time and no systematic differences in the responses or data produced by different sub-groups of the population, apart from those due to differing energy expenditure levels. Where comparisons between different groups or over time are important, relative values can be as useful as absolute values. Categories can be used, to assess differences in the proportions in each category over time or by sub-group, although that will not describe fully changes over time. (This is analogous to predicting the future prevalence of diabetes based on the rising percentage of the population who are obese without also factoring in the rising mean BMI among those who are obese).

1.1.2 Assessment of energy expenditure at the individual level

The principal purpose of estimating energy expenditure in a dietary survey is so that energy intake can be analysed together with energy expenditure, to assess the degree of under-reporting of food and drink intake (or, less commonly, over-reporting). Ideally, this requires accurate absolute values at the individual level, so that 'energy in / energy out' assessments can be made for each individual participant. Where energy expenditure cannot be measured accurately enough, it is possible to assess under-reporting by grouping participants into a few categories of energy expenditure and comparing reported intake between these groups.

1.2 Methods for measuring energy expenditure

1.2.1 Doubly-labelled water (DLW)

The DLW method involves drinking a known, weighed dose of water containing two labelled, non-radioactive and natural 'stable' isotopes – doubly-labelled water. Data are obtained from one pre-dose urine sample and ten spot urine samples from ten consecutive days. This is considered the gold standard for accurate measurement of energy expenditure. Assuming no laboratory errors in the preparation of the dose or measurement of the isotopes in the urine samples, inaccuracies can arise if: the dose (calculated on a per kg body weight basis) is not completely consumed; there is contamination from the dose bottle of a urine sample; or samples are wrongly labelled (wrong person, day, or time).

Additionally, if the DLW collection period does not overlap fully with the time period in which other measures are taken, results may be accurate but misleading in those whose energy expenditure varies considerably from day to day.[1] For example, the food diary is kept for four days, the physical activity questionnaire refers to the previous seven days, the Actigraph is worn for seven days, and the urine samples for DLW measurement are collected over 10 days. In most cases, the differences in these time periods will not affect the results but in a few, more extreme cases, they may do so.

1.2.2 Rationale for measuring physical activity

Physical activity is measured in NDNS primarily not to measure the population's activity levels per se but to estimate energy expenditure, as physical activity together with the resting metabolism accounts for 90% of energy use on average.[2] The remaining 10% is used for diet-induced thermogenesis.[2]

A number of direct and indirect methods of measuring physical activity in the general population have been used in different settings. Direct methods are:

- Self (or interviewer) administered questionnaires;
- Self (or observer) recorded activity diary or log;
- Self conducted (or remotely recorded) mechanical or electronic monitoring.

Each of these methods could be used to monitor levels of physical activity within the general population. Both the diary method and objective monitoring are costly to administer and place a burden upon the participant. Observer-recorded diaries or logs are seldom feasible in a free-living population as they require the participant to be accompanied during all waking hours by an observer, while self-recorded diaries and logs pose a substantial burden on participants. Practical objective measures that could be used in a large, free-living population have not been available until recently. Questionnaires, by contrast, are quick and relatively easy to administer but are less accurate.

1.2.3 Subjective measures of physical activity

Activity diaries have been used in previous NDNS. They appear to be a valid method for estimating physical activity level (PAL) and total energy expenditure (TEE) when compared with results from DLW measurements, for example in adolescents.[3] However, they add considerably to the respondent burden compared with answering a single, interviewer-administered questionnaire. As NDNS is already a considerable load for participants, which can cause both issues of research ethics and depressed response rates, it was decided to use questionnaires rather than activity diaries in the rolling programme NDNS to measure physical activity.

Questionnaires are relatively quick and relatively easy to administer. However, self-reported physical activity can be over-estimated.[4, 5] There are a number of areas where error may be introduced.

- Participants may experience difficulty in recalling all activities (this can lead to either under-reporting, if they omit activities, or over reporting, if they “telescope” and include activities beyond the recall period).
- Individuals’ assessment of the duration and intensity of physical activity may be inaccurate.
- Social desirability biases – participants may overestimate their levels of activity to provide socially desirable answers.

Self-reported questionnaires have been found to be of limited usefulness for measuring energy expenditure to enable ‘energy in/energy out’ comparisons at an individual level. Assessment of energy expenditure requires more complete recording of physical activity than assessment of physical activity in relation to health. If an individual undertakes 30 minutes of vigorous activity, six other periods of walking briskly for five minutes at a time or a slow three-mile stroll that takes an hour and a half affects their energy expenditure much more than it affects their fitness or whether they meet the recommendations.

However, despite their limitations, questionnaires have been used widely in the past to assess energy expenditure as this is an inexpensive and convenient method.[6] Studies using questionnaires to assess energy expenditure date back to the early 1970s[7] and there are many physical activity questionnaires specifically designed to assess energy expenditure, e.g. the Questionnaire d'Activité Physique Saint-Etienne,[8] Tecumseh Community Health Study Questionnaire,[9] Tecumseh Occupational Activity and past month Minnesota Leisure Time Questionnaire,[10] and the Tecumseh Occupational Activity and past year Minnesota

Leisure Time Questionnaire.[1 11] For all these reasons and because the Tender Specification required measurement of physical activity per se and expected that a questionnaire would be the primary method for this data collection, that was what was proposed.

Energy expenditure (EE) assessment using questionnaires is still a common technique to assess free living energy expenditure. A recent literature review in the American Journal of Clinical Nutrition provides a comprehensive account of the validity of questionnaires to measure EE and the issues associated with its measurements.[11]

1.2.4 Objective measures of physical activity

Monitoring of physical activity can be undertaken by registration of heart rate or movement, or both.[12] Accelerometers, which measure movement and are most commonly worn around the waist, were introduced in the early 1980s.[13] They have been used for a number of purposes including: assessing physical activity for population surveillance (eg in the National Health and Nutrition Examination Survey, NHANES, in the USA and the Health Survey for England 2008); investigating the correlates of physical activity; measuring outcomes of interventions, and validating self-reported physical activity in surveys.[13] Accelerometer data are objective and standardized; direct monitoring reduces recall bias and other problems of subjectivity. However, it is not error-free. Participants' co-operation is required to wear the monitor during waking hours and, in some studies, to record activities when it is not worn, such as during swimming. The Actigraph detects movement in the vertical plane so does not record activities like cycling or rowing.

Accelerometers can measure the frequency, intensity, and duration of some aspects of physical activity. In addition, objective data can have greater precision, so the same statistical power can be obtained with lower sample sizes.

Accelerometers are very good at recording time spent being inactive and time spent on dynamic activity, such as walking or running. They are less good at recording static activity, which could prove to be important when assessing energy expenditure. However, accelerometer counts increase consistently with ambulatory speed, a finding replicated in studies conducted at different times in different place using different methods and on different populations.[13]

The advantages of using an accelerometer to collect activity data (compared with self-reported activity) are those of being objective and standardized measures. Direct monitoring reduces recall bias and other problems of subjectivity. However, the method is not error-free. Participants' co-operation is required to wear the monitor during waking hours and to record activities when it is not worn, for example while swimming. Some accelerometers that detect movement in the vertical plane, such as the Actigraph worn on the hip, do not record many counts during activities like cycling or rowing, despite significantly increase activity energy expenditure. This highlights the main limitation of single-site accelerometry; the great heterogeneity of the count-PAEE relationship in biomechanically diverse activities. Although a further potential imitation can be the lack of contextual information (the type or purpose of the activity), this is not an issue in NDNS, as physical activity is being assessed primarily as a proxy for energy expenditure.

Although accelerometers have been calibrated to assess energy expenditure, this has generally been conducted by walking and running at different speeds on a treadmill but attempts have been made to simulate “daily activities” in the laboratory, during which simultaneous measurement of oxygen consumption and accelerometric intensity have produced very different calibration equations. More recently, ‘free-living’ equations have also been generated. Conversion of counts per minute to energy expended is assumed to be constant for a given level of counts per minute, regardless of the activity undertaken that generated that level of counts.

2 Overview of method

The NDNS Comparison Study fieldwork took place in May – July 2007 (see Comparison Study for the NDNS RP). The aims were to:

- test the repeatability of the new questionnaires;
- validate the estimated energy expenditure calculated from the new questionnaires against DLW,; and
- to validate the new questionnaires' ability to assess activity patterns against the Actigraph monitor.

A 20% bias was expected for self-reported methodologies compared with DLW energy expenditure. A 20% bias was also expected within each age-sex-dietary recording method in assessing energy intake:total energy expenditure using DLW.

An additional aim was to assess the feasibility of including and objective measure of physical activity within the main NDNS.

Data on energy expenditure was collected in the NDNS Comparison Study so that energy intake can be analysed together with energy expenditure. As well as self-assessed data from the questionnaire on physical activity, which was collected from all participants aged 11 and over, the comparison study recorded an objective measure of energy expenditure (EE) using the doubly-labelled water method (DLW) and data from Actigraphs. The same eligibility rules applied to participation in the energy expenditure part of the study as did to the rest of the survey.

A sub-sample of participants who had completed the questionnaire and had height and weight measured were eligible and were asked to participate in this energy expenditure measurement part of the study. The intention was to estimate EE from both DLW and Actigraph data, in a total of 160 people, selected by quota sampling. As shown in the table below, each quota cell had eight people, to cover men and women, five age groups (4-10, 11-15, 16-49, 50-64 and 65+ years), and both dietary methods.

The pilot study would also be used to develop statistical methods that address measurement error in the energy expenditure estimates of the questionnaires and of the Actigraphs. A tool would then be prepared, based on the data from the DLW sub-sample participants, to predict likely energy expenditure in an individual. This tool would then be used to identify inadequate reporters by applying the Goldberg cut-off technique,[14] or similar methods suggested by Huang *et al*[15] and by Rennie *et al*.[16]

Table 1. Composition of quota sample for energy expenditure sub-study

	Dietary data collection method			
Age	**Diary**		**24hr recall**	
	Male	**Female**	**Male**	**Female**
4 to 10	8	8	8	8
11 to 15	8	8	8	8
16 to 49	8	8	8	8
50 to 64	8	8	8	8
65+	8	8	8	8

In addition to the DLW and Actigraph data collection for the quota of 160 people, all those in the main Comparison Study aged 11-15yrs were asked to wear an Actigraph, to validate the revised physical activity questionnaire developed for this age group. The Actigraph procedures were the same as for the quota sample, except that DLW was not administered.

The sub-study was introduced at the final interviewer visit of the main Comparison Study and verbal agreement obtained to participate. Once consent to participate in the study had been obtained, the Actigraph sub-study followed the protocol set out below (Table 2). For children, consent was sought from both the parent (or legal guardian) of the child and the child him/herself.

There were two additional household visits for participants in the energy expenditure sub-study. At the first visit, the interviewer:

- explained the DLW and Actigraph procedures;
- collected one pre-dose urine sample;
- administered the DLW dose;
- initialised the Actigraphs; and
- gave consenting participants an Actigraph.

At the second visit, the interviewer:

- collected the 10 post-dose samples (one collected daily by the participant for a total of 10 days);
- collected the Actigraphs; and
- re-administered the physical activity module of the face-to-face questionnaire.

study.

Table 2. Process for energy expenditure sub-study

Process	Adults (aged 16+) and children aged 11-15	Children aged 4-10
Final NDNS Comparison Study visit	Interviewer administered the NDNS physical activity questionnaire and introduces energy expenditure sub-study	Interviewer introduced the energy expenditure sub-study
First additional visit	Interviewer placed the Actigraph on the participant, collected the pre-dose urine sample, & administered the DLW dose.	
Telephone call 1 (mid week)	Interviewer makes telephone call to participant to check there are no problems and remind participant to continue to wear the Actigraph.	
Second additional visit (10 days later)	After seven full days of accelerometric monitoring, the interviewer visited to collect the Actigraph. Participants were then sent a voucher to thank them for their co-operation so far.	
	The physical activity questions are administered again so as to compare objective and subjective data and also to assess the test/retest reliability of the subjective questionnaire.	
	The daily urine samples were collected.	
	Returned Actigraphs were sent to NatCen's Operations Department for data download and recharging. Actigraphs were then returned to the interviewer for future use.	

3 Doubly labelled water

The doubly labelled water (DLW) method involved a subset of participants in the NDNS Comparison Study drinking a known, weighed dose of water labelled containing two non-radioactive and natural 'stable' isotopes (oxygen-18 and deuterium). The data were obtained from one pre-dose urine sample and ten spot urine samples from ten consecutive days. The method and results have been reported elsewhere.[17]

4 Physical activity questionnaire

4.1 Development of the questionnaires

There are significant problems in trying to capture accurate self-reported information on physical activity. These issues are more acute for children, where parents of children aged 2-12 answer the questions on the child's behalf. This report builds upon a body of work that the Joint Health Surveys Unit have undertaken since Autumn 2006 to improve the capture of physical activity data within national surveys of health and health-related behaviour. This work and the key outputs are summarised below. More detail on the development and cognitive testing of the questionnaires is provided in Appendix A; a report has already been provided to the FSA on the cognitive testing of the questionnaires.[18]

Desk-based review of all physical activity questionnaires used on other UK and international studies was conducted April-May 2006. One conclusion was that physical activity in children under 11 could not be measured in sufficient detail by the use of questionnaires to estimate energy expenditure. It was therefore agreed that objective measures would be used for all children aged 4-10yrs in NDNS, subject to parental consent and the child's assent and to this being shown feasible in the Comparison Study. No further work was done on questionnaire-based assessment of physical activity for children under 11.

Development of a new surveillance questionnaire is time-consuming, requiring both qualitative and quantitative assessment.[19] The former was conducted through cognitive testing and expert panel review for both the adult and child questionnaires. The latter requires testing for validity.

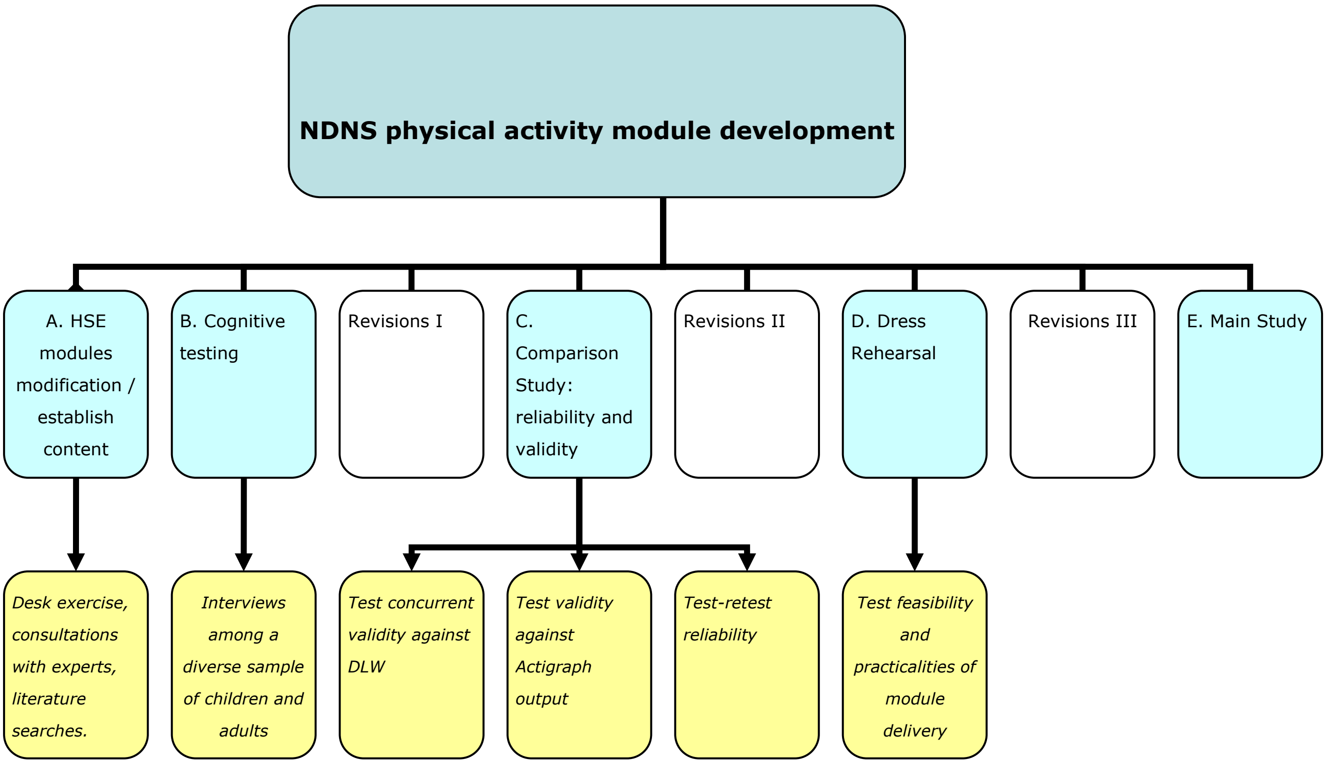
In summary, the aim was to cover all major types of activity that can contribute significantly to energy expenditure: occupational activity, active transport (walking and cycling), sports and leisure, housework/DIY/gardening, and active play for children. Questions in the existing or recently revised HSE questionnaires for adults and children were reviewed for appropriateness.

Existing questionnaires were reviewed and discussions held with experts across England, some of whom also commented on drafts of the questionnaires.

Questions to obtain important details required for accurate estimation of energy expenditure but not covered in the HSE questionnaires were sought in other, validated questionnaires. Where these did not exist, new questions were developed. The results of the cognitive testing of the draft NDNS physical activity questionnaires have been reported previously.[18]

The developed NDNS questionnaires were long, detailed, and elaborate in order to include all types and domains of activity, with the aim to achieve face validity for EE assessment. The recent review of the use of physical activity questionnaires makes the point repeatedly that the main reason most questionnaires have had limited validity for assessing EE is because the questionnaire enquires only about specific activity domains, particularly focusing on structured exercise and recreational activity.[11]

Figure 1 Sequence for development of physical activity questionnaires



Two new physical activity questionnaires were developed May-June 2006. The questionnaire modules were developed aimed at collecting better quality data about participants' physical activity in sufficient detail to measure voluntary energy expenditure. One was designed for use with adults and with young people aged 16 to 18, as some young people in this age-group are in work and patterns of activity in those remaining in education tend to be more similar to adults than to younger children. This is referred to in the rest of this report as the *Adult's questionnaire.* The other, referred to in this report as the *Children's questionnaire* was developed for use with children aged 11 to 15.

4.2 Content and administration of the questionnaires

4.2.1 Content of the questionnaires

The adults' questionnaire included questions on walking, cycling, occupational activity, domestic activity, and sports and exercise. Questions on walking and cycling elicited information on physically-active commuting as well as leisure activities. Occupational activity asked whether the participant was mostly sedentary or standing; walking or being otherwise active; or a mixture of the two; and also asked about duration and frequency of moderate activity and of vigorous activity at work.

The questions from the children's questionnaire covered all aspects of their day, including time at school, active travel (walking or cycling) to and from school, as well as the time spent being active during school breaks.

In both questionnaires, questions were asked about the number of days the specific activity was undertaken in the preceding seven- or 14-day period, how much time was spent per occasion, and the intensity level at which the activity was performed. This last was assessed by asking about becoming sweaty, out of breath, or raising their heart beat for cycling and sports and exercise and on walking speed for walking.

4.2.2 Administration of the questionnaires

Both questionnaires were prepared in two versions for the Comparison Study. The version used with participants randomly allocated to assessing dietary intake using an unweighed diary was administered at the second interviewer visit, when the diary was collected. This version asked about physical activity in the preceding seven days. For those whose dietary intake was assessed using four 24hr recalls, the physical activity questionnaire was administered at the end of the last dietary interview (the fourth, except where participants stated at the third interview that they would not be prepared to have a fourth such interview, in which case it was administered after the third dietary recall). This version asked participants about their physical activity over the preceding 14 days. In both cases, the information collected about physical activity covered the time period during which dietary intake was assessed. Both adults and children (aged 11 and over) were asked the questions themselves, although a parent or guardian was present when a child was interviewed.

4.3 Estimation of energy expenditure from questionnaire responses

The variables derived from the questions on walking accounted for the pace at which the informant was walking. Likewise, the sports and activities variables were derived according to the intensity (i.e. whether informant's heart beat was raised) and the time spent in each sport of exercise. For some activities to be considered moderate or vigorous, the informant had to have a raised heart beat while engaging in the activity. In general, most activities were counted as light, moderate, or vigorous based on the intensity coding scheme developed by Ainsworth and her colleagues.[20]

Occupational activity assumed a seven-hour work day for those in paid work and a four-hour work day for those in education or unpaid or voluntary work. Participants were asked on how many days they had been vigorously active[i] and how much time they usually spent on each one of those days being vigorously active at work (for at least five minutes at a time). Vigorous activity at work was assigned a MET value of 4.0METs. Similar questions were asked about moderately vigorous activity[ii]; this time was ascribed a value of 3.0 METs.[iii]

After allowing for moderate and vigorous activity, the remainder of the time at work (up to a maximum of 7hrs for those in paid employment or self-employed and a maximum of 4hrs for those in education or doing voluntary work, including the reported moderate and vigorous activity) was then given METs according to how sedentary they reported they were at work. For those reporting spending most of their work time walking or being otherwise active, this time was assumed to be spent in light activity and was given a MET value of 2.0. Time spent sitting or standing was considered to be sedentary and was given a MET value of 1.5. Those reporting spending half their time sitting or standing and half their time walking or being active were allocated half the remaining time at 1.5 and half at 2.0.

Except for the limit on time spent sitting/standing or walking at work, no time limits were imposed on the time spent on one day in any activity provided the responses were feasible both for that single item and when considered along with time reported on other activities.

---

[i] "I'd like you to think about vigorous activities which take hard physical effort that you did as part of your work. Vigorous activities make you breathe much harder than normal. These may include things like heavy lifting, digging, or heavy building work. With vigorous activities you may get out of breath and you may start sweating."

[ii] "Moderate physical activities make you breathe somewhat harder than normal. These may include activities like carrying light loads or walking briskly while at work."

[iii] These MET values were suggested by Professor Wareham after our inspection of the individual data suggested overestimates of the vigorousness of occupational activity, given the reported occupation.

The questionnaires were analysed to estimate the mean daily time spent in light, moderate, or vigorous physical activity, being sedentary, or sleeping. These were then converted to METs (metabolic equivalents) and therefore mean daily energy expenditure, using basal metabolic rate calculated using the modified Schofield estimates.[21]

Variables from different sections of the questionnaire were then combined to produce composite variables. For adults, the composite variables measure the following:

- Total Physical activity-derived energy expenditure (ie energy expenditure above BMR during active time)
- Questionnaire-derived total energy expenditure, including BMR for all 24hrs

In the derivation of all the variables, the base for whether or not the informant was included was established by whether the informant could have theoretically participated in the particular activity in question. For example, the base for the domestic activity variables includes everyone who answered the question on housework. Similarly, the base for the variables on active travelling to and from school includes everyone who answered affirmatively to the question on whether the child had gone to school in the last seven days.

No adjustment was made for age, except in derivation of BMR, where the following equations were used to estimate BMR in MJ/d[21 22]:

- Male aged 10-17: BMR= (0.068* wt) + (0.574*ht) + (2.157)
- Female aged 10-17: BMR= (0.035* wt) + (1.948* ht) + (0.837)

- Male aged 18-29: BMR= (0.063* wt) - (0.042* ht) + (2.953)
- Female aged 18-29: BMR= (0.057* wt) + (1.184* ht) + (0.411)

- Male aged 30- 59: BMR= (0.048* wt) - (0.011* ht) + (3.670)
- Female aged 30-59: BMR= (0.034* wt) + (0.006* ht) + (3.530)

- Male aged 60+: BMR= (0.038* wt) + (4.068* ht) - (3.491)
- Female aged 60+: BMR= (0.033* wt) + (1.917* ht) + (0.074)

4.4 Validation of physical activity questionnaire in 11- to 15-year-olds

4.4.1 Face validity

Face validity is one aspect of internal validity. It assesses whether questions measure what they set out to measure. Cognitive testing was used during the development of the questionnaires to check whether the draft questions were interpreted by participants to measure what the researchers intended, prior to the validation study. The new questionnaires were cognitively tested July 2006.[18] They were revised September 2006 based on the recommendations from the cognitive testing.

4.4.2 Test-retest reliability

Test-retest reliability is repeatability, i.e the likelihood of the same participant providing the same responses when the same instrument is used on different occasions.

Repeatability was tested by asking each participant in the Actigraph sub-study to answer the physical activity questionnaires twice within the fieldwork period. Comparisons of the two sets of results for each participant were then made to assess repeatability. This design is the standard method for assessing test-retest reliability of physical activity questionnaires and is based on the assumption that physical activity levels are relatively constant between the periods the different questionnaire administrations enquired about. All participants were asked to answer the physical activity questions at the outset of the project. For both adults and children aged 11 and over, the questionnaire was administered for the second time when the Actigraph was collected. The second questionnaire therefore covered most of the week in which the Actigraph was worn. If urine samples (for DLW analysis) were collected at the same time, this was at least 10 days after the Actigraph was handed out It should be noted that non-identical responses could be correct on both occasions, as the activities undertaken may vary at different times.

Table 3 Variables used in energy expenditure analysis

Variable name	Description
Children's questionnaire: *used for participants aged 11-15*	
BMR_ch	BMR
ChPAEEMJ	Mean daily physical activity energy expenditure (MJ/d) Week 1
ChQnTEEMJ	Mean daily questionnaire-derived total energy expenditure (MJ/d) Week 1
ChPAEE2MJ	Mean daily physical activity energy expenditure (MJ/d) Week 2
ChQnTEE2MJ	Mean daily questionnaire-derived total energy expenditure (MJ/d) Week 2
Adults' questionnaire: *used for participants aged 16+*	
BMR	BMR
AdPAEEMJ	Mean daily physical activity energy expenditure (MJ/d) Week 1
AdQnTEEMJ	Mean daily questionnaire-derived total energy expenditure (MJ/d) Week 1
AdPAEE2MJ	Mean daily physical activity energy expenditure (MJ/d) Week 2
AdQTEE2MJ	Mean daily questionnaire-derived total energy expenditure (MJ/d) Week 2

4.4.3 Statistical methods

Repeatability, an aspect of measuring agreement, is an important issue especially when new instruments are being developed. Data on moderate to vigorous physical activity time and mean daily energy expenditure were compared as continuous variables. All statistical analyses were conducted for males and females separately where numbers allowed. Analyses were also stratified by age. For analysis of the children's questionnaire, the analyses were for those aged 11-15. For the adults' questionnaire, the age groups used were 16-49, 50-64, and 65 and over.

Intraclass correlation (ICC) is used to measure inter-rater reliability for two or more raters for continuous variables. It may also be used to assess test-retest reliability and for assessing the agreement between two different methods of assessment, as in this study. ICC may be conceptualized as the ratio of between-groups variance to total variance. ICC is interpreted similarly to the Kappa statistic. A value of 1 indicates perfect relative agreement and a value of 0 indicates no agreement. A rule of thumb is shown in Table 4 below for assessing the relative agreement between two different methods of assessment.[23]

Table 4. Interpretation of kappa values[23]

Kappa value	Description
≤ 0.09	Poor agreement
0.10 – 0.20	Slight agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement
0.81 – 1.0	Almost perfect agreement

5 Actigraph

5.1 Introduction

Given the known potential for error with subjective reports, it was essential to test the new physical activity questionnaires for adults and for children aged 11-15 for repeatability and for validity against an objective measure, the Actigraph. It was also important to test the feasibility of using an objective measure in a population-wide survey, to determine whether it would be practical to use it to measure activity in children aged 4-10yrs in the main NDNS.

5.2 Equipment

The Actigraph (model GT1M) is a small (smaller than a matchbox) and lightweight digital uniaxial accelerometer that is worn on the waist using a clip or belt, and detects vertical accelerations from 0.05 to 2.50g, with a frequency response of 0.25-2.50Hz. Movements outside of normal human motion are filtered out electronically.[24] The filtered acceleration signal is digitized and summed over a specified time interval (one minute) to produce a number of 'counts' for that time period, which is stored and the summing restarts.

The Actigraph was chosen on the grounds of cost, ease of use and compliance, and a wide comparison base. Actigraphs have been used successfully on many large-scale surveys including the US National Health and Nutrition Examination Survey (NHANES 2003/04). This study used accelerometers among more than 1,000 participants (aged 6+) showing that objective physical activity measures are feasible in large-scale population studies. Actigraphs are currently being recommended for use within the British Millennium Cohort Study and the National Study of Diet and Nutrition (NDNS). It has been shown to be reliable and adequate for assessing physical activity.[25 26]

Consenting participants were asked to wear the device on their waist, positioned above the right hip, for seven continuous days. At the end of the seven day period, the Actigraphs were returned to NatCen's Operations Department for downloading the data and for the Actigraphs to be cleaned and recharged. Once this was completed, the Actigraphs were allocated a new id number and returned to the interviewer for use on a different participant.

5.3 Timelog

Each participant (or his/her parent/guardian) was asked to keep a log, recording when the Actigraph was put on and taken off, the reason for this, and activities undertaken that are poorly recorded by the Actigraph (eg cycling) with its duration.

5.4 Estimating energy expenditure from Actigraph recordings

5.4.1 Preliminary analyses

Initial Actigraph analyses were conducted by UCL. Data processing by NatCen using Mahuffe software (written by the MRC Epidemiology Unit, Cambridge) provided average accelerometric intensity, as well as time spent at different intensity levels, here labelled as sedentary, light, moderate, vigorous, or very vigorous activity. These aggregated results were converted to energy expenditure. Actigraph data were included from participants who wore the Actigraph for at least four days for at least 500 minutes per day but any wear after the end of seven days was ignored.

Estimates of energy expenditure were made both including and excluding activity and time data from the Actigraph logbook. Inclusion of logbook information required lengthy and often manual processing, however, estimation of energy expenditure including and excluding this information affected the overall results very little.

Expert advice received afterwards noted that such information is generally poorly recorded, particularly after the first day, and adds little useful information in relation to the burden it imposes on the participants and the resources required for entering and analysing such data. No further information is therefore presented about how these data were used.

5.4.2 Definitive analyses by MRC Cambridge

Raw Actigraph data were processed to take account of periods not worn, based on the length of strings of zero movement. Strings of 45 min or less were considered 100% valid (monitor worn), whereas for strings over 75 min the probability that the Actigraph was worn was set at 0%. Intermediate lengths of zero strings were assigned a wear probability between 0 and 100%, using linear interpolation, e.g. a zero string of 60 min is assigned a wear probability of 50%. As the sampling procedure with respect to the diurnal rhythm was biased towards awake time (monitors could be taken off at night), a second flagging machinery was imposed to detect sleep periods, for which the registered zero movement, despite the inference that the monitor with great certainty is not worn, is again considered valid. Sleep probability was defined using a 5-6hrs initial ramp-up detection with 10-12hrs roll-off. Summing wear and sleep probability yields the combined belief in every observation in the raw Actigraph time-series, which can subsequently be “repaired” using the formula:

$ACC_{repair}(t) = \text{WearSleep probability} * ACC_{raw}(t) + (1-\text{WearSleep probability}) * ACC_{inferred}(t)$.

The inferred ACC used for this imputation was calculated as the probability-weighted average counts per minute (cpm) for that time of day, based on hour-specific data from all days on which that individual wore the Actigraph, for example all data between 3 and 4 pm across all days. For this imputation, only days with at least 360 minutes of wear (integrated wear probability > 360) were included.

The complete (repaired) ACC time-series was converted to minute-by-minute physical activity energy expenditure (PAEE) from four different published equations, which were all derived using the old generation of Actigraph (model 7164). Therefore, a (conservative) inter-generation monitor correction factor of 0.93 was applied before conversion to PAEE.[24 27] For the published equations yielding estimates of total (BMR + PA) energy expenditure, 1 MET was subtracted corresponding to the resting oxygen consumption of 3.5 O_2/min/kg for adults (e.g. Freedson[28] Swartz[29]) and 5.5 O_2/min/kg for children (e.g. Trost[30] Puyau[31]). Similarly, an energetic value of 20.35 J/ml O_2 was used[32]. For example, the equation developed by Swartz et al[29] was rephrased to yield PAEE from GT1M-based information as:

$PAEE_{Swartz}(t) = (2.606 - 1 + 0.0006863 * (ACC_{repair}(t) / 0.93)) *3.5*20.35$

Positive intercepts are by-products of the linear regression procedure used to derive the equations but commonly represent non-valid extrapolations. It is well-recognised in the field that assigning a non-zero value of PAEE to minutes with no movement is counter-intuitive, especially considering the high prevalence of inactive periods in daily living; for example Corder et al[33] noted that the Trost and colleague's equation[25] yielded a six-fold overestimation of PAEE, most likely because about 90% of the data collected during free-living lie outside the range of the laboratory data used for deriving this equation. It is generally accepted, therefore, that when PAEE is plotted against accelerometry output (cpm) and a conversion equation is derived, the equation should go through zero. However, different researchers use different methods to adjust the (often linear) equations to ensure this with no consensus on the best method. In the present analyses, the equations were adjusted with the flex movement point method,[34] using a flex point 100 cpm. This means that for all equations, an alternative equation was used for the interval between 0 and 100 cpm; the relationship was forced through the origin (0,0) and the PAEE at 100 cpm).

The estimation of PAEE in kJ/day/kg was done for participants of all ages using two different approaches; first the time-integral of the min-by-min PAEE time-series, and second the average instantaneous PAEE (J/min/kg) multiplied by an assumed awake period per day (aw.P), with each valid day weighted by the square-root of the fraction of time the Actigraph was worn that day. Only participants who accumulated at least 24 hours of valid data (integrated wear probability > 24*60) were included in the analysis.

The Actigraph-based PAEE estimates were compared to the PAEE estimate from DLW, derived using the total daily energy expenditure (TEE) based on the carbon dioxide production with a food quotient of 0.85 and Schofield's estimate of awake BMR, combined with a correction for the slightly lower (5%) metabolism during sleep:

$$PAEE_{DLW} = 0.9 * TEE_{DLW} + (aw.P + [24-aw.P]*0.95)*BMR_{Schofield} / 24$$

In the present analyses, children aged seven or less was assumed to sleep 12 hours per day and adolescents/adults aged 15 years and over were assumed to sleep eight hours per day, with intermediate ages being assigned intermediate values using linear interpolation.

In order to obtain an Actigraph-derived estimate for total daily energy expenditure, the estimates of physical activity energy expenditure were added to the 24-hr estimate of resting metabolic rate (Schofield BMR[21] with sleep correction), and the total divided by 0.9 (to allow for the 10% of energy expenditure due to diet-induced thermogenesis), and finally multiplied by body weight. The TEE estimates from Actigraph were compared to the estimate from DLW (in kJ/day). The DLW method estimate TEE within 10% accuracy, which translates to an accuracy of 20% for the PAEE estimate for active individuals (PAL=2.0) under the assumption of zero error in the BMR estimate. Since error in BMR does exist on the individual level and most have a PAL<2.0, accuracy of the PAEE estimate from DLW must be assumed worse than 20%.

5.5 Analysis

Actigraph data were analysed:

- to compare physical activity energy expenditure and total energy expenditure as determined by the second questionnaire with energy expenditure as determined from the Actigraph data; and
- to compare total energy expenditure as determined by DLW with total energy expenditure derived from Actigraph data.

The same statistical methods were used as for analysing test-retest repeatability of the questionnaire (see section 4.4 and Table 4 above).

6 Results

6.1 Actigraph

6.1.1 Adherence to Actigraph wear

Children

64 children aged 4-15 agreed to take part in the Actigraph sub-study. Twenty wore the Actigraph for at least 500 minutes on all seven days. An Actigraph was worn for at least four days for at least 500 minutes per day by 30 boys and 31 girls (95% of those agreeing to take part). On average, the Actigraph was worn for 722 minutes a day by these 30 boys and for 730 minutes per day by these 31 girls, on days when it was worn for at least 500 minutes[i].

There were some non-significant variations in the length of wearing an Actigraph by age (mean 12.6hrs per day for children aged 11-15yrs, 11.4hrs per day for children aged 4-10yrs). This is to be expected, as the younger children were likely to have gone to bed earlier.

30 girls and 28 boys wore the Actigraph enough (>24 hours) to be included in the present analyses, covering on average 51% of each monitored day (slightly higher for girls than boys, p=0.097).

Adults

Of the 90 participants aged 16 and over who agreed to wear the Actigraph, 60 wore the Actigraph on seven days for at least 500 minutes per day.[i] 88 (44 males and 44 females, 98% of those agreeing to take part) wore the Actigraph for at least 500 minutes per day for at least four days.

On average, the Actigraph was worn for 13.6 hours per day by those aged 16 and over, considering days on which it was worn for at least 500 minutes. As with children, variations by age were not significant.

45 men and 45 women wore the Actigraph enough (>24 hours) to be included in the present analyses, covering on average 59% of each monitored day (higher for women than men, p=0.023).

6.1.2 Comparison between DLW and Actigraph results

Physical Activity Energy Expenditure (PAEE)

In those taking part in the DLW study, average (SD) accelerometric intensity during the time the Actigraph was worn was 678 (202) cpm for children aged 4-10yrs, 480 (139) cpm for children aged 11-15yrs, and 323 (142) cpm for adolescents/adults aged 16 yrs and over. Corresponding values for 24-hr estimates were 581 (154), 374 (131), and 243 (113) cpm respectively.

There was no interaction by sex in the overall agreement analyses, so results are combined for males and females. On average, Actigraph equations derived using treadmill activity underestimated PAEE (by about 50%) but the equation based on lifestyle activities by Swartz[29] estimated PAEE reasonably well on the group level, with the Puyau equation[31] offering intermediate level of accuracy. All four equations explained about 26% of the

---

[v] Variable: w1regtim.

variation in bodymass-specific PAEE (kJ/day/kg). The time-integration approach was generally more accurate than the awake time sampling method.

Table 5. Comparison of PAEE results: Time-integration approach

Age-group	N	PAEE (kJ/day/kg)				
		DLW (SD)	Freedson (SD) [RMSE]#	Swartz (SD) [RMSE]#	Trost (SD) [RMSE]#	Puyau (SD) [RMSE]#
4-10	28	73.0 (28)	38.5 (8)** [42.7]	72.5 (11) [24.7]	43.8 (11)** [38.6]	52.5 (8)** [32.3]
11-15	30	63.6 (25)	30.7 (8)** [40.7]	61.6 (8) [24.3]	33.9 (8)** [38.1]	44.2 (6)** [31.0]
16+	90	47.8 (17)	24.3 (8)** [28.2]	52.7 (12)** [16.9]	25.9 (8)** [27.2]	37.3 (9)** [18.8]
All	*148*	*55.8 (24)*	28.3 (9)** [34.1]	58.2 (14) [20.2]	30.8 (12)** [32.0]	41.6 (10)** [24.6]
			$R^2 = 0.26$	*$R^2 = 0.26$*	*$R^2 = 0.25$*	*$R^2 = 0.27$*

RMSE (root mean square error) is the geometrical average distance from the estimate to the criterion, thus a measure of overall accuracy (lower is better).

*p<0.05, **p<0.01 different from DLW estimate

Table 6. Comparison of PAEE results: Awake-time sampling approach

Age-group	N	PAEE (kJ/day/kg)				
		DLW (SD)	Freedson (SD) [RMSE]#	Swartz (SD) [RMSE]#	Trost (SD) [RMSE]#	Puyau (SD) [RMSE]#
4-10	28	73.0 (28)	57.5 (13)** [30.6]	102.3 (15)** [39.9]	66.9 (17) [28.2]	74.7 (12) [26.3]
11-15	30	63.6 (25)	36.0 (8)** [36.5]	67.8 (12) [24.9]	40.9 (10)** [33.0]	49.1 (9)** [28.1]
16+	90	47.8 (17)	25.1 (8)** [27.7]	52.2 (12)* [16.7]	27.1 (10)** [26.2]	37.2 (9)* [18.8]
All	*148*	*55.8 (24)*	33.4 (15)** [30.2]	64.8 (23)** [24.4]	37.4 (19)** [28.1]	46.7 (17)** [22.5]
			$R^2 = 0.26$	*$R^2 = 0.27$*	*$R^2 = 0.26$*	*$R^2 = 0.27$*

RMSE (root mean square error) is the geometrical average distance from the estimate to the criterion, thus a measure of overall accuracy (lower is better).

*p<0.05, **p<0.01 different from DLW estimate

Table 7. Derived PAEE equations

Age group	PAEE equation in kJ/day/kg	R^2	RMSE#
Awake estimates			
4-10	$0.03944 \cdot AG(cpm)_{awake}$ - $9.3 \cdot$male sex + 50.9	0.120	27.1
11-15	$0.04874 \cdot AG(cpm)_{awake}$ - $7.6 \cdot$male sex + 43.7	0.139	24.4
16+	$0.04524 \cdot AG(cpm)_{awake}$ - $0.9 \cdot$male sex + 33.7	0.139	16.4
All	$0.04564 \cdot AG(cpm)_{awake}$ - $0.15 \cdot$age - $4.2 \cdot$male sex + 44.2	0.287	20.2
24-hr estimates			
4-10	$0.06888 \cdot AG(cpm)_{24hr}$ - $10 \cdot$male sex + 38.1	0.184	26.1
11-15	$0.03676 \cdot AG(cpm)_{24hr}$ - $10 \cdot$male sex + 54.6	0.116	24.7
16+	$0.05239 \cdot AG(cpm)_{24hr}$ - $0.9 \cdot$male sex + 35.5	0.115	16.6
All	$0.05378 \cdot AG(cpm)_{24hr}$ - $0.14 \cdot$age - $4.3 \cdot$male sex + 45.2	0.290	20.2

RMSE (root mean square error) is the geometrical average distance from the estimate to the criterion, thus a measure of overall accuracy (lower is better).

Total Energy Expenditure (TEE)

On average, the treadmill-based Actigraph equations, combined with Schofield estimates of resting energy expenditure, underestimated TEE, whereas the Schofield-Swartz estimate of TEE was reasonable (Table 8). As with PAEE, intermediate accuracy was observed for the Puyau estimate. Each of the four TEE estimates could explain about 80% of the variance in DLW-based TEE.

Table 8. Comparison of TEE results: Time-integration approach

Age-group	**N**	**TEE (MJ/day)**				
		DLW (SD)	**Freedson (SD) [RMSE]#**	**Swartz (SD) [RMSE]#**	**Trost (SD) [RMSE]#**	**Puyau (SD) [RMSE]#**
4-10	28	7.01 (1.5)	6.00 (1.0)** [1.2]	6.98 (1.3) [0.7]	6.16 (1.1)** [1.1]	6.40 (1.1)** [0.9]
11-15	30	10.7 (2.3)	8.79 (1.6)** [2.4]	10.7 (2.1) [1.4]	8.97 (1.7)** [2.2]	9.61 (1.8)** [1.7]
16+	90	11.4 (2.5)	9.43 (1.9)** [2.4]	11.9 (2.5)** [1.5]	9.56 (2.0)** [2.3]	10.6 (2.2)** [1.6]
All	*148*	*10.5 (2.9)*	8.65 (2.1)** [2.2]	10.7 (2.9) [1.4]	8.80 (2.2)** [2.1]	9.59 (2.5)** [1.5]
			$R^2 = 0.82$	$R^2 = 0.80$	$R^2 = 0.80$	$R^2 = 0.81$

RMSE (root mean square error) is the geometrical average distance from the estimate to the criterion, thus a measure of overall accuracy (lower is better).

*p<0.05, **p<0.01 different from DLW estimate

Table 9. Comparison of TEE results: Awake-time sampling approach

Age-group	N	TEE (MJ/day)				
		DLW (SD)	Freedson (SD) [RMSE][#]	Swartz (SD) [RMSE][#]	Trost (SD) [RMSE][#]	Puyau (SD) [RMSE][#]
4-10	28	7.01 (1.5)	6.53 (1.2)** [0.9]	7.80 (1.4)** [1.1]	6.80 (1.4) [0.9]	7.02 (1.2) [0.7]
11-15	30	10.7 (2.3)	9.07 (1.6)** [2.1]	11.0 (2.1) [1.4]	9.35 (1.6)** [1.9]	9.86 (1.8)** [1.6]
16+	90	11.4 (2.5)	9.50 (1.9)** [2.4]	11.9 (2.5)** [1.5]	9.68 (2.0)** [2.2]	10.6 (2.1)** [1.6]
All	*148*	*10.5 (2.9)*	8.85 (2.1)** [2.1]	10.9 (2.7)** [1.4]	9.07 (2.1)** [2.0]	9.75 (2.3)** [1.5]
			R^2 = 0.80	*R^2 = 0.79*	*R^2 = 0.77*	*R^2 = 0.80*

[#] RMSE (root mean square error) is the geometrical average distance from the estimate to the criterion, thus a measure of overall accuracy (lower is better).

*p<0.05, **p<0.01 different from DLW estimate

6.2 Test-test repeatability of the questionnaires

6.2.1 Children's questionnaire

The questionnaire shows good repeatability in girls and adequate correlation in boys for measures of total activity time and physical activity energy expenditure (PAEE). For total energy expenditure (TEE), there was acceptable correlation for boys and excellent correlation for girls. Numbers were small.

Table 10. Test-test repeatability of children's questionnaire, boys and girls aged 11-15

		Intraclass Correlation	95% Confidence Interval		F Test with True Value 0				Base
			Lower Bound	Upper Bound	Value	df1	df2	Significance	
Average daily total physical activity time	Boys	0.526	-0.198	0.812	2.108	19.0	19	0.056	20
	Girls	0.621	0.087	0.843	2.639	21.0	21	0.016	22
Average daily physical activity energy expenditure	Boys	0.492	-0.283	0.799	1.969	19.0	19	0.074	20
	Girls	0.706	0.293	0.878	3.405	21.0	21	0.004	22
Average daily total energy expenditure	Boys	0.562	-0.107	0.827	2.281	19.0	19	0.040	20
	Girls	0.808	0.537	0.920	5.206	21.0	21	0.000	22

6.2.2 Adults' questionnaire

The following analyses were conducted using a slightly different set of assumptions for analysing the occupational activity but were consistent between the first and second questionnaire.

Comparing the second questionnaire[i] with the first questionnaire[ii], the ICC for mean daily total activity time in men was 0.29 (fair agreement); in women it was 0.18 (slight agreement). These figures hide marked variation by age in men. The ICC was 0.29 in males aged 16-49 (p=0.0008) but 0.615 (p=0.000) in men aged 65+, representing substantial agreement in the oldest men but fair agreement in the youngest group.

Comparing the second questionnaire[iii] with the first questionnaire[iv], the ICC for daily physical activity energy expenditure in men was 0.76 and in women was 0.71 (each p<0.001), representing substantial agreement. There was also significant, substantial agreement for each age-group in women; for men aged 16-49, there was almost perfect agreement (ICC 0.94, p<0.001, n=15).

Comparing the second questionnaire[v] with the first questionnaire[vi], the ICC for daily total energy expenditure showed almost perfect agreement in men aged 16-49 (ICC0.95 , p<0.001), in all men aged 16 and over (ICC 0.82, p<0.001) and in women in each group except those aged 65 and over (ICC 0.82, p<0.001 for all women). There was also substantial agreement for men aged 50-64.

6.3 Comparison of questionnaire data with DLW data

6.3.1 Children's questionnaire

Energy estimates from the questionnaire for 31 children aged 11-15 was compared with DLW and Actigraph estimates.

There was no significant difference between the questionnaire and the Actigraph for PAEE (Table 11). The mean was not significantly different from the DLW estimate, though individual results were very variable, but overall the questionnaire was deemed "wildly inaccurate" as it explained only 10% of the variance of the DLW PAEE estimates across the two age groups (1% within the younger group). RMSE values for the questionnaire estimates were about twice as high as for the Actigraph estimates in the comparison with DLW.

For TEE estimates, the questionnaire results were not compared with the Actigraph estimates as they use the same constants and adjustments, so only the PAEE element differs in its estimation. The questionnaire results estimated TEE with no significant bias, compared with the DLW results; however the RMSE was 3MJ/d, indicating low level of accuracy for individual estimates (Table 12).

6.3.2 Adults' questionnaire

Energy estimates from the questionnaire for 91 participants aged 16+ was compared with DLW and Actigraph estimates.

---

[i] Variable: AdTotT2

[ii] Variable: AdTotT

[iii] Variable: AdQn PAEE2MJ

[iv] Variable: AdQnPAEEMJ

[v] Variable: AdQnTEE2MJ

[vi] Variable: AdQnTEEMJ

There was no significant difference between the questionnaire and the Actigraph for PAEE (Table 11). The mean was significantly lower than the DLW estimate, though individual results were variable. The questionnaire explained 12% of the variance of the DLW PAEE estimates. RPAQ produces slightly better results and is significantly shorter.[44]

The TEE estimates were again not compared with the Actigraph estimates as they use the same constants and adjustments, so only the PAEE element differs in its estimation. The questionnaire did not differ significantly from the DLW estimates (Table 12) and produced TEE values with an accuracy (RMSE) of about 3MJ/d. RPAQ provides similar results.[44] The DLW method, on the other hand, provides TEE estimates within ±1MJ/d (10%).

Table 11. Comparison of physical activity energy expenditure results

Age-group	N	PAEE (kJ/day/kg)		
		DLW (SD)	$\text{Actigraph}_{\text{Swartz}}$ (SD) [RMSE]#	PAQ (SD) [RMSE]#
11-15	31	62.6 (26)	61.6 (8) [24.3]	60.9 (53) [55.9]
16+	91	47.2 (18)	52.7 (12)** [16.9]	36.5 (33)**++ [33.7]
All	*122*	*51.1 (21)*	54.9 (12) [19.0]	*42.7 (41)* ++* *[40.5]*
			R^2 = 0.19	*R^2 = 0.09*

RMSE (root mean square error) is the geometrical average distance from the estimate to the criterion, thus a measure of overall accuracy (lower is better).

*p<0.05, **p<0.01 different from DLW estimate

+p<0.05, ++p<0.01 different from Actigraph (Swartz) estimate

Table 12. Comparison of total energy expenditure results

Age-group	N	TEE (MJ/day)		
		DLW (SD)	$\text{Actigraph}_{\text{Swartz}}$ (SD) [RMSE]	PAQ (SD) [RMSE]
11-15	31	10.6 (2.3)	11.0 (2.1) [1.4]	10.4 (2.1) [3.0]
16+	91	11.4 (2.4)	11.9 (2.5)** [1.5]	10.5 (5.0)** [2.9]
All	*122*	*11.2 (2.4)*	11.6 (2.5)* [1.5]	10.5 (3.8)** [2.9]
			R^2 = 0.70	*R^2 = 0.42*

RMSE (root mean square error) is the geometrical average distance from the estimate to the criterion, thus a measure of overall accuracy (lower is better).

*p<0.05, **p<0.01 different from DLW estimate

7 Discussion

7.1 Expected lack of correlation between questionnaire data and objective measures

7.1.1 Allowing for estimated BMR

When deriving Physical Activity Energy Expenditure (PAEE) from physical activity questionnaire data, we used Schofield's regression equation for BMR.[21] However, it is only a model - an approximation that fits at the population level. At the individual level, those with extremes of age and body mass and composition can appear to have a greater BMR per minute than their estimated TEE during periods of activity, leading to a negative estimate for PAEE.

It has been proposed that to avoid this that PAEE should be recalculated using

- (MET-1) * Minutes being active

instead of

- (MET * Minutes being active) – (BMR per minute * minutes being active).

This could then be compared with DLW-derived PAEE.

However, Neilson and colleagues have pointed out that PAEE is derived by subtracting Resting Metabolic Rate (RMR, generally equated with BMR) and thermogenesis. Studies have found that the proportion of healthy adults with estimated RMR from prediction equations within 10% of RMR measured by calorimetry varied from 45% to 81% in non-obese individuals and from 38% to 70% in studies of obese individuals.[11] Neilson *et al* point out that a physical activity validation study using such prediction equations to predict PAEE would inevitably suggest disagreement, which may be wrongly attributed to the questionnaire.[11] Nonetheless, PAEE is by definition never negative and it would seem reasonable to truncate item-based estimations if they should return negative values.

7.2 Other comments

The figures for correlation depend on the actual sample, so age and sex-stratified analyses are not recommended within the same instrument.

Senior researchers at the MRC Epidemiology Unit in Cambridge were very impressed with the results of the comparison of the Actigraph and DLW data and with the quality of the underlying fieldwork.

7.3 Other options for estimating energy expenditure in participants aged 11+ in future NDNS years

7.3.1 Children aged 11-18yrs

Children's Physical Activity Questionnaire (C-PAQ)

The C-PAQ was recommended for NDNS if we change to a shorter questionnaire to assign NDNS participants aged 11-15yrs into three to five categories of physical activity. This eight page questionnaire has been used with children aged 12-13yrs and 16-17yrs and with parents of children aged 4-5yrs. Questions cover physically active transport, vigorous intensity activities, moderate intensity activities, muscle-strengthening activity, and screen-based (sedentary) activities in the previous month. However, the validation results were poor: although there were no mean bias for CPAQ in 12-14 year olds, the error was substantial.[35]

Physical Activity Questionnaire for Children (PAQ-C).

The PAQ-C is a **seven-day** recall self-report questionnaire designed to assess daily activity in the moderate to vigorous range. General physical activity scores were calculated as an average physical activity score (PA score) in a continuous range from 1 (low active) to 5 (high active). Validity of the PAQ-C was determined previously using the aerobic step fitness test as the criterion measure. Test-retest reliability of the PAQ-C was also assessed during a seven-day period: they averaged the five measurements of the PAQ-C to represent the physical activity level for each child.[36] The PAQ-C has been used in previous studies with children of a similar age.[37] However, this instrument yields a physical activity score that may not be convertible to energy expenditure estimates, unless a calibration study is undertaken.

Young people's *Physical Activity Questionnaire (Y-PAQ)*

Y-PAQ has been suggested for NDNS if we change to a shorter questionnaire to assign NDNS participants aged 16-18yrs into three to five categories of physical activity. Y-PAQ uses a seven-day recall period. It has been shown not to be useful in children aged 9-10yrs but has been used in those aged 12-13yrs and 16-17yrs. Similar results were found as for CPAQ: absolute PAEE and MVPA estimated from these self-reports were not valid on an individual level in young people, although YPAQ appeared to rank individuals accurately.[35]

7.3.2 Adults (aged 19+)

Other existing questionnaires were considered during development of the current NDNS questionnaires, including:

- The General Practice Physical Activity Questionnaire (GPPAQ)[38]
- EPIC2 (EPAQ2), developed by the MRC Epidemiology Unit, Cambridge[39, 40]
- The IPAQ (International Physical Activity Questionnaire) long form[41, 42]
- IPAQ short form[43, 42]
- RPAQ (Recent Physical Activity Questionnaire), also developed by the MRC Epidemiology Unit, Cambridge. This has recently been validated against DLW in 50 adults aged 20+.[44]

Results of the validation of RPAQ showed a significant correlation of r=0.43 and a RMSE value of 24.7 kJ/day/kg for PAEE.[44] This compares favourably with the results of the Comparison Study questionnaires, which had r=0.3 (r^2=0.09) overall and RMSE 60.9kJ/d/kg 36.5kJ/d/kg for participants aged 16 and over for PAEE (Table 11).

Soren Brage has therefore recommended that we use RPAQ for NDNS if we change to a shorter questionnaire to assign NDNS adult participants into three to five categories of physical activity. This would ensure comparison with other large cohort studies, such as MRC Fenland Study (n>5,000) and most likely also the enhanced part of UK Biobank (n>100,000).

7.4 Implications of changing the energy expenditure data collection method

7.4.1 Accuracy of estimating energy expenditure

Children aged 11-15yrs

The questionnaire produced an population mean close to that derived from DLW data but the variance in the results was so large that categorisation into meaningful energy expenditure categories is not possible. The Actigraph data resulted in a greater underestimate of EE at the population level but with acceptable variance.

Participants aged 16+

Validation of the RPAQ produced results that were at least as good as those presented in this report for the NDNS adults' questionnaire.[44]

It needs to be understood that EE assessment is not feasible if using a shorter questionnaire that enquires about fewer physical activity domains.[11] Such questionnaires will vary in their ability to classify participants accurately into physical activity categories, based mostly on leisure activity.

7.4.2 Participant burden

Omitting the questionnaire will reduce the third interview by about 20 minutes. This will reduce the respondent burden. It is possible that the length of the third interview, and the number of repetitive questions, is adversely affecting participants' willingness to have a nurse visit.

Wearing an Actigraph was considered acceptable in the Comparison Study: almost all teenagers wore the Actigraph for at least four days for at least 500 minutes per day. The greatest burden is from completing the log book for the days the Actigraph is worn, but other studies have shown this is completed inaccurately, with completeness falling as the week progresses. Larger population studies using accelerometers do not ask participants to complete a log book. Our initial findings were little different when analyses included or excluded data from the log book.

Answering a self-completion questionnaire on physical activity will add five to 10 minutes to the third interviewer visit.

7.4.3 Financial implications

Additional costs incurred by extending Actigraph use to those aged 11-15 years include additional interviewer time for explaining and gaining consent for Actigraph wear; giving the token of appreciation for Actigraph wear; office processing of Actigraph data; and posting out/returning Actigraphs. There will also be additional data entry for the RPAQ. There will also be costs for repairing or replacing Actigraphs.

Savings include reducing the token of agreement for wearing an Actigraph to £10 (for a number of reasons, including no longer completing an Actigraph log book); less office time required as no data entry required from the log; and less interviewer time required for the second interview for all participants aged 11 and over.

8 Proposals for NDNS year 2 onwards (April 2009 onwards)

8.1 What the FSA would like

FSA staff have reiterated that their preference is for 'energy in / energy out' data at an individual level[i] but as this is not possible to achieve, they would accept energy expenditure estimates provided in three to five categories of expenditure level.[ii] They do not wish to abandon the use of questionnaires but would be content with a revised, shorter version that can categorise individuals into those three to five categories.[i]

8.2 Decisions taken December 2008

8.2.1 Children aged 4-10 years

- Continue to use Actigraph, as 2008/2009, but with no log.

[i] Meeting of the NDNS Physical Activity Expert Group, July 2008.

[ii] Email from Melanie Farron-Wilson to Beverley Bates, NatCen, 1 Aug 2008.

- It is suggested (by Soren Brage) to upgrade the firmware to record acceleration in more than a single dimension.
- In addition, movement frequency should be recorded and time resolution maximised (epoch length shortened). Data are back-ward compatible.

8.2.2 Children aged 11-15yrs

- Cease use of the questionnaire (except keep questions on sleep duration).
- Use Actigraph, as in the Comparison Study (but maximising the information as described above), but with no log.

8.2.3 Children aged 16-18 years

- Cease use of the questionnaire (except keep questions on sleep duration).
- Replace with RPAQ, as a self-completion at the diary pick-up interviewer visit. During the course of 2009, a web-based version of RPAQ will be available.

8.2.4 Adults aged 19 and over

- Cease use of the questionnaire (except keep questions on sleep duration).
- Replace with RPAQ, as a self-completion at the diary pick-up interviewer visit.

9 References

[1] Walsh MC, Hunter GR, Sirikul B, Gower BA. Comparison of self-reported with objectively assessed energy expenditure in black and white women before and after weight loss. *Am J Clin Nutr.* 2004;**79**:1013-9.

[2] Montoye HJ, Kemper HCG, Washburn RA, Saris WHM. *Measuring Physical Activity and Energy Expenditure.* Human Kinetics Europe Ltd, 1995.

[3] Bratteby L-E, Sandhagen B, Fan H, Samuelson G. A 7-day activity diary for assessment of daily energy expenditure validated by the doubly labelled water method in adolescents. *Eur J Clin Nutr.* 1997;**9**:585-91.

[4] Klesges RC, Eck LH, Mellion MW, Fulliton W, Somes MW, Hanson C. The accuracy of self-reports of physical activity. *Med Sci Sports Exercise.* 1990;**22**:690-7.

[5] Sims J, Smith F, Duffy A, Hilton S. The vagaries of self-reports of physical activity: a problem revisited and addressed in a study of exercise promotion in the over 65s in general practice. *Family Practice.* 1999;**16**:152-7.

[6] Mathews CE. Use of self-report instruments to assess physical activity. In Welk C (Ed) . *Physical activity assessment for health related research.* Champagne, Illinois: Human Kinetics, 2002.

[7] E.g. Burskik ER, Harris D, Mendez J, Skinner J. Comparison of two assessments of physical activity and a survey method for calorie intake. *Am J Clin Nutr.* 1971; 24: 1119-25.

[8] Bonnefoy M, Normand S, Pachiaudi C, Lacour JR, Laville M, Kostka T. Simultaneous validation of ten physical activity questionnaires in older men : a doubly labeled water study. *J Am Geriatr Soc.* 2001;**49**:28-35.

[9] Philippaerts RM, Weterkerp KR, Lefevre J. Doubly labelled water validation of three physical activity questionnaires. *Int J Sports Med*. 19099;**20**:284-9.

[10] Conway JM, Irwin ML, Ainsworth BE. Estimating energy expenditure from the Minnesota Leisure Time Physical Activity and Tecumseh Occupational Activity questionnaires – a doubly labelled water validation. *J Clin Epidemiol.* 2002;**55**:392-9.

[11] Neilson HK, Robson PJ, Friedenreich CM, Csizmadi I. Estimating activity energy expenditure: how valid are physical activity questionnaires? *Am J Clin Nutr.* 2008;87:279-91.

[12] Haskell WL, Leon AS, Caspersen CJ *et al.* Cardiovascular benefits and assessment of physical activity and physical fitness in adults. *Med Sci Sports Exerc.* 1992;**24**:S201-S220.

[13] Matthews CE. Calibration of accelerometer output for adults. *Med Sci Sports Exercise.* 2005;**37**:S512-S522.

[14] Goldberg GR, Black AE, Jebb SA, *et al*. Critical evaluation of energy intake data using fundamental principles of energy physiology. 1. Derivation of cut-off values to identify under-recording. *Eur J Clin Nutr.* 1991;**45**:569–81.

[15] Huang T, Roberts SB, Howarth NC, McCrory MA. Effect of screening out implausible energy intake reports on relationships between diet and BMI. *Obes Res.* 2005;**13**:1205–17.

[16] Rennie K L, Coward A, Jebb SA. Estimating under-reporting of energy intake in dietary surveys using an individualised method. *Br J Nutr.* 2007;**97**:1169-76.

[17] Bluck L. Energy expenditure: Doubly Labelled Water (DLW) *in NationalDiet and Nutrition Survey Rolling Programme Comparison Study Part 1.* ***A comparison of results by dietary assessment*** *method: repeat 24-hour recall and four-day estimated diet diary (Alison Stephen, Birgit Teucher, Les Bluck et al. unpublished).*

[18] MacKenzie H, Kitchen S. *National Diet and Nutrition Survey. Development of physical activity and sun exposure questions: findings from cognitive interviews.* London: NatCen, 2006.

[19] Yore MM, Ham SA, Aimsworth BE *et al.* Reliability and validity of the instrument used in BRFSS to assess physical activity. *Med Sci Sports Exercise.* 2007;**39**:1267-74.

[20] Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of Physical Activities: an update of activity codes and MET intensities. *Med Sci Sport Exer*. 2000;**32**(suppl):S498-S516.

[21] Schofield WN. Predicting basal metabolic rate. *Hum Nutr Clin Nutr* 1985;**39**C(supp1):5-41.

[22] Henry CJ. Basal metabolic rate studies in humans: measurement and development of new equations. *Public Health Nutr.* 2005;**8**:1133-52.

[23] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;**33**:159-74.

[24] Rothney MP, Apker GA, Song Y, Chen KY. Comparing the performance of three generations of ActiGRaph accelerometers. *J Appl Physiol.* 2008;**105**:1091-7.

[25] Trost SG, Way R, Okely AD. Predictive validity of three ActiGraph energy expenditure equations for children. *Med Sci Sports Exe*. 2006;**38**:380-7.

[26] Pate RR, Almeida MJ, McIver KL, Pfeiffer KA, Dowda M. Validation and Calibration of an Accelerometer in Preschool Children. *Obesity.* 2006;**14**:2000–6.

[27] Corder K, Brage S, Ramachandran A, Snehalatha C, Wareham N, Ekelund U. Comparison of two Actigraph models for assessing free-living physical activity in Indian adolescents. *J Sports Sci*. 2007;**25**:1607-11.

[28] Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications accelerometer. *Med Sci Sports Exerc.* 1998;**30**:777-81.

[29] Swartz et al. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Med Sci Sports Exerc.* 2000;**32** (9,Suppl.):S450–6.

[30] Trost 2001: Physical activity and determinants of physical activity in obese and non-obese children. *Int J Obes Relat Metabl Disord.* 2001;**25**:822-9.

[31] Puyau M, Adolph A, Vohra F, Butte N. Validation and calibration of physical activity monitors in children. *Obes Res.* 2002;**10**:150–7.

[32] Consolazio C, Johnson R and Pecora L. *Physiological measurements of metabolic functions in man*. New York: McGraw Hill, 1963.

[33] Corder K, Ekelund U, Steele RM, Wareham NJ, Brage S. Assessment of physical activity in youth. *J Appl Physiol*. 2008;**105**:977-87.

[34] Brage S, Brage N, Franks PW, Ekelund U, Wong MY, Andersen LB, Froberg K, Wareham NJ. Branched equation modeling of simultaneous accelerometry and heart rate monitoring improves estimate of directly measured physical activity energy expenditure. *J Appl Physiol.* 2004;**96**:343–51.

[35] Corder K, van Sluijs EM, Wright A, Whincup P, Wareham NJ, Ekelund U. Is it possible to assess free-living physical activity and energy expenditure in young people by self-report? *Am J Clin Nutr.* 2009 Jan 14. [Epub ahead of print]

[36] Kowalski K, Crocker P, Faulkner R. Validation of the physical activity questionnaire for older children. *Pediatr Exerc Sci.* 1997;**p**:174-86.

[37] Ahamed Y, Macdonald H, Reed K, Naylor PJ, Liu-Ambrose T, McKay H. School-based physical activity does not compromise children's academic performance. *Med Sci Sports Exerc.* 2007;**39**:371-6.

[38] www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_063812

[39] Physical activity questionnaire of the European Prospective Investigation Into Cancer www.srl.cam.ac.uk/epic/questionnaires/epaq2/epaq2.pdf

[40] Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE. Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire. *Int J Epidemiol.* 2002;**31**:168-174.

[41] Craig CL, Marshall AL, Sjöström M, et al. international physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Ece.* 2003;**35**:1381-95.

[42] International Physical Activity Questionnaire www.ipaq.ki.se/

[43] Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE. International physical activity questionnaire: 12-country reliability and validity *Med Sci Sports Exercise.* 2003;**35:**1381 – 95.

[44] Besson H et al. – *manuscript in preparation*

Appendix A Development and cognitive testing of the new questionnaires

A.1.1 Overview

Full details of the development procedure and pre testing of the new physical activity questionnaires are given in the cognitive report of findings provided to the FSA in August 2006.[i] This section therefore gives a brief overview of the changes to the questionnaire, the questionnaire sources and rationale for inclusion.

A.1.2 Adults

The NDNS questionnaire was developed from scratch but it used elements from other questionnaires.[i]

The table below summarises the main features of the NDNS physical activity questionnaire for adults.

Questionnaire item	Rationale	Source
Occupational activity questions	These questions were adapted from the English version of the long IPAQ (International physical activity questionnaire).	www.ipaq.ki.se/ipaq.htm
Housework and Gardening	Adapted from the Health Survey for England (HSE) questionnaire	
Walking questions	Adapted from the HSE questionnaire, but included bouts lasting 5 minutes or more	
Sports and organised activities - duration	Adapted from the HSE questionnaire, but included bouts lasting 5 minutes or more	
Sedentary activity questions	*No such questions included in the NDNS questionnaire*	

A.1.3 Children

The table below documents the broad question areas developed for the new children's physical activity module, briefly describes the rationale for inclusion and, where appropriate, details original source material.

[i] MacKenzie H, Collins D, Kitchen S. *National Diet and Nutrition Survey. Development of physical activity and sun exposure questions: findings from cognitive interviews.* London: NatCen, 2006.

Questionnaire item	Rationale	Source
School time-related activities	*School breaks and transportation to school offer important opportunities for children to be active.*	
Travel to and from school	A new set of questions were developed to look at active transportation to and from school to capture this data. This has policy importance in that it captures both domain information (activity for transportation) and is a useful recall technique by breaking down potential periods of activity by linking it specifically to a particular purpose or time period.	New questions developed as we did not locate any questionnaires for children that were appropriate for NDNS
Active during school breaks	The rationale for this is the same as for transportational activity in that it offers important opportunities for children to be active and this information has not been captured by any existing questionnaires.	New questions as we did not locate any questionnaires for children that were appropriate for NDNS
Active play	It is widely recognised that play activities contribute to overall activity levels.	New questions as we did not locate any questionnaires for children that were appropriate for NDNS
Non-school based, formal, sports activities	Again, the distinction between this grouping of activities and active play activities has been widely recognised to be important and has been made on most child/adolescent physical activity questionnaires.	Modelled after adults' HSE sports section but tailored for children's activities
Sedentary activity questions	*No such questions for NDNS*	