THE VALIDITY OF HEART RATE AND ACCELEROMETRY IN THE ASSESSMENT OF PHYSICAL ACTIVITY IN PRE-SCHOOL CHILDREN

By

A. HIRJI, B.P.E.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (1995) McMASTER UNIVERSITY
(Human Biodynamics) Hamilton, Ontario

TITLE: The Validity of Heart Rate and Accelerometry in the Assessment of Physical Activity in Pre-School Children

AUTHOR: Arif Hirji, B.P.E. (McMaster University)

SUPERVISOR: Dr. C.J.R. Blimkie

ADVISORY COMMITTEE: Dr. Collin Webber
Dr. Oded Bar-Or
Dr. Storm Russell

NUMBER OF PAGES: viii, 145

Abstract

The purpose of this study was to validate heart rate recording (Polar Vantage XL), and accelerometry (Tritrac-R3D) against a direct observational technique (BEACHES) in the assessment of physical activity in 49 healthy male and female pre-schoolers (3.0 to 5.9 yrs). Activity was assessed in a controlled laboratory environment and in 2 field trials during spontaneous uncontrolled activity. During the laboratory session, subjects wore a heart rate monitor and an accelerometer while performing a choreographed routine of activities. The actual level was coded concurrently on the BEACHES scale by a trained observer. The three measures were significantly correlated during the laboratory (r = 0.76 - 0.85) and field (r = 0.50 - 0.85) conditions. There was no difference in mean scores for activity level of the group between two different days of field observation, however, there was considerable intra-subject variability as reflected by the low to moderate rank order correlations for all techniques. It was found that a measurement interval of greater than 5 min would likely increase the correlations. These findings suggest that the three measures of PA provide similar information about the level of PA in both a controlled and the free play environment, and pre-school children as a group generally maintain a similar level of activity from day to day.

Acknowledgements

I would like to thank Tali Bar-Or and Heather Waters for their time and effort with the data collection. I would like to thank Dr. Elliott for his help with the statistics and the Canadian Fitness and Lifestyle Research Institute for funding this project, and finally I would like to thank my parents for making all this possible.

Table of Contents

List of Tables and Figures

Introduction/literature review

Definition of Physical Activity (PA)

Physical activity (PA) has been operationally defined by Casperson et al. (1985) as "any bodily movement produced by skeletal muscles that results in energy expenditure" (Casperson et al., 1985). Under this broad definition, PA can take the form of active physical leisure, exercise, sport, occupational work and chores, together with any other factors which might influence total daily energy expenditure (EE). From the fitness and health perspective, all determinants of human EE warrant careful and precise investigation.

PA has behavioural, physical and physiological components. PA can be viewed within any of several behavioural contexts, e.g., recreational (Yoesting and Christensen, 1978), rehabilitative (Young, 1990), developmental (Weiss and Bredemeier, 1990), or health (U.S. Department of Health and Human Resources, 1991). Baranowski et al., (1992) have defined this aspect of PA as the volitional movement of the body through space, or exertion of intended force, e.g., isometric exercise (Baranowski et al., 1992). In the measurement of PA, the behaviourist would include the type of activity, the physical location where the activity is being performed (home, school, playground), as well as the interaction with peers, family or teachers as initiators of this behaviour. Even though quantitative analyses of these behavioural factors of PA are not always feasible, these components help us understand the motivation of children toward or away from PA.

The physical measurement of PA involves devices that directly measure motion. The advantage of these motion sensors is that they can measure the amount and intensity of body movement, or a part thereof.

From a physiological perspective, PA is a component of total EE, which also includes resting metabolism, the thermic effects of foods, and growth (Powell et al., 1987). PA can vary in intensity, duration, patterns of intensity by duration, frequency and muscle groups used.

Activity Patterns of Pre-school Children

Throughout childhood and adolescence, girls are less active than boys (Klesges et al., 1986; Kucera, 1985; Saris, 1986). The difference appears to be more one of intensity and type of activity, while differences in overall amount of physical activity are far less evident. It is tempting to assume that the discrepancy in PA is due both to bodily changes (primarily increased adiposity) and to a shift in social priorities. It is an interesting problem to discern at what age gender-related differences in PA first appear. Eaton and Enns (1986) in a study using direct behavioural observation suggest that there is no difference in the activity pattern of girl and boy infants. Noland et al. (1990) through direct observation video analyses of free-play found no gender-related differences in PA of 1-to 4-year-old children. However, when Danner et al. (1991) using the same methods, followed the same sample of subjects a year later, they found a decrease in the girls' activity, but not in the boys'. It is still unclear whether any differences in activity exist between genders at the pre-school (3.0 to 5.9 yr) age.

It has long been recognized that the level of PA among adults declines with age (e.g., Blair et al., 1985). As shown by cross-sectional and longitudinal studies, PA and EE per kg body weight diminish progressively during the first two decades of life, particularly during adolescence (e.g., Russell et al., 1992; Saris, 1986). Hay (1990) has suggested that, as children become more self-sufficient in their leisure time choices, the prominence of physical activity becomes increasingly devalued in their mind. This is due to perceived self inefficiencies developed in early childhood, apparent even before nine years of age. It is not yet clear whether such a decline in PA and EE with age is apparent among pre-schoolers.

As reported by Russell (1992), the most frequently reported spontaneous outdoor activities found during a survey of 433 licensed daycare-centres included: climbing, running races, jumping, cycling, swinging and sliding. Even though the intensity of these activities was not reported in the survey, it seems that free, unstructured activities were higher in intensity than the structured ones. Children also however, appear to spend a lot of time in sedentary activities such as TV watching. Canadian children aged 2-11 years were found in one study to watch TV between 15 and 26 hours a week (Dietz and Gortmaker, 1985). This high degree of TV watching is particularly disheartening in light of the strong association between risk of juvenile obesity and the number of hours of watching TV (Dietz and Gortmaker, 1985). This relationship though, has only been studied in children as young as six years, and has not been investigated in the pre-school population.

As a group, children with a chronic illness or physical disability are less active than their healthy counterparts (Bar-Or, 1983; Rowland, 1990). A child with a physical disability or an illness is likely to be hypoactive, either because the disability itself imposes a limitation, or because of causes that are not directly related to the child's disability such as overprotection, fear ignorance and social isolation (Bar-Or, 1983).

Understanding the possible influence of a family's lifestyle on a child's PA is important for the development of intervention programmes and public health policies. Much research has been conducted on the family aggregation of health-related lifestyle habits such as eating and smoking (Sallis and Nader 1988). Less is known about the effect of the family on a child's PA. When both parents are physically active, young children's activity is considerably higher than when the parents are sedentary (Freedson and Evenson, 1991; Moore et al., 1991).

Some studies show a relationship between parents' and their children's participation in PA, others do not. From the studies that show a relationship, it seems that parents influence mostly the participation of children in organized activities, whereas same-gender peers influence unorganized activity participation (Moore et al., 1991). Even though data on Canadians suggest a near 30% heritability in the general PA pattern of children (Perússe et al., 1989), there seems to be little or no genetic effect on participation in high-intensity activities or in specific sports. Little effort has been made to examine the influence of families on children's perceptions of PA as a health behaviour.

Climatic and seasonal variations may be an important consideration in Canada due both to the range of climatic conditions within the country and the abrupt inter-season climatic variations. The degree of participation in exercise and sport changes noticeably with the season: differences being most evident between winter and summer (Ross and Pate 1987). The effect of limited daylight hours during the winter merits special attention as this is an additional limitation on children's ability to play outdoors. The greater activity level during the summer probably reflects the finding that children are more active when outdoors (Klesges et al., 1990a; Mimura et al., 1991).

Rationale for the assessment of PA in children

Increasing evidence of the health benefits of PA in adults (Paffenbarger et al., 1986; Powell et al., 1987) has stimulated interest among researchers in studying PA in children and adolescents (Simons-Morton and Huang, 1991). Within the health perspective, LaPorte et al. (1985) have suggested that the validity and reliability of PA measurement to determine a) the total amount of elevated EE (above resting EE) or the quantity of PA, and b) the intensity of activity or the quality of the movement are the two most important issues to be addressed in this context. In essence, LaPorte et al. (1985) emphasize that ideally the measurement methods should assess both the quantity and quality of PA.

There are several possible links between childhood levels of PA and health in adulthood. PA during childhood may determine the child's physical fitness which in turn, may affect adult health through adult PA and fitness, or indirectly through its effects on the child's health (Blair et al., 1989; Pate et al., 1990). Interest in the relationship between PA and health has resulted in three major lines of investigation in the pediatric population.

From a health perspective, perhaps the most pursued interest has been the relationship between PA and obesity during childhood. The prevalence of childhood obesity is on the rise in the US (Gortmaker et al., 1987) and is currently reported to range from 10% to 25%, thus making obesity the most common pediatric chronic illness in North America. Given the current and projected future high levels of childhood obesity, it becomes even more important to develop standardized methods for assessing PA and EE which will shed more light on the interrelationships between PA, EE and obesity. More research is needed to further determine the role of hypoactivity compared with other (genetic and environmental) factors in the etiology of juvenile obesity.

The role of PA in obesity stems from the understanding that PA is a major component of total EE. Several studies have reported that obese children are less active than nonobese children (Epstein et al., 1984a; Griffiths and Payne, 1976; Klesges et al., 1984; Klesges et al., 1986), but this relationship is not found consistently (Waxman and Stunkard, 1980). One reason for the inconsistency in findings about the relationship between obesity and hypoactivity is the way in which EE has been reported.

The second major line of research is the relationship between PA in children and cardiovascular disease risk factors. Numerous studies have documented the high prevalence of elevated blood pressure and serum cholesterol in children (Sallis et al., 1988; Simons-Morton and Huang, 1991), and these risk factors predict the development of atherosclerotic lesions in adolescents and young adults (Newman et al., 1986). Physically active and fit children have been shown to have lower levels of blood pressure and more favourable lipoprotein profiles (Montoye, 1985; Sallis et al., 1988; Tell and

Vellar, 1988), so PA may be an important behavioural target for cardiovascular disease primary intervention programmes. Many of the risk factors for coronary heart disease (CHD) appear to begin in childhood and youth (Williams et al., 1981) and as many as 60% of children in the United States exhibit at least one modifiable adult risk factor for CHD by the age of 12 (Berensen et al., 1980).

Although the clinical endpoints of diseases such as CHD and obesity typically occur in adulthood, there is mounting evidence, particularly for CHD, that these diseases have antecedents already during childhood and adolescence (Berensen et al., 1980; Wynder, 1989). Similarly, there is a fairly strong relationship between adult obesity and adolescent obesity (Charney et al., 1976). It appears that these diseases are legitimate pediatric concerns and there is a definite need to better understand their relationship with PA during the childhood years.

The third line of investigation concerns intervention trials to assess the therapeutic effects of PA in the treatment and management of childhood obesity. A number of trials have indicated that PA makes important contributions to the management of obesity (Freedson, 1989; Klesges et al., 1990a).

Identifying antecedents and consequences to PA behaviours will facilitate understanding of the development of aspects of health behaviours in young children. This information may assist behaviour analysts in developing health promotion interventions that are both appropriate and effective for young children. There is a need to understand the factors that affect the type, frequency, duration and intensity of children's PA behaviour because childhood PA may have important health consequences. Research and policy interest in children's PA is driven by the awareness that (a) diseases such as obesity and CHD for which sedentary behaviour is likely a risk factor are lifelong processes with origins during childhood and (b) PA habits established early in life may persist into adult years.

Current understanding of the health effects of PA suggests that, while active children experience some improvements in mental and physical health, the major public health impact of PA is likely to be the reduction of chronic disease risk in adults (Sallis et al., 1992). Better understanding of the determinants of physical activity can guide the development of effective activity promotion, and such efforts hold promise for improving both the present health status of children and their prognosis for health as adults.

While there are some established links between PA in late childhood and adult health, little is known about the relationship between PA during early stages of development, e.g., pre-school years, and health in later childhood or adulthood. This relative paucity of information is due in large part to technical limitations in the assessment of PA in young children.

PA assessment techniques and limitations

Most major contemporary chronic diseases and health problems (e.g., CHD, obesity, and hypertension) appear to be associated with negative lifestyle behaviours, including low levels of PA (Baranowski et al., 1992). Because of these associations, there has been wide spread interest during the last few decades in the development of improved methods for assessing habitual PA and EE in humans under free-living or field conditions. PA measurement techniques have been used in various settings to a) describe PA habits in populations (Montoye and Taylor, 1984), b) classify PA levels for intervention efforts (Sallis et al., 1992), c) assess PA changes over time (Russell et al., 1992; Saris, 1986), and d) identify behavioural correlates of PA (Williams et al., 1989).

The greatest obstacle to validating field methods of assessing habitual PA or total EE has been the lack of adequate criteria to assess accuracy of these techniques. In a recent workshop, convened by the US National Institute of Health, experts identified research topics pertaining to the determinants of PA and the development of strategic interventions in youth, that should be given a priority for support (Sallis et al., 1992). The highest priority was given to the establishment of "gold standards" for the assessment of PA and EE. The group further suggested that a direct observation of the child over the waking hours would emerge as the "gold standard" for monitoring PA. Likewise, the doubly-labeled water technique will probably emerge as the "gold standard" for measuring EE. Because of extremely high costs however, neither of these methods can be used within large-scale surveys. One should therefore seek simpler methods, assess their validity against an acceptable "gold standard," and determine their suitability for large-scale surveys.

There has been extensive research in the past decade on different methods for measuring habitual activity. These investigations have led to four basic classifications of measures of habitual activity:

a. Physiological measures; including:
   - Doubly Labeled Water ($^2H_2{}^{18}O$);
   - Calorimetric Chamber; and
   - Heart Rate Monitors;

b. Physical measures; including:
   - Pedometers and Actometers;
   - Motion Sensors; and
   - Accelerometers;

c. Observational measures; including:
   - Time and Motion analyses; and
   - Behavioural Observation; and

d. Pen and Paper measures; including:
   - Self Report Questionnaires;
   - Interviews;

- Caloric Intake; and
- Personal Log Records.

Table 1 summarizes these measurement classifications and indicates whether the measurement takes place concurrent with or after the activity period. Of the physiological measures, only HR monitoring has the capacity to concurrently estimate EE as well as describe the changes of activity intensities over time. Of the physical measures, only the accelerometer has the same capacity and of the Observational measures, both time and motion analyses and behavioural observation have this capacity. Table 2 summarizes some of the strengths and weaknesses of these measures of habitual activity.

The metabolic rate of a person can be measured directly or indirectly. The direct method is simple to understand but difficult to perform; the subject is placed in a calorimeter, an instrument large enough to accommodate a person, and heat production is measured by the temperature changes in the calorimeter (Figure 1). This is a very accurate method in that it measures heat production directly, but calorimeters are found only in a few research laboratories; accordingly, a simple, indirect method has been developed for widespread use.

Emons et al., (1992) used a respiration chamber, a method of indirect calorimetry to estimate EE over a 24 hour period. This calorimeter measures the volume of the outgoing gas by means of a dry gas meter and gas analysis can be made with an O_2-analyzer. EE can be calculated from measurement of O_2

and CO_2 changes within the chamber. Using the indirect procedure, one measures the subject's oxygen uptake per unit time (by measuring total ventilation and the oxygen content of both inspired and expired air). From this value, one calculates heat production based on the fundamental principle (Figure 2) that the energy liberated by the catabolism of foods in the body must be the same as that liberated when they are catabolized to the same products outside the body. We know precisely how much heat is liberated when 1 L of oxygen is consumed in the oxidation of fat, protein, or carbohydrate outside the body; this same quantity of heat must be produced when 1 L of oxygen is consumed in the body.

Much of the recent research examining the different methods for assessing PA has focused on the use of direct observation, heart rate (HR) monitoring and accelerometry by Caltrac. The observational method describes and quantifies the components of activity behaviour, whereas HR monitoring and accelerometry yield information on EE. HR reflects physiological strain of the cardiovascular system and can be used with gas analyses to estimate EE. The Caltrac measures acceleration of the whole body or body parts (in arbitrary counts) and can also be used to estimate EE; however, the EE estimate equations in the Caltrac are probably not valid for use with children (Bray, 1992) since the equation to estimate basal metabolic rate is based on adult data. Given the labor intensiveness and high costs associated with the direct observational technique it seems likely that any future large scale survey of PA patterns in children will be forced to rely more on alternative but equally valid techniques. HR monitoring and accelerometry, or some combination of the two may provide such an alternative, should they prove to be equally reliable and valid techniques compared to the direct observation approach.

The previous techniques have been used extensively in recent years to describe activity patterns, and less frequently in cross-validation studies within specific populations (ie. healthy children and children with chronic diseases) of school-aged and older children. These techniques have rarely been applied to pre-school children (3.0 to 5.9 years of age) and the reliability and validity of these techniques to describe patterns of activity in this segment of the population remains to be established.

Heart Rate monitoring in the assessment of PA

For exercise scientists, there is no doubt that the HR is a useful indicator of physiological adaptation and intensity of effort. From the Fick equation ($VO_2 = SV \times HR \times a\text{-}v\ O_2$ diff) it is evident that HR is one of the three parameters that can vary in relation to increasing oxygen uptake. The other two variables are stroke volume (SV) and arteriovenous (a-v) oxygen difference. Ideally, cardiac output, since it is a product of the heart rate and SV, should predict EE much more precisely than HR. Unfortunately, no field method to measure cardiac output is available at present (Saris, 1985). The systematic monitoring of HR in the natural environment at work or in play is a relatively recent phenomenon which has arisen from the appearance of new technology and to the greater miniaturization of measuring devices.

Because of technical difficulties encountered in measuring EE by indirect calorimetry in the field, there is interest in finding simpler, less direct methods of recording physiologic data associated with EE. EE is typically estimated from the heart rate/oxygen uptake (HR-VO_2) curve. To use HR as an indicator of EE in this manner, it is usually necessary to bring the subject into the laboratory to establish the individual's HR-VO_2 calibration curve, and then apply field measurements of HR to the curve to derive estimates of VO_2 or EE (Bradfield, 1983; Gretebeck and Montoye, 1992). This procedure is problematic however, in pre-school age children because of their general intolerance for the apparatus required for VO_2 measurements.

Ballor et al. (1989), using individual HR-VO_2 calibration curves of 20 high school students, reported high correlations between HR and EE measured in the lab through gas exchange analysis. In two other studies involving adults where individual HR-VO_2 calibration curves were used, HR estimates of EE in free-living conditions compared favourably with EE determined by doubly labeled water (DLW) measurements (Livingstone et al., 1990; Shulz at al., 1989). Livingstone et al., (1990) found that the daily EE derived from HR in adults was only 2% higher on average than that determined by the DLW technique. The range however varied from -22.2% to +52.1% (SD = 17.9%) among subjects. Shulz et al., (1989) reported correlation coefficients of 0.53 to 0.73 between the DLW and HR estimates of EE in six adults. Emons et al., (1992) compared HR monitoring with an indirect calorimeter (respiration chamber) and the doubly labeled water method for the measurement of 24 hour EE in children between 7 and 11 years of age. This study found that the HR method overestimated 24 hour EE by 10.4% and 12.3% respectively. This study concluded, that in comparison with the inconvenience of the calorimeter and the expense of the doubly labeled water technique, HR monitoring seemed to be the most suitable method for measuring EE in large scale activity studies in children. These studies indicate that although EE can be reliably predicted from HR there seems to be a fairly large difference in EE estimates between HR and other techniques.

In recent years, manufacturers of sport and exercise equipment have attempted to solve the problems of inconvenience of HR monitoring by developing and marketing newer and easier to use instruments. The accuracy of these instruments has, in many cases, been unacceptable or undetermined. Karvonen et al. (1984), compared HR of fourteen adults obtained by means of an electrocardiogram (ECG) to heart rates taken from the memory function of a commercially available radio telemetry HR monitor during exercise and found that the first generation HR monitors overestimated ECG recordings.

Due to technological advances over the last 10 years, HR monitoring has been used ever more commonly in the assessment of PA among children (Freedson, 1989; LaPorte et al., 1985; Saris, 1985; Treiber et al., 1989); there is however, little information on pre-schoolers (Kucera, 1985; Mimura, 1989; Mimura et al., 1991; Mimura et al., 1987). HR can be recorded in the memory of a microcomputer, which can be carried on the wrist as a watch. The device has a transmitter and a receiver. The transmitter is worn on an electrode belt around the chest and transmits ECG signals caused by the heart beat into the receiver and into the memory of the microcomputer, where they can be read later. Aspects of the telemeter and continuous ECG recordings have been combined in the microcomputer recording of HR. A transmitter connected to electrodes has a wireless contact to the memory of the microcomputer, from which stored data can be analyzed later. Instead of a printed recording the HR is stored in the microcomputer memory. The results of the HR obtained from information fed to the microcomputer and that obtained from ECG recordings for the same time period are highly correlated under both rest and work conditions in adults (Karvonen et al., 1984) and in pre-school children (Bar-Or et al., 1995). Additionally, recently developed HR monitors have been found to be highly valid tools for measuring HR in children as young as four years (Treiber et al., 1989; Tsanakas et al., 1986), both under laboratory conditions and in the field, when compared with direct electrocardiographic (ECG) monitoring.

An important issue in the use of HR monitors is the reliability and accuracy of readings, especially when the instrument is used outside the laboratory. Burke and Whelan (1987) evaluated four different kinds of HR monitors by running two series of tests on each. None of the monitors evaluated were considered accurate or reliable. Movement artifacts appeared to be the biggest problem with the systems, and this limitation needs to be addressed before accuracy and reliability can be improved. Wajciechowski et al. (1991), evaluated the accuracy of the commercially available Polar Electro radio telemetry HR monitor in a group of ten adult females during popular exercises such as walking, jogging, and selected low-impact aerobics regimens by comparing it to standard ECG. The Polar Electro was found (Wajciechowski et al., 1991) to be reasonably accurate with errors within ±8 beats per minute 90% of the time, and the researcher concluded that the Polar Electro, while not as accurate as continuous field assessment of oxygen uptake, was superior to activity logs, pedometers, and questionnaires. Leger and Thivierge (1988) and Thivierge and Leger (1989) in a study involving eight adult men and women found that the AMF Quantum XL (Sport Tester PE 3000) was the most accurate, reliable and valid of 13 commercially available portable HR monitors under laboratory conditions.

A newer generation of the Sport Tester PE 3000 HR monitor, the Polar Vantage 4000 XL, has recently been validated with pre-school children under laboratory conditions of cycle ergometer testing (Bar-Or et al., 1995; Strand and Reeder, 1993; Treiber, 1989; Tsanakas et al., 1986) and has been shown to be even more accurate than its predecessor. Further research with this new and more accurate HR monitor should provide us with a better representation of physiological data from heart rates.

HR monitoring is attractive because it directly measures a physiological parameter known to be related to PA, and because it provides a continuous record that may reflect both intensity and duration of activity. However, it should be noted that several factors other that PA can influence HR. Activation of different muscle groups can cause HR to respond differently (Montoye and Taylor, 1984). For example, the energy cost of leg exercise is greater than arm exercise, but the HR response is greater for arm work. The type of muscular contraction, e.g., isotonic vs. isometric contractions, training status, ambient temperature, dehydration, fatigue, and psychological stress can all influence the HR-EE relationship (Montoye and Taylor, 1984).

Accelerometry in the assessment of PA

Almost all forms of PA require movement of the trunk or extremities. Therefore it is not surprising that many devices have been developed and used to obtain objective measures of body movement over a certain period of time, as an approximation of PA level. The assessment of PA in children by motion sensors mainly has employed the use of two commercially available devices. The first of these and the one used mostly with adults is the Large Scale Integrated Movement Counter (LSI). The LSI, made by GMM Electronics, Inc., (Verona, Pennsylvania) is small (3.8 × 4.5 × 2.2 cm), lightweight (52 g) and is generally worn on the hip. When a motion occurs that results in a minimum of 3 percent inclination or declination of the LSI, a mercury ball activates a microswitch to register a count. The result of the counter can be read by holding a magnet to the side of the instrument, thereby activating a light-emitting diode (LED) display indicating the number of movements. Only the total quantity of movement is measured and no discrimination is made as to the intensity of the movement.

The LSI has been validated in children as young as 1.6 years during a one hour free play observation period (Klesges et al., 1985). Pearson correlations between LSI counts and observed behaviours were low, and at best only 19% of the variance could be explained by the LSI. Klesges et al., (1985) suggested that one reason for the poor validity may have been the inability of the LSI to assess sudden and short bursts of activity. This motion sensor was therefore able to quantify the total movements, but was not able to assess the quality of the movement.

A single plane portable accelerometer, the Caltrac accelerometer (Hemokinetics: Madison, WI) was developed recently to evaluate both the quantity and quality (intensity) of movement. By 1991, the Caltrac had replaced the LSI in PA research. It is relatively inexpensive ($65) and easy to use. The Caltrac, which measures 14×8×4 cm and weighs 400 grams, is worn on the nondominant hip. A change in vertical acceleration causes a ceramic cantilevered beam to twist. Acting like a piezoelectrode, it emits a current proportional to the force acting on it. A small computer in the unit then plots an acceleration curve and integrates the area under that curve. The Caltrac "counts" vertical (vertical to the horizontal axis of the Caltrac) acceleration, and based on an estimation of basal metabolic rate from height, weight, gender and age plus the estimated caloric expenditure from activities, displays a result of total kilocalories expended on a liquid crystal display. Since the equation to estimate basal metabolic rate is based on adult data, the Caltrac, when used with children, is programmed using constants to assess the activity session without the basal metabolic rate component. This results in an output in counts, instead of kilocalories.

In young adults, Montoye et al., (1983) found a correlation of $r = 0.74$ between Caltrac counts and VO_2 with a high test-retest reliability for accelerometer readings of $r = 0.94$. With children in the laboratory, correlation coefficients of $r = 0.8$ to 0.9 were found between the Caltrac and different intensities of walking/running on the treadmill (Kastango and Freedson, 1991; Sallis et al., 1989). Klesges et al., have reported two investigations of the accelerometer with pre-school children. In the first study, adults were observed for one hour of unstructured recreational activity with the FATS observational system while they simultaneously wore the Caltrac accelerometer and the LSI. The second study, evaluated these same two sensors in pre-school children for one hour in a daycare centre during an unstructured free-play situation while their PA was observed and systematically coded also with the FATS observational system (Klesges et al., 1985). Validity of the Caltrac was high with adults ($r = 0.69$) and lower but still significant with the pre-school children ($r = 0.35$). In another study by Klesges and Klesges (1987), 28 pre-school children wore the Caltrac all day (nine hours) while their PA was systematically observed and coded using the FATS observational system (Klesges and Klesges, 1987). The validity coefficient ($r = 0.54$) was somewhat higher than in the previous study, and the correlation was particularly high in the 2.7 to 4-year-old children ($r = 0.80$). These studies indicate that the accelerometer is at least moderately valid in both controlled and all-day applications with pre-school children; the length of monitoring, and the setting, however, may influence the level of validity.

The Caltrac electronic single-plane accelerometer has been found to be useful for estimating habitual PA in adults (Montoye et al., 1983) and children

(Simons-Morton and Huang, 1991) in moderate to vigorous activity. The output of the accelerometer worn at the waist was almost identical to the vertical component of a force platform during bench stepping and knee bends; however, agreement was poor in repeated floor touching tasks (Montoye et al., 1986). Montoye et al., (1986) hypothesized that this was probably due to the fact that the accelerometer in the second task was no longer maintained in a vertical position.

When the Caltrac was validated against EE measured by whole body calorimetry over a 24 hr period in a study by Schutz et al., (1988), a correlation coefficient of $r = 0.92$ was observed. Sallis et al., (1989) recently reported r values of 0.54 and 0.42 between separate days of Caltrac recordings and monitored activity HR in children.

The accelerometer's ability to evaluate the quality (intensity of movement) in addition to the quantity of a movement makes it a useful instrument in the assessment of PA. Motion results in acceleration of the limbs and body, which increases EE. However, even common activities such as walking, bending, and climbing stairs are quite complex and all their associated EE is not reflected in acceleration or deceleration of the mass of the body. Further, during activities such as cycling and rowing, when an accelerometer is attached to the waist, little movement is recorded although the EE may be high.

Activity monitors must be capable of differentiating wide ranges of PA from sedentary to high levels of participation, which characterize the natural patterns of children. Additionally, in patient populations the monitor must be sensitive to increased EE resulting from intervention programs that add considerable PA to an otherwise sedentary lifestyle. The Caltrac accelerometer, although it may lack sensitivity and specificity for subtle movement, provides adequate differentiation of PA for most general uses for groups in the pediatric exercise field.

Meijer et al., (1991) designed an accelerometer based on a piezoelectrical ceramic element placed in a pouch and worn on the back in the L4-L5 region. This accelerometer was compared over seven days with the PA index obtained from DLW and RMR measurements in 21 adults (Meijer et al., 1990). The accelerometer designed by Meijer et al., was reported to be sensitive in three directions, the vertical being the most sensitive. Results from this newly designed instrument looked very encouraging, although technical problems were still evident in the design. Hemokinetics Inc. has recently (May 1993), introduced a new accelerometer, the Tritrac-R3D which is their new generation model to the Caltrac accelerometer. This new accelerometer is capable of measuring acceleration and deceleration in all three ordinal planes and calculating the vector magnitude of the three. The data when transferred to a Personal Computer PC can be printed out and analyzed on a minute by minute basis with the individual X, Y, Z plane counts, or the calculated vector magnitude, making temporal comparisons with HR and direct observation possible. Theoretically this new device can be used readily in population-based studies which require continuous recording of data so that a minute-by-minute profile of activity can be obtained.

While accelerometers have been and are currently used to assess PA, they are not without their limitations. Haskell et al., (1993) have recently identified the major limitations of accelerometers. These include a) their inability to accurately identify increases in EE due to movements up inclines (hills or stairs) or increases in EE due to an increase in resistance to movement (increase in the amount of weight lifted or resistance on a cycle ergometer), b) the inability of single sensors to identify movement that involves various parts of the body (a sensor on the leg will not detect movements of the arms), and c) the inability to detect static exercise.

Direct observation in the assessment of PA

The measurement of EE by VO_2 in field settings is impractical. Consequently, information on EE of young children in the natural setting is limited. The assessment of PA by observation has several advantages over other field techniques. Observation can provide information on the specific type of activity that occurs, results are not limited by recall ability or self reporting biases of individuals, and the technique is particularly useful with young children, from whom accurate and detailed self-report data are more difficult to obtain (Baranowski et al., 1992).

Direct observational instruments have been developed for both children and adults. Observation is usually made by an investigator, who "shadows" the child and systematically records information about the child's behaviour on special forms. Behavioural observation is important for assessing the content of PA, because of its ability to document precisely what the child is doing and for how long. The most important potential value of direct observation is as a "gold standard" for the validation of the other measurements.

The observation of people in their natural setting has been well documented (Hartmann, 1982). Systematic observation is a valid method for the assessment of PA in the natural environment because it requires little inference. Direct observational methods are flexible and allow for the quantification of diverse dimensions related to children's PA. Children engage in a wide variety of unique activity which results partly from a response to many environmental factors, such as presence of others in the environment, physical location, promotion of activity from others, available facilities, availability of toys and equipment, and whether the setting provides for free play or structured activity. Direct systematic observation allows for the study of these environmental factors in conjunction with the assessment of children's PA with limited reactivity or response burden from participants.

The systematic observation of children's PA is an important method in health-related research and curriculum development. Direct observation provides an accurate and comprehensive tool for the assessment of PA that allows the characterization of habitual PA directly or after review of film or videotape. Survey techniques are limited in terms of recall or self-reporting biases by individuals and it is here that direct observation has an advantage over these techniques. Direct observation does not require any equipment that might hinder the movement of subjects, it can be used in a variety of settings, and can be used for either short or long periods of time.

For all its merits, direct observation is not a practical approach for large-scale epidemiological research because it can be "reactive," time-consuming, costly, intrusive, and require extreme diligence on the part of the observer (Ainsworth et al., 1994). Reactivity to observation refers to the possible alterations of behaviour due to the presence of an observer. All measurement techniques, however, are likely to be reactive to some extent. Because observation requires both time and diligence on the part of the observer, its usefulness is limited to small-scale studies (LaPorte, 1985) and for validating survey questionnaires (Klesges et al., 1984; Klesges et al., 1990b; O'Hara et al., 1989; Puhl et al., 1990).

With this technique, events studied must be observable and codeable, and records are thereby limited to events that are seen. Observers or video-recorders need to be in the environment where behaviours of interest occur, thereby limiting situations where data can be collected. Observers must be properly trained to reduce subject reactivity caused by observers being present in the observational environment. Observers must be properly trained to be rigidly objective, and not to be evaluative or judgemental about what they see (McKenzie, 1991). Substantial time is required to train observers to be consistent in their coding methods to ensure high inter-observer reliability. Associated with validity is the concept of inter-observer reliability, or agreement on observed events by independent observers. A primary constraint to the reliability of any instrument is the extent of the inter-observer agreement. When too many activity categories exist and distinctions between them are unclear, inter-observer agreement is reduced.

Systematic observation is a direct measure of behaviour and PA that requires little subjective inference or interpretation and it does not depend on child or parental memory or interpretation. This makes it an often used criterion measure, or "gold standard" for validating other measures of PA, such as mechanical and electronic monitoring.

Many health-related researchers are interested in energy expenditure and most of the observational instruments have been validated against measures of energy expenditure or heart rates. This has resulted in a paradox whereby physiological, physical, and behavioural methods of assessing PA have been validated against each other. While PA can be assessed directly through observational methodology, EE can only be inferred. For this reason the observational method is often used as the "gold standard" measure of PA against which the other methods are validated.

Four observational instruments have been designed to observe pre-school children in diverse environments (Klesges et al., 1984; Klesges et al., 1990a; McKenzie et al., 1991; Puhl et al., 1990), and all permit simultaneous coding of several events associated with PA and are useful for comparing PA levels in different locations. Table 3 compares these four instruments. All four have been carefully constructed, and field tested for children. Reported reliabilities are high and the four instruments are generalizable for use in studying PA as it relates to health. The Children's Activity Rating Scale (CARS) reported by Puhl et al., (1990) and the Behaviour of Eating and Activity for Children's Health Evaluation System (BEACHES) reported by McKenzie et al., (1991) are the only two that have had their activity levels validated against HR or estimates of children's caloric expenditure.

The CARS is a five-level children's activity rating scale designed to categorize the intensity of PA and discriminate between levels of EE in young children. In a study by Puhl et al., (1990) the CARS was used by trained observers over a 12-month period to assess PA of 3-4 year-old children during field observation. The EE for each level was assessed by measuring VO_2 and HR of 5-6 year-old children while they performed eight activities representing the five CARS levels (Puhl et al., 1990). VO_2 and HR at each level were significantly different from each other. The data demonstrated that the CARS encompassed a wide range of EEs, discriminated between levels of EE, and could be used by trained observers to reliably evaluate PA and estimate EE of young children (Puhl et al., 1990).

The BEACHES is a comprehensive and complex observational system designed to simultaneously code children's PA and eating behaviours and related environmental events. It can be used at home, at school, and in most other settings in which a targeted child might be found. The system was developed within the framework of behaviour analyses and includes coding for 10 separate dimensions (McKenzie et al., 1991). This observational system requires at least 50 hours of training and practice for the observers. Even though only 19 children took part in the study by McKenzie et al., (1991) and no significant differences between several levels of the five observation classes were observed, the BEACHES demonstrated high inter-observer reliability and reasonable results with HR monitoring.

Combined methods in the assessment of PA

Researchers have begun to evaluate the degree to which the various techniques assess common factors or characteristics of PA. The LSI has been compared with various survey techniques in several populations; however, correlations with recall procedures were poor. Table 4 summarizes studies to date that have employed a combination of methods in the assessment of PA in children. As is evident, there have been very few studies that have combined different techniques for the assessment of PA in this population. The use of criterion measures seems to be a constant in these studies, but the criterion measure is not consistent. Eight of the fourteen studies in Table 4 used direct observation as the **criterion** measure (Danner et al., 1991; Epstein et al., 1984b; Klesges and Klesges, 1987; Klesges et al., 1990b; McKenzie et al., 1991; Mukeshi et al., 1990; Puhl et al., 1990; Saris et al., 1977). Two of the fourteen studies (Klesges et al., 1985; O'Hara et al., 1989) used direct observation as the **tested** measure. The studies using direct observation as the criterion, validated the Caltrac and heart rate assessment of PA. The studies using direct observation as the tested measure, validated this technique against the Caltrac and heart rate assessment of PA. The need for a standard criterion measure is vital if research in this area is to advance.

The use of combined methods has generally employed only two of the four classifications for the assessment of PA. Some studies (Klesges and Klesges,

1987) have validated accelerometry against direct observation (physical vs. observational), others (Sallis et al., 1990) have validated accelerometry with HR (physical vs. physiological), whereas others (O'Hara et al., 1989) have validated direct observation with HR (observational vs. physiological). To date, there has been only one study (Klesges et al., 1990b) that has attempted to assess PA by all three of these accepted measure of PA, namely physical, physiological and observational. Klesges et al., (1990b) using HR as the criterion measure failed to find significant correlations with either the Caltrac accelerometer or SCAN-CATS direct observational method.

There are many unresolved questions regarding the methodology of PA testing, particularly in pre-schoolers (defined here as age range 3.0 to 5.9 years). The techniques discussed previously have been used abundantly in recent years to describe activity patterns, and less frequently in cross-validation studies within various populations (ie. healthy children and children with chronic diseases) for school-aged and older children. These techniques have rarely been applied with pre-school children and the reliability and validity of these techniques to describe patterns of activity in this segment of the population remains to be established.

Summary Review of the Methods of Assessing PA

Numerous methods have been developed for measuring the patterns and extent of PA as well as EE related to PA in children. The variety of methods indicates that there is no single method that is feasible, reliable and valid for describing and quantifying all components of a child's PA and EE.

The most valid and accurate method of total EE measurement over several days is the doubly-labelled water dilution technique. This method is however, extremely expensive and requires sophisticated and specialized equipment for its analysis. Similarly, direct observation (by an investigator or through video analyses), is the most valid method for describing the nature of activities as they actually occur. This method is labor intensive, thereby making it expensive and impractical for large scale studies. In contrast, recall questionnaires are inexpensive and, as such can be used in large scale surveys. However, they depend on the child's own perception and memory or alternatively, on first-hand knowledge by a parent, teacher or caregiver of the child's activity pattern. In either case, the objectivity and validity of recall questionnaires is only fair, particularly when the recall period spans more than a few days. At the present time no acceptably valid and reliable self-report measure of habitual PA is available for use among a pediatric population. However, recall questionnaires are by far the most reasonable for studies of large populations. These instruments are capable of providing good information regarding the frequency, duration, and intensity of activities when the categories being used are kept distinct, broad in definition and few in number. Those techniques and instruments that have demonstrated promise are worthy of further validation. Combining interviews with self-report questionnaires may increase the validity of the latter to a small degree (Sallis, 1991). However the cost and time involved in this procedure limit its feasibility. Furthermore, the reactive effects of these procedures have not been examined.

A diary and other means of keeping real-time self-records of activities are more valid than recall questionnaires, but they are impractical for young children. Furthermore, the need to record periodically one's activities may interfere with the child's daily routine and spontaneity. As well, the technique is limited to children with sufficient cognitive skills to complete the diary. Riddoch et al (Riddoch et al., 1992) found a low correlation between activity diary data and those obtained from HR monitoring, among 11 to 16-year-old girls and boys. The association was particularly low for low-intensity activities.

HR monitoring and accelerometry are reasonably economical, and objective, and yield reliable quantitative data. Accelerometers monitor the cumulative number of body movements, as well as their intensity. Results can be converted to arbitrary activity counts, and described as energy units. The Caltrac accelerometer has been found to be acceptable for children of all ages. Its liability though, is that it senses motion in only one plane, disregarding movements in other planes or those executed by body segments that are distant from the site of the sensor. As a result, it has low validity when used with pre-schoolers and other young children who habitually perform a substantial repertoire of body movements. A recently developed accelerometer, the Tritrac-R3D, is sensitive to movement in three planes and can yield min-by-min information. It's mode of operation and validity, when used with children, has yet to be determined.

Miniaturized HR monitors can be carried for 24 hours or longer, and they provide continuous information about the metabolic level at different parts of the day. This technology however, has not yet been incorporated in studies involving pre-school age children (Bar-Or et al., 1995). The main drawback of this instrument is that HR depends not only on the child's metabolic level, but also on climatic heat and cold stresses, emotional state, body posture, fatigue, dehydration, etc. As a result, HR monitoring yields values of EE that are some 10-15% higher than the actual EE. Much more research is needed to determine the extent to which each of the above factors affect a child's HR-metabolism relationship. At its current stage of development, this method is very valuable for yielding information on EE in groups, and less so in individuals.

As is apparent, each of the currently available methods for the measurement or assessment of PA and EE has many strengths and weaknesses. Future studies, large scale surveys in particular, should use a combination of methods. This is especially true if the intent is to yield information on the three domains of PA, that is, the physical, physiological, and behavioural components. Much more research and development are needed to refine each of the above methods and make them suitable for young children.

Objectives

The study is intended to determine in pre-school children:

1. The relationship of activity levels determined from heart rate (Polar Vantage XL 4000), accelerometry (Tritrac-R3D) and direct observation (BEACHES) methods in a controlled laboratory environment (see methods for operational definition of controlled laboratory environment).

   Question. Do the three methods provide similar information about the level of PA in a controlled environment?

2. The relationship of activity levels determined from heart rate (Polar Vantage XL 4000), accelerometry (Tritrac-R3D) and direct observation (BEACHES) methods in field trials when the child performs uncontrolled spontaneous activity (see methods for operational definition of spontaneous activity in an uncontrolled environment).

Question. Do the three methods provide similar information about the level of PA in an uncontrolled environment?

3. The consistency of activity patterns on different occasions in the natural environment during free play as determined by the different approaches (see methods for operational definition of free play in the natural environment).

Question. Do the three approaches provide similar information of activity patterns during free play in the natural environment on two different days?

Materials and Methods

Subjects

Forty-nine healthy male and female pre-schoolers, ages 3.0 to 5.9 years participated in various stages of this study. Subject characteristics, including mean age and sex, are described in Table 5. Subjects were residents of the Hamilton region who attended one of three local daycare centres. This was a cross-sectional study of a convenience sample. No attempt was made to obtain a representative, random sample of children. The subjects were recruited by sending invitations to participate in the study to parents of the daycare centres. Parents then consented to volunteer their children for participation in the study. This study was approved by the McMaster University, Faculty of Health Sciences Research Ethics Board.

Methods

Study Design. During recruitment, subjects were assigned into either group 1 or group 2. Group 1 subjects participated in both the laboratory and field conditions of the study, whereas group 2 subject were only observed in the field.

The study consisted of three components, including (1) measurements taken in the child's normal daycare or home environment, (2) measurements taken during outdoor free play time at the child's daycare, and (3) measurements taken in the laboratory under controlled conditions,

In the field phase of the study, direct observation, accelerometry and HR monitoring (detailed methodological descriptions follow) were made simultaneously on each of two days, either at the child's daycare (*n* = 64 observation sessions) or home (*n* = 4 observation sessions) for a period of 1.5 hours each day for both groups of subjects. A total of 34 subjects participated in this component of the study. Subjects participated in normal daycare or home activities without interference from the observer. These activities may have been indoor or outdoor depending on the program set down by the daycare centre administrators for that day. This type of activity was defined as "spontaneous activity in an uncontrolled environment".

Daycare centre data were examined to extract blocks of time in which the child was engaged in outdoor free play. This involved activity in which the children were free to participate in any activity or apparatus they wished. This "free play in the natural environment", usually lasted about 1 hour each day.

In addition, each child from group 1 made one visit to the Children's Exercise and Nutrition Centre Laboratory at Chedoke Hospital. During the laboratory visit, measurements of height, weight, four skinfolds and bioimpedance were taken. An assessment of resting metabolic rate (RMR) was also conducted. Following these measurements in the laboratory for the Group 1 subjects, the Tritrac and HR monitor were attached. The child then performed a choreographed routine consisting of several "staged" activities (each lasting about 4-5 minutes) including; stair climbing, lying down, colouring, playing with a ball, playing with blocks while sitting, and performing simple calisthenics. Total time in the laboratory was about 1.5 hours. This type of activity was defined as being in a “controlled lab environment”. A totoal of 17 subjects participated in the laboratory component of the study.

Altogethor, 3 boys and 2 girls participated in the lab phase only, 8 boys and 4 girls participated in the lab and field phase (i.e., either in the daycare or the home setting), and 16 boys and 16 girls participated in the field phase only. In total 27 boys and 22 girls participated in the study. In terms of the three objectives of this study, 24 boys and 20 girls participated in the “spontaneous activity in an uncontrolled environment” component, 17 boys and 16 girls participated in the “free play in the natural environment” component, and 11 boys and 6 girls participated in the controlled laboratory component of the study.

Measurement Techniques

Heart Rate Monitoring. Monitoring of HR was carried out using the Polar Vantage XL (Figure 3). The transmitter was strapped to the chest. The receiver was placed in an opaque pouch and placed on the wrist to minimize reactivity and to prevent unnecessary handling by the child. Information accumulated in the receiver (stored sequentially over time) was downloaded subsequently to a PC for analyses of EE. The data collected by the HR monitor were in the form of beats per minute sampled every five seconds. A HR curve for the 1.5 hour period was generated (Appendix A). A HR distribution was also generated showing the percent time spent within certain ranges of HR (Appendix B). The raw data (taken every five seconds) for a one minute interval were averaged to give a mean HR for that minute (Figure 4).

The HR monitor was worn on three separate occasions for group 1. At the laboratory, the monitor was placed on the subject prior to measurement of basal metabolic rate and resting (laying) HR. ECG HR data as well as Polar Vantage XL data were collected for a period of 15 min while the subject was laying down watching a movie. Subsequently, and while still wearing the monitor, the subject then performed a choreographed routine which consisted of several levels of activity ranging from recumbant resting to sprinting. Each level of activity was maintained for 4-5 min.

Group 1 and 2 subjects wore the Polar HR monitor for a period of 1.5 hrs on two separate occasions at the daycare while participating in normal activity. Daycare testing sessions were separated by at least 1 day.

Accelerometry. Motion analysis, using the Tritrac-R3D was measured for the same time periods during which the HR data were collected. The apparatus was affixed through a belt to the child's right hip just above the iliac crest (Figure 5).

The Tritrac collected movement counts in the three planes and then calculated a "vector magnitude" (a mathematical resultant vector of the three component ordinal plane vectors) for each minute.

During the lab visit, the Tritrac was attached to the subject immediately prior to the choreographed routine. The accelerometer was rested motionless for a minimum of 5 min prior to attachment to establish a baseline count.

During the daycare visits, the Tritrac was attached to the subject at the same time as the HR monitor and observational data were collected for a period of 1.5 hrs (Figure 6).

Direct Observation. Direct observation of the child's activities during the waking hours was also conducted during the three occasions (laboratory and daycare visits) in which HR and accelerometry data were accumulated. This was done by an abbreviated version of the BEACHES (McKenzie et al., 1991). This version which was developed together with T. McKenzie, the developer of the BEACHES, focused on the child's physical activities and their determining factors (Figure 7).

The San Diego BEACHES is a comprehensive direct observational system designed to simultaneously code the physical activity and eating behaviour of children and related environmental factors. It can be used at home, at school and in any other setting that a targeted child might go. The system was developed from a behaviour analytic (operant psychology/social learning theory) viewpoint and includes coding for 10 separate categories. The abbreviated version of the BEACHES observational system focusses on physical activity and the environmental factors related to PA.

Direct observation involved listening to a tape which prompted the observer to "observe" the child for a 25-second interval. The tape then prompted the observer to "record" the activities; the observer was given 35 seconds to complete the recording of information. In this manner, observed activity data were recorded at one minute intervals recorded instantaneously at the 25th second of the minute.

The observations were coded into five mutually exclusive categories that provided for a continuum of all types of activity. Inter-observer agreement (reliability) scores for PA are extremely high (kappa 0.91) because body position (lying down, sitting, standing, and walking) is the determining factor for four of the categories. Only for category five (very active) do observers have to make fine discriminations.

The activity level provides an estimate of the intensity of the child's PA. Codes 1 to 4 (lying down, sitting, standing, walking) describe the body position of the child and code 5 (very active) describes when the child is expending more energy than he or she would during ordinary walking. For example, code 5 (very active) would be used to indicate the child is wrestling with a peer (even though he or she is lying on his or her back) or pedaling a moving tricycle or stationary bike (even though sitting).

When the child was in transition from one category to another, the code for the higher category was entered. For example, if at the observational signal the child was partially lying down and partially sitting up, the code '2' (sitting) was entered; if the child was getting up from either sitting or lying down, the code '3' (standing) was entered. Some sample activity classifications were;

kneeling (weight on knees only = standing)
kneeling (weight on knees and buttocks = sitting)
inactive "on all fours" = standing
seated swinging (arms producing no momentum) = sitting
seated swinging (arms producing momentum) = very active
carrying, pushing, pulling objects (obviously struggling) = very active

The lab visit also included a period of observation during the choreographed routine. These data were accumulated concurrently with accumulation of HR and accelerometry data. During the daycare visits, subjects were observed during normal activity for the same 1.5 hr period that HR and accelerometry data were recorded.

Coding of Data. Group 1 had a five digit, and group 2 had a six-digit identification number based on the following convention. For the five digit identification number (Group 1 subjects), the first 2 digits represented the subject's decimal age in years (multiplied by 10) on their first visit to the laboratory or daycare centre for the purposes of this study. The 3rd digit represented the sex of the subject (1-male, 2-female). The 4th and 5th digits were numerically assigned numbers based on the number of subjects in the study to date. For example subject I.D. 36116 represents a 3.6 year old male who was the 16th subject recruited into the study.

The six digit identification number (Group 2), represents subjects recruited into the second phase of the study which involved observation only in the daycare centres. The extra digit was added after the digit identifying sex, and represented the specific daycare centre from which the subject was recruited.

The files used to save the data were saved as the subject identification number with an extension denoting the session. The laboratory visits had the extension -Lab, while the field sessions were given extensions, -D1, -D2, -H1, or -H2, indicating whether the observation session was in the daycare or at home (-D vs. -H), and whether it was the first or second observation session (1 vs. 2).

Data Merging and Management Procedures. A Windows application program was written by the computer programmer for CENC (Children's Exercise and Nutrition Centre, Chedoke Hospital) to allow the merging of Tritrac (accelerometry), Polar Vantage XL (heart rate) and BEACHES (observational) data. A sample of merged data can be found in appendix C. A summary of the status of merged files can be found in appendix D.

Tritrac Data. Initially these data were down loaded from the Tritrac-R3D into the PC using manufacturer provided (Tritrac, Hemokinetics: Madison, WI) software. The data were then saved and subsequently retrieved into Word Perfect 5.1. The data collected before and after the observational session were then deleted. This edited data file was then saved as a DOS text file and retrieved into a Windows application for the merging of data. The Tritrac display's data with a one minute lag in the time base so that for example minute 10:24 of real time is recorded by the Tritrac as 10:23. The observation comments were scanned and where it was found that the subject played with the

Tritrac, or where the Tritrac was agitated a value of -1 (for missing data) was entered for that minute.

Heart Rate Data. The data from the HR monitor were down loaded through the manufacturer provided (Polar Vantage XL) software to a PC and saved. They were then retrieved into the Windows application for data merging. The individual 5 second recordings were averaged to obtain a mean one minute HR, and the start points for these data were synchronized with the same minute recorded by the Tritrac. The time bases were adjusted so that the same minute was displayed for both the Tritrac and the HR monitor. For example, the HR recording for minute 10:24 was aligned with the Tritrac record at time 10:23 which represented minute 10:24 of real time. The individual 5-second raw values were then scanned manually to look for anomalous records. All data below 40 bpm or above 220 bpm were discarded and a new average for that minute calculated, because these are known to be outside the physiological range of normal values for heart rate for children. If six or more of the twelve data points for the minute were found to be anomalous then a symbol (-1) for missing data was entered into the data merging program for that minute interval. If outlier data were found (e.g., one data point at 60 bpm while the other 11 data points read 110-120), then that value was discarded and a new average for that minute was calculated and entered into the merged data. The following protocol was followed for handling outlier data. If three or fewer suspect data points were found in a one minute interval then these data were discarded. If four or more outlier data points were found in a one minute interval then these outlier data were not discarded and were included in the minute average. Outlier data was defined as a HR recording within a given 1 min interval which differed by ± 40 bpm.

BEACHES. The BEACHES data were manually entered into the Windows data merging program again lining up each minute of real time with Tritrac and HR records. Minutes where an observation was not recorded were given a symbol (-1) for missing data.

Once all the data had been edited and merged, the entire file was saved in ASCII file format. The file was then retrieved into Quattro Pro 5.0 for windows where a correlation matrix for TT-HR, TT-Obs and HR-Obs was calculated.

Data analyses strategy

Question 1 "*Do the three methods provide similar information about the level of PA in a controlled environment*" and Question 2 "*Do the three methods provide similar information about the level of PA in an uncontrolled environment*": Data Analyses strategy;

Once all the data for each laboratory observation session were merged, a correlation was performed among the three measures of PA. These correlations were then transformed to *z*-values utilizing a standard table for *r* to *z* transformations. One of the basic assumptions of ANOVA is that the data is normally distributed. Correlational analyses render values ranging from -1 to +1 resulting in range of data not normally distributed. By transforming the correlations to *z*-scores we are converting these values to those that are normally distributed, which then allows us to perform other statistical analyses without violating any of the fundamental assumptions. Each session was then placed in a table (Tables 6 and 7) represented by three *z*-scores, one for each of the three sets of relationships (TT-HR, TT-Obs and HR-Obs). The *z*-scores were then used for all further analyses. Mean *z*-scores were transformed back to *r*-values to report correlations. In the laboratory, a one way analyses of variance was performed whereas in the field a two-way repeated measures analysis of variance was performed to test for significant differences in the three relationships (TT-HR, TT-Obs, and HR-Obs) as well as differences between two days of observation (Tables 8, and 9). A Tukey A post hoc analysis was performed to assess which of the relationships significantly differed from each other. An overall summary of these statistics can be found in Table 10.

Question 3 "*Do the three approaches provide similar information on activity patterns during free play in the natural environment*": Data Analyses Strategy;

A comparison was made between two different days of free play outdoor observation for each subject. The average heart rate, Tritrac and observation scores for each subject were compared from day 1 to day 2 (Tables 11 and 12a). The correlations (*z*-scores, Table 7) from each day for each relationship were also compared from one day to the next for all the various combinations to see if there was a difference in the degree to which the three measures of PA agreed with each other over different days for the same subject. Finally a rank order correlation was performed between days for each of the three measures to assess the consistency of observed activity levels for each of the evaluation procedures (Table 12b).

Results

Table 6 summarizes the correlation matrix between the various combinations of dependent measures for all the subjects who came to the laboratory. Correlations between TT and HR were significant and ranged from $r = 0.68$ to $r = 0.95$ (overall $r = 0.85 \pm 0.05$; $n = 17$). Correlations between TT and Obs were significant and ranged from $r = 0.51$ to $r = 0.86$ (overall $r = 0.76 \pm 0.07$; $n = 17$). Correlations between HR and Obs were significant and ranged from $r = 0.33$ to $r = 0.88$ (overall $r = 0.76 \pm 0.09$; $n = 17$). Dependent t-tests were run on the z-transformed correlations to determine whether there were any differences in the strength of the relationships among the various combinations. Tables 8 summarizes the results of the one-way ANOVA and t-tests in the laboratory. There was no significant difference (Table 10) in the correlations between HR-Obs and TT-Obs; however, the TT-HR correlation was significantly greater than the other two, HR-Obs and TT-Obs. The TT-Obs and HR-Obs techniques appear to provide similar information about the level of PA in a controlled laboratory setting.

The correlation matrices for all subjects who were observed in the daycare (home data of $n = 5$ were pooled with daycare data, as no significant differences between them were found) are presented in Table 7. Correlations between TT and HR (Figure 8) were significant and ranged from $r = 0.20$ to $r = 0.89$ (overall $r = 0.69 \pm 0.12$; $n = 68$). Correlations between TT and Obs (Figure 9) were significant and ranged from $r = 0.05$ to $r = 0.88$ (overall $r = 0.58 \pm 0.11$; $n = 68$). Correlations between HR and Obs (Figure 10) were significant and ranged from $r = 0.05$ to $r = 0.77$ (overall $r = 0.50 \pm 0.15$; n = 68). Dependent *t*-tests were run to determine if any differences existed in the strength of the relationship among the various combinations. Tables 9 summarizes the results of the 2-way repeated measures ANOVA and the *t*-tests. Significant differences were found (Table 10) between TT-HR and TT-Obs, TT-HR and HR-Obs, and TT-Obs and HR-Obs. The results suggest that the various combinations of techniques measure different aspects of PA in an uncontrolled field setting.

Inter-method comparisons (Table 10) of correlations (*z*-scores) indicated that there was a significant difference in the TT-HR relationship between lab and field conditions ($t = 5.46$; $P < 0.001$). Similarly there was a significant difference in TT-Obs correlations between lab and in the field conditions ($t = 6.51$; $P < 0.001$), and there was a significant difference in HR-Obs between the lab and in the field($t = 6.88$; $P < 0.001$), the lab values being higher than field correlations in all cases. These results suggest that all three methods of assessing PA are fairly well correlated with each other and are all valid tools for PA assessment in controlled or uncontrolled environments, but perhaps are more valid in the laboratory than in the field.

The results for the consistency study of activity patterns are summarized in Tables 11 and 12. The repeated measures analysis of variance shows no significant main effect for day. As indicated in Table 12a, there were no significant differences in group activity scores between testing sessions within any of the separate techniques. This indicates that the groups were fairly consistent in the amount of activity performed between the two observed sessions. To ascertain the degree to which the subjects maintained their level of activity, a rank order correlation was performed for each of the three instruments between the two days of observation. The correlations ranged from 0.39 - 0.68 and these results suggest that there is considerable individual variability in activity levels between testing days. These data suggest that as a group subjects tended to maintain a similar level of activity from day to day, but that the reliability of individual activity levels between days is low to moderate.

Discussion

Accelerometry and Direct Observation Relationship

Mukeshi et al. (1990) tested the validity of the Caltrac accelerometer for use with pre-school children during unstructured free play. Values of EE derived by the Caltrac were compared to values of EE derived from direct observation of subjects' PA every 5 seconds for an hour in the daycare or on the playground. The observation method used by this group was the Fargo Activity Timesampling Survey - FATS (Klesges et al., 1984). The correlation coefficient found by Mukeshi et al. (1990) between EE derived from the Caltrac and EE derived from the observational method was $r = .62$ ($p < .05$). When the children's weight, height, age and sex were factored out of the Caltrac scores, however, the correlations fell to $r = .25$ (n.s.). The conclusion of this study was that the Caltrac was not of much value as a means of measuring PA in children 2-3 years of age. The present study found a correlation coefficient of $r = .58$ ($p < .05$) between arbitrary activity counts of the Tritrac and the BEACHES direct observation method in the field (Figure 9). The use of arbitrary counts in the present study factors out the effects of weight, height, age and sex from the analyses. The higher correlation (compared to Mukeshi et al. 1990) found in the present study reflects recent improvements in the development of methods of assessing PA. The Tritrac, by virtue of the fact that it measures activity in three planes and has the capability to display min-by-min changes in activity level, makes it a more reliable and valid tool than the Caltrac for the assessment of PA and EE in epidemiological research.

Klesges and Klesges (1987) also assessed the convergent validity and sources of error of the Caltrac and validated this device against observed all day physical activity levels of 2-4 year old children in their natural environment. Spearman rank order correlations between hourly readings of Caltrac and FATS ranged from $\rho = 0.62$ to 0.95, with an all day correlation of 0.54. The hourly interval correlations in Klesge's study were generally higher than the observed correlation ($r = 0.58$) over a one and a half hour period between Tritrac and BEACHES in the present study, but our value is comparable to the daily correlation of 0.54 reported by Klesges and Klesges (1987).

In studies with adults, we see much higher correlations between direct observation of levels of PA and accelerometry. Ballor et al. (1989), evaluated the Caltrac's ability to estimate caloric expenditure during mixed activities and compared it with estimation of EE by video analyses. Video analysis involved a rating scale based on energy expenditure estimates from the literature (American College of Sports Medicine, 1986; Consalazio, 1963; McArdle et al., 1986; Morehouse and Miller, 1976; Passmore and Durnin, 1955). The videotapes were analyzed and activities were assigned MET values (activity expressed in multiples of resting energy expenditure, 3.5 ml $O_2 \cdot kg^{-1} \cdot min^{-1}$) in 15 second intervals. Ballor et al. (1989) reported a correlation of $r = 0.95$ between caloric expenditure estimations determined via Caltrac and video analyses in a high school basketball game. A simulation of a basketball game was also performed in the laboratory to act as a validation of the methods used to estimate EE. The validation procedure was necessary because the basketball class data revealed that three measures (Caltrac, HR and Observation by video analyses) used to estimate EE were highly correlated but statistically different from each other. Correlations for the laboratory sessions were not reported. In the present study, a correlation of $r = 0.76$ was found between Tritrac and direct observation in the laboratory choreographed routine. This higher correlation in the laboratory is closer to values reported by Ballor et al. (1989) for the basketball game and as reported by them similar to results for the laboratory simulation. In the present study an observational method was employed that used 1 minute time sampling. Sampling for such a short period will result in a greater number of misclassifications which will be equally distributed as over and underestimations of actual activity levels. The potential for greater misclassifications over short observational periods when compared to real time assessments of PA by the Tritrac will result in lower correlations between the observational and accelerometry techniques. This sampling artifact accounts in part for the slightly lower correlations between the observational and accelerometry techniques in the present study compared to those of Ballor et al., (1989). With a high EE activity like basketball, the HR range was probably greater in the study of Ballor et al (1989) than in the present study, which will result in higher correlation coefficients due to the increased range of data. The smaller range of HR measured may account for the lower correlations found in the present study.

The relatively low correlations found in this study (compared to similar comparisons in adults) between observational and accelerometry assessments of PA may also be partially explained by the mechanisms of age-related energy cost differences between populations. The discrepancy in the strength of the relationships may be attributed to acceleration differences for the same activities between the children and adults. Differences in energy cost between children and adults appear to be related to differences in the economy of locomotion. The O_2 cost during weight bearing activities (i.e., walking and running) is higher in children when expressed per kilogram bodyweight (Bar-Or, 1983). Furthermore, the higher O_2 cost in young children cannot be explained merely by a difference in resting metabolism, which is only 1 to 2 ml/kg/min greater than adult values (Guyton, 1986). The most logical explanation for the increased energy cost of weight bearing activities in children is their relative 'wasteful' gait while walking or running (Bar-Or, 1983; Spurr et al., 1984). It has been suggested that young children are less economical due to the higher stride frequency necessitated by their shorter limbs (Astrand, 1952). Davies (1980) suggests that the frequency of leg movement in children is not optimally matched to the force required to produce the most economical conversion of aerobically produced energy to mechanical work. Compared to adults, children are less efficient in weight bearing activities and primarily exhibit a greater variability in EE for a given activity. The greater variability in the younger populations may account for the lower correlations found in the present study.

A study by Danner et al., (1991) examined the validity of the Caltrac motion sensor for measuring PA in young children and described changes over time in PA among 47 pre-school children (aged 2.5 to 5.4 yrs). Children were videotaped in a controlled unstructured environment while wearing the Caltrac motion sensor. Caltrac readings and observational measures (CARS) of PA were significantly related in both year one ($r = 0.86$) and year two ($r = 0.83$). Using a similar physical environment to that of the present study, Danner et al., (1991), reports higher correlations ($r = 0.58$) between Tritrac and observation in the field. Highly significant correlations between behaviourally determined mean activity scores and Tritrac readings in both observational sessions were found in our study. We would expect higher TT-Obs correlations if we discounted level 1 obervation data and the corresponding TT data, since the TT measurements would not change when the subject is in a lying, resting states.

Together, these results suggest that the Tritrac monitor provides a valid measure of PA in a natural setting.

The results from the present study suggest that the Tritrac-R3D may be a useful device for assessing PA in young children in future epidemiological research. Klesges and Klesges (1987) found correlations from 0.57 to 0.95 between Caltrac scores and direct observation on an hour-to-hour basis whereas Mukeshi et al., (1990) found a correlation of 0.62 for a 2-hour observation period. Danner et al., (1991) found correlations of 0.86 in year one and 0.83 in year two for one hour of observation. Our correlation (r = 0.58) for 90 minutes of field observation was somewhat lower than in these studies, but the Tritrac provides several technical advantages over the Caltrac device. In contrast to the Caltrac, the subject's actual height, weight, age and gender are entered during the computer initialization of the device. These factors do not enter into the calculation of the vector magnitude, and therefore were not used in the present study. Additionally, the device can be programmed to analyse results for desired measurement intervals rather than solely providing one result for the entire measurement period, as is the case with the Caltrac. When downloading, the data are presented as raw acceleration counts in each of the three ordinal planes, the vector magnitude of the three accelerations, and kilocalories burned per interval. The Tritrac therefore appears to provide considerably more information about the dynamics and temporal nature of PA than the Caltrac, with about the same degree of concurrent validity.

Finally, this study used a completely different observation system (BEACHES) compared to the aforementioned studies. Klesges and Klesges (1987) used the FATS system while Mukeshi et al., (1990) used an adaptation of the FATS. Danner et al., (1991) used the CARS system. The BEACHES system has five categories based on body position. The CARS system also has five categories based on EE, with two levels allocated for stationary activity and three levels for dynamic activity involving movement from place to place (translocation). The FATS system, on the other hand, includes eight types of activities (e.g., laying, sitting, standing, walking, etc.) and a 3-point scale for ranking the intensity of each activity (low, moderate, high). The BEACHES system whose focus is on behavioural aspects of the child's health may not be the ideal system when assessing PA patterns and intensities. It may be that the types of activities young children engage in are better suited to the CARS type of ranking as opposed to trying to categorize movements into activities that can be named.

Klesges et al., (1990b) in a study on pre-school children performed some factor analyses (on a multi-method approach to the measurement of childhood PA) and found that motion sensor activity correlated only with itself, indicating that the Caltrac activity was not related to other measures of PA including direct observation (SCAN-CATS). These investigators concluded that perhaps the measures of childhood PA used in their study may have been inadequate. This argument is consistent with previous research that has shown the difficulty of measuring childhood activity (Klesges and Klesges, 1987; Saris 1986). The use of direct observation and accelerometry in their study argues against the interpretation that the measures were inadequate, as these measures have been shown to be quite reliable and valid in this population. In fact these assessment methods are generally considered criteria against which other measures are judged (Klesges and Klesges, 1987; Baranowski et al., 1984). However, given the variability of physical activity in children, perhaps even the best measures of activity are inadequate unless they continue for extended periods.

Heart Rate and Direct Observation Relationship

O'Hara et al., (1989) used min-by-min HR during physical education periods to validate an observational method for quantifying intensity of activity, the Children's Physical Activity Form (CPAF). This observational system employs partial-time sampling (1-minute intervals) with four categories representing different levels of intensity of movement. This method has only been validated on third-fifth grade students. The mean Pearson correlation reported by O'Hara et al., (1989) between the activity points and heart rates was $r = 0.64$. The present study found a mean Pearson correlation of $r = 0.50$

between observed levels of activity as measured by the BEACHES and HR in the field (Figure 10). The CPAF observational method has yet to be established as a reliable system for the assessment of PA in children and has to date not been used with preschoolers. Although the correlation in the present study was lower than that reported by O'Hara et al., (1989), the observational system in the present study has been used extensively in recent years and has been validated in the pre-school population. It is known that there are many factors affecting HR besides the physiological response to exercise, such as excitement and climate. These factors may play a more important role in the HR variation in younger subjects, because it is known that HR at a given work rate increases with increasing temperature and level of arousal in children (Haymes et al., 1974; Haymes et al., 1975), thereby accounting at least in part for the lower correlations in our study of pre-schoolers.

A study by Ballor et al., (1989) used HR recording, caloric expenditure curves and an activity rating scale to determine the relationship between HR recording and video analyses estimates of energy expenditure. Video analysis involved a rating scale based on energy expenditure estimates from literature (American College of Sports Medicine, 1986; Consalazio, 1963; McArdle et al., 1986; Morehouse and Miller, 1976; Passmore and Durnin, 1955). The videotapes were analyzed and assigned MET values (activity expressed in multiples of resting expenditure, 3.5 ml $O_2 \cdot kg^{-1} \cdot min^{-1}$) in 15 second intervals.

Ballor et al. (1989) reported a correlation of $r = 0.89$ between caloric expenditure estimations determined via HR and video analysis in a high school basketball game. A simulation of a basketball game was also performed in the laboratory to serve as a validation of the methods used to estimate EE. Correlations of these lab data were not reported, but the investigators did report that the Caltrac and video estimates of EE as a percentage of HR expenditure were very close in both the basketball class (Caltrac = 84%, video analyses = 63%) and simulation (Caltrac = 81%, video analyses = 63%). In the present study, a correlation of $r = 0.76$ was found between HR and direct observation in the laboratory choreographed routine.

The CARS observational system was validated by Puhl et al., (1990) with the use of HR monitoring. Although no correlation between HR and observation was reported, Puhl et al., (1990) reported that the activities chosen to represent the five different CARS levels elicited heart rates within expected ranges and that HR values were significantly different between levels. The present study found correlation coefficients of $r = 0.76$ (laboratory) and $r = 0.50$ (field) between HR as measured by the Polar Vantage XL and observation as measured by the BEACHES.

McKenzie et al., (1991) developed the BEACHES observational system and validated it against HR. The mean heart rates and energy expenditures were derived for each activity category. This activity coding system and validation study are similar to those of Epstein et al., (1984b), but the BEACHES categories were associated with a more progressive increase in HR than the Epstein et al., (1984b) system. The correlation coefficients between BEACHES observation scores and HR were not reported by McKenzie et al., but are thought to be similar to the correlation found in the present study of $r = 0.50$ (Figure 11), as the same observational system, a similar HR monitor and the same age range of subjects were used. The HR found at each activity level, were consistently higher than that reported by McKenzie et al., this may simply be due to sampling biases.

The validity of the APEE (Activity Patterns and Energy Expenditures), observational measurement system used by Epstein et al., (1984b) was established by correlating activity scores for each child each day with the changes in HR from rest to activity associated with free play. The Pearson product moment correlation between HR change and activity scores was $r = 0.82$ ($p < 0.001$). There are several differences between the study by Epstein et al., (1984b) and the present one. First, the subjects in the present study were normal males and females between the ages of 3 and 5 years, whereas Epstein et al., (1984) used overweight female children between 5 and 8 years. Secondly, the observational system used by Epstein et al., (1984b) has not been validated in any other age group and has since not been used as an observational system in epidemiological studies.

Accelerometry and Heart Rate Relationship

Conceptually, it appears that by simultaneously recording both HR and motion from body accelerations in each of the ordinal planes, it should be possible to obtain an accurate estimate of the profile of EE due to PA performed throughout the day.

A study by Meijer et al., (1989) used a portable accelerometer equipped with a three-directional sensor and a HR monitor to assess PA and EE in adults under laboratory conditions and during normal daily life. In their study, the relationship between HR and EE was particularly weak at low activity levels. No correlation was given for the relationship between HR and accelerometry, however, an estimate based on the data presented from the field condition suggests a correlation of $r = 0.47$. This value is lower than that found in the present study for identical variables (Figure 8) under both the field ($r = 0.69$) and laboratory conditions ($r = 0.85$). Miejer et al., (1989) defended their low values by explaining that the weak relationship found between HR and EE at low activity levels corresponding with earlier findings (Saris, 1986), in combination with the influence of anxiety, heat, and posture on HR, has a large negative effect on the accuracy of the HR method as a measure of EE of the individual in normal life. The present study found significant correlations which were fairly strong in a much younger population than used by Miejer et al., (1989).

Haskell et al., (1993) recently described a procedure which uses simultaneous recording of HR and two accelerometers to provide an accurate profile of PA. In their study which involved adults, the simple linear regression for predicting oxygen uptake from HR during various exercises resulted in an R^2 of 0.81. The study also demonstrated the potential for increasing the accuracy of the estimation of EE using this procedure by having separate regression equations for different exercises. As for the Caltrac, the results indicated that the motion sensor quite accurately tracks the changes in speed, but does not respond in any quantitative way to changes in slope. The inability of the Caltrac to detect increasing intensity of some types of exercises has apparently been overcome with the Tritrac-R3D (Hemokinetics Inc.) used in the present study. Haskell et al., (1993) made no attempt to correlate the accuracy of the accelerometer's prediction of oxygen uptake, with that of the HR monitor. Rather, the investigators performed a multiple regression analysis incorporating both methods to predict oxygen uptake. The R^2 derived from this analyses was 0.89. The addition of information from the accelerometer seemed to account for an additionalsigificant component(8%) of the variability in this relationship.

Under controlled testing situations in the present study, the Tritrac accelerometer was found to be a highly reliable instrument, and the comparison with HR monitoring data supported its usefulness. A correlation of $r = 0.85$ between the two measures in the laboratory was obtained in the present study.

The coefficients reported here are higher than those reported for a laboratory study (Montoye et al., 1983) with adults $r = 0.74$. The present study's correlation of $r = 0.69$ between accelerometry and HR in the field is substantially higher than coefficients ($r = 0.49$) reported by Sallis et al., (1990) for school age children in the field. The higher correlations in the present study may be explained in part by improvements in accelerometer technology since the Tritrac now gathers data from three planes compared to the single plane assessed by the Caltrac, as well as its new ability to synchronize min-by-min data with HR. The results of adult and child studies suggest that accelerometry is a promising method for assessing PA in children. Given the findings of higher validity with adults (Klesges et al., 1985) it seems surprising that the correlations in the present study have been higher than those found in previous studies with adults (Montoye et al., 1983) and school-age children (Sallis et al., 1990).

Ballor et al., (1989) reported a correlation of $r = 0.92$ between Caltrac and heart rate estimates of caloric expenditure in high school students during a basketball class. This correlation is substantially higher than that of the $r = 0.69$ found in the present study. The higher range of HR elicited in a high activity sport such as basketball may have accounted for some of the difference in the strength of correlations between the accelerometer and the HRM. Additionall, the difference in the strength of the correlations between studies can be attributed in part to the use of different accelerometers, since the heart rate monitors were similar, and to differences in age groups. The correlation coefficients found by these investigators are unusually high compared to those reported in similar studies (Montoye et al., 1983; Sallis et al., 1990). Ballor et al., (1989) did report that the HR estimate of caloric expenditure was approximately 20% higher than the Caltrac estimate, and that this difference was statistically significant.

Reliability of Activity Patterns

The daycare provides a semi-structured environment for activity during the course of the day. Although children have many options and free choices of activity, certain activities are promoted in different rooms, at different times throughout the day. Outdoor play is completely unstructured and the children are allowed to participate in any activity they wish. Comparison of this uncontrolled free activity time should provide an index of the reliability of activity on a day to day basis. In the present study, no significant day-to-day differences were found for free-time activity assessed with any of the three measurement techniques (Table 12a). These results indicate that activity patterns of the children as a group are fairly consistent from day-to-day if the same opportunities to participate in uncontrolled free-play are provided. The low *r* values (Table 12b) however indicate that while as a group subjects tended to maintain a similar day-to-day level of activity, there was a fair deal of individual variability in activity level between testing days.

In a study involving high-school students, Sallis et al., (1990) found that neither mean activity HR nor Caltrac activity counts per hour were significantly different between separate days involving more than ten hours of monitoring each day. Orenstein et al., (1993) in a study of school-age children also found consistent patterns of PA as reflected by Caltrac measures over three days. In their study (Orenstein et al., 1993) the intraclass correlation coefficient for PA measured by Caltrac between days was $r = 0.69$. Both of these studies (Orenstein, 1993; Sallis et al., 1990) report similar findings to ours, and together these findings suggest that PA patterns are fairly consistent from one day to the next as a group, regardless of how PA is measured, but that individual activity patterns are less stable and more variable over time. Orenstein et al., (1993) and Sallis et al., (1990) conducted their studies on school-age children (7.0 to 18.0 years), and this difference may call into question comparisons to the present study on pre-school children.

Durant et al., (1992) looked at the daily HR patterns and the between day reliability of HR monitoring in 3-5 year old children, and found that the reliability of HR for two days of observation separated by 3-6 months ranged from $r = 0.65$ to $r = 0.66$. The present study found a correlation of $r = 0.68$ between two days of HR monitoring in a group of similar age. Durant et al.,

(1992) monitored heart rate for over 12 hours each day whereas the present study was restricted to 90 minutes of monitoring each day. Durant et al., (1992) went on to hypothesize that at this level of reliability, just over four days of recording would be necessary to achieve a reliability of 0.80.

With regard to accelerometers, there have been very few studies examining the reliability of activity between different days. Washburn et al., (1989) compared the output of the Caltrac ($kcal \cdot kg^{-1} \cdot h^{-1}$) worn by 35 U.S. postal carriers on two workdays. The day-to-day correlations for activity varied by season; $r = 0.62$ summer, vs. $r = 0.47$ winter. Gretebeck and Montoye (1992) in their study with adults found a 2 day Spearman-Brown (rank order) correlation of 0.76 for a 3-dimensional Caltrac system measuring METs. The reported correlations in both of these studies (Gretebeck and Montoye, 1992; Washburn et al., 1989) were higher than the correlation of $r = 0.43$. The present study was conducted on a population in which activity is believed to be more spontaneous and varied. This variability and unpredictability of activity patterns in children may account for the lower correlations in the present study.

To examine long-term consistency of PA, Danner et al., (1991) compared correlations for PA assessed by Caltrac and the CARS on two separate occasions separated by 1 year. Observations were made in a confined space (room) on children 2.5 to 5.4 years of age. The children were allowed free choice of activity, and each wore a Caltrac accelerometer while activity was simultaneously videotaped. Danner et al., (1991) failed to find any consistency of PA levels from year 1 to year 2 when measured by either the Caltrac or activity ranking. The low correlations from year 1 to year 2 may reflect an actual change in activity levels or simply a situational response due to lack of control over influential variables. For example, children were not observed in the same play groups in both years. Other variables such as time of day and whether the child was rested or tired could not be controlled and could also have affected consistency of scores from year to year. The results from the present study suggest that if we were concerned about the consisency of activity level from day to day in a group, then the three methods described in the present are valid techniques for future studies. However, the assessment of individual activit pattern consistency is not reccomended with the techniques presented in this study.

Data Synchronization and Manipulation

The cardiac output increase during exercise is associated with greater sympathetic activity and less parasympathetic activity to the heart. Thus, heart rate and stroke volume rise, causing an increased cardiac output. The heart rate changes are usually much greater than the stroke volume changes. However, this change in autonomic activity, and its resulting effect on heart rate are not instantaneous. A physiological lag occurs in response to changes in levels of activity before the heart stabilizes at its' new rate, matched to the new exercise intensity. To account for this physiological lag, and any possible lag in the mechanical device used to measure heart rate, a one minute time shift was applied to the heart rate data for a subsample of 12 observational sessions. These adjusted data were then correlated with the original Tritrac and observational scores. The results for the data manipulation involving adjustment for the one minute time lag in HR are summarized in Table 13. The manipulation had no significant effect on the TT-HR results, but the HR-Obs correlation decreased significantly after the adjustment. The statistical summary of these data is presented in Table 14.

The HR-Obs correlation (*z*-scores) with the 1 min HR lag was significantly different from original correlations and can be explained through qualitative analyses of observation records. As recorded in the comments section of the observation sheet, we know for example that an activity may have been terminated during a particular interval. Comparison of that interval with HR recordings indicate that HR may take up to several minutes to reach physiological levels corresponding to the new activity level. The Tritrac-Heart rate correlation, although not significantly different after introduction of the HR lag was very close to reaching significance ($t = 1.96$; $p = 0.08$). This finding also suggests that the time lag may have an important influence on the strength of the inter-correlations among the various measurement techniques.

A qualitative analysis of the behavioural observation data revealed that during the field phase of the study, when subjects were in an uncontrolled environment, most of the activity was at a low-medium level. Subjects in general were found to spend most of their time sitting or standing (2 and 3 on the BEACHES activity scores). Very little time was spent at a high level of activity (5-very active on BEACHES) or low level of activity (1-lying on BEACHES). This resulted in a skewed distribution of the data. To assess the effect of this non-normal distribution on the strength of the inter-method correlations, a subsample of 10 observation sessions was manipulated through the removal of random sections of data. The aim of the manipulation was to leave each file with a balanced range of activities representing a normal curve. This distribution represented 5 min Low Level Activity (LLA), 10 min Medium Level Activity (MLA), and 5 min High Level Activity (HLA), where LLA = 1-2, MLA = 3-4, and HLA = 5 on the BEACHES five point activity scale. This adjustment had no significant effect on the magnitude of the correlations among the three comparisons (Tables 14 and 15). To test the possibility that correlations among tasks might be influenced by activity level, further analyses were done, separating subjects into tertiles based separately on all three methods. With this manipulation, there were no significant differences in the magnitude of the correlations for the three measures of PA among the high, moderate or low activity groups. The results from these analyses suggest that level of activity, or unequal distribution of activity did not affect the degree to which the three measures of PA correlated with each other.

Upon analysis, many of the recorded measures occurred in transition between different levels of activity. To obtain a clearer understanding or representation of the interrelationship among the three measurement techniques during clearly defined levels of activity it was decided to discard or disregard transitional data from the analyses. In the laboratory, subjects performed a choreographed routine which was designed to sustain exercise intensity for a minimum of four minutes at several different levels, to permit stabilization of the physiological and physical measures of PA. With this design, it was possible to remove data from transitional intervals between these distinct and stable levels of activity. This manipulation provided several minutes of data with which we can say with fair certainty that subjects were at a specific level of activity. This removes the effects of potential lags in HR and accelerometry in the transition from one level of activity to the next. The results for the data manipulation involving removal of transitional exercise phases during the laboratory session are summarized in Tables 14 and 16. For all combinations of techniques, the correlations significantly improved after removal of transitional phases. The magnitude of the increase in correlation coefficients was similar for all comparisons. The statistical summary of these data is presented in Table 14. This suggests that the relationship between the different measures of PA are affected by transitional lags in HR and accelerometry from one discrete level of activity to another.

These findings suggest that although the assessment of PA is fairly valid and reliable even on a minute by minute basis, a longer measurement interval would probably dampen momentary differences arising from transition from one activity to another and contribute to higher correlations. An interval of 5 or perhaps even 10 minutes would attenuate the effects of the physical and physiological lag resulting in a higher agreement between different measures.

Conclusion

Measuring physical activity is difficult, particularly in children. Over 30 different techniques have been used and none are fully satisfactory. Daily physical activities involve numerous actions that are difficult to quantify without interfering with the subject's normal patterns of exercise. As a rule, the easier and more convenient the technique for assessing these activities, the less valid it is likely to be. The best means of quantifying physical activity is by measuring energy expended during the exercise. This calls for measuring oxygen consumption, a procedure entailing the use of a mouthpiece, tubing, gas analyzers, ventilometers, and recording devices. Clearly this is not a practical technique for establishing activity levels in large groups of children. At the other extreme, questionnaires and diaries completed by a parent, teacher, or child are easy to administer but are of variable reliability.

The development of small, portable, unobtrusive HR monitors and accelerometers has greatly advanced efforts to measure habitual physical activity. These techniques offer the advantage of measuring exercise intensity without interfering with normal activity patterns. The development of observational methods of assessing activity levels and their relationship to HR and energy expenditure has further increased our ability to more validly and reliably measure actual energy expenditure during activity.

The present study examined the relationship between three different measures of PA and the degree to which they agreed in the laboratory as well as in the field. The three measures reflected the physical (accelerometry), physiological (heart rate) and behavioural (direct observation) categories or dimensions of PA, so that a more rounded analysis of PA could be made. This study found that the three measures were significantly correlated, especially during the laboratory condition. It was also found that measurement intervals of greater than 5 minutes would likely increase the correlation, because the longer interval would dampen the effect of a physical and physiological lag of the measuring devices. From a biological perspective, this study found that group level of physical activity was fairly stable or consistent between two different days of observation. However, there appears to be a fair degree of individual variability in activity levels from day to day, and individual activity level of preschool children appears to be less consistent than group activity level.

Promoting physical activity should be a high-priority for the future because of the repeated demonstrations in epidemiologic studies of an association between inactivity fostered by modern lifestyles and risk of coronary heart and other diseases. Physical activity is generally viewed as having a favourable influence on the growth, biological maturation, and fitness of children and youth. Inferences about the role of PA in health and disease are based largely on short-term experimental studies. Caution in needed in extending these observations from small samples to the general population. Longitudinal studies that span childhood and adolescence and that control for PA are limited. The expense in terms of money and time has created a need for better techniques to assess PA in both short and long-term population studies.

However, there are still many problems relating to the assessment of PA in children. There is as yet no agreed upon "gold standard" measurement technique. Furthermore, existing validation criteria, where various instruments account for different aspects of PA, only explain part of the variability in daily PA habits. There is a need in the future to refine techniques and conduct more studies that combine different aspects of PA that will permit better application and generalizability to all socioeconomic, ethnic, gender, and age strata of the population.

BIBLIOGRAPHY

Ainsworth B.E., H.J. Montoye, and A.S. Leon. Methods of assessing physical activity during leisure and work. *In: Physical Activity, Fitness and Health: International proceedings and consensus statement,* C. Bouchard, R.J. Shephard, and T. Stephens (Eds.). New York: Human Kinetics, 1994, pp.146-159.

American College of Sports Medicine. *Guidelines for exercise testing and prescription.* Philadelphia: Lea & Febiger, 1986.

Astrand P.O. *Experimental Studies of Working Capacity in Relation to Sex and Age.* Copenhagen: Munksgard, 1952.

Ballor, D.L., L.M. Burke, D.V. Knudson et al. Comparison of three methods of estimating energy expenditure: Caltrac, heart rate, and video analyses. *Res. Q. Exerc. Sport.* 60(4):362-368, 1989.

Bar-Or O. *Pediatric Sports Medicine for the Practitioner.* New York: Springer Verlag, 1983.

Bar-Or, T., O. Bar-Or, H. Waters et al. Validity and social acceptability of the Polar Vantage XL as a means of measuring heart rate during rest, exercise and recovery in 3- to 5-year-old children. *in press.* 1995.

Baranowski, T., C. Bouchard, O. Bar-Or et al. Assessment, prevalence, and cardiovascular benefits of physical activity and fitness in youth. *Med. Sci. Sports Exerc.* 24(6) Suppl:S237-S247, 1992.

Baranowski, T., R. Dworkin, C.J. Cieslik et al. Relibility and validity of children's self-report of aerobic activity: Family Health Project. *Res. Q. Exerc. Sport.* 55:309-317, 1984.

Berenson G.S., C.A. McMahon, and A.W. Voors. *Cardiovascular Risk Factors in Children: The Early Natural History of Atherosclerosis and Essential Hypertension..* New York, NY: Oxford University Press, 1980.

Blair S.N., D.B. Clark, K.J. Cureton et al. Exercise and fitness in childhood: implications for a liftime of health. *In: Perspectives in Exercise Science and Sports Medicine. Vol 2. Youth, Exercise and Sport,* C.V. Gisolf and D.R. Lamb (Eds.). Indianapolis: Benchmark Press, 1989, pp.401-430.

Blair, S.N., W.L. Haskel, P. Ho et al. Assessment of habitual physical activity by a seven-day-recall in a community survey and controlled experiment. *Am. J. Epidemiol.* 122:794-804, 1985.

Bradfield, R.B. A technique for determination of usual daily energy expenditure in the field. *Am. J. Clin. Nutr.* 37:461-467, 1983.

Bray, M.S., Morrow, Jr., J.M. Pivarnik et al. Caltrac Validity for estimating caloric expenditure with children. *Ped. Exer. Sci.* 4:166-179, 1992.

Burke, M. and M. Whelan. The accuracy and reliability of commercial heart rate monitors. *Br. J. Sports Med.* 21:29-32, 1987.

Casperson, C.J., K.E. Powell, and G.M. Christenson. Physical activity, exercise and physical fitness: Definitions and distinctions for health related research. *Publ. Hlth. Rep.* 100:126-131, 1985.

Charney, E., H.C. Goodman, M. McBride et al. Childhood antecedents of adult obesity. Do chubby infants become obese adults?. *N. Eng. J. Med.* 295:6-9, 1976.

Consolazio P., R.E. Johnson, and L. Pecora. *Physiologic measurements of metabolic functions in man.* New York: McGraw-Hill, 1963.

Danner, F., M. Noland, M. McFadden et al. Description of the physical activity of young children using movement sensors and observation methods. *Ped. Exer. Sci.* 3:11-20, 1991.

Davies, C.T.M. Metabolic cost of exercise and physical performance in children with some observations on external loading. *Eur. J. Appl. Physiol.* 45:95-102, 1980.

Dietz, W.H. and S.L. Gortmaker. Do we fatten our children at the television set? Obesity and television viewing in children and adolescents. *Pediatrics.* 75:807-812, 1985.

Durant, R.H., T. Baranowski, H. Davis et al. Reliability and variability of heart rate monitoring in 3-, 4-, or 5-yr-old children. *Med. Sci. Sports Exerc.* 24(2):265-271, 1992.

Eaton, W.O. and L.R. Enns. Sex differences in human motor activity level. *Psychology Bulletin.* 100:19-28, 1986.

Emons, H.J.G., D.C. Groenenboom, K.R. Westerterp et al. Comparison of heart rate monitoring combined with indirect calorimetry and the doubly labeled water (2H218O) method for the measurement of energy expenditure in children. *Eur. J. Appl. Physiol.* 65:99-103, 1992.

Epstein, L.H., R. Koeske, and R.R. Wing. Adherence to exercise in obese children. *Journal of Cardiac Rehabilitation.* 4:185-195, 1984.

Epstein, L.H., C. McGowan, and D. Woodall. A behavioral observation system for free play in young overweight female children. *Res. Q. Exerc. Sport.* 55(2):180-183, 1984.

Fenster, J.R., P.S. Freedson, R.A. Washburn et al. The relationship between peak oxygen uptake and physical activity in 6- to 8-year old children. *Ped. Exer. Sci.* 1(2):127-136, 1989.

Freedson, P.S. Field monitoring of physical activity in children. *Ped. Exer. Sci.* 1:8-18, 1989.

Freedson, P.S. and S. Evenson. Familial aggregation in physical activity. *Res. Q. Exerc. Sport.* 62:384-389, 1991.

Gortmaker, S.L., W.H. Dietz, A.M. Sobol et al. Increasing pediatric obesity in the United States. *Am. J. Dis. Child.* 141:535-540, 1987.

Gretebeck, R.J. and H.J. Montoye. Variability of some objective measures of physical activity. *Med. Sci. Sports Exerc.* 24(10):1167-1172, 1992.

Gretebeck, R.J., H.J. Montoye, D.L. Ballor et al. Coment on heart rate recording in field studies. *J. Sports Med. Phy. Fit.* 31(4):629-631, 1991.

Griffiths, M. and P.R. Payne. Energy expenditure in small children of obese and nonobese parents. *Nature.* 260:698-700, 1976.

Guyton A.C. *Textbook of Medical Physiology.* Philadelphia: Saunders, 1986.

Hartmann D.P., ed. *Using Observers to Study behavior.* San Francisco, Calif: Jossey Bass, 1982.

Haskell, W.L., M.C. Yee, A. Evans et al. Simultaneous measurements of heart rate and body motion to quantitate physical activity. *Med. Sci. Sports Exerc.* 25(1):109-115, 1993.

Hay J.A. *Children's self perceptions about activity. Confidence and enjoyment.* Unpublished doctoral dissertation. McMaster University, 1990.

Haymes, E.M., E.R. Buskirk, J.L. Hodgson et al. Heat tolerance of exercising lean and heavy prepubertal girls. *J. Appl. Physiol.* 36:566-571, 1974.

Haymes, E.M., R.J. McCormick, and E.R. Buskirk. Heat tolerance of exercising lean and obese prepubertal boys. *J. Appl. Physiol.* 39:457-461, 1975.

Karvonen, J., J. Chwalbinska-Moneta, and S. Saynajakangas. Comparison of heart rate measured by ECG-recorder and microcomputer. *Phys. Sportsmed.* 12(6):65-69, 1984.

Karvonen, J. and T. Vuorimaa. Heart rate and exercise intensity during sports activities: Practical Application. *Sports Med.* 5:303-312, 1988.

Kastango, K.B. and P.S. Freedson. Validation of a kilocalorie conversion of Caltrac activity counts. *Med. Sci. Sports Exerc.* [Abstract]. 23:S61 1991.

Klesges, L.M. and R.C. Klesges. The assessment of children's physical activity: a comparison of methods. *Med. Sci. Sports Exerc.* 19(5):511-517, 1987.

Klesges, R.C., T.J. Coates, L.M. Klesges et al. The FATS: An observational system for assessing physical activity in children and associated parent behaviour. *Behavioral Assessment.* 6:333-345, 1984.

Klesges, R.C., L.H. Eck, C. Hanson et al. Effects of obesity, social interactions, and physical environment on physical activity in preschoolers. *Health Psychol.* 9(4):435-449, 1990.

Klesges, R.C., K. Haddock, and L.H. Eck. A multimethod approach to the measurement of childhood physical activity and its relationship to blood pressure and body weight. *J. Pediatr.* 116:888-893, 1990.

Klesges, R.C., L.M. Klesges, A.M. Swenson et al. A validation of two motion sensors in the prediction of child and adult physical activity levels. *Am. J. Epidemiol.* 122:400-410, 1985.

Klesges, R.C., J.M. Malott, P.F. Boschee et al. The effects of parental influences on children's food intake, physical activity and relative weight. *International Journal of Eating Disorders.* 5:335-346, 1986.

Kucera M. Spontaneous physical activity in preschool children. *In: Children and Exercise XI,* R.A. Binkhorst, H.C.G. Kemper, and W.H.M. Saris (Eds.). Champaign, IL: Human Kinetics Publishers, Inc., 1985, pp.175-182.

LaPorte, R.E., J.A. Cauley, C.M. Kinsey et al. The epidemiology of physical activity in children, college students, middle-aged men, menopausal females and monkeys. *J. Chron. Dis.* 35(10):787-795, 1982.

LaPorte, R.E., H.J. Montoye, and C.J. Casperson. Assessment of physical activity in epidemiological research: Problems and prospects. *Publ. Hlth. Rep.* 100:131-146, 1985.

Leger, L. and M. Thivierge. Heart rate monitors: validity, stability, and functionality. *Phys. Sportsmed.* 16(5):143-146, 1988.

Livingstone, M.B.e., A.M. Prentice, W.A. Coward et al. Simultaneous measurements of free-living energy expenditure by the doubly labeled water method and heart-rate monitoring. *Am. J. Clin. Nutr.* 52:59-65, 1990.

McArdle W.D., F.I. Katch, and V.L. Katch. *Exercise Physiology: Energy, Nutrition and Human Performance.* Philadelphia: Lea & Febiger, 1986.

McKenzie, T.L. Observational measures of children's physical activity. *J. Sch. Health.* 61(5):224-227, 1991.

McKenzie, T.L., J.F. Sallis, and P.R. Nader. SOFIT: System for observing fitness instruction time. *J. Teach. Phys. Ed.* 11:195-205, 1991.

McKenzie, T.L., J.F. Sallis, P.R. Nader et al. BEACHES: An observational system for assessing children's eating and physical activity behaviours and associated events. *Journal of Applied Behaviour Analyses.* 24:141-151, 1991.

Meijer, G.A., K.R. Westerterp, H. Koper et al. Assesment of energy expenditure by recording heart rate and body acceleration. *Med. Sci. Sports Exerc.* 21(3):343-347, 1989.

Meijer, G.A., K.R. Westerterp, F.M.H. Verhoeven et al. Methods to assess physical activity with special references to motion sensors and accelerometers. *IEEE Trans. Boim. Eng.* 38:221-229, 1991.

Meijer G.A., K.R. Westerterp, L. Wouters et al. Validity of the accelerometer in the field: A comparison with the doubly labeled water technique.. *In: Physical activity implications for human energy metabolism,* Anonymous (Eds.). Meijer, G.A. Thesis: University of Maastricht, 1990, pp.51-61.

Mimura, K. Physiological study on the physical activity of children in kindergarten. *Osaka Research Journal of Physical Education.* 27:21-29, 1989.

Mimura, K., H. Hebestreit, and O. Bar-Or. Activity and heart rate in preschool children of low and high motor ability: 24-hour profiles. *Med. Sci. Sports Exerc.* 23:S12 1991.

Mimura, K., M. Kazuya, Y. Tatsumoto et al. Study on the activity of kindergarten children (3) - at kindergarten, at home, and during sleeping. *Memoirs of Osaka Kyoiku University, Series IV.* 36:57-65, 1987.

Montoye H.J. Risk indicators for cardiovascular disease in relation to physical activity in youth. *In: Children and Exercise, XI,* Binkhorst et al. (Eds.). Champaign, IL: Human Kinetics, 1985, pp.3-25.

Montoye H.J., S.B. Servais, and J.G. Webster. Estimation of energy expenditure from a force platform and an accelerometer. *In: Sports Science,* J. Watkins, T. Reilly, and L. Burwitz (Eds.). London: Spon, 1986, pp.375-380.

Montoye, H.J. and H.L. Taylor. Measurement of physical activity in population studies: A review. *Hum. Biol.* 56(2):195-216, 1984.

Montoye, H.J., R. Washburn, S. Servais et al. Estimation of energy expenditure by a portable accelerometer. *Med. Sci. Sports Exerc.* 15(5):403-407, 1983.

Moore, L.L., D.A. Lombardi, M.J. White et al. Influence of parents' physical activity levels on activity levels of young children. *J. Pediatr.* 118:215-219, 1991.

Morehouse L.E. and A.T. Miller. *Physiology of Exercise.* St. Louis: C.V. Mosby, 1976.

Mukeshi, M., B. Gutin, W. Anderson et al. Validation of the Caltrac movement sensor using direct observation in young children. *Ped. Exer. Sci.* 2:249-254, 1990.

Newman, W.P., D.S. Freedman, and A.W. Voors. Relation of serum lipoprotein levels and systolic blood pressure to early atherosclerosis: the Bogalusa Heart study. *N. Eng. J. Med.* 314:138-144, 1986.

Noland, M., F. Danner, K. DeWalt et al. The measurement of physical activity in young children. *Res. Q. Exerc. Sport.* 61(2):146-153, 1990.

O'Hara, N.M., T. Baranowski, B.G. Simons-Morton et al. Validity of the observation of children's physical activity. *Res. Q. Exerc. Sport.* 60(1):42-47, 1989.

Orenstein, D.M., P.A. Nixon, R.A. Washburn et al. Measuring physical activity in children with Cystic Fibrosis: A comparison of four methods. *Ped. Exer. Sci.* 5:125-133, 1993.

Paffenbarger, Jr., R.T. Hyde, A.L. Wing et al. Physical activity, all-cause mortality, and longevity of college alumni. *N. Eng. J. Med.* 314:605-613, 1986.

Passmore, R. and J. Durnin. Human energy expenditure. *Hsiological Reviews.* 35:801-840, 1955.

Pate, R.R., M. Dowda, and J.G. Ross. Associations between physical activity and physical fitness in American children. *Am. J. Dis. Child.* 144:1123-1129, 1990.

Perusse, L., A. Tremblay, A. Leblanc et al. Genetic and environmental influences on habitual physical activity and exercise participation. *Am. J. Epidemiol.* 129:1012-1022, 1989.

Powell, K.E., P.D. Thompson, C.J. Casperson et al. Physical activity and the incidence of coronary heart disease. *Annu. Rev. Public Health.* 8:253-287, 1987.

Puhl, J., K. Greaves, M. Hoyt et al. Children's activity rating scale (CARS): Description and calibration. *Res. Q. Exerc. Sport.* 61(1):26-36, 1990.

Riddoch C., C. Mahoney, N. Murphy et al. Validation of an activity diary to assess children's moderate and vigorous physical activity. *In: Children and Exercise XVI. Pediatric Work Physiology,* J. Coudert and E. VanPraagh (Eds.). Paris: Masson, 1992, pp.115-117.

Ross, J.G. and R.R. Pate. The National Children & Youth Fitness Study: A summary of findings. *JOPERD.* 58:51-56, 1987.

Rowland T.W. *Exercise and Children's Health.* Champaign, IL: Human Kinetics, 1990.

Russell S.J., C. Hyndford, and A. Beaulieu. *Active living for Canadian children and youth: a statistical profile.* Ottowa: Canadian Fitness and Lifestyle Research Institute, 1992.

Sallis, J.F. Self-report measures of children's physical activity. *J. Sch. Health.* 61:215-219, 1991.

Sallis, J.F., M.J. Buono, and P.S. Freedson. Bias in estimating caloric expenditure from physical activity in children: Implications for epidemiological studies. *Sports Med.* 11(4):203-209, 1991.

Sallis, J.F., M.J. Buono, J.A. Roby et al. Reliability and Validity of the Caltrac acclererometer as a physical activity monitor for children. *Med. Sci. Sports Exerc.* [Abstract]. 21:S112 1989.

Sallis, J.F., M.J. Buono, J.J. Roby et al. The Caltrac accelerometer as a physical activity monitor for school-age children. *Med. Sci. Sports Exerc.* 22(5):698-703, 1990.

Sallis J.F. and P.R. Nader. Family determinants of health behaviors. *In: Health Behaviors: Emerging Research Perspectives,* D.S. Gochman (Eds.). New York: Plenum, 1988, pp.107-124.

Sallis, J.F., T.L. Patterson, M.J. Buono et al. Relation of cardiovascular fitness and physical activity to cardiovascular disease risk factors in children and adults. *Am. J. Epidemiol.* 127:933-941, 1988.

Sallis, J.F., B.G. Simons-Morton, E.J. Stone et al. Determinants of physical activity and intervention in youth. *Med. Sci. Sports Exerc.* 24(6) Suppl:S248-S257, 1992.

Saris, W.H.M. The assessment and evaluation of daily physical activity in children. A review.. *Acta Pediatr. Scand. Suppl.* 318:37-48, 1985.

Saris, W.H.M. Habitual physical activity in children: Methodology and findings in health and disease. *Med. Sci. Sports Exerc.* 18(3):253-263, 1986.

Saris, W.H.M. and R.A. Binkhorst. The use of pedometer and actometer in studying daily physical activity in man. Part 1: Reliability of pedometer and actometer. *Eur. J. Appl. Physiol.* 37:219-228, 1977.

Schulz, S., K.R. Westerterp, and K. Bruck. Comparison of energy expenditure by the doubly labeled water technique with energy intake, heart rate, and activity recording in man. *Am. J. Clin. Nutr.* 49:1146-1154, 1989.

Schutz, Y., F. Froidevaux, and E. Jequier. Estimation of 24 h energy expenditure by a portable accelerometer. *Proc. Nutr. Soc.* [Abstract]. 47:23 1988.

Simons-Morton, B.G. and I.W. Huang. Heart rate monitor and Caltrac assessment of moderate-to-vigorous physical activity among preadolescent children. *Med. Sci. Sports Exerc.* [Abstract]. 23:S60 1991.

Simons-Morton, B.G., G.S. Parcel, and N.M. O'Hara. Childhood health-related physical fitness: Status and reccomendations. *Annu. Rev. Public Health.* 9:403-425, 1988.

Spurr, G.B., M. Barac-Nieto, J.C. Reins et al. Marginal malnutrition in school-aged Columbian boys: efficiency of treadmill walking in sbmaximal exercise. *Am. J. Clin. Nutr.* 39:452-459, 1984.

Strand, B. and S. Reeder. Using heart rate monitors in research on ftness levels of children in physical education. *J. Teach. Phys. Ed.* 12:215-220, 1993.

Tell, G.S. and O.D. Vellar. Physical fitness, physical activity, and cardiovascular disease risk factors in adolescents: the Oslo Youth Study. *Prev. Med.* 17:12-24, 1988.

Thivierge, M. and L. Leger. Critical review of heart rate monitors. *CAHPER J.* 55:26-31, 1989.

Treiber, F.A., L. Musante, S. Hartdagan et al. Validation of a heart rate monitor with children in laboratory and feild settings. *Med. Sci. Sports Exerc.* 21(3):338-342, 1989.

Tsanakas, J.N., O.M. Bannister, A.W. Boon et al. The "Sport-Tester": A device for monitoring the free running test. *Arch. Dis. Childhood.* 61:912-914, 1986.

U.S. Dept. of Health and Human Services. *Physical activity and fitness. Healthy People, 2000, National Healt Promotion and Disease Prevention Objectives.* DHHS Pub no (PHS) 91-50212. Washington, D.C.: U.S. Government Printing Office, 93-110, 1991.

Vander A.J., J.H. Sherman, and D.S. Luciano. *Human Physiology: The Mechanisms of Body Function.* New York: McGraw-Hill, 1985.

Wajciechowski, J.A., R.C. Gayle, R.L. Andrews et al. The accuracy of radio telemetry heart rate monitoring during exercise. *Clin. Kines.* Summer:9-12, 1991.

Washburn, R.A., T.C. Cook, and R.E. LaPorte. The objective assessment of physical activity in an occupationally active group. *J. Sports Med. Phy. Fit.* 29:279-284, 1989.

Waxman, M. and A. Stunkard. Caloric intake and expenditure in obese boys. *J. Pediatr.* 96:187-193, 1980.

Weiss, M.R. and B.J.L. Bredemeier. Moral development in sport. *Exerc. Sport Sci. Rev.* 18:331-378, 1990.

Williams, C.L., B.J. Carter, and E.L. Wynder. Prevalence of selected cardiovascular and cancer risk factors in a pediatric population: the "Know Your Body" Project, New York. *Prev. Med.* 10:235-250, 1981.

Williams, E., R.C. Klesges, C.L. Hanson et al. A prospective study of the reliability and convergent validity of three physical activity measures in a field research trial. *J. Clin. Epidemiol.* 42(12):1161-1170, 1989.

Wynder, E.L. Coronary artery disease prevention: Cholesterol, a pediatric perspective. *Prev. Med.* 18:323-409, 1989.

Yoesting, D.R. and J.E. Christensen. reexamining the significance of childhood recreation patterns on adult leisure behavior. *Leisure Sciences.* 1:219-229, 1978.

Young A. Exercise, fitness, and recovery from surgery, disease or infection. *In: Exercise, Fitness and Health: A Concensus of Current Knowledge,* C. Bouchard, R.J. Shephard, T. Stephens et al. Champaign, IL: Human Kinetics Books, 1990, pp.589-600.

Figure 1. Direct method for measuring metabolic rate. The water flowing through the calorimeter carries away the heat produced by the person's body. The amount of heat is calculated from the total volume of water and the difference between inflow and outflow temperatures.
From Vander et al., 1985.

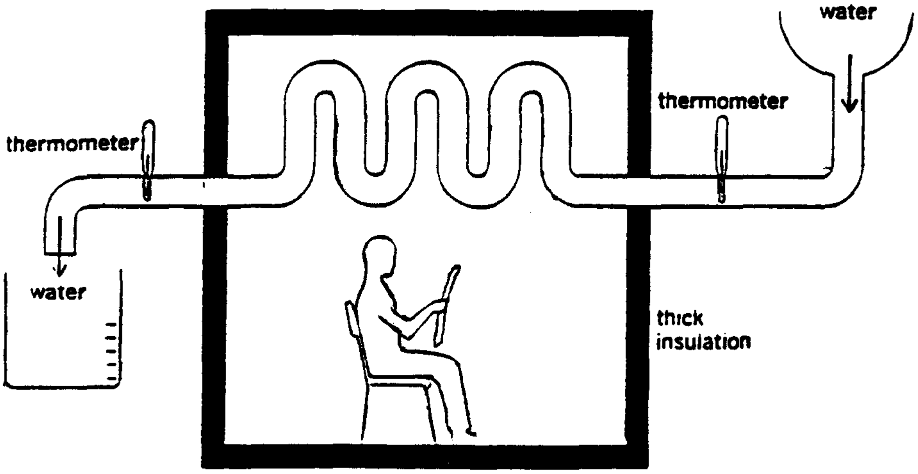


Fig 1

Figure 2. Indirect method for measuring metabolic rate. The calculation depends upon the basic principle that when 1 L of oxygen is utilized in the oxidation of organic nutrients, approximately 4.8 kcal is liberated.
From Vander et al., 1985 (103).

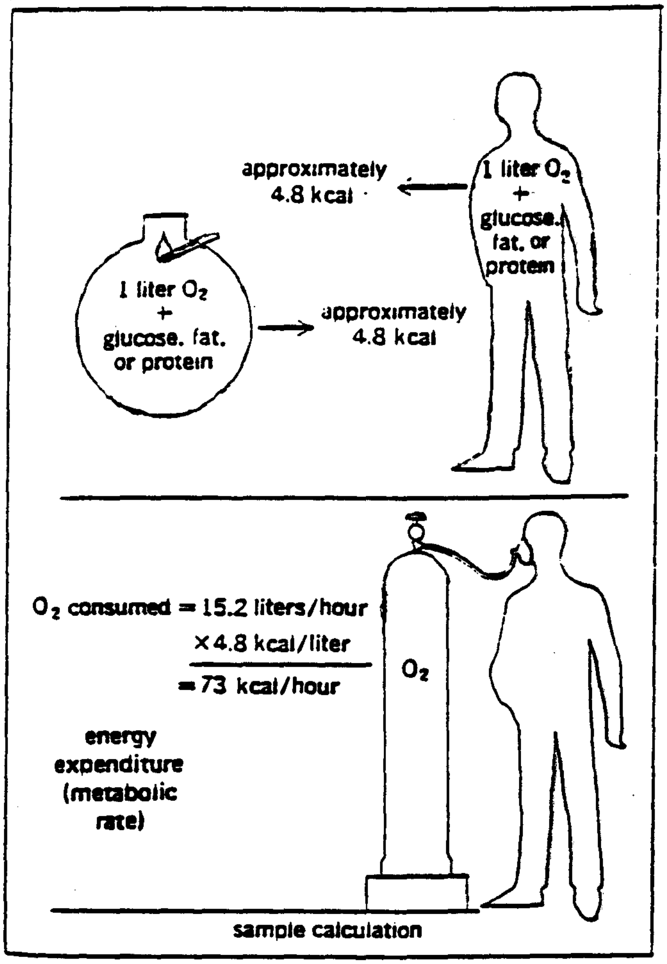

approximately
4.8 kcal
1 liter O₂
+
glucose, fat, or protein
1 liter O₂
+
glucose, fat, or protein
approximately
4.8 kcal
O₂ consumed = 15.2 liters/hour
× 4.8 kcal/liter
= 73 kcal/hour
O₂
energy expenditure (metabolic rate)
sample calculation

Fig 2

Figure 3. The electrodes of the microcomputer used to measure heart rate are placed on the chest. The microcomputer displaying the heart rate is on the wrist.
Adapted from Karvonen and Vuorimaa, 1988 (38).

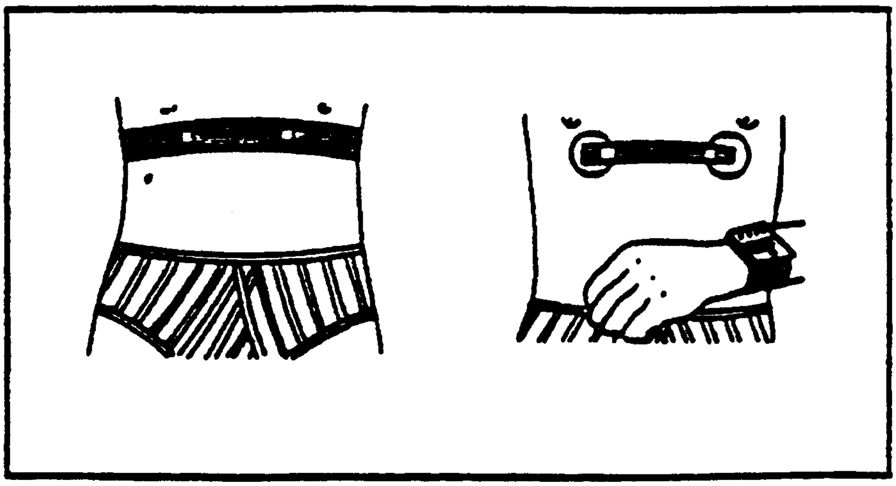


Fig 3

Figure 4. A heart rate listing of the raw data accumulated by the Polar Vantage XL during a 1 hr 38 min session. The recordings are taken at 5 second intervals. Each row in the listing represents 12 values for each minute of HR recording.

```
------------------------------------------------------------------------------
                              HEART RATE LISTING
------------------------------------------------------------------------------
  Copyright by POLAR ELECTRO              SOURCE FILE: Noname
 Time    Heart Rate Values
00:00    142    133    132    123    120    120    116    113    102    101    105    112
00:01    114    116    119    114    115    123    128    120    115    113    115    114
00:02    116    121    121    118    106    101    103    103    111    123    136    138
00:03    143    150    152    156    148    132    129    118    115    115    111    108
00:04    105    102    107    120    128    129    140    141    139    142    138    139
00:05    123    115    111    114    124    113    128    126    116    113    105    131
00:06    137    141    154    157    159    156    154    154    153    151    143    138
00:07    145    120    113    120    131    139    151    157    163    167    169    168
00:08    166    161    152    132    121    114    116    128    144    155    165    176
00:09    182    182    182    182    184    179    174    174    158    134    125    115

                              Final Time:  01:38:25.6, HR 113
-Intermediate Times:----------------------------------------------------------
```

Fig 4

Figure 5. The orientation of the x, y and z directions of the Tritrac-R3D when worn on a subjects's belt.

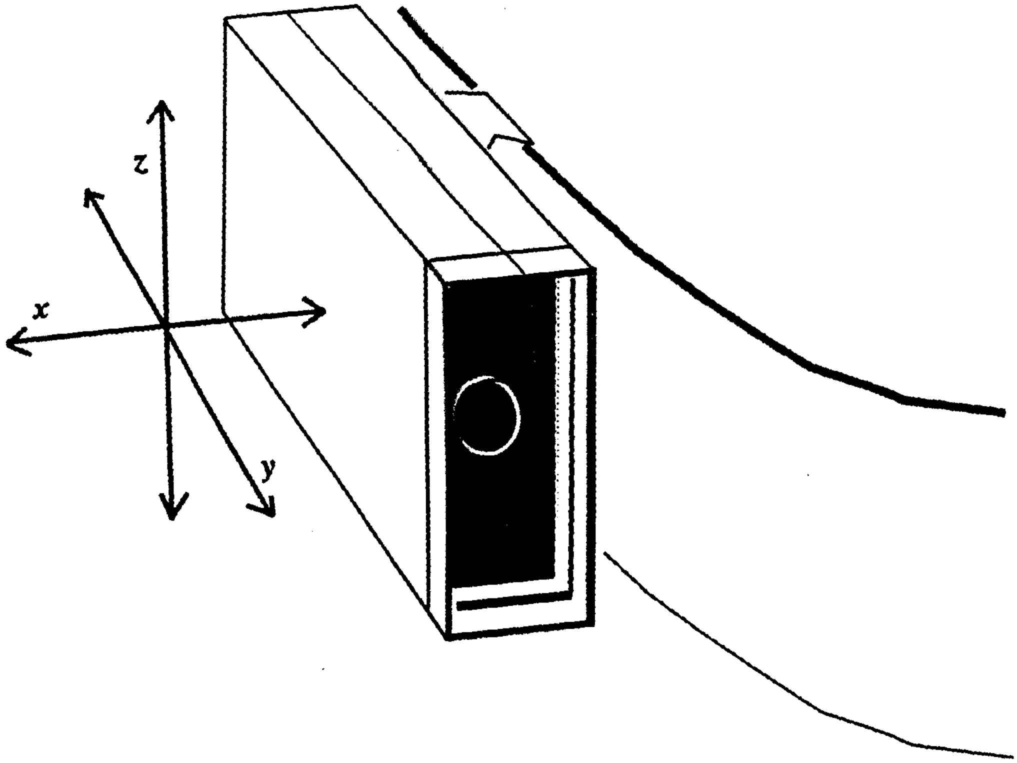


Fig 5

Figure 6. An example of the resulting DOS printout produced by the Tritrac-R3D. The first section of the printout will be the user parameters. Next there are seven collumns: date and time of each data point, X activity (Cnt1), Y activity (Cnt2), Z activity (Cnt 3), vector magnitude (Vec.Mag.), activity kcals (Act.Cals), and metabolic kcals (Met.Cals). The activity count in each of the X, Y and Z planes represent acceleration data in each of the three ordinal planes. These three counts of acceleration data are strictly at right angles to each other, and do not necessarily represent the frontal, sagittal and transverse planes. The vector magnitude represents the calculated composite movement for all three directions; this was the criterion value used for all analyses in this study. Activity kcals represent the calculated energy expenditure of the subject per interval based on the subject's age, height, weight and gender. Metabolic kcals represent the computed basal metabolic energy expenditure per interval.

```
                    ID 46129
                 Name TT2 46129
           Time Stamp 22-NOV-93 09:19
                    Age 05
                  Gender M
                Height (cm) 107
                Weight (kg) 018
          Minutes per Interval 01
           Comment 1 Lab Visit, Arif's Subject
             Comment 2 Tali Assissting
        Initial Battery Reading 138   92%
       Current Battery Reading 140   93%
                 Error Flag N

  Date & Time        Cnt1    Cnt2    Cnt3  Vec.Mag  Act.Cals  Met.Cals
22-NOV-93 10:05        46     113     127      176      1.92      0.50
22-NOV-93 10:06       247     575     792     1009     11.30      0.50
22-NOV-93 10:07       662    1226    2039     2469     27.79      0.50
22-NOV-93 10:08      1130    1472    2244     2911     32.88      0.50
22-NOV-93 10:09      1428    1456    2072     2907     32.77      0.50
22-NOV-93 10:10       426     341     581      797      8.92      0.50
22-NOV-93 10:11       166      86     315      366      4.06      0.50
22-NOV-93 10:12        33      88      38      101      1.13      0.50
22-NOV-93 10:13         3      60       8       60      0.67      0.50
22-NOV-93 10:14       110     134     117      209      2 26      0.50
22-NOV-93 10:15        74      74      64      122      1.35      0.50
22-NOV-93 10:16        42      56      35       78      0.79      0.50
22-NOV-93 10:17        24      48      34       63      0.67      0.50
22-NOV-93 10:18        89     139     180      244      2.71      0.50
22-NOV-93 10:19       402     699     528      963     10.84      0.50
22-NOV-93 10:20       583     758     871     1293     14.57      0.50
22-NOV-93 10:21       748    1065     775     1514     17.06      0.50
```

Fig 6

Figure 7. A sample of a BEACHES coding sheet. The only category used in this study was the activity level section. This coded for activity between 1 and 5.

	1	2	3	4	5	6	7	8	9	10
1.0 ENVIRONMENT										
1 ALONE										
2 MOTHER										
3 FATHER										
4 SIBLING(S)										
5 PEER(S)										
6 TEACHER(S)										
7 OTHER ADULT(S)										
2.0 PHYS.LOCATION										
1 INSIDE HOME										
2 OUTSIDE HOME										
3 OUTSIDE GENERAL										
4 PLYGRND/PLYSPC										
5 INSIDE SCHOOL										
6 CAFETERIA										
7 OUTSIDE SCHOOL										
3.0 ACTIVITY LEVEL										
1 LYING DOWN										
2 SITTING										
3 STANDING										
4 WALKING										
5 VERY ACTIVE										
	1	2	3	4	5	6	7	8	9	10

Fig 7

Figure 8. Relationship between accelerometry as measured by the Tritrac-R3D and heart rate as measured by the Polar Vantage XL for a 5.4 year old female (subject I.D. - 542225-D2) in a daycare setting.

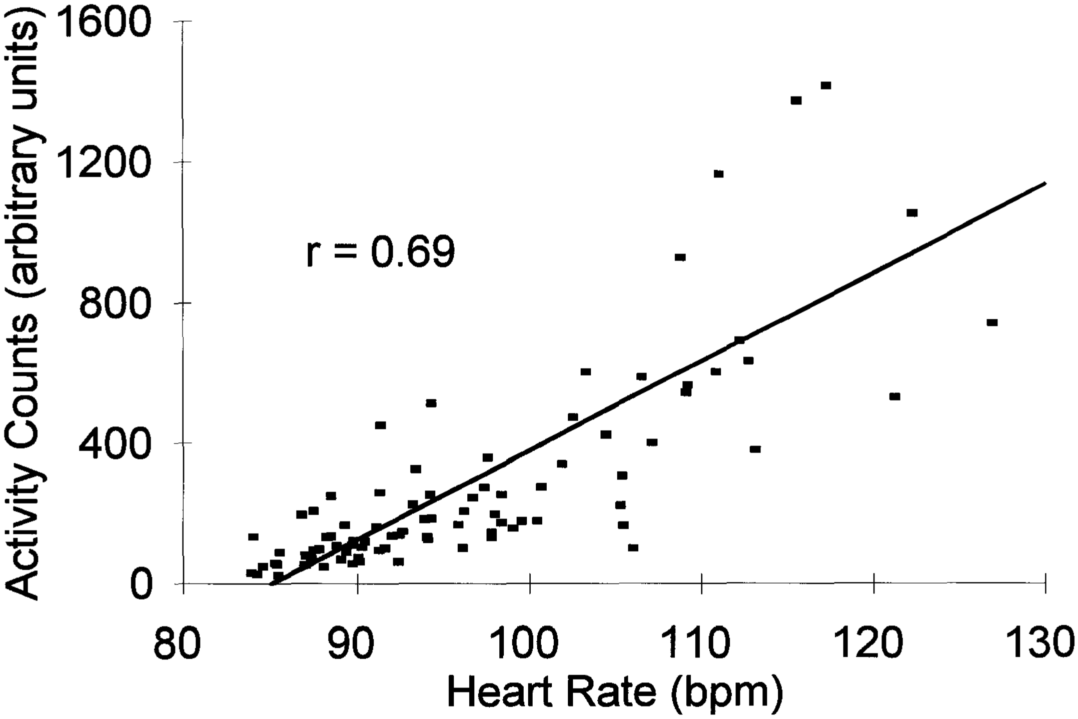

1600
1200
800
400
0
Activity Counts (arbitrary units)
r = 0.69
80
90
100
110
120
130
Heart Rate (bpm)

Fig 8

Figure 9. Relationship between accelerometry as measured by the Tritrac-R3D and direct observation as measured by the BEACHES observational system in a daycare setting.

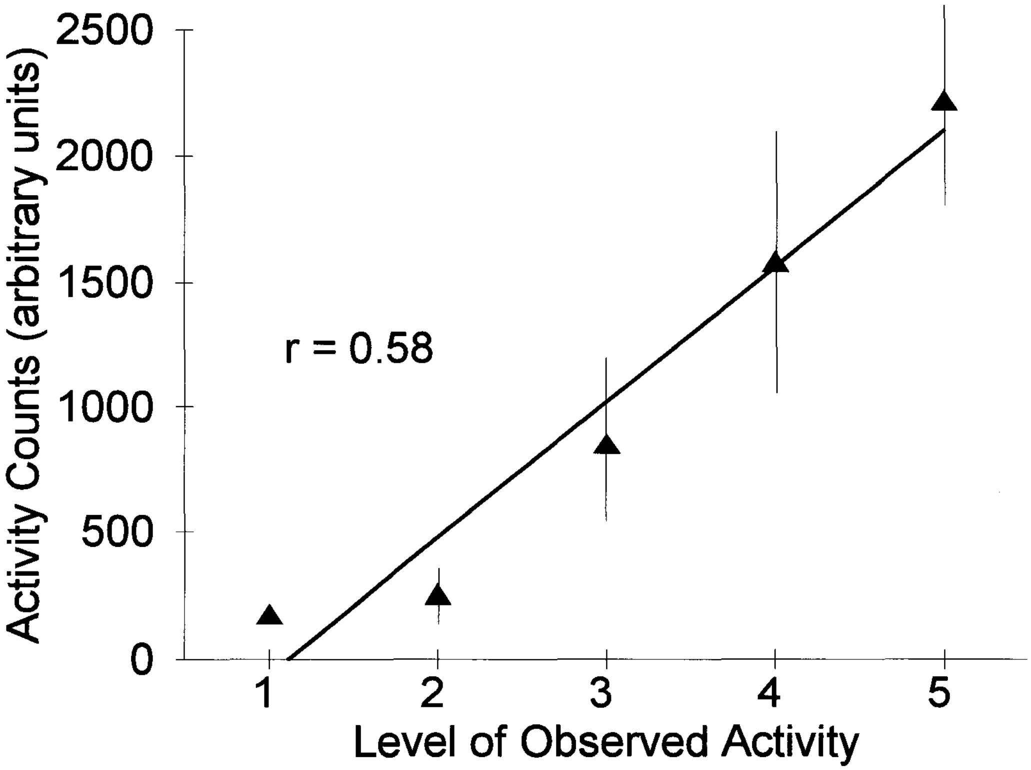

2500
2000
1500
1000
500
0
Activity Counts (arbitrary units)
r = 0.58
1
2
3
4
5
Level of Observed Activity

Fig 9

Figure 10. Relationship between heart rate as measured by the Polar Vantage XL and direct observation as measured by the BEACHES observational system in a daycare setting.

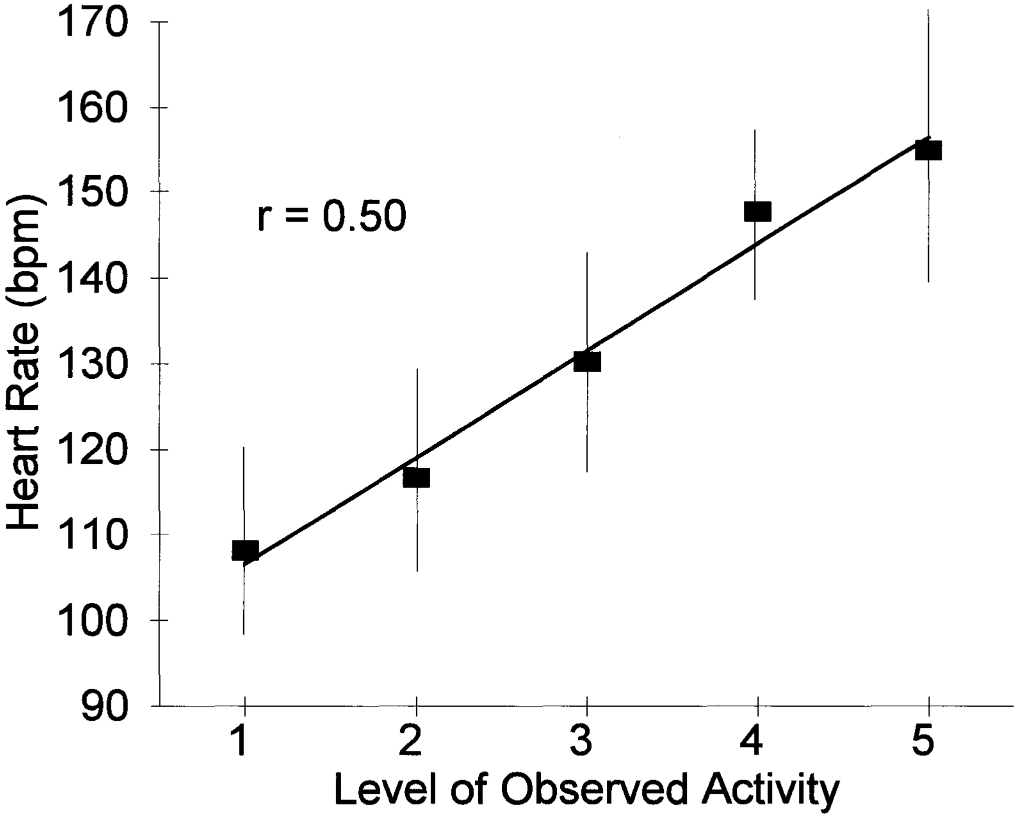

170
160
150
140
130
120
110
100
90
Heart Rate (bpm)
r = 0.50
1
2
3
4
5
Level of Observed Activity

Fig 10

Figure 11. A comparison of the relationship between HR and Obs found in the present study to that found by McKenzie et al., 1991.

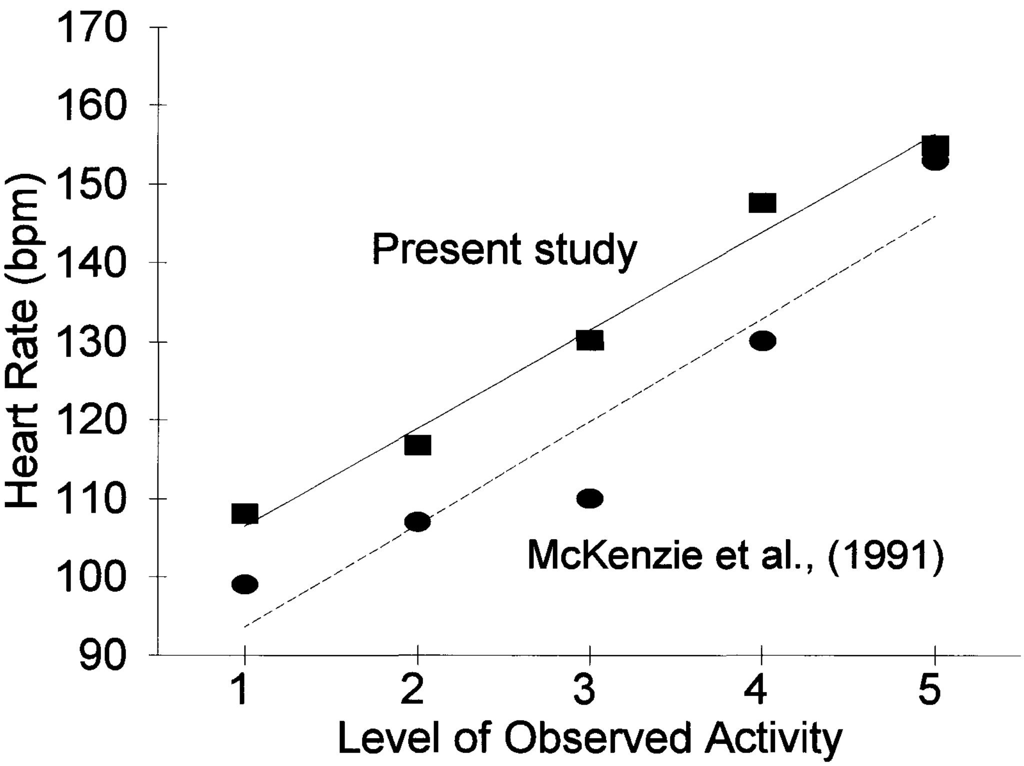

Heart Rate (bpm)
170
160
150
140
130
120
110
100
90
1
2
3
4
5
Level of Observed Activity
Present study
McKenzie et al., (1991)

Fig 11

Table 1. Commonly used methods for assessing habitual physical activity

Measures of Habitual Activity		**EE Estimates**		**Activity Pattern Description**	
Measure Classification		*Concurrent*	*Retroactive*	*Concurrent*	*Retroactive*
Physiological	Heart Rate Monitors	✔		✔	
	Doubly labeled Water	✔			
	Calorimetric Chamber	✔			
Physical	Pedometers, Actometers			✔	
	Motion Sensors			✔	
	Accelerometers	✔		✔	
Observational	Time and Motion	✔		✔	
	Behavioral Observation	✔		✔	
Pen and Paper	Self-Report		✔		✔
	Interviews		✔		✔
	Caloric Intake		✔		
	Self-Keeping Logs	*	✔	*	✔

Table 2. Strengths and weaknesses of selected methods of habitual activity assessment

Measure	***Advantages***	***Disadvantages***
Heart Rate Monitoring	• Fairly linear relationship with VO_2 • Measures duration, intensity • Continuous, long term monitoring possible • fairly low reactivity • Objective • Reliable • Reasonably Economical	• Need to establish individual HR-VO_2 regressions • HR-VO_2 relationship altered with climate, muscle group employed or emotional state • HR-VO_2 relationship poorest at low HR • Type of activity not recorded
Accelerometry	• Objective • Measures duration, intensity • Continuous, long term monitoring possible • fairly low reactivity • Reliable	• Nonspecific, no discrimination of different forces • Validity dependant on placement
Behavioral Observation	• High validity for activity pattern description • Allows measures of type, duration, frequency and subjective intensity • Ability to record complex activities	• Reactive (unless covert) • High manpower cost • Need to establish high inter- and intra-observer reliabilities • Time consuming • Activities hard to quantify

Table 3. Characteristics of selected instruments designed for observing children's PA

Location: Identifies settings where the instrument has been used. "Diverse" indicates the system may be used in any location.

Observation Strategy: Identifies the specific observation method the system uses. Momentary time-sampling refers to coding events occurring at the end of the observation interval. Partial time-sampling refers to coding events that have occurred during a specific time interval, usually 5-20 seconds.

Activity Categories: Specifies the number and briefly identifies the activity categories coded.

Validation: Identifies specific child variables correlated with the activity codes to permit EE to be estimated from observed data.

Data Summary: Identifies the primary summary variables. Additional summary variables may be calculated for some instruments.

Associated Variables: Identifies associated variables coded simultaneously with the PA levels.

Training Time: Identifies the relative amount of time necessary for initially training observers to use the instruments reliably. Low, 10-15 hours; medium, 15-25; and high, 25-40.

Test Site: Identifies the setting where subjects were observed during field tests of the instrument.

Subjects: Identifies characteristics of subjects in the field test.

Reliability: Identifies types and results of reliabilities reported in field tests. reliability in systematic observation typically refers to the degree two or more persons simultaneously viewing an activity using the same behaviour definitions and coding conventions record the same codes. I-I refers to interval-by-interval reliabilities. Kappa is a statistic that takes chance agreement into consideration. I-I scores are typically higher than Kappa scores.

Adapted From McKenzie, 1991.

	FATS (Klesges et al., 1984) (Fargo Activity Timesampling Survey)	CARS (Puhl et al., 1990) (Children's Activity Rating Scale)	BEACHES (McKenzie et al., 1991) (Behaviors of Eating and Activity for Children's Health Evaluation System)	SCAN CATS (Klesges et al., 1990a) (Studies of Children's Activity and Nutrition Children's Activity Timesampling Survey)
Location	home	diverse	diverse	diverse
Observation Strategy	partial interval recording: 10-second obs/rec intervals	partial time-sampling; 1-minute intervals	momentary time-sampling; 1-minute intervals	momentary time-sampling; 10-second obs/rec intervals
Activity Categories	8 sleeping; lying down; sitting; crawling; climbing; standing; walking; and running: each with 3 intensity levels - minimal; moderate; and extreme.	5 stationary, no movement; stationary, movement; translocation easy; translocation, moderate; and translocation, strenuous.	5 lying; sitting; standing; walking; and very active.	4 stationary; minimal activity; slow movement; and rapid movement.
Validation	none; correlated with LSI readings	heart rate; VO_2	heart rates	none
Data Summary	% intervals	% intervals; Kcal/Kg/min	% intervals/time; Kcal/Kg/min	% intervals/time
Associated Variables	location; persons present; interactors; interactions; and child response.	location; others present; television; eats; interactors; prompts; and consequences	location; TV; interactors; eating; prompts; consequences; and child responses.	location; persons present; interactors; and prompts
Training Time	high	high	high	high
Test Site	home	diverse	diverse	home
Subject	children, ages 20-48 months	children, ages three-six	children ages four-nine	children, ages three-six
Reliability	I-I (91%-98%) Kappa (0.90)	I-I (84.1%)	Kappa (0.91)	field scores, NA; videotape Kappa (0.91)

Table 4. Validity and reliability of combined methods to assess activity in children Adapted from Freedson, 1991

Study	N	Ages (years)	Criterion Measure	Tested Measure	Reliability	Validity
Saris et al. (1977)	11	4-6	direct observation and teacher questionnaire	pedometer actometer	NR	r = .93* r = .97*
Fenster et al. (1989)	5 males 13 females	6-8	peak oxygen uptake	LSI	NR	r = .59*
LaPorte et al. (1982)	22 males	12-14	three-day food intake Taylor Leisure Time Activity Survey	LSI	NR	r = .30 r = .16
Klesges et al. (1985)	18 males 12 females	1.7-5.7	LSI Caltrac	direct observation	NR	r = .40* (LSI) r = .35* (Caltrac)
O'Hara et al. (1989)	18 males 18 females	8-10	heart rate	direct observation	NR	r = .64*
Freedson and Evenson (1991)	13 males 17 females	5-9	NR	Caltrac	r = .38* (D1 vs. D2) r = .79* (D1 vs. D2) r = .42* (D1 vs. D2)	NR
Klesges and Klesges (1987)	17 males 13 females	2-4	direct observation	Caltrac	r = .90 inter-observer correlation (direct observation)	r = .57 - .95* (hourly) r = .54* (whole day)
Mukeshi et al. (1990)	11 males 9 females	2.4-3.3	direct observation	Caltrac	NR	r = .62*
Sallis et al. (1990)	20 males 15 females	8-13	heart rate	Caltrac	r = .10 (heart rate) r = .30 (Caltrac)	r = .49*
			recall of previous 10 hours	Caltrac heart rate	r = .06 (recall)	r = .49* (day 1 Caltrac) r = .39* (day 2 Caltrac) r = .54* (day 1 heart rate) r = .42* (day 2 heart rate)
Danner et al. (1991)	29 males 22 females	2.5-5.4	direct observation	Caltrac	NR	r = .83 - .86*
Klesges et al. (1990b)	122 males 100 females	3-6	direct observation	Caltrac	NR	insignificant
Puhl et al. (1990)	12 males 13 femals	5-6	direct observation	oxygen uptake heart rate	I-I (84.1%) interval-by-interval reliability	NR
McKenzie et al. (1991)	19	4-9	direct observation	heart rate	Kappa (0.91)	NR
Epstein et al. (1984)	19 females	5-8	direct observation	heart rate	93.3% (mean inter-observer percent agreement)	r = .82*

Table 5. Physical characteristics of subjects. Height and weight data were only collected for subjects that visited the laboratory, these subjects are presented in the parentheses.

Characteristic	Male	Female	Pooled
Age	4.3 ± 0.9	4.3 ± 0.7	4.3 ± 0.8
Height	106.2 ± 9.2	106.0 ± 8.2	106.2 ± 8.7
Weight	18.8 ± 5.6	17.8 ± 1.7	18.5 ± 4.7
n	30 (11)	19 (6)	49 (17)

Table 6. Individual subject correlations (Perason product moment) and r to z transformed values between techniques for assessing PA in the laboratory. TT-HR indicates Tritrac-heart rate relationship. TT-Obs indicates Tritrac-observation relationship. HR-Obs indicates heart rate-observation relationship. Correlations represent simultaneous real time (per minute) measures from the various techniques made during the observational period. Length of observational period varied between 24 min and 30 min among subjects. Tritrac values are counts/min; HR = bpm; observation values between 1 and 5.

	Correlation Matrix			**z-score Matrix**		
Subject ID	*TT-HR*	*TT-Obs*	*HR-Obs*	*TT-HR*	*TT-Obs*	*HR-Obs*
31114	0.86	0.78	0.77	1.29	1.05	1.02
32132	0.74	0.79	0.70	0.95	1.07	0.87
33118	0.78	0.79	0.72	1.05	1.07	0.91
33120	0.76	0.66	0.72	1.00	0.79	0.91
34122	0.86	0.74	0.71	1.29	0.95	0.89
37219	0.88	0.83	0.83	1.38	1.19	1.19
39128	0.88	0.69	0.83	1.38	0.85	1.19
39223	0.84	0.77	0.83	1.22	1.02	1.19
43206	0.85	0.81	0.81	1.26	1.13	1.13
43221	0.68	0.79	0.33	0.83	1.07	0.34
45127	0.86	0.51	0.45	1.29	0.56	0.49
46125	0.85	0.86	0.88	1.26	1.29	1.38
48112	0.95	0.69	0.74	1.83	0.85	0.95
48230	0.78	0.71	0.80	1.05	0.89	1.10
53107	0.87	0.83	0.85	1.33	1.19	1.26
55131	0.91	0.80	0.73	1.53	1.10	0.93
56215	0.87	0.76	0.81	1.33	1.00	1.13
mean z-score				1.25	1.00	0.99
n				17	17	17
Overall Correlation				r = .85	r = .76	r = .76
S.D.				0.05	0.07	0.09

Table 7. Individual subject correlations (Perason product moment) and r to z transformed values between techniques for assessing PA in the field. TT-HR indicates Tritrac-heart rate relationship. TT-Obs indicates Tritrac-observation relationship. HR-Obs indicates heart rate-observation relationship. Correlations represent simultaneous real time (per minute) measures from the various techniques made during the observational period. Length of observational period varied between 47 min and 90 min among subjects. Tritrac values are counts/min; HR = bpm; observation values between 1 and 5. 7a indicate first day of observation. 7b indicates second day of observation.

7a.

	Correlation Matrix			z-score Matrix		
Subject ID	***TT-HR***	***TT-Obs***	***HR-Obs***	***TT-HR***	***TT-Obs***	***HR-Obs***
31114-D1	0.55	0.55	0.55	0.62	0.62	0.62
321202-D1	0.74	0.46	0.28	0.95	0.50	0.29
321205-D1	0.54	0.73	0.55	0.60	0.93	0.62
322203-D1	0.59	0.32	0.09	0.68	0.33	0.09
32265-D1	0.54	0.58	0.52	0.60	0.66	0.58
33120-D1	0.78	0.56	0.58	1.05	0.63	0.66
34122-D1	0.48	0.64	0.19	0.52	0.76	0.19
352201-D1	0.74	0.59	0.62	0.95	0.68	0.73
352216-D1	0.58	0.60	0.70	0.66	0.69	0.87
37219-D1	0.63	0.72	0.49	0.74	0.91	0.54
372209-D1	0.76	0.59	0.64	1.00	0.68	0.76
38110-D1	0.88	0.69	0.74	1.38	0.85	0.95
38168-D1	0.74	0.70	0.60	0.95	0.87	0.69
39108-D1	0.72	0.57	0.50	0.91	0.65	0.55
39126-D1	0.49	0.43	0.61	0.54	0.46	0.71
39223-D1	0.52	0.55	0.42	0.57	0.62	0.45
39286-D1	0.67	0.60	0.05	0.81	0.69	0.05
41150-D1	0.20	0.64	0.53	0.20	0.76	0.59
41153-D1	0.75	0.56	0.64	0.97	0.63	0.76
422160-D1	0.89	0.62	0.49	1.42	0.73	0.54
44213-D1	0.79	0.69	0.77	1.07	0.85	1.02
45127-D1	0.61	0.66	0.32	0.71	0.79	0.33
46125-D1	0.35	0.58	0.18	0.37	0.66	0.18
46257-D1	0.58	0.35	0.08	0.66	0.37	0.08
47176-D1	0.56	0.49	0.44	0.63	0.54	0.47
48112-D1	0.51	0.41	0.48	0.56	0.44	0.52
48230-D1	0.39	0.42	0.23	0.41	0.45	0.23
49228-D1	0.77	0.64	0.57	1.02	0.76	0.65
53107-D1	0.67	0.59	0.47	0.81	0.68	0.51
542225-D1	0.87	0.70	0.67	1.33	0.87	0.81
48205-H1	0.66	0.42	0.61	0.79	0.45	0.71
43206-H1	0.83	0.54	0.60	1.19	0.60	0.69
56215-D1	0.84	0.65	0.51	1.22	0.78	0.56
58100-D1	0.69	0.68	0.31	0.85	0.83	0.32

7b.

	Correlation Matrix			z-score Matrix		
Subject ID	***TT-HR***	***TT-Obs***	***HR-Obs***	***TT-HR***	***TT-Obs***	***HR-Obs***
31114-D2	0.85	0.59	0.67	1.26	0.68	0.81
321202-D2	0.87	0.57	0.64	1.33	0.65	0.76
321205-D2	0.83	0.72	0.65	1.19	0.91	0.78
322203-D2	0.61	0.32	0.38	0.71	0.33	0.40
32265-D2	0.45	0.42	0.24	0.49	0.45	0.25
33120-D2	0.83	0.72	0.62	1.19	0.91	0.73
34122-D2	0.65	0.60	0.38	0.78	0.69	0.40
352201-D2	0.73	0.57	0.58	0.93	0.65	0.66
352216-D2	0.81	0.62	0.50	1.13	0.73	0.55
37219-D2	0.71	0.72	0.53	0.89	0.91	0.59
372209-D2	0.71	0.63	0.71	0.89	0.74	0.89
38110-D2	0.48	0.49	0.35	0.52	0.54	0.37
38168-D2	0.64	0.05	0.41	0.76	0.05	0.44
39108-D2	0.48	0.49	0.35	0.52	0.54	0.37
39126-D2	0.55	0.35	0.56	0.62	0.37	0.63
39223-D2	0.62	0.66	0.63	0.73	0.79	0.74
39286-D2	0.70	0.54	0.70	0.87	0.60	0.87
41150-D2	0.82	0.65	0.56	1.16	0.78	0.63
41153-D2	0.80	0.62	0.56	1.10	0.73	0.63
422160-D2	0.84	0.65	0.65	1.22	0.78	0.78
44213-D2	0.78	0.55	0.63	1.05	0.62	0.74
45127-D2	0.87	0.69	0.67	1.33	0.85	0.81
46125-D2	0.43	0.50	0.22	0.46	0.55	0.22
46257-D2	0.72	0.55	0.48	0.91	0.62	0.52
47176-D2	0.65	0.58	0.62	0.78	0.66	0.73
48112-D2	0.60	0.41	0.42	0.69	0.44	0.45
48230-D2	0.77	0.38	0.32	1.02	0.40	0.33
49228-D2	0.36	0.44	0.06	0.38	0.47	0.06
53107-D2	0.53	0.62	0.29	0.59	0.73	0.30
542225-D2	0.82	0.64	0.59	1.16	0.76	0.68
48205-H2	0.44	0.12	0.51	0.47	0.12	0.56
43206-H2	0.79	0.57	0.63	1.08	0.65	0.74
56215-D2	0.52	0.67	0.38	0.58	0.81	0.40
58100-D2	0.53	0.88	0.49	0.69	1.38	0.54
mean z-score				0.84	0.65	0.55
n				68	68	68
Overall Correlation				r = 0.69	r = 0.58	r = 0.50
S.D.				0.12	0.11	0.15

Table 8. Summary of the statistical analyses of the difference in mean z-scores between various combinations of physical activity assessment techniques determined in the laboratory. 8a is the analyses of variance. 8b is the post hoc t-tests. *indicates significance $p < 0.05$

8a.

Analysis of Variance

Summary

Groups	*Count*	*Sum*	*Average*	*Variance*
TT-HR	17	21.26	1.25	0.06
TT-Obs	17	17.06	1.00	0.03
HR-Obs	17	16.85	0.99	0.07

Source of Variation

Total	*SS*	*df*	*MS*	*F*	*P-value*	*F-crit*
Between Methods	0.73	2	0.36	7.03	0.00	3.19
Within Methods	2.48	48	0.05			
Total	3.21	50				

8b

Technique Combinations	**Combination Pairs**	
	TT-HR	TT-Obs
TT-Obs	*t = 3.03; P < 0.005	
Hr-Obs	*t = 3.47; P < 0.001	t = 0.16; P = 0.87

Table 9. Summary of the statistical analyses of the difference in mean z-scores between various combinations of physical activity assessment techniques determined in the field. 9a is the analyses of variance. 9b is the post hoc t-tests.
*indicates significance $p < 0.05$

9a.

Analysis of Variance

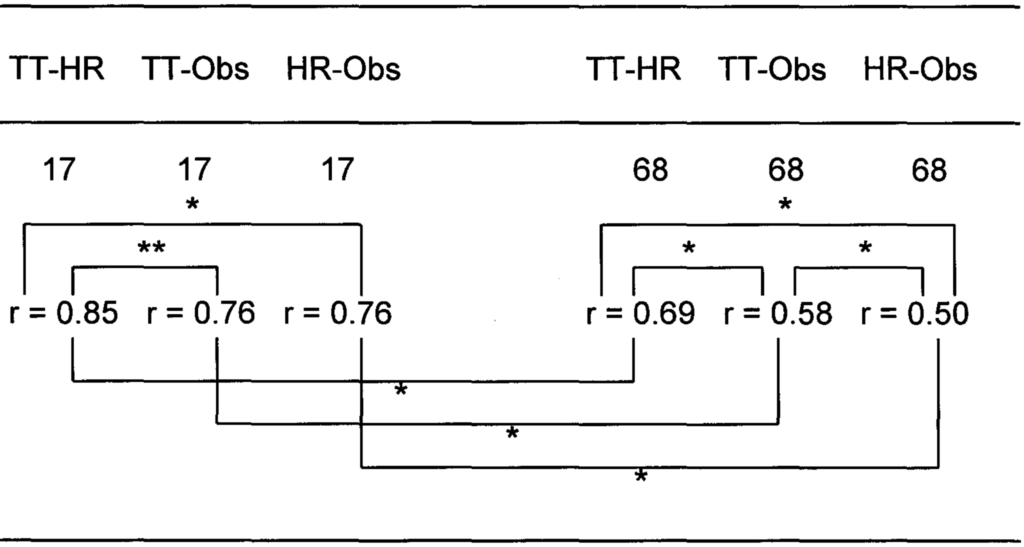
Total	*TT-HR*	*TT-Obs*	*HR-Obs*
Count	68.00	68.00	68.00
Sum	57.17	44.52	37.64
Average	1.68	1.31	1.11
Variance	0.16	0.08	0.10

Source of Variation	***SS***	***df***	***MS***	***F***	***P-value***	***F-crit***
Day	0.02	1.00	0.02	0.29	0.59	3.89
Method	2.89	2.00	1.44	24.63	0.00	3.04
Interaction	0.05	2.00	0.03	0.43	0.65	3.04
Within	11.60	198.00	0.06			
Total	14.55	203.00				

9b.

Technique Combinations	**Combination Pairs**	
	TT-HR	TT-Obs
TT-Obs	*t = 4.38; P < 0.001	
Hr-Obs	*t = 6.49; P < 0.001	*t = 2.74; P < 0.001

Table 10. Individual subject correlations (Pearson Product Moment) between techniques for assessing PA in the laboratory and in the field. TT - indicates Tritrac, HR - indicates heart rate, and Obs - indicates direct observation. Individual correlations from each observational period were transformed to z-scores; the mean z-score for the group was then transformed back to an r-value to obtain the group correlation. Length of observational periods varied between 24 - 30 min (lab) and 47 - 90 min (field) among subjects.
Tritrac values are counts/min; HR = bpm; observation values between 1 and 5.
* indicates significance $p < 0.001$, ** indicates significance $p < 0.005$.



TT-HR	TT-Obs	HR-Obs	TT-HR	TT-Obs	HR-Obs
17	17	17	68	68	68
r = 0.85	r = 0.76	r = 0.76	r = 0.69	r = 0.58	r = 0.50

Table 11. A comparison of physical activity levels from Tritrac, heart rate and observational technique on two separate occasions in the field.

	Tritrac-R3D		Polar Vantage		BEACHES	
	Day 1	Day 2	Day 1	Day 2	Day 1	Day 2
31114	293.27	551.88	127.86	122.70	2.57	3.11
321202	505.21	457.56	128.19	123.96	2.94	2.81
321205	529.61	444.19	134.55	111.28	3.31	2.82
322203	107.11	360.92	104.75	128.12	2.74	3.36
32265	553.35	259.17	138.61	129.54	3.18	3.04
33120	300.49	539.26	108.00	121.46	2.78	3.34
34122	152.39	279.81	118.92	118.16	2.04	2.61
352201	812.57	371.97	139.45	134.54	3.04	2.97
352216	284.14	212.63	112.05	98.90	2.90	2.67
37219	242.00	343.56	118.96	122.81	2.48	2.59
372209	452.05	280.49	126.27	123.36	2.95	3.04
38110	799.31	677.13	132.94	120.37	3.10	3.10
38168	265.15	781.00	126.39	146.73	2.63	3.62
39108	710.14	676.43	123.23	120.35	2.84	3.11
39126	362.74	491.78	103.66	121.84	2.85	2.97
39223	532.57	553.23	135.26	130.51	2.91	3.25
39286	745.73	491.74	142.69	139.74	2.82	3.13
41150	211.49	427.92	98.15	121.77	2.71	2.84
41153	553.79	598.54	113.79	108.95	3.21	2.99
43206	631.00	705.93	115.74	122.27	3.08	3.20
44213	480.54	455.23	126.93	125.68	3.04	2.90
45127	254.16	520.39	115.61	107.42	2.35	2.68
46125	276.13	201.06	120.50	117.97	2.63	2.46
46257	482.89	357.25	125.79	126.79	3.07	2.98
47176	686.11	552.14	137.97	130.73	3.17	3.06
48112	435.89	495.03	126.37	128.90	3.02	2.82
48205	211.89	274.50	105.16	102.94	2.34	2.64
48230	323.76	318.99	117.59	116.86	3.02	3.01
49228	363.07	211.19	119.02	112.34	2.57	2.39
53107	390.63	679.46	102.58	114.90	3.07	3.03
542225	173.75	259.32	104.95	95.67	2.47	2.61
56215	428.30	235.79	97.61	92.48	2.71	2.60
58100	168.95	209.92	85.43	94.28	2.70	2.87
Mean	415.76	432.59	119.24	119.22	2.83	2.93
S.D.	195.97	166.25	13.93	12.62	0.29	0.28
n	33	33	33	33	33	33

Table 12a. Two day comparison of level of activity for each of the three measures of PA during free outdoor play.

Instrument	Day 1 mean	Day 2 mean	t-test
Tritrac	415.76	432.59	t = 0.50; P = 0.62 (n.s.)
Heart Rate	119.24	119.22	t = 0.01; P = 0.99 (n.s.)
Observation	2.83	2.93	t = 1.87; P = 0.07 (n.s.)

Table 12b. Spearmen Brown rank order correlations between 2 days of observation for each of the three measures of PA measured during free outdoor play.

Instrument	Correlation
Tritrac	0.43
Heart Rate	0.68
Observation	0.39

Table 13. Subsample of subject z-scores with a one minute time lag in the heart rate time base. A comparison of correlations before and after manipulation.

Original z-Scores				*1 min HR time Lag*			
Subject ID	*TT-HR*	*TT-Obs*	*HR-Obs*	*Subject ID*	*TT-HR*	*TT-Obs*	*HR-Obs*
352201-D1	0.95	0.67	0.73	352201-D1	1.00	0.67	0.66
352201-D2	0.89	0.65	0.66	352201-D2	0.76	0.65	0.55
39126-D1	0.69	0.46	0.71	39126-D1	0.95	0.46	0.56
39126-D2	0.62	0.37	0.63	39126-D2	0.93	0.37	0.50
39223-D1	0.59	0.62	0.45	39223-D1	0.89	0.62	0.56
39223-D2	0.74	0.79	0.74	39223-D2	0.83	0.79	0.59
39223-Lab	1.22	1.02	1.19	39223-Lab	1.13	1.02	1.00
41153-D1	0.95	0.63	0.76	41153-D1	1.05	0.63	0.47
41153-D2	1.26	0.73	0.63	41153-D2	1.05	0.73	0.62
53107-D1	0.91	0.68	0.51	53107-D1	0.69	0.68	0.38
53107-D2	0.59	0.73	0.30	53107-D2	0.66	0.73	0.26
53107-Lab	1.33	1.19	1.26	53107-Lab	1.42	1.19	1.10
mean	0.90	0.71	0.71	***mean***	0.95	0.71	0.60
n	12	12	12	***n***	12	12	12
Correlation	r = 0.72	r = 0.61	r = 0.61	***Correlation***	r = 0.74	r = 0.61	r = 0.54
S.D.	0.11	0.13	0.14	***S.D.***	0.05	0.13	0.14

Table 14. Summary of the effects of data manipulation on correlations.

Manipulation	TT-HR	TT-Obs	HR-Obs
Removal of Transitional Minutes	*t = 5.27 P < 0.001	*t = 4.02 P < 0.001	*t = 6.34 P < 0.001
1 min HR Time Lag	t = 0.95 P = 0.36		*t = 3.81 P < 0.005
Balanced Range of Activities	t = 0.47 P = 0.65	t = 0.87 P = 0.41	t = 0.58 P = 0.58

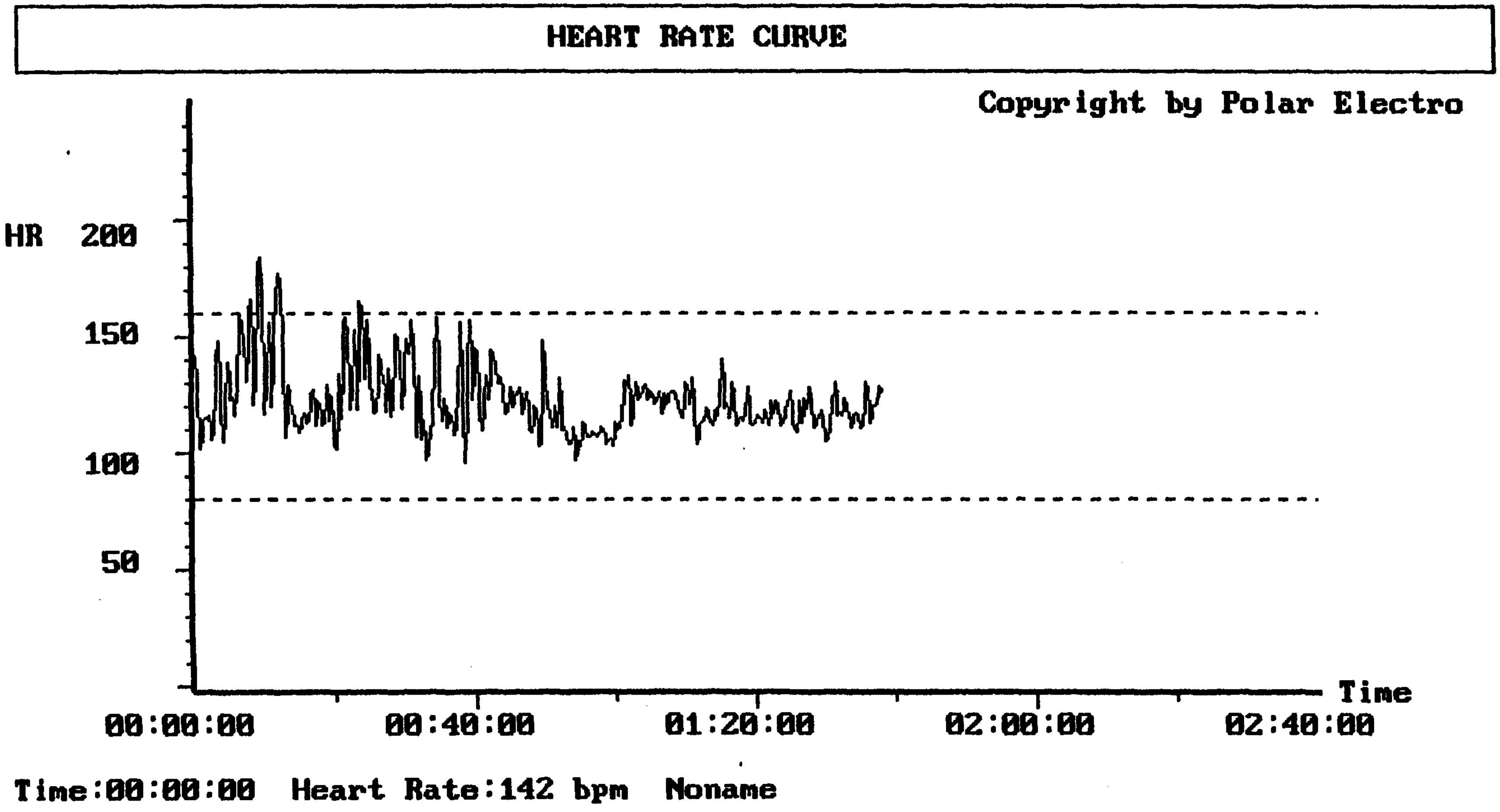
Table 15. Subsample of subject z-scores having data manipulated to achieve a balanced range of activities. A comparison of correlations before and after manipulation.

Original Z-Scores				*Balanced Range of Activities*			
	TT-HR	*TT-Obs*	*HR-Obs*		*TT-HR*	*TT-Obs*	*HR-Obs*
352216-D1	0.66	0.69	0.87	352216-D1	0.29	0.50	1.13
352216-D2	1.13	0.73	0.55	352216-D2	1.29	0.44	0.35
372209-D1	1.00	0.68	0.76	372209-D1	1.19	0.63	0.89
372209-D2	0.89	0.74	0.89	372209-D2	1.10	0.95	1.29
38168-D1	0.95	0.87	0.69	38168-D1	0.81	0.66	0.45
38168-D2	0.76	0.05	0.44	38168-D2	0.59	-0.04	0.34
44213-D1	1.07	0.85	1.02	44213-D1	1.00	0.78	0.81
44213-D2	1.05	0.62	0.74	44213-D2	1.00	1.00	0.73
542225-D1	1.33	0.87	0.81	542225-D1	1.47	0.49	0.55
542225-D2	1.16	0.76	0.68	542225-D2	1.66	1.05	0.91
mean	1.00	0.69	0.75	***mean***	1.04	0.64	0.74
n	10	10	10	***n***	10	10	10
Correlation	r = 0.76	r = 0.60	0.63	***Correlation***	r = 0.78	r = 0.57	r = 0.63
S.D.	0.07	0.13	0.09	***S.D.***	0.11	0.18	0.16

Table 16. Subsample of subject z-scores having data manipulated to remove the transitional minutes of activities. A comparison of correlations before and after manipulation.

Original z-Scores				***Elimination of Transitional Values***			
Subject ID	*TT-HR*	*TT-Obs*	*HR-Obs*	*Subject ID*	*TT-HR*	*TT-Obs*	*HR-Obs*
31114-Lab	1.29	1.05	1.02	31114-Lab	1.74	1.05	1.66
33118-Lab	1.05	1.07	0.91	33118-Lab	1.26	1.83	1.38
33120-Lab	1.00	0.79	0.91	33120-Lab	2.09	1.83	2.65
34122-Lab	1.29	0.95	0.89	34122-Lab	1.22	1.47	2.65
37219-Lab	1.38	1.19	1.19	37219-Lab	1.74	1.74	1.95
39128-Lab	1.38	0.85	1.19	39128-Lab	1.95	1.95	2.09
39223-Lab	1.22	1.02	1.19	39223-Lab	1.83	1.47	1.66
43206-Lab	1.26	1.13	1.13	43206-Lab	1.74	1.19	1.38
43221-Lab	0.83	1.07	0.34	43221-Lab	1.19	1.19	0.91
45127-Lab	1.29	0.56	0.49	45127-Lab	2.09	0.60	0.79
46125-Lab	1.26	1.29	1.38	46125-Lab	1.74	1.66	2.65
53107-Lab	1.33	1.19	1.26	53107-Lab	1.42	1.29	2.09
55131-Lab	1.53	1.10	0.93	55131-Lab	1.95	1.53	1.74
56215-Lab	1.33	1.00	1.13	56215-Lab	1.42	1.02	1.66
mean	1.24	1.02	1.00	***mean***	1.67	1.42	1.80
n	14	14	14	***n***	14	14	14
Correlation	r = 0.85	r = 0.77	r = 0.76	***Correlation***	r = 0.93	r = 0.89	r = 0.95
S.D.	0.04	0.07	0.1	***S.D.***	0.03	0.06	0.03

Appendix A. Heart rate curve generated with the software of the Polar Vantage XL. This session lasted approximately 1 hr 40 min, and shows the general trends of the HR during this period.



HEART RATE CURVE

Copyright by Polar Electro

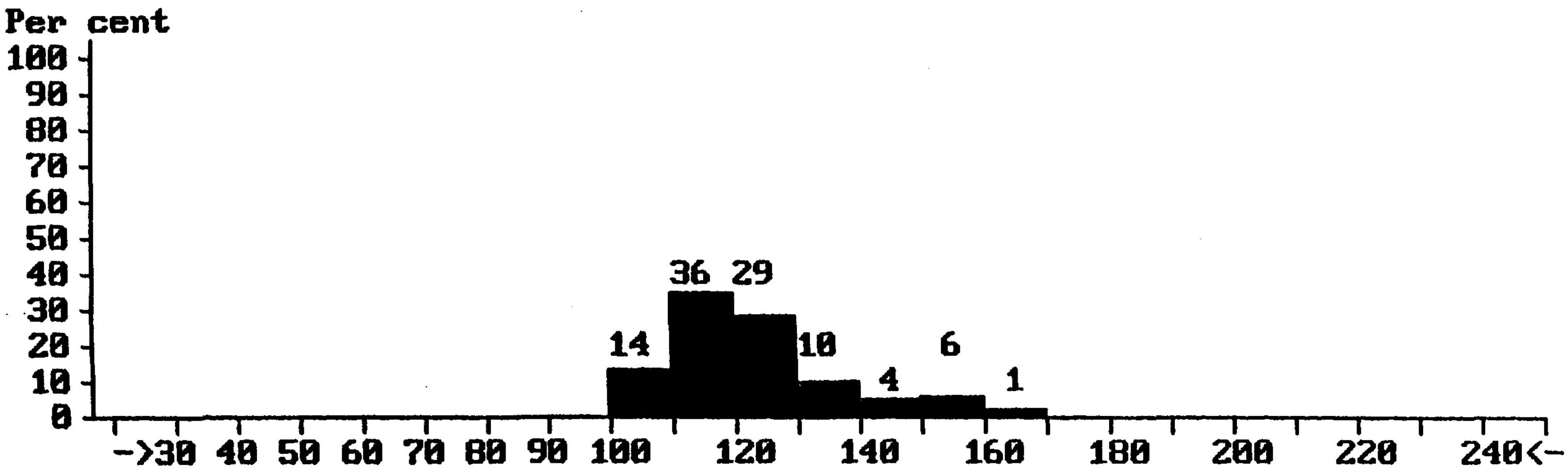
Time:00:00:00 Heart Rate:142 bpm Noname

Appendix B. Heart Rate distribution as generated by the Polar Vantage XL software. This session lasted 1 hr 38 min, and the distribution shows what percent of the total time interval the subject spent within discrete HR ranges.

HEART RATE DISTRIBUTION

The length of the exercise : 01:38:25.6
Copyright by Polar Electro

The selected time interval is Whole Test



SOURCE FILE: Noname

Appendix C. A sample of merged data. Data have been merged sequentially over time with each minute having a Tritrac (vector magnitude) acceleration count, an average heart rate for that minute and a BEACHES activity level.

31114 - Laboratory Observation

	Test Time	Tritrac Value	Heart Rate	BEACHES Value
1	17:54	1651	136.4	5
2	17:55	2701	154.3	5
3	17:56	2550	175.3	4
4	17:57	2030	152.7	5
5	17:58	478	124.3	1
6	17:59	134	116.7	1
7	18:00	145	114.2	1
8	18:01	59	111.6	1
9	18:02	431	125.5	2
10	18:03	176	115.9	2
11	18:04	84	113.9	2
12	18:05	135	120.1	2
13	18:06	213	117.7	2
14	18:07	1112	133.9	5
15	18:08	775	141	5
16	18:09	757	137.9	3
17	18:10	1191	144.3	5
18	18:11	321	130.7	2
19	18:12	406	126.1	2
20	18:13	107	123.5	2
21	18:14	64	124.4	2
22	18:15	2834	135.2	5
23	18:16	3592	174.6	5
24	18:17	959	151.2	5

Appendix D. A summary of merged files showing the status of each individual subject's data.

Summary of merged files as at 11 April 94

Subject #	Lab	Obs 1	Obs 2	Comments
37101	☹ ☹ TT invalid HR invalid	☹ ☹ TT invalid HR invalid	☹ ☹ TT invalid No HR values	Hr Values invalid to 1 minute sampling TT: Pre-recall
54102	☹ ☑ TT invalid	☹ ☹ TT invalid HR file missing	☑	TT: Pre-recall
64103	☹ ☑ no HR data	☒ file closed	☒ file closed	ST stopped - No HR data in Lab visit
43204	☹ ☑ TT invalid	☑ Very Short observation	🖫 TT file missing	TT: Pre-recall
48205	?	☑	☑	Home Observation
43206	☑	☑	☑	Home Observation
53107	☑	☑	☑	
39108	☹ ☑ TT invalid	☑	☑	TT: Pre-recall
44209	🕘 Not observed yet	🕘 Not observed yet	🕘 Not observed yet	
38110	☑	☑	☑	TT: Pre-recall
55111	🖫 TT file missing	☑	☑	
48112	☑	☑	☑	
44213	🖫 TT file missing	☑	☑	
31114	☑	☑	☑	
56215	☑	☑	☑	
36116	☹ ☑ Hr values useless	🖫 TT file missing	🕘 Not observed yet	Start time of ST not recorded (Lab)

Subject #	Lab	Obs 1	Obs 2	Comments
42117	☹ No TT values	🕒 Not observed yet	🕒 Not observed yet	TT units recalled
33118	☑	🕒 Not observed yet	🕒 Not observed yet	
37219	☑	☑	☑	
33120	☑	☑	☑	
43221	☑	☒ file closed	☒ file closed	Subject no longer in Daycare
34122	☑	☑	☑	
39223	☑	☑	☑	
xxx24				
46125	☑	☑	☑	
39126	?	☑	☑	
45127	☑	☑ Short Observation	☑	
39128	☑	🕒 Not observed yet	🕒 Not observed yet	
46129		🕒 Not observed yet	🕒 Not observed yet	
48230	☑	☑	☑	
55131	☑	☑	?	
32132	☑	?	☑	Home Observation
58100	☯	☑	☑	
41153	☯	☑	☑	
32265	☯	☑	☑	
38168	☯	☑	☑ Very Short observation	
47176	☯	☑	☑	

Subject #	Lab	Obs 1	Obs 2	Comments
39286	☯	☑ Very short observation	☑	
58185	☯	☑	💾 TT file missing	
561227	☯	☑	Need to print out ST file	
322203	☯	☑	☑ Short	
41150	☯	☑	☑	
46257	☯	☑	☑	
321202	☯	☑	☑	
321205	☯	☑	☑	
352216	☯	☑	☑ Short Observation	
372209	☯	☑	☑	
352201	☯	☑	☑	
492228	☯	☑	☑ Short	
542225	☯	☑	☑	
371215	☯	☹ TT invalid	☹ TT invalid	TriTrac values invalid due to nonsence data

Legend

☑ - Merged file

💾 - Component file missing

☒ - File closed

☹ - Flawed file

🕒 - Not observed yet

? - Unidentified file problem

☯ - Phase 2 subject, no lab visit