IMPROVING METHODS OF MEASURING PHYSICAL ACTIVITY FOR HEALTH RESEARCH: VALIDATION OF THE LIFETIME PHYSICAL ACTIVITY QUESTIONNAIRE (L-PAQ)

by

MARY ABIGAIL DE VERA

B.Sc., University of British Columbia, 2001

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Health Care and Epidemiology)

THE UNIVERSITY OF BRITISH COLUMBIA

August 2006

ABSTRACT

Lifetime physical activity plays an important determinant role in health outcomes including physical fitness, well-being, and development of disease, and the evaluation psychometric properties of instruments measuring this construct are important in epidemiologic studies. This study provided a comprehensive evaluation of the Lifetime Physical Activity Questionnaire (L-PAQ), a novel, Internet-based, self-administered instrument measuring lifetime exposure to physical activity across three domains: sports/recreation, occupation, and domestic. Validation of the L-PAQ involved a test-retest study to evaluate instrument reliability and a two-part construct validity study involving comparison to two questionnaires measuring similar constructs and hypotheses testing of L-PAQ constructs. The L-PAQ demonstrated good reliability with intraclass correlation coefficients for average lifetime hours/week of physical activity ranged from 0.58 to 0.83. The highest reliability was seen for sports (0.83) and lowest for total activity (0.58). Intermediate correlation coefficients were obtained for occupational (0.72) and domestic (0.60) activity. Based on the results of comparison with the LT-PAQ, the Lifetime Total Physical Activity Questionnaire, and the CT-PAQ, the Chasan-Taber Physical Activity Questionnaire, the L-PAQ demonstrated very good validity for domestic activity and moderate validity for sports and occupational activity. The robustness of the measurement of lifetime physical activity by the L-PAQ was further supported by the results of hypotheses tests between constructs measured by the questionnaire and descriptors of survey respondents. The relationship between lifetime physical activity measured by the L-PAQ and gender agree with previous findings. In addition, this was the first study to show relationships between lifetime physical activity

and education level. Overall, validation studies of the L-PAQ contribute to the existing knowledge of measurement of lifetime physical activity and provide evidence for the utility of this instrument.

TABLE OF CONTENTS

List of Abbreviations

OA	osteoarthritis
PA	physical activity
PAQ	physical activity questionnaire
PAJH	Physical Activity and Joint Health
PAJHS	Physical Activity and Joint Health Study
L-PAQ	Lifetime Physical Activity Questionnaire
LT-PAQ	Lifetime Total Physical Activity Questionnaire
CT-PAQ	Chasan-Taber Physical Activity Questionnaire
ARC	Arthritis Research Centre
MET	Metabolic Energy Equivalents
TLH	Total Lifetime Hours of Physical Activity
TLHS	Total Lifetime Hours Sports Activity
TLH**O**	Total Lifetime Hours Occupation Activity
TLH**D**	Total Lifetime Hours Domestic Activity
LAHW	Lifetime Average hours per week of Physical Activity
LAHWS	Lifetime Average hours per week Sports
LAHW**O**	Lifetime Average hours per week Occupation
LAHW**D**	Lifetime Average hours per week Domestic
METHW	Lifetime MET•Hr/Wk of Physical Activity
METHWS	Lifetime MET•Hr/Wk Sports
METHW**O**	Lifetime MET•Hr/Wk Occupation
METHW**D**	Lifetime MET•Hr/Wk Domestic

List of Tables

List of Figures

Acknowledgements

I would like to acknowledge my thesis committee members, Dr. Jacek Kopec, Dr. Heather McKay, and Dr. David Wilson. I would like to say a special thank you to my supervisor, Dr. Kopec for his guidance and wisdom throughout this project.

I would like to thank the BC Medical Services Foundation for awarding me with a 2005 Summer Studentship Award for this project.

I would like to acknowledge staff support at the Arthritis Research Centre of Canada, especially Mr. Paul Doerfling for his tremendous help with the PAJH dataset and Mr. Eric Sayre, who designed the on-line data collection system for ARC. My gratitude also to Dr. Christine Friedenreich and her research team at the Alberta Cancer Board for allowing me use of the Lifetime Total Physical Activity Questionnaire and providing me with materials for interview preparation and questionnaire scoring.

This project has allowed me to meet men and women who shared their time with me to participate in my project. Some of the greatest lessons learned during my Masters training were learned from these interviews.

I would like to acknowledge support from the Department of Orthopaedics at BC Children's Hospital, especially Dr. Christine Alvarez and Dr. Stephen Tredwell, who introduced me to the world of research. To Ms. Angie Perdios, thank you for all your encouragement and advice.

Finally, I would like to extend my love and gratitude to my family and friends for their tremendous support throughout the Masters years and beyond.

Dedication

This thesis represents hard work, passion, and my life's motivation for helping others. All these things I learned from my parents, thank you mom and dad.

I. INTRODUCTION

1.1 Health Research and Measurement of Physical Activity

Physical activity (PA) plays an important determinant role in health outcomes including physical fitness, well-being, and development of disease (1, 2). The relationship between physical activity and lowered incidence of disease has been studied in numerous epidemiologic investigations including studies of cardiovascular disease, diabetes, osteoporosis, and cancer (3) with many of these studies measuring exposure as *current* levels of activity (4, 5). Since many chronic diseases have a long developmental period, it is the *lifetime* or historical contribution of physical activity that is of important consideration when assessing risk factors. Knowing lifetime physical activity exposures also potentially allows for determination of etiologically relevant periods of disease development.

In order to determine associations between physical activity and health-related outcomes, it is important to use a physical activity measurement instrument that is accurate, reliable, and practical. The questionnaire – either self-administered or interviewer-administered – has been used as the medium of measurement for lifetime physical activity as it is most convenient for large-scale epidemiologic studies. Given the primary use of questionnaires to measure lifetime physical activity exposure, it is important to delineate the problems associated with them in order to develop methods for improvement of measurement.

1.2 Measurement of Physical Activity

Caspersen et al. defined physical activity as "any bodily movement produced by skeletal muscles that result in energy expenditure" (6, 7). Energy expenditure encompasses basal metabolic rate (which accounts for 50-70% of total energy expended), the thermic effect of food

(7-10% of total energy expenditure), and physical activity (8). As the last component of energy expenditure, physical activity has the greatest variability between individuals and is multi-factorial, comprised of activities of daily living, sports and leisure, occupation, and domestic activities (8). Thus, it is intuitive that most health research investigations on the relationship between physical activity and disease are focused on the contribution of physical activity to energy expenditure. It is also this tremendous variability that lends to the complexity of measurement of physical activity.

In attempting to measure physical activity, it is important to consider both the properties of physical activity and properties of the measurement instrument itself. In terms of physical activity properties, one must consider the health-related dimensions of physical activity that have the most relevance to the disease or condition under study. The health-related dimensions of physical activity range from caloric expenditure and aerobic intensity, which may be relevant to physical fitness and cardiovascular health, to weight-bearing and strength, which may be relevant to bone mass or osteoporosis, to joint loading, which may be relevant to arthritis (6, 7). To further highlight the differential effects on the prevention or risk for disease of these physical activity properties, 100 calories expended in swimming may have beneficial effects to cardiovascular health while 100 calories expended in weight training may have a more favourable effect on bone mass or osteoporosis. Health researchers recognize that it is also important to focus on the properties of physical activity that are most likely to be associated with the specific outcome of interest when evaluating the relationship between physical activity and a disease or condition. Perhaps, there is greater emphasis in the literature on properties of the measurement instrument rather than the properties of physical activity. Psychometric or measurement properties of an instrument pertain to the utility and quality of the instrument and also to the level of evidence provided by results of the study applying the instrument. That is, the ability to detect any significant associations between physical activity and disease outcomes is largely determined and/or limited by the quality and accuracy of the physical activity measurement instrument.

Laporte et al (9) reported that over 30 different methods have been used in epidemiologic studies assessing physical activity and these could be grouped into the following seven main classifications: 1) calorimetry, 2) job classification, 3) physiological markers, 4) behavioural observation, 5) mechanical and electronic monitors, 6) dietary measures, and 7) survey procedures (9). Calorimetry, which measures energy expenditure through heat production using both direct and indirect methods, provides an objective and accurate assessment of physical activity. However, it is labour-intensive and expensive, making it unsuitable for large epidemiologic studies. Job classification is based on ranking jobs according to levels of activity and assumes that all individuals within each occupational category exhibit similar levels of activity. This assumption leads to limitations due to variability within jobs, misclassification of occupational intensity, and seasonal and temporal variations in occupations. In addition, assigning physical activity levels based on occupational activity alone, omits other domains of activity including sports and recreational activities and household activities. The next three classifications of physical activity assessment, physiological markers (such as measuring subjects' maximum oxygen consumption or using doubly labelled water technique), behavioural observation (monitoring individuals and rating observed activity), and use of mechanical and electronic monitors (including heart rate monitors, pedometers, and electronic motion sensors) have also been used in epidemiologic studies but at great study costs (9). In addition, they impart considerable respondent burden, leading to problems with compliance.

Dietary measures and survey procedures involve asking respondents to record and/or relay retrospective information (9). Use of dietary measures involves asking subjects to record previous dietary intakes using food recall diaries. The caloric value of food intake is then used as an estimate of energy expenditure, and hence, indirectly, physical activity. Errors associated with this method include the fact that total caloric intake for a person is not only influenced by activity levels but also by total body weight. Further assumptions must be then made in order to relate total caloric intake to physical activity. Finally, dietary measures do not allow for delineation of types of activity performed, their duration, frequency, and intensity. These are important properties of physical activity used in describing and quantifying participation in different types of activities. If the purpose of a study is to measure physical activity and survey methods are going to be used, then it is more intuitive to directly ask participants about their actual physical activity participation patterns rather than infer physical activity from indirect measures.

Given the limitations of physical activity assessment methods described – ranging from high study costs to incomplete descriptions of physical activity to flaws in assumptions used – the last classification of physical activity assessment methods, survey or questionnaire procedures, are justifiably the most widely used method in epidemiologic investigations of the relationship between physical activity and disease.

1.2.1 Physical Activity Questionnaire

There are many characteristics of the physical activity questionnaire (PAQ) that make it suitable for use in population-based studies. One of these characteristics is non-reactiveness, meaning that respondents' behaviours are not modified by completing this measure. Second, the relative ease of administration and reasonable cost make questionnaires practical for implementation in large epidemiological studies. Finally, questionnaires are applicable since investigators have the ability to adapt questionnaires to suit their target populations (8).

Physical activity questionnaires have four elements: 1) *nature and detail of the activity* (subjects may be asked to provide the frequency, length, and intensity of specific activities or they may simply be asked whether or not they performed such an activity); 2) *means of collecting information* (subjects may be asked to fill out a form or be interviewed face-to-face or via telephone); 3) *summary index of physical activity* (a continuous score estimating energy expended or ordinal scale ranking subjects by level of activity); and 4) *(time) span of physical activity measured* (ranging from past day to past year to lifetime) (9).

The nature and detail of activity surveyed captures the complexity of the questionnaire and the domains of physical activity covered. Complexity refers to the level of detail asked about physical activity. As demonstrated by questionnaires used in previous health research studies, these ranged from single-item questionnaires asking individuals about their level of activity relative to others of their age and sex (10) or whether they performed an activity long enough to begin perspiring (11) to multi-item questionnaires asking detailed questions of participation in physical activity (3, 4, 5, 12). The details of such multi-item questionnaires include the type or domain of physical activity to describe the nature of the activity, elements of time to quantify the amount of participation in each activity (or performance of the activity), indicators of intensity, and further descriptors of the activity that are relevant to the health research question. The type or domain of physical activity measured in physical activity questionnaires have ranged from occupational activity performed at work, domestic activity performed in and around the home, and sports or recreational activity for leisure, fitness, or competitive pursuits. While early epidemiological studies based estimates of physical activity on occupational job titles (13, 14), it has been recognized that the amount and level of physical activities has declined in most occupations (5). Thus, in the 1990's there was a greater emphasis on physical activity questionnaires querying sports and leisure time pursuits (5). A second important recognition by authors in the 1990's is that most physical activity questionnaires did not represent activities performed by women through the omission of questions on household or domestic activity (3). Together, the three domains of sports/recreation, occupation, and household capture the best estimates of physical activity performed by a person (and/or populations) and thus, a comprehensive physical activity questionnaire should include all three domains. The elements of time to quantify the amount of participation in each activity include questions on the overall duration of participation (the number of years that respondents has performed the activity) and the frequency of participation (the number of months per year, the number of days per week, and the length of each occasion of participation).

The second element of PAQs is the *means of administration* of the questionnaire, and this has important implications for study designs in terms of costs and potential respondent burden and/or compliance. There are a number of ways a questionnaire may be administered including face-to-face or telephone interviews and self-administration. Self-administration may involve participants completing the questionnaire at the study site or at their homes through mail outs. Technological advances have allowed investigators to administer questionnaires using computer adaptive systems (15). Aside from study feasibility considerations, one must also consider the biases associated with each type of administration. When one or more interviewers are involved in studies, extensive training and quality control must be employed to ensure standardization and consistency of techniques. However even with such cautionary steps, there are still inherent biases in agreement between interviewers. In their report of the Historical Leisure Activity Questionnaire (HLAQ), Kriska described discrepancies between interviewers' interpretation of questions (5). Telephone interviews are prone to the same bias as face-to-face interviews. In addition, the lack of visual cues or face-to-face interactions may make it difficult to maintain the flow of the interview. Of all questionnaire administration methods described, self-administration is most frequently used in large epidemiologic studies due to their low costs (the costs of mailing are much less than salaries for interviewers) and minimal burden imparted on participants (they do not have to commute to the study site). However, this method is also prone to a number of biases associated with self-report. These biases include socially desirable response bias in which people tend to over-report what is perceived as good behaviour. Other biases include the opposite of socially desirable response bias, deviation or faking bad, yea-saying or acquiescence bias, end-aversion, positive skew, and halo bias (16). In addition, even with explicit instructions some respondents may not properly interpret questions.

The *summary score* is the reported or calculated estimate physical activity based on the questionnaire. For questionnaires simply asking respondents about their levels of physical activity, the summary score is an ordinal score (light, moderate, or heavy). However, for questionnaires asking greater detail about physical activities, summary scores could be calculated by applying formulae to reported frequency and duration of physical activities. According to Kriska et al. (8), the two most common units used to quantify physical activity questionnaire data are obtained by 1) calculating time spent performing the reported physical activity or 2) taking the time spent in activity and weighting it by an estimate of the intensity of that activity. Total time is calculated as the product of the frequency of the activity (for example *2x/week*) by duration or length of time that the activity is performed (for example *2hours/x*) to yield the number of hours per week spent in that activity. Taking this one step further and multiplying the number of hours per week spent in the reported activity by the average intensity of this activity yield an estimate of the energy expenditure for that activity. The average intensity of an activity is usually expressed as the metabolic cost for that activity or MET. One MET or metabolic equivalent represents the metabolic rate of a resting individual and is set at 3.5 mL of oxygen consumed per kilogram of body mass per minute (1 kcal/kg/h) (17). Compendiums of MET values assigned for specific activities have been compiled by Ainsworth et al. in 1989 and further revised in 2001 and are used in calculations of energy expenditure in physical activity research (17, 18). Figure 1.1 summarizes calculation steps for the common units of quantifying physical activity questionnaire data.

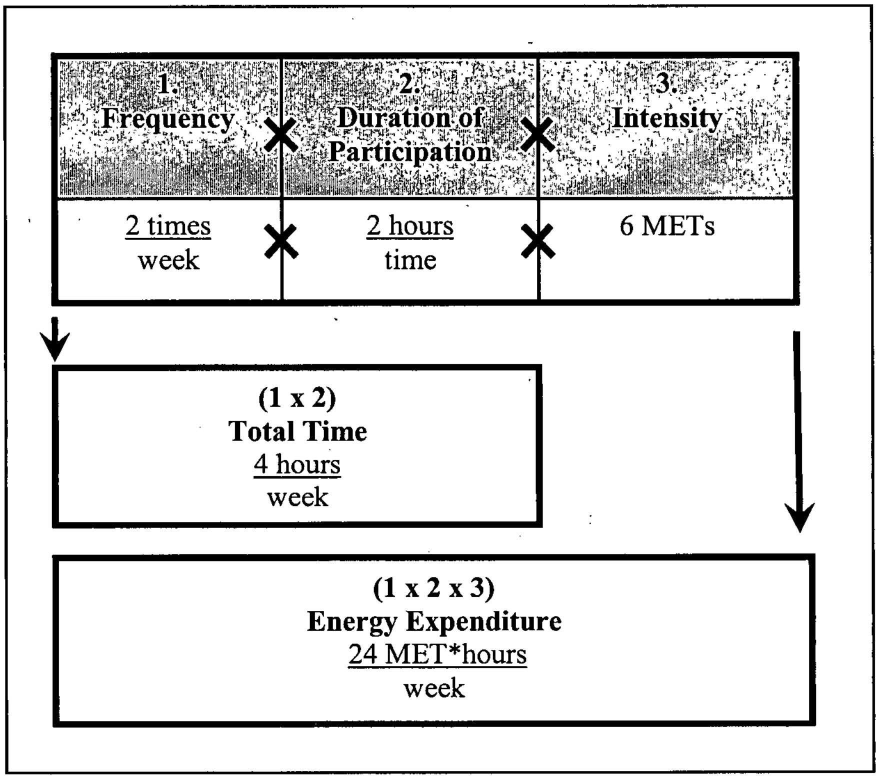


Figure 1.1 Calculation of Two Commonly Used Summary Scores for Physical Activity Questionnaires

While all four elements of physical activity questionnaires described are important, perhaps the one that is most relevant to the validation studies of this thesis is the *time span of physical activity measured.* It is important to make the distinction between the *time span of physical activity measured* in a questionnaire and the *elements of time to quantify the amount of participation in each activity.* While the latter concept capture variables that are used in the calculation of summary scores to quantify physical activity reports, the first concept captures the timeframe of overall questionnaire measurement and ranges from past day, past week, past year or lifetime physical activity.

Overall, the elements of physical activity questionnaires, *nature and detail of the activity, means of collecting information, summary index of physical activity, and span of physical* *activity measured,* form the basis of a 4-type classification of physical activity questionnaires. *Physical activity diary surveys* involve the shortest time frame of recall, usually 24 hours or less, and are self-administered. *Physical activity recall surveys* ask questions of physical activity over a longer period, usually the past 1 – 7 days, and are administered by subjects themselves or by interviewers. *Quantitative history physical activity questionnaires* are similar to recall surveys by mode of administration but differ in that they ask questions over a much longer time frame, usually past year or longer. Finally, *general surveys* do not ask any specific information about the nature and detail of physical activities performed (9).

By the definition of measurement of physical activity over long time frames of past year or longer, *quantitative history physical activity questionnaires* encompass lifetime physical activity questionnaires. The interest of measurement of lifetime or historical exposure to physical activity in health research studies was prompted by the rationale that many diseases have a long developmental period. Thus measuring *current* levels of physical activity or physical activity over shorter time frames may not provide etiologically relevant associations between physical activity and disease.

The complexity of measurement of lifetime physical activity exposures is based on the time span of measurement and associated problems. One of these problems is the potential for recall bias (2, 19). This follows because respondents are asked to report on activities that they performed anywhere from 10 to 50 years prior. Attempts by many investigators to address the issue of recall bias include use of cognitive tools to aid in recall (19). An equally problematic issue with lifetime physical questionnaires is that despite reports of reliability, there are few reports of validity studies of existing instruments. Validity is the ability of the questionnaire to measure precisely what it has been designed to measure while the reliability of a questionnaire refers to its ability to produce consistent results (20). These psychometric or measurement properties, which together define the accuracy and utility of a questionnaire will be further described in the following section.

1.3 Measurement Theory

Epidemiology is a science grounded on the measurement of disease and determinants of disease or exposures. Exposure is defined as "any of a subject's attributes or any agent with which he or she may come in contact that may be relevant to his or her health" (20). Such a wide definition include: agents that may cause physiological effects; agents that may cause or protect from disease; agents that may confound the association between another agent and a disease or physiological effect; agents that may modify the effects of other agents; and agents that may determine outcome of disease, such as screening procedures or treatment.

The term *measurement instrument* or *instrument* denotes a method or a set of methods used to measure the variables of interest in an epidemiological study (20). For purpose of this thesis and through its remainder, the term *instrument* will be used to refer to the measurement tool and the term *construct* to refer to the item or concept that the instrument is measuring. In its specific usage, instrument refers to tools used in research studies including questionnaires, personal interviews, biochemical analysis of blood or other biological specimens, and physical or chemical analysis of the environment (20). However, in its broader use, the term encompasses all parts of the measurement process and includes not only the actual tool itself but also instructions for its use as well as coverage of the procedures involving data collection, analysis and presentation of the results. In Armstrong et al.'s synthesis of the specific and broad definitions of the term instrument they state that "a measurement instrument deserves to be so called, with the connotation of accuracy and reproducibility that the word 'instrument' implies, only when all of the procedures outlined above are written down in such detail that one set of investigators could, within the limits of biological or physical variability, reproduce the measurements obtained by another using only this written description" (20).

1.3.1 Instrument Validation

There is no objective "gold standard" available for measuring total lifetime PA. In theory, development of this gold standard would involve prospectively following a cohort of subjects and measuring their PA levels using objective methods. The lack of such a gold standard lends to the complexity of validation of an instrument measuring lifetime PA with necessary steps including: 1) application of sound methods of instrument development to ensure face and content validity; 2) thorough pilot testing; 3) evaluation of the instrument's reliability or ability to produce consistent results; and 4) most importantly, establishing construct validity or the ability of the instrument to measure what it has been designed to measure. The first two steps represent instrument development while the last two steps involve the evaluation of the measurement properties of the instrument, and the objectives of this thesis.

The *reliability* of an instrument refers to its reproducibility or ability to produce consistent results under the same circumstances. Instrument reliability describes the extent to which the measurement is free from random errors (20) and is defined by the following equation:

Formula 1.1 $$\text{Reliability} = \frac{\text{Subject Variability}}{\text{Subject Variability} + \text{Measurement Error}}$$

In more practical terms, the reliability of an instrument refers to the consistency of measurement when it is repeated in the same subjects (20). In most areas of research, studies that assess the reliability of an instrument are often referred to as *intramethod* studies. Intramethod reliability studies may be further categorized into one of two types: *test-retest studies* establish the reliability of a measure when administered to the same subjects at two or more separate occasions and *inter-rater studies* establish the reliability of an instrument when it is administered by two or more raters to the same subjects (20).

The *validity* of an instrument refers to its ability to measure precisely the construct that it has been designed to measure or the true value of the "construct," the characteristic or object being measured by the instrument (20). Thus, validity refers to the ability of an instrument to measure the true exposure in a population of interest. Armstrong further described that estimates of the validity of an instrument as estimates of the exposure measurement error in a population (20). While reliability study designs are quite straightforward, validity studies have a wider range of complexity.

According to Cronbach, validity studies are a process of determining the degree of confidence that one can place on inferences that are made about people based on their scores on an instrument of interest (21). In essence, assessing the validity of an instrument is a process of hypothesis testing, where one evaluates rationalized relationships between constructs measured by the instrument. When assessing or describing validity, authors refer to one of the three C's of validity – *content validity, criterion validity,* and *construct validity.* An instrument is shown to have good content validity if it samples all the relevant or important domains of the construct it is measuring (16, 21). However, good content validity does not necessarily mean that the instrument is measuring what it is supposed to measure. Criterion validity has traditionally been defined as the correlation of the instrument with a criterion measure - some other instrument measuring the construct and usually the 'gold standard' or most widely accepted instrument (16). Criterion validity is further divided into two processes of concurrent and predictive validity. While concurrent validity studies administer the instrument of interest and the comparative instrument (criterion measure) at the same time, in predictive validity studies, the criterion measure is obtained in future studies (16, 21). Commonly cited examples of predictive validity are college admission tests or diagnostic tests which are compared to future outcomes such as completion of a degree program or confirmation and progression of predicted disease (16).

When there is no gold standard for the measurement of a construct, the most rigorous and important test of validity is construct validity. Construct validity studies are used when the construct measured by the instrument of interest is not readily observable or cannot be "operationally defined" (21). To illustrate with an example, physical characteristics such as height or weight are readily observed. By contrast, psychological attributes such as anxiety and intelligence, and many health measures including the construct of interest for this thesis, lifetime physical activity, are not readily observed.

There are two main reasons for instrument development and construct validation of an instrument: 1) the construct is novel and there are no existing instruments for its measurement; 2) current instruments are felt to insufficiently measure the construct by missing key components or domains (16). It is important to distinguish these two reasons for construct validation with the rationale for criterion validity which is usually to replace existing instruments with one that is shorter, cheaper, or minimizes respondent burden (16). Ultimately, the purpose of construct validation is to develop an improved instrument for the measurement of a construct which would then translate into an improved ability to explain a broader range of findings when the instrument is applied in an epidemiological study.

The methods of construct validation are well-established in psychology and social science research, especially in the development of scales measuring psychological and personality attributes. One method of assessing the construct validity of an instrument is to test its *convergent validity* with instruments measuring the same or similar construct (to which it should be related). Construct validity can also be evaluated by testing hypotheses about relationships between the construct being measured to other indices of that construct (16) or between the construct and characteristics of the population to which the instrument is applied. A classic example of hypothesis testing of an instrument measuring a psychological trait or disorder is to apply the instrument to a group of people with known diagnoses of the disorder and compare scores of people who do not have the diagnoses. A more relevant example to this thesis is comparing differences between groups based on hypotheses about activity levels across different physical activity domains. For example, one could test the hypothesis that women would report greater domestic activity than men or the opposite relationship that men would have greater occupational activity scores than women. While content and criterion validity can often be evaluated in a single study, assessment of construct validity is a process which usually involves a series of studies. Conducting such studies strengthens the 'nomological network' of interlocking beliefs or theories about the construct being measured and the validity of the instrument (21).

II. LITERATURE REVIEW: Questionnaires Measuring Lifetime Physical Activity

To identify and describe published questionnaires measuring lifetime physical activity as well as limitations in current measurement of this construct, a review of the validation of quantitative physical history questionnaires was conducted. This review was conducted as part of the Systematic Overviews Course offered at the Department of Health Care and Epidemiology at the University of British Columbia (HCEP 516). Prompted by previous author reports that "despite their wide use, little is known about the validity and reliability of quantitative history physical activity questionnaires" (9) and the "validation problems of history physical activity assessment" (8), the rationale for this review was to highlight issues behind these statements and provide a description of the lifetime physical activity in the current literature. The review was centered on the specific question, "What is the quality of quantitative history physical activity questionnaires measuring lifetime or historical physical activity?" from which the following objectives were formulated: 1) to identify lifetime physical activity questionnaires (published and unpublished) and 2) to review validity and reliability studies in terms of methods used and reported results.

The literature search strategy first involved identifying questionnaires measuring lifetime physical activity using electronic bibliographic databases including Medline (1966 - March 2005) as the main source and the supplementary sources, SPORTDiscus (1949 – March 2005) and PsychINFO (1987 – March 2005) to identify any additional studies or check for duplicates. The following terms, which mapped to MeSH headings, were used in the search strategy: questionnaire, health surveys, physical activity, exercise, activities of daily living, recreation, leisure activity, reproducibility of results, and psychometrics. "Physical activity" was used in a separate search strategy as a keyword since this term did not map to a MeSH heading. To ensure that search results were inclusive, a term indicating the time span of measurement (for example use of the terms, "lifetime" or "historical") was not used. Another justification for exclusion of these terms was that they did not map to any MeSH headings and their application as keywords in a search strategy might eliminate questionnaires that did not use these terms. The second step of the search strategy was to identify validation studies for identified questionnaires. In some cases, the results of the initial search yielded the report of a validation study while in some cases it yielded a study reporting use of the questionnaire. Thus, names of identified questionnaires were separately entered into individual searches to identify any further validation studies. Finally, bibliographies of selected references were also hand-searched for any additional questionnaires or validation studies.

Questionnaires were included in the review if they measured lifetime or historical physical activity, used in or developed for use in adult populations (18 yrs and older), and written in the English language. A validity and/or reliability study must have been reported (single publication or separate publications) in order for the instrument to be included in the review. Though not an inclusion criterion, the availability of the questionnaires was also considered in this review. Many questionnaire development and validation studies report on such studies but do not append a copy of the questionnaire. When possible, an available version of the questionnaire was obtained for this review to allow for a qualitative summary. Since there are a number of physical activity questionnaires that measure past year physical activity and the focus was on questionnaires that measure physical activity over a longer timeframe, past-year questionnaires were not included in the review.

Descriptive information extracted from the questionnaires were based on the elements of physical activity surveys described in Laporte's review and include: time frame of physical activity measured, nature and detail of the activity (domains of activity plus time unit or level of detail), method of collection and summary index of physical activity (9). Data extracted from reliability studies include the type of study, sample size, study details (for example the time between administration of questionnaires in a test-retest study), and main results (reported coefficients). Of interest were authors' interpretation of reliability study results and their reported assessments of the results (excellent reliability, good reliability, moderate reliability). These data were also tabulated if available in the text. Data extraction of validity studies were done in a similar fashion and information obtained include the type of validation study, sample size, study details, main results, and authors' reported assessment and interpretation of results. Finally, information on whether the instrument was applied in further research studies was also reviewed.

The search strategy resulted in 183 publications in which 55 physical activity questionnaires were reported. Abstracts were reviewed for the 55 papers reporting on physical activity questionnaire to determine if the questionnaire met the inclusion criteria. One of the challenges of the search was that it did not allow for distinct separation of questionnaire types and thus a number of questionnaires were excluded because they were measuring past week physical activity (10 questionnaires) or measuring long-term activity but in adolescents or children (4 questionnaires). Twenty-eight papers were also eliminated as they reported the use of the questionnaire and not a development or validation of the questionnaires. One paper evaluated the validity of different measures of lifetime physical activity, which included an interviewer-administered questionnaire (22). However there was insufficient description of the questionnaire for extraction of data relevant for this review. Thus, this questionnaire was excluded from this review. Overall, seven questionnaires met all inclusion criteria outlined in the search strategy and thus, included in this review. Table 2.1 provides a summary of the names and references of the included questionnaires as well as information on whether a copy of the questionnaire was readily obtained. Abbreviated names will be used in subsequent descriptions and discussions of the questionnaires included in this review. Table 2.2 summarizes each questionnaire according to the four elements of physical activity surveys described by Laporte (9).

Table 2.1 Quantitative History Physical Activity Questionnaires Included in the Review

Name of Questionnaire	Abbreviated Name	Primary Reference(s)	Obtained Copy of Questionnaire
Retrospective Physical Activity Survey	**RPAS**	Kriska 1988 (4)	Y
Historical Leisure Activity Questionnaire	**HLAQ**	Kriska 1990 (5)	Y
Historical Physical Activity Questionnaire	**HPAQ**	Winters-Hart 2004 (23)	N
Lifelong Physical Activity Questionnaire	**LLPAQ**	Dan 1990 (24)	N
Lifetime Total Physical Activity Questionnaire	**LT-PAQ**	Friedenerich 1998 (12)	Y
Quantification de l'Activite Physique	**QUANTAP**	Vuillemin 2000 (15)	N
Chasan-Taber Lifetime Physical Activity Questionnaire	**CT-PAQ**	Chasan-Taber 2002 (3)	Y

Table 2.2 Description of Questionnaires Included in the Review

Abbreviated Questionnaire Name	Time frame of physical activity measured	Domains of Activity	Time unit or level of detail	Method of Administration	Summary Index of Physical Activity
RPAS	Historical PA	Leisure	Duration and frequency of participation in 20 listed sports/recreational activities	Self	Kilocalories/week
HLAQ	Historical, Past-Year, Past-Week PA	Leisure Occupational	Duration and frequency of participation in activity	Interview	MET*hours/week
HPAQ	Historical PA	Leisure	Duration and frequency of sports/recreational activities and specific questions about walking	Interview	Kilocalories/week
LLPAQ	Lifelong PA	Occupation Leisure Household	Participants asked to choose one of four activity levels to rate their physical activity in occupation, leisure, and household for each decade in their working life	Self	Ordinal scale of physical activity rating: Level 1 (light activity) to Level 4 (heavy activity)
LT-PAQ	Lifetime PA	Occupation Household Exercise/Sports	Type of activity, duration, frequency, and intensity	Interview	MET*hours/week
QUANTAP	Lifetime PA	Sports at school Leisure sport Occupation Daily Activities	Duration and frequency of participation in activity	Interview	MET*minutes/year
CT-PAQ	Lifetime PA (Past year PA)	Recreational Household	Frequency and duration of participation in 27 listed recreational activities and 5 household activities	Self	MET*hours/week

This review was conducted to summarize information on current lifetime physical activity questionnaires in the literature with seven questionnaires measuring lifetime physical activity meeting the inclusion criteria established for this review. Looking at the nature and detail of physical activity measured, there is a variation in the physical activity domains captured by the questionnaires. However, consistent with all questionnaires is the inclusion of leisure and/or sports/recreational activity. This follows the recognition by many researchers of the decline in physical activity in most occupations and the significant contribution of exercise and leisure pursuits to physical activity in today's society (8). Three questionnaires – LLPAQ, LT-PAQ and QUANTAP - cover the most comprehensive physical activity domains by including occupation, leisure, and household activity.

Six questionnaires asked details on duration and frequency of activities performed allowing for calculation of summary scores to quantify physical activity (RPAS, HLAQ, HPAQ, LT-PAQ, QUANTAP, and CT-PAQ). The LLPAQ asked respondents to rate level of physical activity in three domains of occupation, leisure, and household for each decade of their working life. With this questionnaire, physical activity was rated on an ordinal scale ranging from Level 1 (light activity) to Level 4 (heavy activity). Four of the questionnaires have a self-administered format and three are administered by interviewers. The LT-PAQ emphasized the incorporation of cognitive interview techniques to help improve respondent recall and the QUANTAP combined an interviewer-administered format with QUANTAP, a structured, computer-assisted interview tool.

Reliability studies and validity studies for included questionnaires were also reviewed. Table 2.3 summarizes information on reported reliability studies for questionnaires and includes information on the study sample size, the time span between questionnaire administration, and main results. Table 2.4 summarizes information on validity studies including sample size, type of validity study, and main results.

Table 2.3 Summary of Reliability Studies of Included Questionnaires

	Test-Retest Span	N	Age Range	Main Results	Authors Rating of Reliability
RPAS	2–3 mo	23	Post-menopausal women	Kappa statistics ranging from 0.39 to 0.47 **Authors did not provide further information about which statistic corresponded to which PA domain. However, authors stated that these values represented fair agreement beyond chance.*	Moderately Reliable
HLAQ	1–3 wk	69	10-59 yr	Historical Activity: ρ = 0.94 Past year Activity: ρ = 0.89	Reliable
HPAQ	*Not reported				*Not reported
LLPAQ	2 wk	59	33-62 yr	Total score: ρ = 51 Occupation: ρ =46 Leisure: ρ = 48 Household: ρ = 57 **Authors only specified statistics as "test-retest correlations."*	Not commented
LT-PAQ	6-8 wk	115	Not specified	Lifetime PA ρ = 0.74 Occupational PA ρ = 0.87 Household PA ρ = 0.77 Exercise/sports PA ρ = 0.72	Highly Reliable
QUANTAP	2 wk	30	13-90 yr	Sport at school ρ = 0.64 Leisure sport ρ = 0.83 Occupation ρ = 0.85 Daily activities ρ = 0.81 **Reported most relevant correlations*	Not commented
CT-PAQ	1 yr	131	39-65 yr	Total PA: ICC = 0.82 Recreational PA: ICC = 0.87 Household PA: ICC = 0.71	Highly Reliable

Table 2.4 Summary of Validity Studies of Included Questionnaires

	Description of Validity Study	Type of Validity Study	N	Age Range	Main Results
RPAS	Comparison of most recent time period of survey with Paffenbarger Survey	Not specified	223	Post-menopausal women	Sport Index ρ=0.09 Kcal/week ρ=0.38
HLAQ	Comparison with 7-day activity monitor (Caltrac accelerometer)	Not specified	69	21-36 yr	ρ = 0.62 **Validity for past-week PA component of the questionnaire*
HPAQ	Correlation with 4 administrations of the past year physical activity questionnaire over a 17-yr period	Construct validity	163	70–79 yr	1982 PAQ: ρ=0.39 1985 PAQ: ρ=0.45 1995 PAQ: ρ=0.57 1999 PAQ: ρ=0.62
LLPAQ	Correlation with cardiorespiratory fitness in different age groups	Not specified	9 to 21	33-62 yr	Total PA ρ=0.16
LT-PAQ	Not reported				
QUANTAP	Correlation with % body fat at time of survey administration	Construct validity	419	13-90 yr	Males: ρ=-0.17 Females: ρ=-0.30 **By years prior to assessment*
	Hypothesis testing of expected differences between gender in daily energy expenditure	Construct validity	419	13-90 yr	Male>Female sport Male>Female leisure Male>Female occupation Male<Female in ADL
CT-PAQ	Correlation with four 1-week Physical Activity Logs administered 4 times over 1-year study period	Construct validity	131	39-65 yr	Total PA : ρ=0.26 Moderate PA: ρ=0.15 Vigorous PA ρ = 0.52 **Validity for past-year PA component of the questionnaire*

The reliability of the questionnaires in this review were all evaluated with test-retest studies, done at varying intervals from the shortest interval of 1 week (HLAQ) to the longest interval of 1 years (CT-PAQ). One questionnaire (HPAQ) did not have a reported reliability study. However, this questionnaire was adapted from the Paffenbarger Questionnaire which has been validated in a number of previous studies (25, 26). The test-retest study is an appropriate method of establishing the reliability of the questionnaire. In principle, a questionnaire that is reliable should yield similar results when applied on separate occasions. Authors used varying statistics to describe the reliability of their questionnaires, with some studies reporting more than one statistic. Pearson correlation coefficients were reported for 3 studies, the kappa statistic for 1 study, and the intraclass correlation coefficient (ICC) for one study. Two studies reported unspecified test-retest correlations. Good to excellent reliability were described by authors for most questionnaires.

Evaluating the validity of physical activity questionnaires is a somewhat greater challenge than evaluating reliability. Shephard reported that a number of authors who have developed physical activity questionnaires have limited their studies of psychometric properties to the evaluation of reliability, often neglecting the more important issue of validity (2). Of the seven questionnaires included in this review, six had corresponding validity studies reported. However two of these validity studies were conducted for versions of the questionnaires that measure shorter time spans of physical activity. Specifically the validity study for the HLAQ was for the past-week version and the validity study for CT-PAQ was for the past-year version of the questionnaire. The RPAS, which divided life into four time periods (14-21, 22-34, 35-50, and 50+ years), was validated by comparing subject response to the last age period (50+) against the Paffenbarger Questionnaire (26), an instrument measuring current physical activity. One questionnaire (LT-PAQ) did not have a reported validity study. However, it was the only questionnaire that had detailed descriptions of pre-testing and incorporation of cognitive methods to enhance subject recall (12). Finally, also of interest in this review was whether authors noted the type of validity study applied. Three authors did not specify the type of validity study while three authors specified the application of construct validity studies.

Equally important to studies of instrument validity and reliability were any further studies applying to measure associations between lifetime physical activity and disease or health outcomes. Such reports provide an indicator of the utility of the questionnaires. Three of the seven included questionnaires were used in at least one published study (RPAS, HLAQ, and QUANTAP) (4, 27, 28). One questionnaire (LT-PAQ) was used in two epidemiologic studies (29, 30). These papers were further reviewed to determine whether authors related psychometric properties of their questionnaires with findings from studies in which they were used.

The reliability of RPAS was actually incorporated in the report of its application in a study of the association of historical physical activity and adult bone parameters in post-menopausal women (4). This was the only paper that related psychometric properties of the questionnaire to results of a study in which it was used. The authors commented that since the RPAS was shown to be reliable, estimates of kilocalories for their study population were correlated with bone area and density (4). The authors also reported that bone parameters were related to historical physical activity, which provided further support for the utility of their questionnaire (4). There were no comments about the psychometric properties of HLAQ and their impact on study findings in Kriska et al.'s study of physical activity and non-insulin-dependent diabetes mellitus in Pima Indians (5). The QUANTAP was used to study the relationship of lifetime physical activity and bone mineral density of the lumbar spine and femoral neck bone (28). Authors did not comment about the measurement properties of the questionnaire but instead cited memory bias as one of the limitations of the study which may have been reduced through the design of the QUANTAP (28). Finally, the LT-PAQ was used in two separate case-control studies: lifetime physical activity and breast cancer (29) and lifetime physical activity and prostate cancer (30). Similar to the application of the QUANTAP, the authors did not comment about the measurement properties of the LT-PAQ but rather cited some of the biases associated with their studies. One of these was nondifferential misclassification bias (of physical activity exposure). The authors interpreted the effect of this bias as an underestimation of the true effect of physical activity and cancer.

Overall, this review of quantitative history physical activity questionnaires identified current instruments measuring lifetime physical activity in the literature as well as summarized reported validity and reliability studies. The application of these questionnaires in studies of the association between historical physical activity and disease were also reviewed. While there are a number of review articles on the measurement of physical activity using questionnaires (2, 31, 32, 33), this review is the first that has focused on instruments measuring lifetime or historical physical activity and their validation and utility.

III. OBJECTIVES

3.1 Summary of Research Objectives

The utility of an instrument is defined by its psychometric properties. Instrument reliability is the ability to produce replicable results while instrument validity is the ability to measure what it is supposed to measure. The purpose of this study was to evaluate the measurement properties of the Lifetime Physical Activity Questionnaire (L-PAQ) and thus, provide a comprehensive validation study of a quantitative history physical activity questionnaire.

The following objectives provided the framework for this study:

1. To establish the reliability of the Lifetime Physical Activity Questionnaire through test-retest study

2. To establish the construct validity of the Lifetime Physical Activity Questionnaire through comparison with similar measures of lifetime physical activity and testing of hypotheses pertaining to the constructs measured by the questionnaire

3.2 Application of the Research

This thesis project was an adjunct to the Physical Activity and Joint Health (PAJH) cohort study, a large study investigating the relationship between lifetime physical activity and osteoarthritis at the Arthritis Research Centre of Canada (ARC). Osteoarthritis (OA) is a chronic joint condition characterized by gradual disintegration of articular cartilage (34). It is the most common form of arthritis and causes significant burden in terms of pain and reduced quality of life (34). Symptomatic and radiographic knee OA occurs in about 6% of the population age 30 years and older and hip OA in 3% (35).

There is no known cure for OA (35). Thus, an understanding of risk factors for OA is important in delaying onset, modifying disease course, or preventing occurrence. Despite wide acceptance that high level physical activity is a risk factor for OA, the effect of lifetime (habitual or moderate) physical activity on OA development remains unknown. Previous studies that have attempted to evaluate this relationship have been hindered by methodological issues of both exposure (physical activity) and outcome (OA) measurement. McAlindon et al. assessed physical activity using a questionnaire which asked subjects how many hours they spent performing light, moderate, and heavy activities, a method susceptible to misclassification because the same subject could be assigned as having high levels of moderate activity (many hours) and low levels of heavy activity (few hours) at the same time (36). Other studies have based measures of physical activity on occupational activities, thus providing very limited estimates of physical activity since household and leisure activities were not considered (37). Another issue that has not been addressed by earlier studies is consideration of the biology of OA pathogenesis. Previous assignment of physical activity has not involved any account of the cumulative effect of different activities or the role of joint forces on articular degeneration. A study involving canines suggested that damage to cartilage consistent with OA required repeated impacts with peak stress above a certain level. The authors proposed that impact damage is cumulative and stress-rate dependent (38). Finally, no attempts at quantifying moderate lifetime physical activity have been done in previous studies.

The PAJH cohort study addresses many of the methodological issues of physical activity measurement that were evident in previous studies, both in exposure (physical activity) and outcome (OA) measurement. Of particular relevance to this thesis is the measurement of exposure or lifetime physical activity.

IV. MATERIALS AND METHODS

4.1 Overview of Methods

The Lifetime Physical Activity Questionnaire (L-PAQ) is the instrument used to measure lifetime physical activity exposure in the PAJH cohort study at the Arthritis Research Centre of Canada (ARC). Validation studies of the L-PAQ for this thesis involved the assessment of instrument reliability and construct validity. A test-retest study comparing measurements of two separate administrations of the questionnaire was used to evaluate the reliability of L-PAQ. Two studies were conducted to assess L-PAQ construct validity. First, the L-PAQ was compared with previously described and validated questionnaires measuring similar constructs but using different methods of administration - one questionnaire required administration with an interviewer and the other instrument was administered by subjects themselves. Second, the robustness of the L-PAQ was assessed through testing of a-priori hypotheses about constructs measured by the questionnaire.

The validation studies were approved by the University of British Columbia's Behavioural Research Ethics Board (BREB). A copy of the certificate of approval is included in Appendix I (A1.1). A number of different data sources involving both prospectively collected data and previously collected data were synthesized for the validation studies. Figure 4.1 provides an overview of these datasets as well as the methods used in the validation studies.

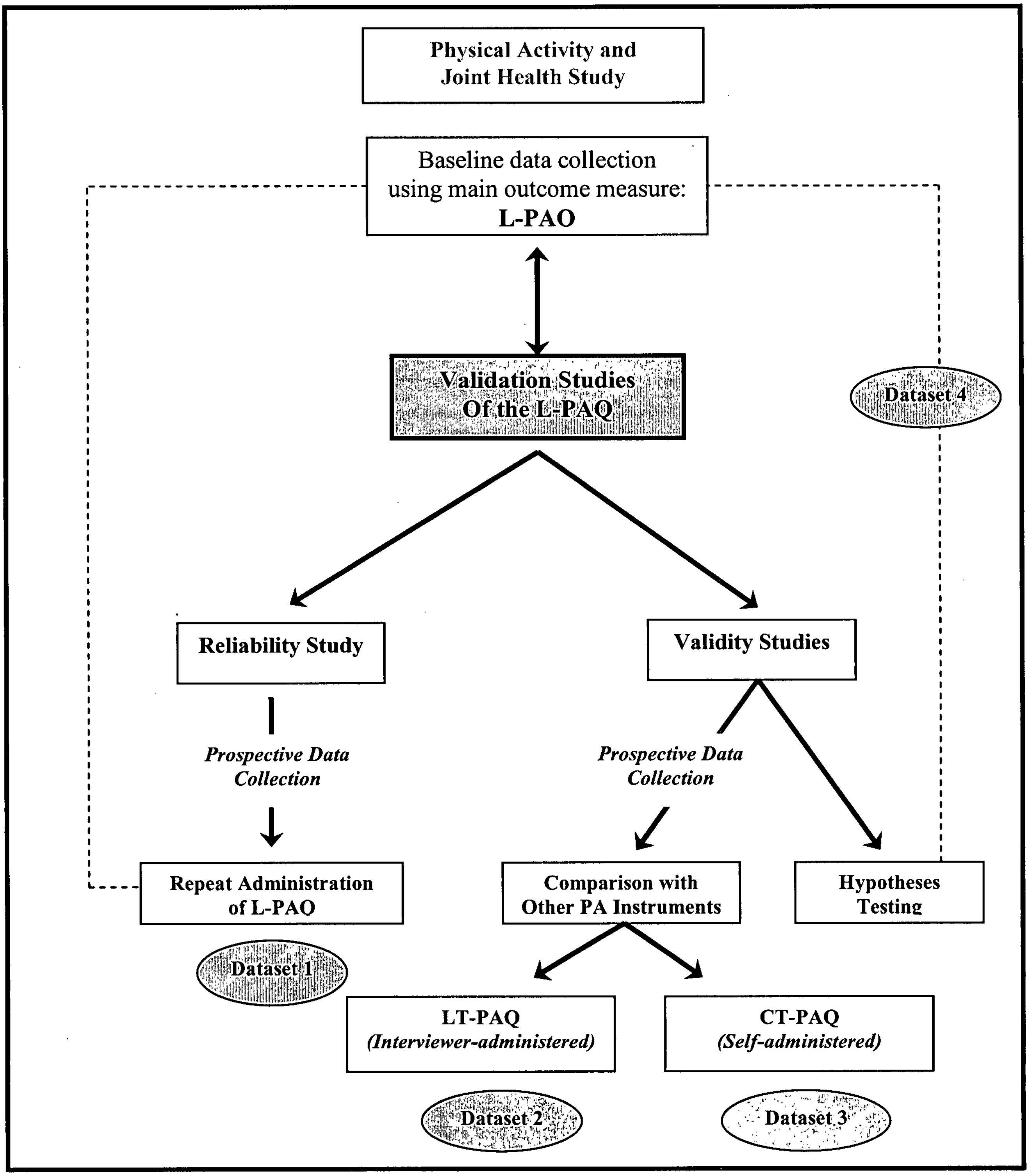


Figure 4.1 Overview of Validation Studies

4.2 Lifetime Physical Activity Questionnaire (L-PAQ)

The Lifetime Physical Activity Questionnaire (L-PAQ) is an Internet-based, self-administered instrument measuring lifetime physical activity. The L-PAQ was administered as part of a larger survey, the Physical Activity and Joint Health Survey (PAJHS) during baseline data collection for the PAJH cohort study. The PAJHS consisted of five sections: 1) Demographic Information (6 items) consisted of questions about gender, date of birth, marital status, education level, annual household income, and ethnic background; 2) General Health Questions (178 *potential* items) asked about any existing health conditions, height, weight, walking ability during the past four weeks, pain, and quality of life-related questions; 3) Knee and Hip Health (34 *potential* items) consisted of specific questions about knee and hip pain or any previous OA diagnoses; 4) Smoking History (14 *potential* items) asked questions about use of tobacco and tobacco products; and 5) L-PAQ (728 *potential* items).

While the number of *potential* items in each section of the L-PAQ was listed, it is important to note that subjects were not required to answer all items. Administration of the PAJHS over the Internet allowed for the use of computer adaptive technology to increase efficiency of data collection and minimize respondent burden. This was done by setting negative responses to specific questions to prompt the questionnaire to move forward, past questions and/or sections that were logically irrelevant to the respondent. For example, a person reporting that they do not smoke or have never used a tobacco product will not be prompted to answer further questions in the Smoking History section of the PAJHS. Similar question logic was applied to questions in the L-PAQ so that respondents only completed questions relevant to physical activities that they reported participating in.

The L-PAQ defined lifetime physical activity across three main domains: 1) sports/recreation; 2) occupation; and 3) domestic. In the sports/recreational section, respondents were provided a list of 64 possible sports activities ranging from aerobics to weightlifting. These are presented in Appendix 1.1 along with assigned metabolic energy equivalents used in calculations to quantify L-PAQ data, which will be described in Chapter 4.4. Respondents were asked to check sports activities that they have performed at least 100 times in their lifetime, prompting the computer adaptive system to move forward to more detailed questions about the activity. For each sport reported, respondents were asked the following string of questions: duration of participation, frequency of participation, and the length of time of participation during each occasion. Respondents were also asked to provide a rating of their perceived exertion while performing the activity using a scale that ranged from 1 as "Not Active" to 10 "Very Active."

The occupation section of the L-PAQ used an open format in which respondents indicated jobs that they had over their lifetime. For each occupation reported, respondents were asked to report the job title, the duration of the job, the number of hours per week in each job, and whether the job was full-time, part-time or seasonal. Subjects were also asked to rate their perceived exertion while performing the occupational activity.

Domestic activity was the final section of the L-PAQ. Respondents were asked questions on four general areas of household activity: 1) caring for children; 2) caring for elderly or disabled individuals; 3) gardening; and 4) housework. An additional question on other household PA was provided as an option to subjects who may have performed other household activities not captured by the four general areas of household activity provided by the L-PAQ. For each domestic activity that respondents indicated that they have performed, respondents were also asked about duration of participation, frequency of participation, the length of time of participation during each occasion, and a rating of their perceived exertion while performing the activity. Table 4.2 summarizes specific questions asked for *each* sports, occupation, and domestic activity that respondents reported participating in.

Table 4.1 L-PAQ Questions Used to Quantify Participation in Three Domains of Physical Activity

Sports / Recreation
(Using the first item in the L-PAQ sports domain, "Aerobics," as an example)

Purpose of Questions	Specific Questions		Units
Questions on duration of participation in each sports activity	**Q1**	At what age did you start participating in Aerobics?	**YOP:** Years of participation
	Q2	At what age did you stop participating in Aerobics? If you are still participating in Aerobics, please fill in your current age.	
Questions on frequency of participation in each sports activity	**Q3**	How many months per year did you participate in Aerobics?	**WPY:** Months per year converted to Weeks per year
	Q4	How often did you participate (per week, per month, per year)?	**OPW:** Occasions per week (all units converted)
Questions on length of time of participation in one occasion of sports activity	**Q5**	On average, how long did you participate on each occasion (minutes, hours)?	**HPO**: Hours per occasion (all units converted)

Occupation

Purpose of Questions	Specific Questions		Units
Identify occupation	**Q1**	Please list Job #	
Questions on duration of participation in each occupation	**Q2**	At what age did you start participating in Job#?	**YOP:** Years of participation
	Q3	At what age did you stop participating in Job#? If you are still participating in Job#, please fill in your current age.	
Questions on frequency of participation in each occupation	**Q4**	What type of employment was Job# (full-time, part-time, seasonal)?	**WPY:** Weeks per year (all units converted)
	Q5	How long was a season on average?	
Questions on length of time of participation in one occasion of occupation	**Q6**	How many hours per week did you work on average?	**HPW**: Hours per week

Domestic

Purpose of Questions	Specific Questions		Units
Questions on duration of participation of domestic activity	**Q1**	At what age did you begin caring for children?	**YOP:** Years of participation
	Q2	At what age did you stop caring for children? If you are still caring for children, please fill in your current age.	
Questions on frequency of participation of domestic activity	-	**Assumed at 52 weeks per year*	
Questions on length of time performing domestic activity	**Q3**	How many hours per week did you care for children on average?	**HPW**: Hours per week

Cognitive methods were incorporated in the administration of the L-PAQ to improve respondent recall. One method was providing a comprehensive list of sports/recreation activities to facilitate responses in this section. Similar methods were used in two other self-administered physical activity questionnaires measuring lifetime physical activity, the Retrospective Physical Activity Survey (RPAS) and the Chasan-Taber Physical Activity Questionnaire (CT-PAQ). Showing a list of possible sporting activities was also one of the cognitive tools used in the Lifetime Total Physical Activity Questionnaire interviews (LT-PAQ). Aside from the list of activities, L-PAQ questions were structured to capture average activity levels and not specific life events. This was done because previous researchers have shown that generic memory of usual or common patterns is more readily recalled than episodic memory (19).

Relevant to the PAJH cohort study are additional questions attached to each physical activity asking respondents to estimate the frequency and/or duration of several major body movements involving knee and hip. Question strings were asked for all three physical activity domains with differences reflecting the nature of the activity. In the sports/recreational section, respondents were asked to estimate the *duration* (number of minutes in an hour) of 4 bodily movements (sitting, standing, walking, and running/jogging) and estimate the *frequency* (number of times per hour) of 3 bodily movements (squatting, squatting with lifting, and jumping). In both occupational and household activity sections, respondents were asked to estimate the *duration* (number of hours) of 9 bodily movements (sitting, standing, standing while holding objects >50 lbs, walking, walking while carrying objects >50 lbs, pushing objects over 75 lbs, using heavy tools, squatting continuously, and kneeling continuously). In both sections, respondents were also asked to estimate the *frequency* (number of times in an 8-hour period) of 4 bodily movements (squatting, squatting with lifting objects <50 lbs, squatting with lifting objects >50 lbs, and climbing or descending a flight of stairs or a ladder). Though not the main theme of these validation studies, this data will be useful in quantifying lifetime PA exposure in terms of biomechanical forces transmitted through the joint (for the PAJH cohort study). In addition, the duration estimates of the bodily movements in the occupational physical activity section were used in assigning metabolic equivalents and intensities to occupational activities reported (described in Chapter 4.4.1)

4.3 Data Collection

4.3.1 Subject Recruitment

Subjects for the validation studies were recruited from PAJH cohort study participants following baseline data collection. This involved mass electronic mail recruitment and administration of the PAJHS, including the L-PAQ over the Internet.

The following inclusion criteria were used to identify eligible PAJH cohort study participants for the validation studies: 1) completion of the L-PAQ; 2) provision of contact information (telephone number or contact address; and 3) residency in the Lower Mainland of British Columbia. Since the validation studies were applied to the baseline L-PAQ data, it was important that only subjects with complete L-PAQ datasets were included in the recruitment list. Residency in the Lower Mainland, British Columbia area was imperative as one aspect of data collection for the validation studies involved face-to-face interviews at the Arthritis Research Centre (ARC) in Vancouver, British Columbia.

After applying the inclusion criteria to the baseline PAJH cohort study participants, a mailing list for recruitment for the validation studies was generated. Letters of invitation were sent to potential subjects in five mailing waves, consisting of approximately forty individuals in each wave. Mailing waves were sent one-month apart; recruitment and data collection took six months. Information letters sent to potential subjects encouraged them to contact the investigator if they were interested in participating or to anticipate a telephone call from the investigator within two weeks of the mailing. Study procedures, specifically the completion of three physical activity questionnaires, were described during the telephone conversations with potential participants. Validation studies were designed such that LT-PAQ interviews were administered first, followed by subject self-administration of the CT-PAQ and the retest version of the L-PAQ. This was done to facilitate recruitment since the LT-PAQ interviews involved considerable commitment from participants due to the duration of interviews and the requirement for subjects to commute to ARC. Subjects for the validation study received an honorarium for their participation.

4.3.2 Reliability Study: Repeat Administration of the L-PAQ

The reliability study involved a repeat administration of the L-PAQ. Participants in the validation studies were provided with the Internet address and login access to the retest version of the L-PAQ. The questionnaire contained identical questions as the original L-PAQ administered as part of the PAJHS. Sections of the PAJHS on general health, smoking history, and knee and hip health were not included in the retest administration as the reliability tests were focused on the L-PAQ. This also minimized participant burden.

Subject completion of the L-PAQ was monitored using the ARC on-line questionnaire system. Reminder emails and telephone calls were used to ensure completion of the repeat version of the L-PAQ. The time interval between completion of the baseline L-PAQ and the repeat administration of the L-PAQ was recorded.

4.3.3 Validity Study Part I: Interview Administration of the Lifetime Total Physical Activity Questionnaire (LT-PAQ)

The Lifetime Total Physical Activity Questionnaire (LT-PAQ) is a previously validated interviewer-administered lifetime physical activity questionnaire developed by Christine Friedenreich and colleagues at the Alberta Cancer Board. The development and reliability study of the Lifetime Total Physical Activity Questionnaire (LT-PAQ) was published in 1998 (12). The LT-PAQ questionnaire was subsequently used in two separate investigations: a case-control study of lifetime physical activity and breast cancer risk (29) and more recently, a case-control study of lifetime physical activity and prostate cancer risk (30). This questionnaire was designed to assess the amount and level of physical activity that respondents have done over their lifetime. Physical activity was defined into three domains: occupational, household and exercise/sports. The amount of activity involved measurement of the duration and frequency while the level of activity involved asking respondents to rate each reported activity as sedentary, light, moderate, or heavy. Authors of this questionnaire reported a high degree of reliability with Pearson correlation coefficients of 0.72 for exercise/sports, 0.87 for occupation, 0.77 for household, and 0.74 for total physical activity (12).

The LT-PAQ was selected for the validation studies for two main reasons: 1) it is a comprehensive instrument and was developed to include cognitive methods to aid recall of past physical activity, and 2) it defines lifetime physical activity into similar domains of occupation, household, and exercise/sports consistent with the L-PAQ. In addition, its application in two separate investigations of lifetime physical activity and cancer has shown its utility as a PA measurement instrument.

Administration of the LT-PAQ involved face-to-face interviews with subjects. Development of the interviewer script and data collection forms for the validation studies of this thesis involved communication and consultation with members of Dr. Friedenreich's research team. Dr. Friedenreich's research team members also kindly provided materials used during training of their interviewers in Alberta. Thorough pilot testing of the LT-PAQ interviews with a sample of fifteen volunteers from The Arthritis Society (TAS) was conducted prior to actual interviews conducted for the validation studies. LT-PAQ items used for the thesis validation studies are included in Appendix 1.5.

As part of the LT-PAQ interview process participants were mailed a Life Events Calendar prior to each interview. The Life Events Calendar was also described by Friedenreich as memory aid developed to improve recall in subjects (12). Subjects were encouraged to complete the calendars to help facilitate the interview process. All interviews for administration of the LT-PAQ were conducted by the thesis author at ARC. Use of a single interviewer for this study attempted to minimize potential biases that may have been introduced by different interviewers. The duration of all interviews were recorded. To ensure that a consistent interview technique was used throughout the duration of data collection, a random number of interviews were recorded and each recording was reviewed.

4.3.4 Validity Study Part II: Administration of the Chasan-Taber Physical Activity Questionnaire (CT-PAQ)

The second instrument used in the validation studies is a single-page, self-administered lifetime physical activity questionnaire developed by Chasan-Taber et al (3). This instrument, henceforth referred to as CT-PAQ, was adapted from the Historical Leisure Activity Questionnaire (HLAQ), a previously validated interviewer-administered questionnaire measuring historical physical activity (5).

The CT-PAQ was designed to assess the duration, frequency, and intensity of lifetime physical activity across two domains: leisure time activities and household activities. The CT-PAQ incorporated a limited list of twenty-seven leisure time activities (e.g. walking, swimming, biking) and five household activities (gardening, grooming children, playing with children, light housekeeping, and heavy housekeeping). The design of the CT-PAQ divided respondents' lifespan across four age periods of 14-21, 22-34, 35-50, and >50 years. For activities performed more than 10 times in their lifetime, respondents were asked to provide an estimate of the number of years, the months per year, and hours per week spent in each activity during each age period. The CT-PAQ was shown to be highly reproducible with intraclass coefficients of 0.82 for lifetime total PA, 0.87 for lifetime recreational PA, and 0.78 for lifetime household PA (3).

The CT-PAQ was selected for our validation studies for two main reasons: first, it is an easily administered instrument and second, it was also developed to include cognitive methods to aid respondents with recall of their past physical activity.

Validation study participants were given the CT-PAQ following the LT-PAQ interview. Following an explanation of how to complete the questionnaire, participants were given the option of either completing the entire questionnaire at ARC or taking the questionnaire home to complete. While most participants opted to complete the CT-PAQ immediately following their interview, a small number of participants chose to fill the questionnaire at home. A copy of the CT-PAQ is included in Appendix 1.6 (page 145).

4.3.5 Validity Study Part III: Hypothesis Testing of L-PAQ Constructs

The second part of the validity studies involved testing of a priori hypotheses of L-PAQ constructs. Hypotheses were formulated based on known and/or rationalized relationships between physical activity constructs as measured by the L-PAQ and demographic variables collected as part of the PAJH cohort study. Since baseline PAJH cohort study data was used in this analysis, there was no subject recruitment or data collection involved.

4.4 Data Analysis Part I: Questionnaire Scoring

4.4.1 Lifetime Physical Activity Questionnaire Scoring (L-PAQ)

Variables derived from L-PAQ data to describe lifetime physical activity were summarized into three scoring units: 1) total lifetime hours of physical activity (TLH); 2) lifetime average hours per week of physical activity (LAHW); and 3) lifetime MET•hours per week of physical activity (METHW).

For all three domains of physical activity (sports, occupation, and domestic), outcome variables were estimated as the total number of hours spent in each type of activity over each person's lifetime or total lifetime hours (TLH) of physical activity. This was the base unit, which was further used to derive the lifetime average hours per week (LAHW) spent in each type of activity. Data were also expressed as units of energy expenditure by multiplying the hours spent by the estimated metabolic cost of that activity over lifetime to obtain MET•hours per week (METHW).

Scoring algorithms for the L-PAQ were developed based on questions asked in each of the three PA sections (Table 4.2). The generic formula for total lifetime hours of participation is:

Formula 4.1 Total Lifetime Hours = Duration * Frequency * Length of Occasion of Participation

Based on this, the formulae for total lifetime hours spent in *each* reported activity for sports, occupation, and domestic domains were derived and summarized in Figure 4.2.

I. Total Lifetime Hours of Sports PA
(for each reported sports activity)

Formula 4.2 TLHSn = YOPSn * WPYSn * OPWSn * HPOSn

$$\textbf{TLHSn} = \textbf{YOPSn} * \left[\frac{x\ \textbf{months}}{\textbf{year}} * \frac{(4.33\ \text{weeks})}{\textbf{month}}\right] * \frac{x\ \textbf{occasions}}{\textbf{week}} * \frac{x\ \textbf{hours}}{\textbf{occasion}}$$

TLHSn = total lifetime hours of sports/exercise PA
YOPSn = years of participation in sports (*duration*)
WPYSn = weeks per year participation in sports *(frequency)*
OPWSn = occasions per week participation in sports
HPOSn = hours per occasion participation in sports *(length of occasion of participation)*

II. Total Lifetime Hours of Occupational PA
(for each reported occupation)

Formula 4.3 TLHOn = YOPO * WPYO * HPWO

TLHOn = total lifetime hours of occupation
YOPOn = years of participation in occupation (*duration*)
WPYOn = weeks per year participation in occupation *(frequency)*
HPWOn = hours per week participation in occupation *(length of occasion of participation)*

III. Total Lifetime Hours of Domestic PA
(for each reported domestic activity)

Formula 4.4 TLHDn = YOPD * WPYD * HPWD

$$\text{WPYD} = \left[\frac{52\ \text{weeks}}{\text{year}}\right]$$

TLHDn = total lifetime hours of domestic PA
YOPDn = years of participation in domestic PA(*duration*)
WPYDn = weeks per year participation in domestic PA *(frequency)* * *Treated as a constant of 52 hr/wk)*
HPWDn = hours per week participation in domestic PA *(length of occasion of participation)*

Figure 4.2 Formulae for Total Lifetime Hours Participation for *Each* Reported Activity in Three Domains of Physical Activity Measured by the L-PAQ

Calculation of total lifetime hours of participation for each physical activity domain involved three steps. These steps are summarized in Figure 4.3. First, Formulae 4.2, 4.3, and 4.4 were applied for *each* reported sport, job, and domestic activity to obtain total lifetime hours of participation for *each reported activity*. Second, total lifetime hours for *each physical activity* *domain* was obtained by summing total lifetime hours for each reported activity (in that domain). Third, total lifetime hours of *total physical activity* was calculated by adding total lifetime hours across all three domains.

Sports PA	**Occupational PA**	**Domestic PA**
Step 1: Calculating Total Lifetime Hours for *Each Reported Activity*		
$\text{Sports}_n \xrightarrow{\text{Formula4.2}} \text{TLHS}_n$	$\text{Job}_n \xrightarrow{\text{Formula 4.3}} \text{TLHO}_n$	$\text{Domestic}_n \xrightarrow{\text{Formula4.4}} \text{TLHD}_n$
Step 2: Calculating Total Lifetime Hours for *Each Physical Activity Domain*		
Formula 4.5 **TLHS** $= \Sigma\,(\text{TLHS}_1 + \text{TLHS}_2 + \ldots \ldots + \text{TLHS}_n)$	**Formula 4.6** **TLHO** $= \Sigma\,(\text{TLHO}_1 + \text{TLHO}_2 \ldots\ldots + \text{TLHO}_n)$	**Formula 4.7** **TLHD** $= \Sigma\,(\text{TLHD}_1 + \text{TLHD}_2 \ldots\ldots + \text{TLHD}_n)$
Step 3. Calculating Total Lifetime Hours for *Total Physical Activity*		
Formula 4.8 TLHPA = TLHS + TLHO + TLHD		

Figure 4.3 Steps in Calculating Total Lifetime Hours at 3 Levels: 1) Individual Activity 2) Physical Activity Domain; 3) Total Physical Activity

With the base unit of total lifetime hours, further calculations were applied to quantify physical activity in terms of average hours per week of activity over lifetime. This was done by dividing total lifetime hours for each reported activity by the constant of 52 (weeks per year) and by respondent age as shown in Formula 4.9.

Formula 4.9 Lifetime Average Hr/Wk $= \dfrac{\text{Total Lifetime Hours}}{52 * \text{Age}}$

Calculations for average hours per week of physical activity over a lifetime are summarized in Figure 4.4. Similar to calculation of total lifetime hours, these also involved three steps: 1) calculation of lifetime average hours/week for *each reported activity*; 2) calculation of lifetime average hours/week for *each physical activity domain*; 3) summation of lifetime average hours/week across sports, occupation, and domestic domains to obtain lifetime average hours/week of *total physical activity*.

Step 1: Calculating Lifetime Average Hours/Week for *Each Reported Activity*		
Sports PA	**Occupational PA**	**Domestic PA**
$\mathbf{LAHWS_n} = \dfrac{TLHS_n}{52 \bullet Age}$	$\mathbf{LAHWO_n} = \dfrac{TLHO_n}{52 \bullet Age}$	$\mathbf{LAHWD_n} = \dfrac{TLHO_n}{52 \bullet Age}$
Step 2: Calculating Lifetime Average Hours/Week for *Each Physical Activity Domain*		
Formula 4.10	**Formula 4.11**	**Formula 4.12**
$\mathbf{LAHWS}=\Sigma(LAHWS_1 + LAHWS_2 + \ldots + LAHWS_n)$	$\mathbf{LAHWO}=\Sigma(LAHWO_1 + LAHWO_2 \ldots + LAHWO_n)$	$\mathbf{LAHWD}=\Sigma(LAHWD_1 + LAHWD_2 \ldots + LAHWD_n)$
Step 3. Calculating Lifetime Average Hours/Week for *Total Physical Activity*		
Formula 4.13 LAHWPA = LAHWS + LAHWO + LAHWD		

Figure 4.4 Steps in Calculating Lifetime Average Hours/Week at 3 Levels: 1) Individual Activity 2) Physical Activity Domain; 3) Total Physical Activity

Finally, to convert data into energy expenditure, average hours per week in each reported activity was multiplied by the estimated metabolic cost of that activity using activity codes and metabolic equivalents (MET) reported in the Compendium of Physical Activities (18). The resulting unit of physical activity is MET•hours per week.

Formula 4.14 MET•hour/week = Average hours/week • MET (metabolic equivalent)

For each sports and domestic activity, representative MET codes were assigned using the Compendium of Physical Activities (18). These are summarized in Appendix 1.2 and Appendix 1.4 respectively. Assignment of MET values for occupational activity had the potential for extreme variability based on highly varied occupations and occupational titles. Thus, rather than assigning MET values on job titles, they were assigned on the reported duration of nine body movements during a typical eight-hour working day. These movements were: 1) sitting; 2) standing; 3) standing and holding or moving objects over 50 lbs; 4) walking; 5) walking and carrying objects over 50 lbs; 6) moving or pushing objects over 75 lbs; 7) using heavy tools; 8) squatting continuously; 9) kneeling continuously. MET codes and values for these body movements during occupation are also summarized in Appendix 1.2.

Similar to previously described calculation for total lifetime hours and lifetime average hours/week, calculation of MET•hours/week also involved three steps: 1) calculation of MET•hours/week for *each reported activity*; 2) calculation of MET•hours/week for *each physical activity domain*; 3) summation of MET•hours/week across sports, occupation, and domestic domains to obtain MET•hours/week of *total physical activity*. These steps are summarized in Figure 4.5

Step 1: Calculating Lifetime MET•Hours/Week for *Each Reported Activity*		
Sports PA	**Occupational PA**	**Domestic PA**
$\mathbf{METHWS_n} = \text{LAHWS}_n \bullet \text{MET}$	$\mathbf{METHWO_n} = \text{LAHWO}_n \bullet \text{MET}$	$\mathbf{METHWD_n} = \text{LAHWD}_n \bullet \text{MET}$
Step 2: Calculating Lifetime MET•Hours/Week for *Each Physical Activity Domain*		
Formula 4.15	**Formula 4.16**	**Formula 4.17**
$\mathbf{METHWS} = \Sigma(\text{METHWS}_1 + \text{METHWS}_{\backslash 2} + \ldots + \text{METHWS}_n)$	$\mathbf{METHWO} = \Sigma(\text{METHWO}_1 + \text{METHWO}_2 + \ldots + \text{METHWO}_n)$	$\mathbf{METHWD} = \Sigma(\text{METHWD}_1 + \text{METHWD}_2 + \ldots + \text{METHWD}_n)$
Step 3: Calculating Lifetime MET•Hours/Week for *Total Physical Activity*		
Formula 4.18 METHWPA = METHWS + METHWO + METHWD		

Figure 4.5 Steps in Calculating Lifetime MET•Hours/Week of Physical Activity at 3 Levels: 1) Individual Activity 2) Physical Activity Domain; 3) Total Physical Activity

4.4.2 Lifetime Total Physical Activity Questionnaire (LT-PAQ) Scoring

Methods for scoring of the LT-PAQ were developed using literature describing this questionnaire (12, 29, 30) and communication with members of Dr. Friedenreich's research team. Formulae for average number of hours per week spent in occupational activity, household activity, and exercise/sports activity were obtained from the manuscript reporting on the LT-PAQ's development and are summarized in Appendix 1.5. Scoring and summation steps to obtain total physical activity in each of the three sections were similar to those used in the L-PAQ calculations. LT-PAQ total physical activity was estimated as the sum of occupational, household, and exercise/sports activities in hours/week. Energy expenditure was also calculated for the data using MET codes generously provided by Dr. Friedenreich's research team.

4.4.3 Chasan-Taber Life Physical Activity Questionnaire (CT-PAQ) Scoring

Methods for scoring of the CT-PAQ were developed using a variety of literature related to this questionnaire including: 1) the report of the development and reliability study (3); and 2) the report of the development and validation of the Historical Leisure Activity Questionnaire from which the CT-PAQ was adapted (5). Formulae for average number of hours per week spent in leisure activity and household activity were obtained from the manuscript reporting on the CT-PAQ's development and are summarized in Appendix 1.6. Scoring and summation steps to obtain total physical activity in each of the three sections were similar to those used in the L-PAQ calculations.

Scores for sports PA (and consequently total PA) were calculated with and without the inclusion of respondent self-reports of walking. This follows because CT-PAQ authors reported on the lack of reliability of self-reports of walking (3).

4.5 Data Analysis Part II: Analysis of Validation Studies

Section 4.1 provided a brief overview of the methods applied in this thesis project as well as a flowchart summarizing datasets used for analysis. In sum, four separate datasets were analyzed – one dataset for the reliability study and three datasets for the validity studies. Not all subjects completed all three components of the validation studies, requiring for creation of these separate datasets (such that they contained data for only subjects who completed that specific study). All datasets were compiled following application of scoring algorithms to the questionnaires within each dataset. Table 4.2 summarizes the datasets used in the analysis in terms of data contained and type of study. All analyses were performed using SPSS Version 12.0 (SPSS Inc, Chicago, Illinois).

Table 4.2 Summary of Datasets Used in Analyses of Validation Studies

Dataset	Validation Study	Specific Study Details	Sample Size	Questionnaires Contained in Dataset	
1	Reliability	Test-retest administration of L-PAQ	76*	Baseline administration of L-PAQ	Repeat administration of L-PAQ
2	Construct Validity	Correlation between LPAQ and LTPAQ	84*	Baseline administration of L-PAQ	LT-PAQ interview-administered
3	Construct Validity	Correlation between L-PAQ and CT-PAQ	80*	Baseline administration of L-PAQ	CT-PAQ subject-administered
4	Construct Validity	Hypothesis testing of constructs measured by L-PAQ	4269	Baseline administration of L-PAQ	

**These three datasets essentially consisted of the same subjects. However, not all subjects recruited for the validation studies completed all three parts, thus the variation in sample size.*

4.5.1 L-PAQ Reliability

The application of statistical techniques for the analysis and interpretation of a reliability study depend on specific aspects of the study design (20). The reliability of continuous measures in intramethod studies is estimated by the *intraclass correlation coefficient (ICC)*, which is based on analysis of variance (ANOVA) techniques and defined by the equation:

Formula 4.19 $$\rho_x = \text{ICC} = \frac{\sigma^2_{\text{subjects}}}{\sigma^2_{\text{subjects}} + \sigma^2_{\text{error}}}$$

Formula 4.19 is the general form of the ICC, which is a ratio of the variance of the subject expected scores on the instrument over total variance. There are four versions of the intraclass correlation and the choice of which version to use depends on the design of the study (20). ICC1 is often used for simple replication studies in which there is no characteristic that differentiates the first and second measure across all subjects. Two versions of the intraclass correlation, the ICC2 and ICC3, are based on two-way ANOVA models and are applicable to instruments that are administered by raters or interviewers. ICC2 is used when the raters in the reliability study are the same as the raters in the (larger) epidemiologic study for which the instrument will be used. Thus, the observers are said to be fixed, and a *fixed* effects ANOVA model applies. ICC3 is used when the raters in the reliability study are a *sample* from the population of raters in the (larger) epidemiological study. In contrast to ICC2, the observers are now random and the 2-way *random* effects ANOVA model applies (20). The final version, ICC4 excludes measurement effects as a source of variations in the measure, X. This is based on a 2-way fixed effects ANOVA model. The formulae for ICC 4 are:

Formula 4.20 $$\mathrm{ICC4} = \frac{\sigma_{\mathrm{subject}}^2}{\sigma_{\mathrm{subject}}^2 + \sigma_{\mathrm{error}}^2}$$

Formula 4.21 $$\mathrm{ICC4} = \frac{(\mathrm{SMS} - \mathrm{EMS})}{\mathrm{SMS} + (k - 1)\mathrm{EMS}}$$

Of interest in validation studies of this thesis is the reproducibility of the L-PAQ when it is administered in the same fashion to the same respondents at two or more occasions. A test-retest study which establishes the reliability of an instrument administered to the same subjects at two separate occasions was conducted. ICC4 was calculated as the most appropriate descriptor of L-PAQ reliability. Coefficients were calculated between L-PAQ administrations for each scoring unit (total lifetime hours, average hr/wk, MET•hr/wk) and across all PA domains (sports, occupation, domestic, and total PA). For each coefficient, 95% confidence intervals were also obtained.

Further analyses were conducted to evaluate L-PAQ reliability. Questionnaire scores for the baseline and repeat administration of the L-PAQ were compared using Wilcoxon signed-rank test to test differences among pairs. Non-parametric methods were used due to the skewed distribution of the scores. Finally, Bland-Altman plots were constructed (39). This method describes the agreement between two quantitative measurements, in this case, the baseline and repeat administration of the L-PAQ. Bland-Altman plots were constructed for each scoring unit (total lifetime hours, average hours/week, MET•hours/week) and across all PA domains (sports, occupation, domestic, and total PA) by graphing the difference between administrations of the L-PAQ against the mean of both administrations for all paired values (scores for each respondent across administrations). The standard error of measurement (SEM) was also calculated using the equation:

Formula 4.22 $$\text{SEM} = \frac{\text{Standard Deviation of Differences (Between Administrations)}}{2^{1/2}}$$

4.5.2 L-PAQ Validity Part I: Comparative Studies of L-PAQ and LT-PAQ

Correlation coefficients were obtained between similar domains of physical activity measured by the L-PAQ and the LT-PAQ for similar scoring units. Corresponding baseline L-PAQ data for subjects completing the LT-PAQ interviews were scored for the correlation analysis. A total of 8 correlations were calculated:

Units: Lifetime Average Hours/Week

1. L-PAQ sports/recreation versus LT-PAQ sports/exercise
2. L-PAQ occupation versus LT-PAQ occupation
3. L-PAQ domestic versus LT-PAQ household
4. L-PAQ total PA versus LT-PAQ total PA

Units: MET•Hours/Week

5. L-PAQ sports/recreation versus LT-PAQ sports/exercise
6. L-PAQ occupation versus LT-PAQ occupation
7. L-PAQ domestic versus LT-PAQ household
8. L-PAQ total PA versus LT-PAQ total PA

The distribution of physical activity scores and measures of skewness were obtained to determine the appropriate correlation coefficient to report. Since questionnaire scores had a skewed distribution, non-parametric methods were used for the analyses. Specifically, Spearman correlation coefficients were obtained to determine the relationship between L-PAQ and LT-PAQ scores.

Questionnaires scores for the L-PAQ and LT-PAQ were also compared by conducting Wilcoxon signed-rank tests for differences among pairs, setting the significance level at p=0.05 a priori. Finally, sensitivity analyses were performed by repeating the comparison using L-PAQ data from the repeat administration of the questionnaire.

4.5.3 L-PAQ Validity Part II: Comparative Studies of L-PAQ and CT-PAQ

Correlation coefficients were obtained between similar domains of physical activity measured by the L-PAQ and CT-PAQ. Corresponding baseline L-PAQ data for subjects who completed the CT-PAQ interviews were scored for the correlation analysis. Since CT-PAQ scores for sports (and consequently total physical activity) were calculated with and without the inclusion of subject reports of walking, separate correlations were performed to determine the effects of the inclusion of walking. Therefore, a total of 12 correlations were calculated:

Units: Lifetime Average Hours/Week

1. L-PAQ sports/recreation versus CT-PAQ leisure time activity (walking included)
2. L-PAQ domestic versus CT-PAQ household
3. L-PAQ total PA versus CT-PAQ total PA (walking included)
4. L-PAQ sports/recreation versus CT-PAQ leisure time activity (walking excluded)
5. L-PAQ domestic versus CT-PAQ household
6. L-PAQ total PA versus CT-PAQ total PA (walking excluded)

Units: MET•Hours/Week

7. L-PAQ sports/recreation versus CT-PAQ leisure time activity (walking included)
8. L-PAQ domestic versus CT-PAQ household
9. L-PAQ total PA versus CT-PAQ total PA (walking included)
10. L-PAQ sports/recreation versus CT-PAQ leisure time activity (walking excluded)
11. L-PAQ domestic versus CT-PAQ household
12. L-PAQ total PA versus CT-PAQ total PA (walking excluded)

Similar analyses applied to comparison of the L-PAQ with the LT-PAQ were applied to the comparison of the L-PAQ with the CT-PAQ. The distribution of physical activity scores and measures of skewness were obtained to determine the appropriate correlation coefficient to report. Since questionnaire scores had a skewed distribution, Spearman correlation coefficients were obtained to determine the relationship between L-PAQ and CT-PAQ scores.

Questionnaire scores for the L-PAQ and CT-PAQ were also compared by conducting Wilcoxon signed-rank tests for differences among pairs, setting the significance level at p=0.05 a priori. Finally, sensitivity analyses were performed by repeating the comparison using L-PAQ data from the repeat administration of the questionnaire.

4.5.4 L-PAQ Validity Part III: Hypotheses Testing of L-PAQ Constructs

Hypotheses tests were conducted to explore rationalized relationships between physical activity measured by the L-PAQ and variables descriptive of respondents of the PAJH cohort study. Univariate analyses were performed to determine the relationships between constructs measured by the PAJH survey and L-PAQ physical activity scores. Since baseline L-PAQ data was used for the hypotheses tests and the sample size (n=4269) was large, parametric methods were used for the hypotheses tests.

Hypotheses Based on Gender

We expected a gender difference in sports participation with males having higher sports PA scores than females. This relationship was shown in previous validation studies of the QUANTAP (15). This difference will be tested for duration of participation in terms of total lifetime hours and average hours/week L-PAQ units and for intensity of participation with MET•hours/week.

We also expected a gender difference in occupational activity with males having higher scores than females. This relationship was also shown in the validation study of the QUANTAP (15). This gender difference could also be rationalized by the fact that males would likely have higher number of working years since women would be more likely to take time off work to care for children. Of particular interest with hypotheses testing of occupational PA scores was the MET•hours/week scoring unit which takes into account bodily movement during work. We expected males to participate in more physical-related and/or physically demanding working and thus have higher energy expenditure scores in occupational PA.

Finally, we expected an opposite trend with domestic activity, specifically for females to have higher domestic activity scores than males. Previous researchers have demonstrated that for women, household activity is a major contributor to weekly energy expenditure (40). This is further supported by authors of both the LT-PAQ and the CT-PAQ who incorporated the domain of household activity in their questionnaires in recognizing that previous physical activity questionnaires excluding this domain may have underestimated or misclassified women's physical activity (3, 12).

Independent samples t-tests were used to test gender differences across PA domains. A total of five hypotheses were tested for gender differences in measured lifetime physical activity.

Hypotheses Based on Education Level

In the Demographic Information section of the PAJHS, participants were asked to report the highest level of education they completed. This question was set as a categorical variable with five options: 1) elementary school, 2) high school, 3) technical or trade school, 4) college/university, and 5) post-graduate. There was a sixth option for respondents who did not provide a response – for purposes of this analysis, the small percentage of respondents who did not report their highest level of education (1.1%) was coded as missing data.

Unlike the gender-based hypotheses for which previous authors have related to lifetime physical activity (15), there are no studies to date that have described the relationship of socioeconomic factors to *lifetime* physical activity. Thus the rationale applied to the hypotheses tests of education level and lifetime physical activity were based on social science literature relating participation in *current* physical activities to socioeconomic factors.

One hypothesis tested was the relationship between education level and intensity of occupational activity over lifetime in males. Specifically, we expected males reporting lower education levels (elementary school) to have the highest intensity of occupational activity than males reporting higher education levels (university/college and post-graduate study). We also expected males who attended trades and technical to have intermediate intensity of occupational activity. Fogelman et al. studied the relationship between socioeconomic and health factors and physical activity using the Baeke Physical Activity Questionnaire, an instrument measuring past month activity and showed that males with lower education levels had greater physical activity at work (41). This relationship was also shown by He et al. (42). Relevant L-PAQ scoring unit for this hypothesis test is MET•hours/week as it provides an indicator of intensity of physical activity.

Another hypothesis tested was the relationship between education level and sports activity. The hypothesis that people with lower levels of education will have lower sports activity participation was tested in both males and females. Droomers et al. showed this relationship in their study of educational differences in leisure-time physical *in*activity (43). Similar findings of high level of education and greater engagement in sports activities were shown by other authors (41, 44). However as with the rationale for the previous hypothesis on intensity of occupational activity and level of education in males, descriptors using *current* measures of physical activity were extended to our measure of *lifetime* physical activity. A consequent hypothesis that people with lower levels of education will have lower intensity of sports PA was also tested using the L-PAQ MET•hours/week scoring unit.

One way analysis of variance (ANOVA) was used to test hypothesized differences between reported level of education and measured lifetime physical activity. A total of five hypotheses were tested.

Hypotheses Based on General Health Status

In the General Health Section of the PAJH, participants were asked to rate their general health. Three groups were created based on subject responses: Group 1: Excellent/Very Good Health; Group 2: Good Health; and Group 3: Poor/Fair Health.

The specific hypothesis tested was the relationship between reported general health status and sports activity. We expected respondents who reported very good to excellent general health to have higher sports PA scores than participants who reported fair or poor general health. This relationship was tested on two L-PAQ scoring units, lifetime average hours/week and MET•hours/week. The exact wording of the question was "In general would you say your health is" and subjects had the option of reporting one of "Excellent, Very Good, Good, Fair, or Poor." Such wording would likely prompt respondents to think about present health and thus, evaluating the relationship between general health and L-PAQ scores on units of total lifetime hours of activity would not be relevant. Lifetime average hours/week was the most relevant scoring unit to test the hypothesis because its averaged nature takes into account subjects' current activity and it was the most readily interpreted unit. MET•hours/week is also an averaged unit and provides an indicator of the intensity of the activity.

The relationship between reported general health and physical activity measured by the L-PAQ was tested using one-way ANOVA. Analyses were performed separately for males and females and a total of four hypotheses tests were conducted.

Hypotheses Based on Body Mass Index

In the General Health Section of the PAJH survey, respondents were asked to report their current body weight and height. Using these variables, current body mass index (BMI) was obtained for each subject. The BMI is a number calculated from an individual's weight and height and is used as an indirect indicator of body fatness for people.

In their validation of the QUANTAP, Vuillemin et al. looked at the relationship between percent body fat obtained at the time of survey administration and leisure-time physical activity. Authors reported negative correlations with subjects who had lower percent body fat having higher scores for leisure time activity (15). In this study, percent body fat was obtained using dual energy x-ray absorptiometry. While such information was not collected for the PAJH cohort study, information on subject BMI was used as proxy measures for body fatness in participants. Previous research has shown that BMI correlates well with direct measures of body fat including underwater weighing techniques and dual energy x-ray absorptiometry (DXA) (45, 46).

Using Centre for Disease Control standards (CDC Website) (47), subjects' current BMI and were assigned into one of three groups: 1) normal (18.5 – 24.9); 2) overweight (25.0 – 29.9) and 3) obese (>30.0). We expected people who have lower BMI to have higher sports PA scores and people with higher BMI to have lower sports PA scores. Similar to hypotheses tests based on general health status, analyses were done for L-PAQ scoring units of lifetime average hours/week and MET•hours/week as they were relevant scoring units for these tests due to their averaged nature. The relationship between BMI and sports physical activity measured by the L-PAQ was tested using one-way ANOVA. Analyses were performed separately for males and females and a total of four hypotheses tests were conducted. In addition to analyses using ANOVA, correlation coefficients were obtained between BMI and L-PAQ sports PA scores (for comparison with Vuillemin et. al's findings).

Table 4.3 provides a summary of the hypotheses tests conducted on the L-PAQ. Specific hypotheses are numbered numerically (H#) and described along with the concept from which the hypothesis was based and the L-PAQ scoring unit on which the test will be conducted.

Table 4.3 Hypotheses Tests of the L-PAQ

Concept		Specific Hypothesis	L-PAQ Scoring Unit
Hypotheses Based on Gender			
Gender difference in *amount* of sports activity	**H1**	Males will have higher participation in sports than females	Total lifetime hours Average hours/week
Gender difference in *intensity* of sports activity	**H2**	Males will have a higher intensity of sports participation than females	MET•hours/week
Gender difference in *amount* of occupational activity	**H3**	Males will have higher occupational PA scores than females	Total lifetime hours Average hours/week
Gender difference in *intensity* of occupational activity	**H4**	Males will have higher intensity of occupational PA than females	MET•hours/week
Gender difference in *amount* of domestic activity	**H5**	Females will have higher domestic PA scores than males	Total lifetime hours Average hours/week
Gender difference in *intensity* of domestic activity	**H6**	Females will have higher intensity of domestic PA than males	MET•hours/week
Hypotheses Based on Education Level			
Difference in *intensity* of occupational activity based on level of education in males	**H7**	Males with elementary education will have the highest intensity of occupational activity, males with college/university or post-graduate education will have the lowest intensity of occupational activity, and males with trade and technical training will have intermediate intensity of occupational activity	MET•hours/week
Difference in *amount* of sports activity based on level of education (males and females)	**H8**	Males with lowest levels of education will have lowest scores for sports PA and males with highest levels of education will have highest scores for sports PA	Total lifetime hours Average hours/week
	H9	Females with lowest levels of education will have lowest scores for sports PA and males with highest levels of education will have highest scores for sports PA	Total lifetime hours Average hours/week
Difference in *intensity* of sports activity based on level of education (males and females	**H10**	Males with lowest levels of education will have lowest intensity of sports PA and males with highest levels of education will have intensity sports PA	MET•hours/week
	H11	Females with lowest levels of education will have lowest intensity of sports PA and females with highest levels of education will have intensity sports PA	MET•hours/week

Table 4.3 (continued). Hypotheses Tests of the L-PAQ

Concept		Specific Hypothesis	L-PAQ Scoring Unit
Hypotheses Based on General Health			
Difference in *amount* of sports activity based on general health status (males and females)	**H11**	Males reporting excellent/very good health will have the highest scores for sports PA and males reporting poor/fair health will have the lowest scores for sports PA	Average hours/week
	H12	Females reporting excellent/very good health will have the highest scores for sports PA and females reporting poor/fair health will have the lowest scores for sports PA	Average hours/week
Difference in *intensity* of sports activity based on general health status (males and females)	**H13**	Males reporting excellent/very good health will have the highest intensity of sports PA and males reporting poor/fair health will have the lowest intensity of sports PA	MET•hours/week
	H14	Females reporting excellent/very good health will have the highest intensity of sports PA and females reporting poor/fair health will have the lowest intensity of sports PA	MET•hours/week
Hypotheses Based on BMI			
Difference in *amount* of sports activity based on BMI (males and females)	**H15**	Males with low BMI will have the highest scores for sports PA and males with high BMI will have the lowest scores for sports PA	Average hours/week
	H16	Females with low BMI will have the highest scores for sports PA and females with high BMI will have the lowest scores for sports PA	Average hours/week
Difference in *intensity* of sports activity based on BMI (males and females)	**H17**	Males with low BMI will have the highest intensity of sports PA and males with high BMI will have the lowest intensity of sports PA	MET•hours/week
	H18	Females with low BMI will have the highest intensity of sports PA and females with high BMI will have the lowest intensity of sports PA	MET•hours/week

V. RESULTS

5.1 Subjects

Recruitment for baseline data collection for the PAJH cohort study resulted in 6,026 individuals accessing the PAJH survey. With participation in the cohort study defined as completion of the entire PAJH survey, the final number of participants across Canada was 4,269. Subjects' mean age was 61.5 years and median age was 60.8 years. There were more females in the study population (n=2694) than males (n=1575).

Of 4269 PAJH cohort study participants, 1394 did not provide any contact information and 551 provided an indicator of residence in British Columbia. Since it was optional for participants to provide their contact information in the PAJH survey, further sorting was performed to determine participants who resided in Lower Mainland, British Columbia and provided complete information for contact/recruitment for the validation studies. A mail out list of 207 possible validation study participants was generated from this sorting process.

In total, 88 subjects consented to participation in the validation studies (43% recruitment rate), surpassing the intended sample size of 80 participants for the validation studies. The mean age of validation study participants was 64.0 years and there were 45 females and 43 males. Thirty-eight participants (47.5%) contacted the thesis author by telephone to indicate their interest in participating in the study and 29 (32.9%) contacted the thesis author by e-mail. Twenty-nine participants (23.8%) were recruited through telephone calls.

Participants in the validation studies were compared to participants in the PAJH cohort study to ensure that a representative sample has been used. This has important implications for reporting of results of the PAJH cohort study. Comparisons were conducted using independent samples t-test on continuous demographic variables and χ^2 tests on categorical variables.

Table 5.1 provides a comparison between baseline PAJH study and validation study participants for continuous and categorical variables. Continuous variables included age, height, and weight and categorical variables included gender distribution, marital status, ethnic origin, highest level of education obtained, and total household income. Results of independent samples t-tests for continuous variables show that validation study participants (64.0 years) were older than PAJH study participants (61.5 years). The groups did not differ across physical characteristics (height and body weight). Results of χ^2 tests on categorical variables show that there was a more equal gender distribution in the validation study participants with 49% males and 51% females. The PAJH cohort consisted of more females (36% males, 64% females). However, validation study participants did not differ from PAJH cohort study participants in any of the other demographic variables tested (marital status, ethnic origin, highest level of education, total household income) and a representative sample was recruited.

Table 5.1 Comparison of Subjects in the Validation Studies to Subjects in the PAJH Cohort Study

	PAJH Participants (n=4,181)	Validation Studies Participants (n=88)	p-value
Continuous Variables **p-values for independent samples t-tests*			
Mean Age (years)	61.5 ± 7.6	64.0 ± 7.6	0.02*
Height (inches)	67.3 ± 5.0	67.9 ± 3.6	0.15
Body weight (lbs)	177.4 ± 39.6	173.8 ± 33.1	0.30
Categorical Variables **p-values for χ^2 tests*			
Gender			
Male	36.6	49	0.02*
Female	63.4	51	
Marital Status			
Single	4.7	9.1	0.06
Married/Common Law	58.0	69.9	
Divorced/Separated/Widowed	33	24.9	
Other/Did not provide	0	0.6	
Ethnic Origin			
White	93.6	90.9	0.61
Other	6.3	9.1	
Highest Level of Education			
Elementary / High School	33.8	29.5	0.87
Technical or Trade School	17.4	18.2	
College/University/Post Graduate Study	47.7	51.1	
No Answer	1.1	1.1	
Total Household Income			
$0 - $19,999	8.5	5.7	0.47
$20,000 - $39,999	24.8	34.3	
$40,000 - $59,999	24.4	25.7	
$60,000 - $79,999	18.1	15.7	
$80,000 - $99,999	10.6	7.1	
$100,000 or more	11.8	11.4	
Don't know / Did not provide	1.9	0.0	

5.2 Baseline L-PAQ Scores

Variables derived from L-PAQ data to describe lifetime physical activity were summarized into three L-PAQ scoring units (metrics): 1) total lifetime hours of physical activity (TLH); 2) lifetime average hours per week of physical activity (LAHW); and 3) lifetime MET•hours per week of physical activity (METHW). L-PAQ summary scores were calculated for the baseline administration of the L-PAQ for sports, occupation, domestic, and total PA.

Following initial calculations, some steps were applied in the scoring process of the L-PAQ to clean the data prior to analyses for the validity and reliability studies. Since the purpose of these data adjustments was to clean and *not* manipulate the data, a conservative approach was used in rationalizing and applying the following steps:

Sports/Recreational Physical Activity Section

(1) Data from "Activity 1" to "Activity 4" which were open-ended questions asking participants to report on sports/recreational activities not provided in the list were excluded. Responses ranged from "hundreds of activities, question too broad" to "mostly standing" to "sitting" to repeated reports of activities covered in the list (for example reports of curling, walking, swimming). Such variation also precluded assignment of MET values for calculating MET•hours/week of activity and thus, these data were not included in the calculation of L-PAQ scores for analysis of the validation studies.

Occupational Physical Activity Section

(2) An adjustment was applied in calculation of MET•hours/week for occupational activity. As described, assignment of MET intensities was based on reported duration of

nine body movements during a typical eight-hour working day. Some respondents over reported on combination of movements, resulting in hours of occupational activity exceeding 168 hours/week (despite these respondents having reported typical work-week hours in a question that specifically asked how many hours they worked in a week). Since it not possible to have greater than 168 hours/week of combined bodily movements in occupational PA, these data were recoded as missing.

<u>Domestic Physical Activity Section</u>

(3) In the domestic PA section, respondents were asked to report on "other" household activity that they may perform. Similar to the "other activities" open-ended questions in the sports/recreational section, responses ranged from "walking dog" to "mopping" to repeated reports of activities covered in the list. Such variation also precluded assignment of MET values for calculating MET•hours/week of activity and thus, these data were recoded as missing.

(4) Respondents were asked to *report* the number of hours per week spent doing 4 types of household activity (caring for children, caring for elderly, gardening, and housework). A number of subjects reported a maximum 168 hours/week of childcare, mis-interpreting the literal meaning of the question. Since this is an impossibility that would significantly affect calculation of physical activity scores, *reported* activities greater than 126 hours per week were recoded as missing data. This was chosen as a conservative ceiling by study investigators as an appropriate time spent in domestic activity that would still allow for minimal self-care and sleep. In addition, experience by

the principal investigator during the face-to-face interviews encountered mothers who reported up to 18 hours/day for 7 days/week (126 hour/week) caring for children (especially with infants and in families with multiple children). These subject reports remained consistent despite use of probing interview techniques. Finally, social science literature was reviewed for reports or descriptions of time spent by women in domestic activity (48, 49). Thus, it was rationalized that 126 hours/week would be an appropriate ceiling to apply.

All L-PAQ Sections

(5) For calculations of years of participation (YOP) for all three physical activity sections, any negative values for YOP were recoded as system missing (data not used) and any values of zero were recoded to "1". Since YOP was calculated as the difference between the age that participation has ended and the age that participation began, possible entry errors included input of the ages in the wrong order, resulting in a negative value for YOP. YOP directly feeds into the calculation for total lifetime hours (TLH) and a negative value for YOP would result in a negative TLH value which is not possible. Thus, it was rationalized that data would not be used in the analysis. Recoding YOP values of zero to 1 would apply to participants who reported starting and ending participation at the same age (which essentially equates to 1-year of participation).

(6) For some respondents, *calculated* lifetime average hours per week of physical activity greater were greater than 168 hours/week. Since it is not possible to have physical activity greater than 168 hours/week, data for the appropriate section were

recoded as missing if *calculated* average hours per week was greater than168. Since data from each PA section contributes to the calculation of lifetime average hours per week for total physical activity, it is intuitive that this data for total average hours per week should also be recoded as missing if any of the PA sections were coded as missing (>168 hours/week).

(7) A parallel situation to #6 is *calculated* average hours per week for total lifetime PA that is greater than 168, despite average hours per week in each PA domain being lower than 168 hours. In other words, the sum of the sections is greater than the number of hours in a week. In this case, data for all PA sections and total PA were recoded as missing.

Baseline L-PAQ scores presented are those obtained following application of the described steps of data cleaning. Table 5.2 summarizes calculated physical activity scores for sports, occupation, domestic, and total physical activity for three scoring units of the L-PAQ: 1) total lifetime hours, 2) lifetime average hours/week, and 3) lifetime MET•hours/week. Rows for lifetime average hours/week are highlighted as this is the most readily interpreted L-PAQ scoring unit, and thus will most often be referred to in description of the results (the same format of presentation of results will be utilized throughout the remainder of this chapter). Table 5.2 is followed by figures illustrating the distribution of the scores. Each figure is a group of frequency histograms (with normal curves) corresponding to scores for each of the three physical activity domains and for total PA. Titles and page numbers of corresponding figures for summary scores are tabulated in the last column of Table 5.2.

Similar calculations were applied to L-PAQ datasets used in the different validation studies. To facilitate reporting for the remainder of this chapter, tabulated results for calculated L-PAQ physical activity scores will be presented. For corresponding frequency histograms, please refer to Appendix II. The final column in each table indicates the titles and locations (thesis chapter and page number) of corresponding figures.

Table 5.2 Baseline L-PAQ Summary Scores for All PAJH Cohort Study Participants (n=4269)

	Sports	Occupation	Domestic	Total PA	Frequency Histogram Figure #
L-PAQ Scoring Unit: Total Lifetime Hours of Activity					
Mean	10008	68131	70233	148372	Fig 5.1 (page 71)
Std Dev	16694	38212	75622	83589	
Median	4925	68640	44278	129502	
Minimum	0	0	0	0	
Maximum	339550	471120	562432	612036	
L-PAQ Scoring Unit: Average Hours/Week					
Mean	3.1	21.3	21.9	46.5	Fig 5.2 (page 72)
Std Dev	5.2	11.4	23.1	25.2	
Median	1.6	22.0	14.0	41.1	
Minimum	.0	.0	.0	.0	
Maximum	104.7	141.6	155.4	166.5	
L-PAQ Scoring Unit: MET•Hours/Week					
Mean	18.3	60.1	66.6	146.6	Fig 5.3 (page 73)
Std Dev	31.9	78.5	69.1	111.6	
Median	8.8	38.1	44.1	118.3	
Minimum	.0	.0	.0	.0	
Maximum	623.2	1418.3	510.0	1547.4	

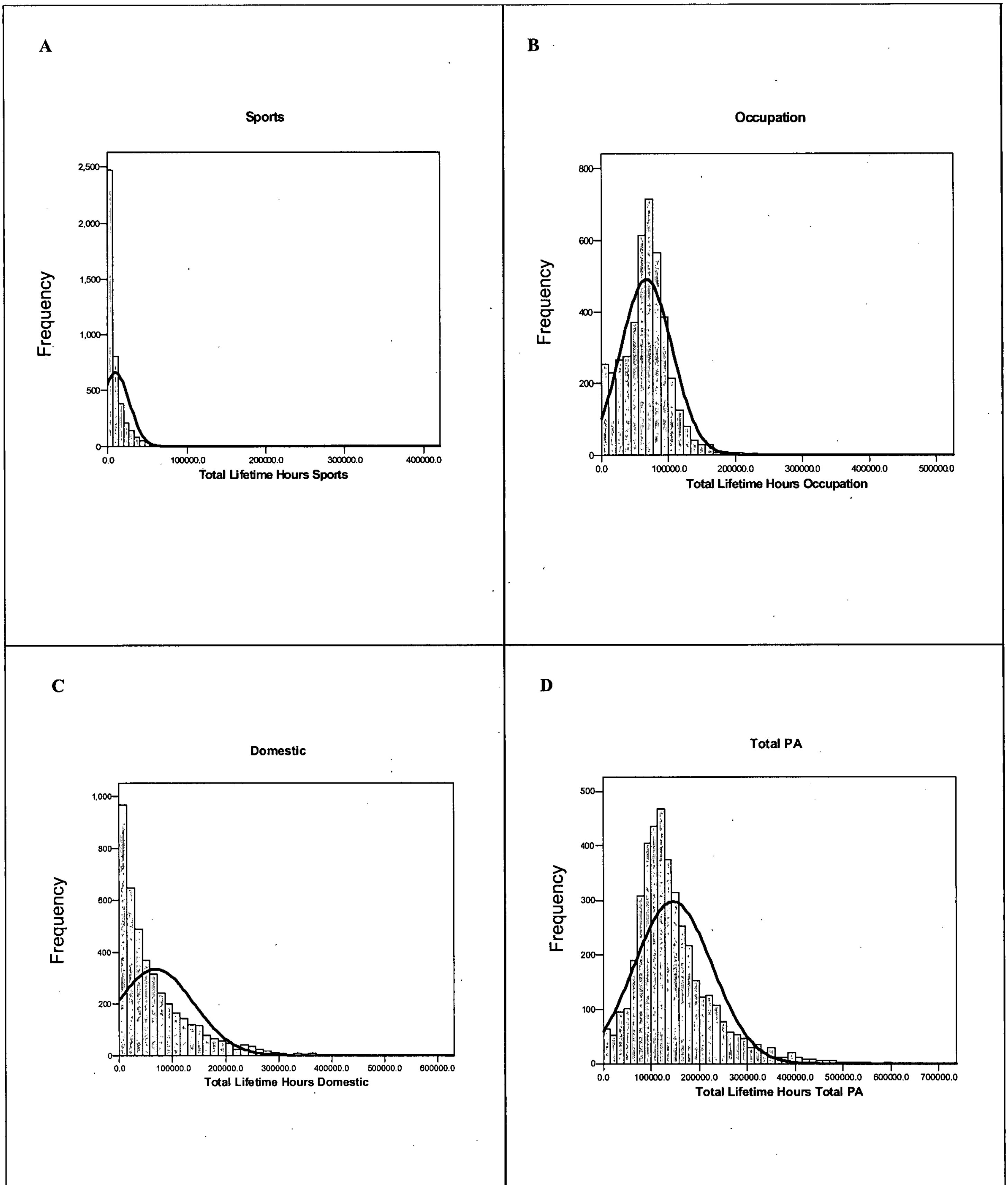


Figure 5.1. Total Lifetime Hours Frequency Histograms for Baseline L-PAQ (n=4269) (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

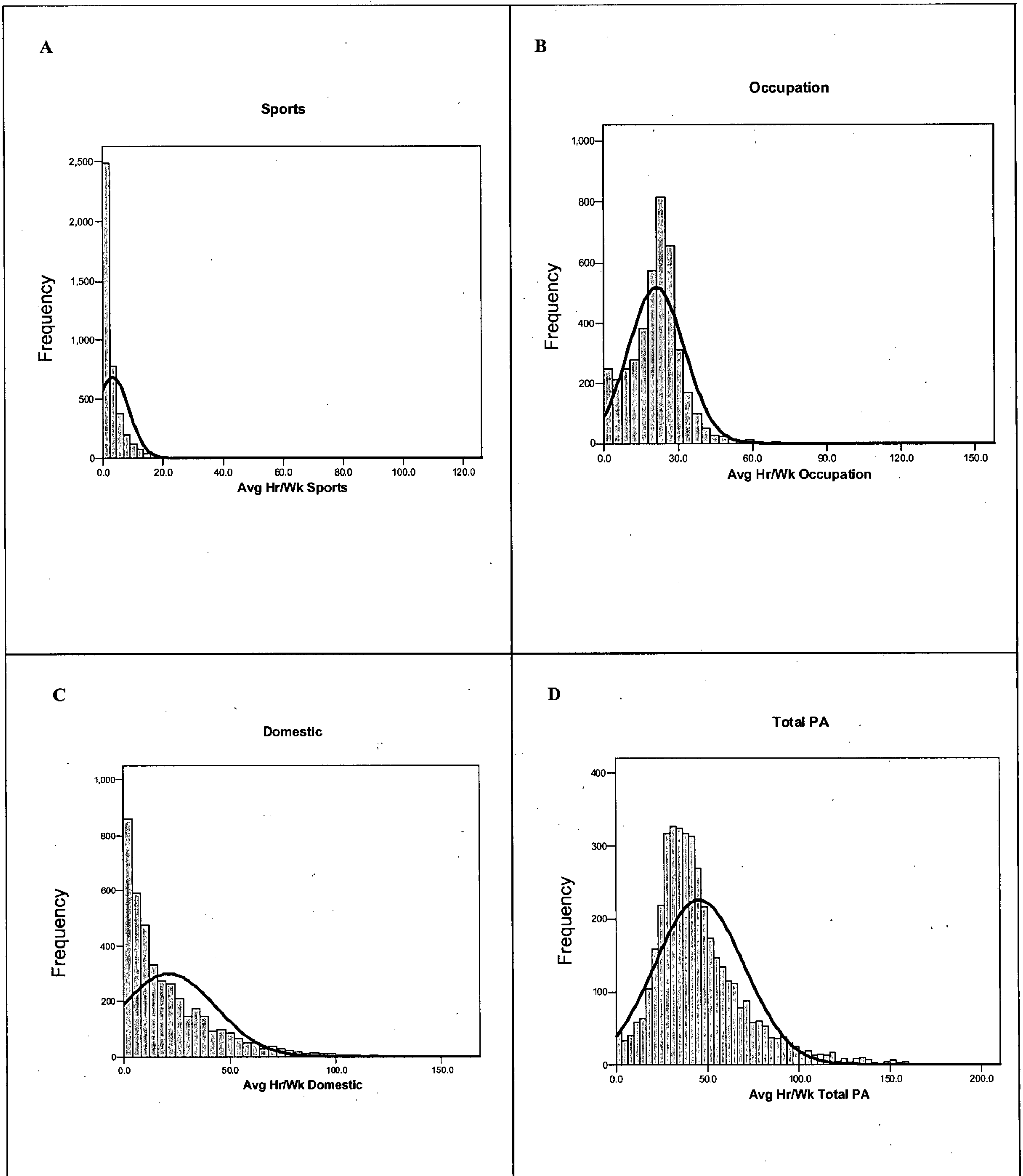


Figure 5.2. Lifetime Average Hours/Week Frequency Histograms for Baseline L-PAQ (n=4269) for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

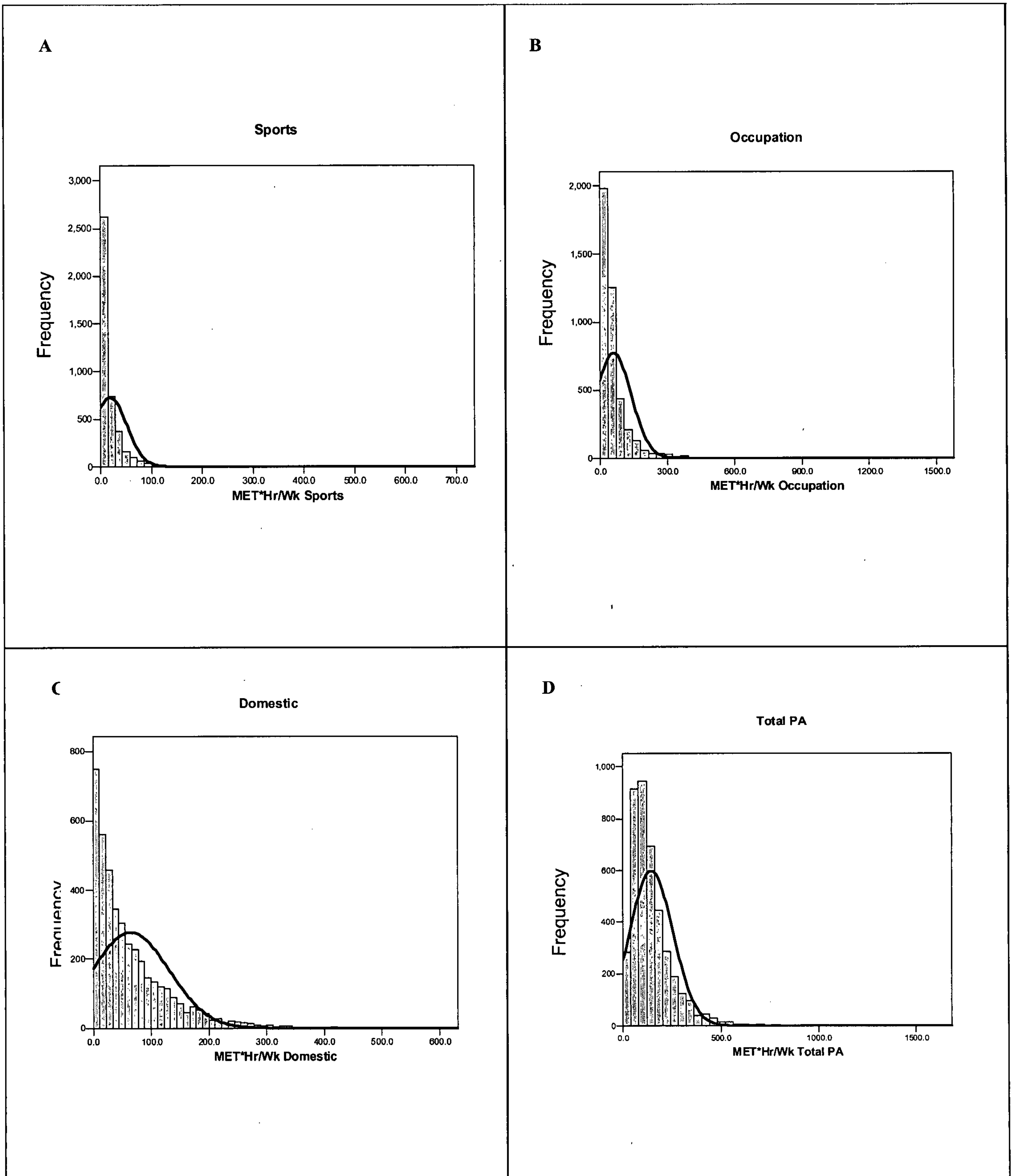


Figure 5.3. Lifetime MET•Hours/Week Frequency Histograms for Baseline L-PAQ (n=4269) for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

5.3 L-PAQ Reliability

The reliability study involved a second administration of the L-PAQ. In total, 76 validation study participants filled out the repeat version of the L-PAQ in its entirety. The average span between administrations of the L-PAQ was 8.01 months. The shortest span between administrations was 2.5 months and the longest span was 10.7 months.

5.3.1 Baseline Administration L-PAQ Summary Scores

Baseline L-PAQ data for the 76 participants who completed the retest administration of the L-PAQ were scored for reliability analysis. Summary baseline L-PAQ scores are presented in Table 5.3.

Table 5.3 Baseline L-PAQ Summary Scores for the Reliability Study (n=76)

	Sports	Occupation	Domestic	Total PA	Frequency Histogram Figure #
L-PAQ Scoring Unit: Total Lifetime Hours of Activity					
Mean	8918	76763	64011	149691	Appendix II Figure A2.1 (page 163)
Std Dev	11479	39684	76913	85381	
Median	4513	74880	31616	130103	
Minimum	0	1559	0	9896	
Maximum	62208	285200	371489	413157	
L-PAQ Scoring Unit: Lifetime Average Hours/Week					
Mean	2.7	23.2	19.2	45.2	Appendix II Figure A2.2 (page 164)
Std Dev	1.5	22.6	10.3	39.8	
Median	3.6	12.4	22.6	25.5	
Minimum	.0	.6	.0	3.4	
Maximum	19.6	98.5	106.4	136.4	
L-PAQ Scoring Unit: Lifetime MET•Hours/Week					
Mean	15.2	56.2	59.1	130.6	Appendix II Figure A2.3 (page 165)
Std Dev	19.2	114.4	70.7	143.4	
Median	8.6	35.4	33.9	89.7	
Minimum	.0	3.0	.0	7.1	
Maximum	85.7	1001.1	358.4	1105.7	

5.3.2 Repeat Administration L-PAQ Summary Scores

Retest L-PAQ data for the 76 participants who completed the second administration of the questionnaire were scored for analysis of L-PAQ reliability. Summary repeat L-PAQ scores are summarized in Table 5.4. All scores were calculated for sports, occupation, domestic, and total physical activity.

Table 5.4 Repeat L-PAQ Summary Scores for the Reliability Study (n=76)

	Sports	Occupation	Domestic	Total PA	Frequency Histogram Figure #
L-PAQ Scoring Unit: Total Lifetime Hours of Activity					
Mean	7709	69640	56384	133734	Appendix II Figure A2.4 (page 166)
Std Dev	9643	28804	55410	57871	
Median	3646	72566	30446	128469	
Minimum	0	5369	0	24089	
Maximum	46267	147678	209664	312344	
L-PAQ Scoring Unit: Lifetime Average Hours/Week					
Mean	2.4	21.1	17.0	40.5	Appendix II Figure A2.5 (page 167)
Std Dev	3.0	8.7	16.8	17.7	
Median	1.1	21.8	9.0	38.2	
Minimum	0	1.9	.0	8.6	
Maximum	14.6	51.0	60.0	100.6	
L-PAQ Scoring Unit: Lifetime MET•Hours/Week					
Mean	13.9	48.1	51.0	113.0	Appendix II Figure A2.6 (page 168)
Std Dev	18.0	54.6	49.5	81.2	
Median	7.3	36.2	28.2	92.1	
Minimum	.0	2.3	.0	19.6	
Maximum	89.8	373.2	189.2	420.1	

5.3.3 Intralass Correlation Coefficients

Intraclass correlations (ICC4) were calculated between baseline administration and repeat administration of the L-PAQ to evaluate instrument reliability. Across all three L-PAQ scoring units, the highest coefficients were obtained for sports activity with 0.82, 0.83, and 0.80 for total lifetime hours, lifetime average hours/week and MET•hours/week respectively. The lowest test-retest correlations were obtained for total PA with 0.58 for total lifetime hours, 0.57 for lifetime average hours/week, and 0.59 for MET•hours/week. Intermediate correlations were seen for occupational and domestic activity. Intraclass correlation coefficients and corresponding 95% CI were obtained for each PA domain and for total PA for all three L-PAQ scoring units. Results of these analyses are summarized in Table 5.5.

Table 5.5 Intraclass Correlation Coefficients Between Baseline L-PAQ and Repeat L-PAQ Administrations

	ICC4 (95% CI)
L-PAQ Scoring Unit: Total Lifetime Hours	
Sports	0.82 (0.72, 0.88)
Occupation	0.72 (0.59, 0.81)
Domestic	0.60 (0.43, 0.73)
Total PA	0.58 (0.40, 0.70)
L-PAQ Scoring Unit: Lifetime Average Hours/Week	
Sports	0.83 (0.75, 0.89)
Occupation	0.72 (0.59, 0.82)
Domestic	0.60 (0.43, 0.73)
Total PA	0.58 (0.40, 0.70)
L-PAQ Scoring Unit: Lifetime MET• Hours/Week	
Sports	0.80 (0.69, 0.86)
Occupation	0.63 (0.47, 0.75)
Domestic	0.60 (0.43, 0.72)
Total PA	0.59 (0.43, 0.72)

5.3.4 Comparison of Questionnaire Scores and Bland-Altman Plots

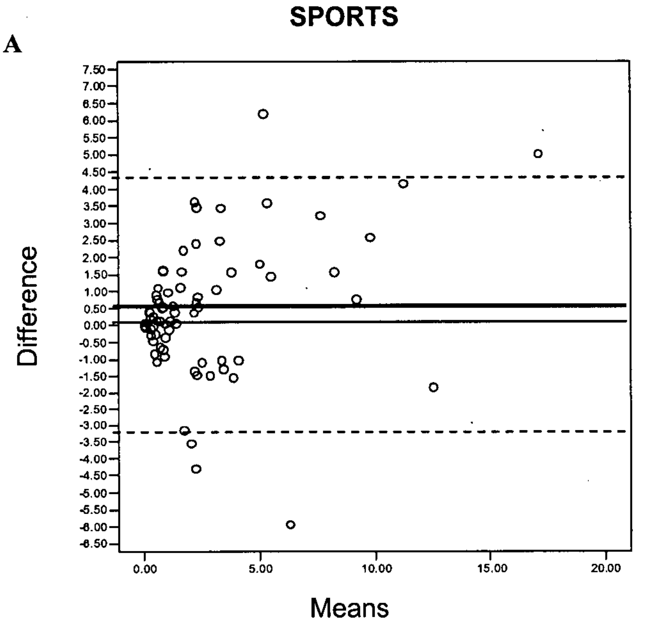
Table 5.6 summarizes results of Wilcoxon-signed rank tests conducted to compare physical activity scores between baseline and repeat administrations of the L-PAQ.

Table 5.6 Comparison of Physical Activity Scores Between Baseline and Repeat Administrations of the L-PAQ

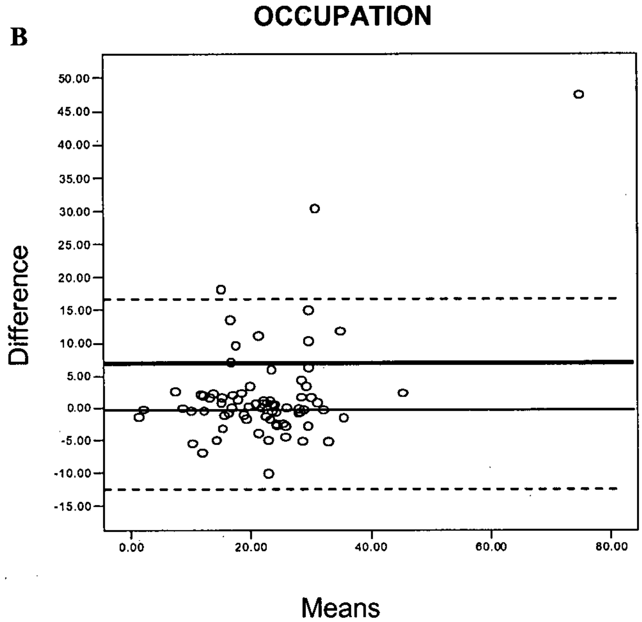
PA Domain	Baseline L-PAQ (mean)	Repeat L-PAQ (mean)	p-value
L-PAQ Scoring Unit: Total Lifetime Hours			
Sports	8918 ± 11479	7709 ± 9643	0.03*
Occupation	76763 ± 39684	69640 ± 28804	0.19
Domestic	64011 ± 76913	56384 ± 55410	0.18
Total PA	149691 ± 85381	133734 ± 57871	0.06
L-PAQ Scoring Unit: Lifetime Average Hours/Week			
Sports	2.7 ± 1.5	2.4 ± 3.0	0.03*
Occupation	23.2 ± 22.6	21.1 ± 8.7	0.20
Domestic	19.2 ± 10.3	17.0 ± 16.8	0.17
Total PA	45.2 ± 39.8	40.5 ± 17.7	0.06
L-PAQ Scoring Unit: Lifetime MET•Hours/Week			
Sports	15.2 ± 19.2	13.9 ± 18.0	0.09
Occupation	56.2 ± 114.4	48.1 ± 54.6	0.51
Domestic	59.1 ± 70.7	51.0 ± 49.5	0.08
Total PA	130.6 ± 143.4	113.0 ± 81.2	0.04*

For total lifetime hours/week and lifetime average hours/week, there were significant differences in questionnaire scores between baseline and repeat administrations of the L-PAQ for sports activity. There were no significant differences for other domains of activity and for total physical activity. For lifetime MET•hours/week, there were no significant differences in questionnaire scores across the three domains but there was significant difference for total physical activity.

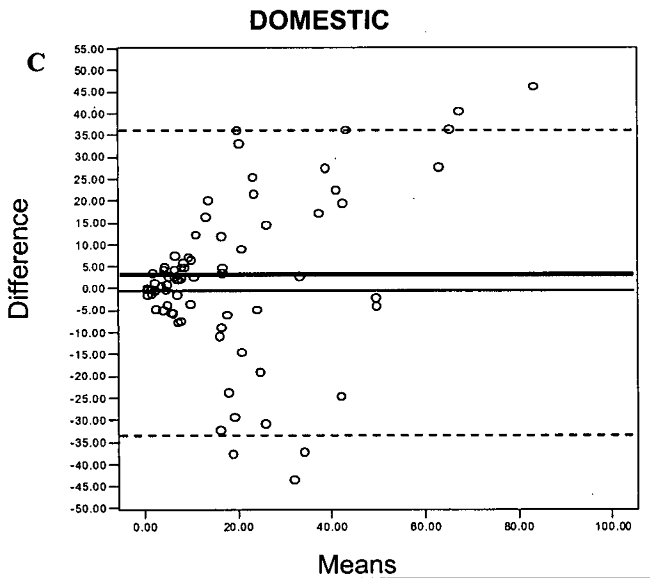
Bland-Altman plots for lifetime average hours/week for sports, occupation, domestic, and total physical activity are shown in Figure 5.4. Bland-Altman plots for total lifetime hours and MET•hours/week are shown in Appendix 2.1. Heavy solid lines indicate the mean of the differences (of pair wise test-retest L-PAQ scores for each respondent) and the dashed lines indicate the limits of agreement (LoA). The upper LoA was calculated by adding 2 standard deviations (SD of the differences) to the mean of the differences and the lower LoA was calculated by subtracting 2 standard deviations (SD of the differences) from the mean of the differences. Tabulated below each Bland-Altman plots are: mean of the differences (and SD), actual LoA values, and the standard error of measurement (SEM). Important to the interpretation of Bland-Altman plots is the size of the interval of the limits of agreement (LOA). For sports, the LOA interval was 7.7 hours/week, for occupation, the interval was 32.8 hours/week, for domestic, 71.2 hours/week, and for total physical activity, 81.4 hours/week.



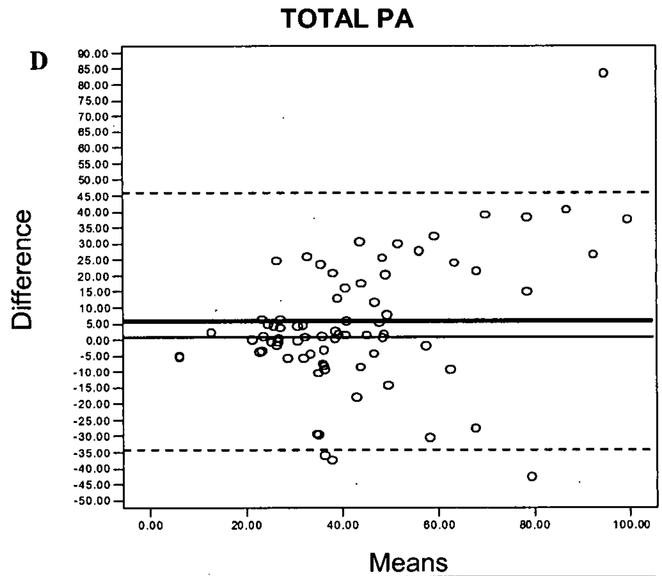
Mean of differences	0.43
SD	1.93
2 x SD	3.86
Limits of Agreement	-3.43 to 4.29
SEM	1.36



Mean of differences	1.96
SD	7.98
2 x SD	15.96
Limits of Agreement	-14 to 17.92
SEM	5.64



Mean of differences	2.41
SD	17.87
2 x SD	35.74
Limits of Agreement	-33.33 to 38.15
SEM	12.64



Mean of differences	4.80
SD	20.35
2 x SD	40.70
Limits of Agreement	-35.90 to 45.50
SEM	14.39

Figure 5.4. Bland-Altman Plots for Lifetime Average Hours/Week (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

5.4 Validity Study Part I: Comparative Studies of the Lifetime Physical Activity Questionnaire (L-PAQ) with the Lifetime Total Physical Activity Questionnaire (LT-PAQ)

A total of 84 interviews were conducted by the principal investigator for the validation studies. Of these, 71 were conducted at the Arthritis Research Centre, 2 were conducted at participants' homes, and 11 were conducted over the telephone. All interviewed participants completed the Life Events Calendars prior to their interviews. The duration of interviews, defined as the time from the introduction of the questionnaire and explanation of the procedures of the interview until the response to the last question, was recorded by the investigator. The mean duration of the interviews was 56 minutes and the median duration was 55 minutes. The shortest interview took 20 minutes while the longest interview took approximately 2 hours.

5.4.1 L-PAQ Summary Scores

Baseline L-PAQ data for the 84 interviewed participants were scored for the comparative studies with the LT-PAQ. Summary L-PAQ scores are summarized in Table 5.7. All scores were calculated for sports, occupation, domestic, and total physical activity.

Table 5.7 Baseline L-PAQ Summary Scores for Comparison with LT-PAQ (n=84)

	Sports	Occupation	Domestic	Total PA	Frequency Histogram Figure #
L-PAQ Scoring Unit: Total Lifetime Hours of Activity					
Mean	8813	74485	65414	148711	Appendix II Figure A2.9 (page 172)
Std Dev	11218	39155	77968	83412	
Median	4621	73840	31603	128464	
Minimum	0	1559	0	9896	
Maximum	62207.8	285199.2	371488.0	413156.7	
L-PAQ Scoring Unit: Lifetime Average Hours/Week					
Mean	2.7	22.5	19.3	44.5	Appendix II Figure A2.10 (page 173)
Std Dev	3.5	12.3	22.4	24.5	
Median	1.5	22.3	10.3	39.7	
Minimum	.0	.6	.0	3.4	
Maximum	19.6	98.5	106.4	136.4	
L-PAQ Scoring Unit: MET• Hours/Week					
Mean	14.8	53.3	59.8	127.8	Appendix II Figure A2.11 (page 174)
Std Dev	18.6	109.9	70.4	138.1	
Median	8.6	34.6	33.0	89.5	
Minimum	.0	3.0	.0	7.1	
Maximum	85.7	1001.1	358.4	1105.7	

5.4.2 LT-PAQ Summary Scores

LT-PAQ summary scores are expressed in two units of measurement: lifetime average hours/week and MET•hours/week. Summary scores for participants who completed the LT-PAQ are summarized in Table 5.8 for lifetime average hours/week and lifetime MET•hours/week. Corresponding frequency histograms with normal curves illustrating the distribution of the scores are summarized in this chapter for lifetime average hours/week in Figure 5.5 (as this is the most readily interpreted scoring unit). Please refer to Appendix II for corresponding frequency histograms for MET•hours/week.

Table 5.8 LT-PAQ Summary Scores for Comparison with L-PAQ (n=84)

	Sports	Occupation	Domestic	Total PA	Frequency Histogram Figure #
LT-PAQ Scoring Unit: Lifetime Average Hours/Week					
Mean	4.2	23.0	12.9	40.1	Figure 5.5 (page 83)
Std Dev	3.7	7.7	9.0	9.3	
Median	3.2	23.1	10.7	40.4	
Minimum	0.2	8.5	0.0	21.4	
Maximum	28.2	44.5	43.9	71.2	
LT-PAQ Scoring Unit: Lifetime MET•Hours/Week					
Mean	21.1	52.7	38.5	112.3	Appendix II Figure A2.12 (page 175)
Std Dev	24.2	27.1	27.4	40.5	
Median	14.6	49.1	31.2	108.3	
Minimum	1.4	14.3	00	53.2	
Maximum	218.8	204.4	129.6	450.4	

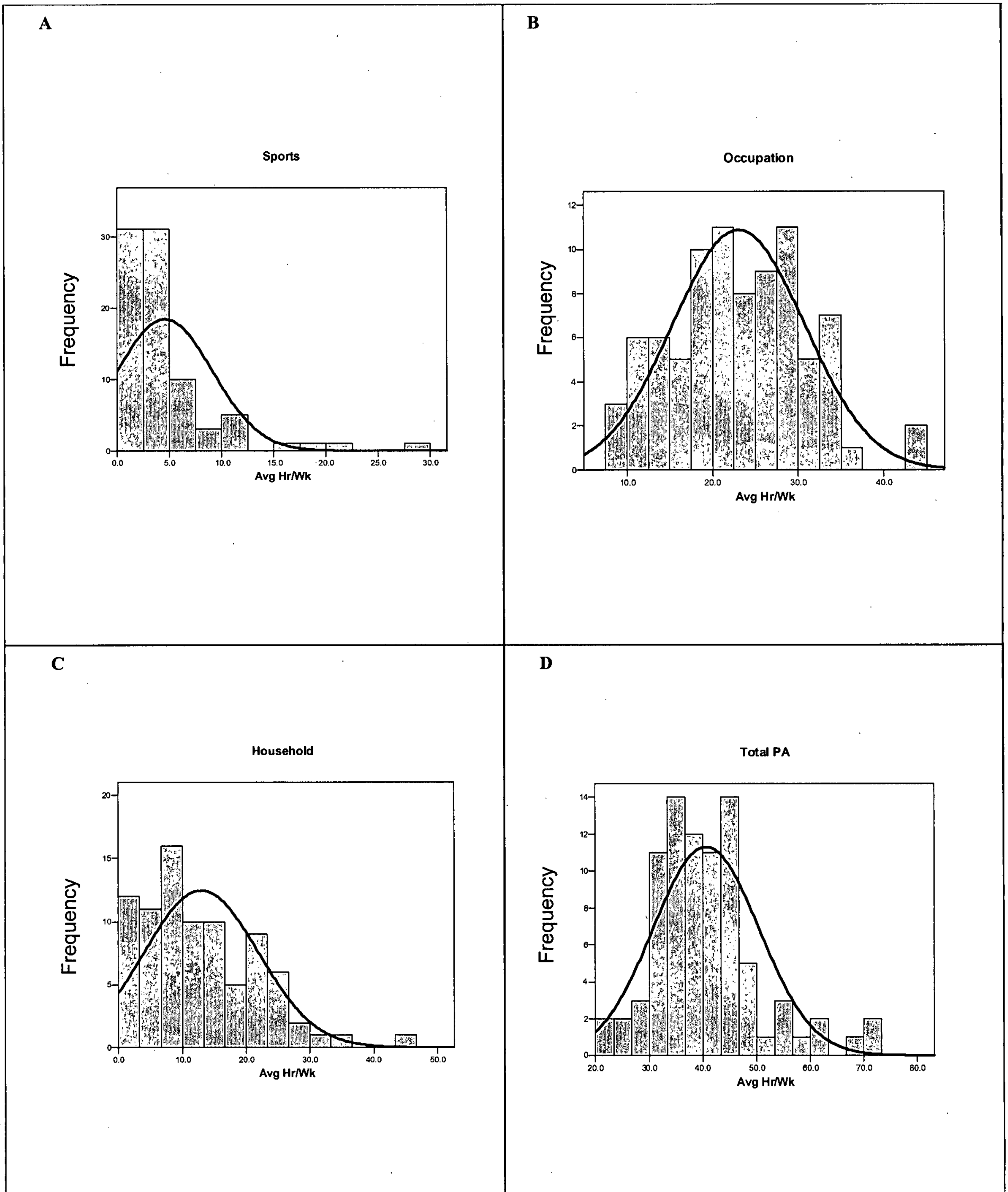


Figure 5.5. LT-PAQ Lifetime Average Hr/Wk Frequency Histograms for (A) Sports, (B) Occupation, (C) Household, and (D) Total Physical Activity

5.4.3 Comparison of the L-PAQ and the LT-PAQ

Spearman correlation coefficients were calculated for similar L-PAQ and LT-PAQ scoring units of lifetime average hours/week and lifetime MET•hours/week across three physical activity domains for total activity. Across both lifetime hours/week and MET•hours/week scoring units, the highest correlations were obtained for domestic activity (0.71 for both scoring units) and the lowest correlations for total physical activity (0.41 and 0.34 for lifetime hours/week and MET•hours/week respectively). Intermediate correlations were seen for sports and occupational activity. Results of the correlation between L-PAQ and LT-PAQ are summarized in Table 5.9. Scatter plots corresponding to the correlations are presented in this chapter for lifetime average hours/week. Please refer to Appendix II for scatter plots corresponding to correlations for MET•hours/week.

Table 5.9 Spearman Correlation Coefficients Between Baseline L-PAQ and LT-PAQ Summary Scores for Similar Scoring Units

Correlation #	L-PAQ	LT-PAQ	Spearman Correlation Coefficient	Scatter Plot Figure #
	Lifetime Avg Hours/Week	**Lifetime Avg Hours/Week**		Fig 5.6 (page 85)
1	Sports PA	Sports PA	0.52	
2	Occupational PA	Occupational PA	0.55	
3	Domestic PA	Domestic PA	0.71	
4	Total PA	Total PA	0.41	
	Lifetime MET• Hours/Week	**Lifetime MET• Hours/Week**		Appendix II Figure A2.13 (page 176)
5	Sports PA	Sports PA	0.60	
6	Occupational PA	Occupational PA	0.50	
7	Domestic PA	Domestic PA	0.71	
8	Total PA	Total PA	0.34	

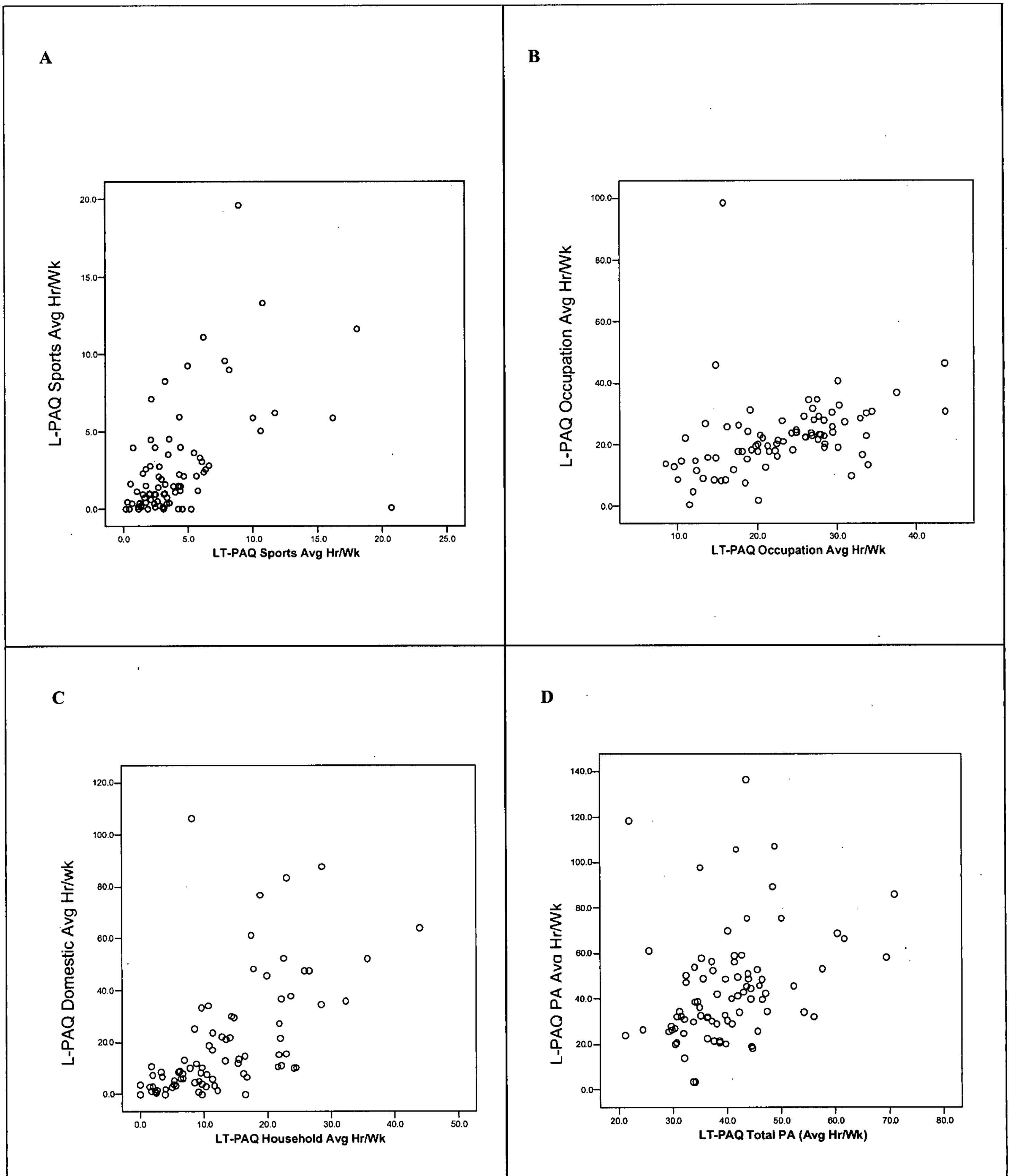


Figure 5.6. Scatterplots showing relationship between L-PAQ Average Hours/Week on the y-axis versus LT-PAQ Average Hours/Week on the x-axis for (A) Sports; (B) Occupation; (C) Domestic; and (D) Total PA

Physical activity scores for the L-PAQ and LT-PAQ were also compared using Wilcoxon signed-rank tests. Analyses were conducted for sports, occupation, and domestic, and total physical activity for similar scoring units between the two questionnaires, lifetime average hours/week and MET•hours/week. Significant differences were seen in the measurement of sports (p=0.0), occupation (p=0.033), and domestic activity (p=0.03) between questionnaires for lifetime average hours/week. Similar trends in results were seen for MET•hours/week. Table 5.10 summarizes the results of Wilcoxon signed-rank tests conducted to compare physical activity scores between baseline L-PAQ and LT-PAQ.

Table 5.10 Comparison of Physical Activity Scores Between Baseline L-PAQ and LT-PAQ

PA Domain	L-PAQ (mean)	LT-PAQ (mean)	p-value
Scoring Unit: Lifetime Average Hours/Week			
Sports	2.7 ± 3.5	4.2 ± 3.7	0.0*
Occupation	22.5 ± 12.3	23.0 ± 7.7	0.033*
Domestic	19.3 ± 22.4	12.9 ± 9.0	0.03*
Total PA	44.5 ± 24.5	40.1 ± 9.3	0.61
Scoring Unit: Lifetime MET•Hours/Week			
Sports	14.8 ± 18.6	21.1 ± 24.2	0.0*
Occupation	53.3 ± 109.9	52.7 ± 27.1	0.0*
Domestic	59.8 ± 70.4	38.5 ± 27.4	0.014*
Total PA	127.8 ± 138.1	112.3 ± 40.5	0.14

For sensitivity analyses, the same methods of comparison were applied between *retest* L-PAQ data and LT-PAQ. Spearman correlation coefficients were calculated for similar L-PAQ and LT-PAQ scoring units of lifetime average hours/week and lifetime MET•hours/week across three physical activity domains for total activity. While values for the correlation coefficients were different, similar trends in findings were seen when *retest* L-PAQ data were correlated with

LT-PAQ data. Across both scoring units, the highest correlations were obtained for domestic activity (0.62 for both units) and the lowest correlations for total physical activity (0.40 and 0.49 for lifetime average hours/week and MET•hours/week respectively). Intermediate correlations were seen for sports and occupational activity. Results of the correlation between retest L-PAQ and LT-PAQ are summarized in Table 5.11. Since these analyses were done as supplementary to analyses for baseline L-PAQ, scatter plots corresponding to the correlations will not be presented.

Table 5.11 Spearman Correlation Coefficients Between Retest L-PAQ and LT-PAQ Summary Scores for Similar Scoring Units

Correlation #	L-PAQ (Scoring Units)	LT-PAQ (Scoring Units)	Spearman Correlation Coefficient
	Lifetime Avg Hours/Week	**Lifetime Avg Hours/Week**	
1	Sports PA	Sports	0.51
2	Occupational PA	Occupational PA	0.58
3	Household PA	Household PA	0.62
4	Total PA	Total PA	0.40
	Lifetime MET• Hours/Week	**Lifetime MET• Hours/Week**	
5	Sports PA	Sports	0.51
6	Occupational PA	Occupational PA	0.55
7	Household PA	Household PA	0.62
8	Total PA	Total PA	0.49

Comparison of questionnaire scores was also conducted between *retest* L-PAQ data and LT-PAQ data. Wilcoxon signed-rank tests were conducted for sports, occupation, and domestic, and total physical activity for similar scoring units between the two questionnaires, lifetime average hours/week and MET•hours/week. Similar to the comparison between *baseline* L-PAQ and LT-PAQ, significant differences were seen in the measurement of sports and occupational activity for lifetime average hours/week. However, there were no differences between for domestic and total activity. Results for Wilcoxon signed-rank tests for MET•hours/week showed similar findings to comparison between *baseline* L-PAQ and LT-PAQ. Specifically, significant differences were seen in the measurement of sports, occupation, and domestic activity. Table 5.12 summarizes results of Wilcoxon signed-rank tests conducted to compare questionnaire scores between *retest* L-PAQ and the LT-PAQ for similar scoring units.

Table 5.12 Comparison of Physical Activity Scores Between Retest L-PAQ and LT-PAQ

PA Domain	L-PAQ (mean)			LT-PAQ (mean)			p-value
Scoring Unit: Lifetime Average Hours/Week							
Sports	2.4	±	3.1	4.1	±	3.7	0.0*
Occupation	21.1	±	8.9	23.2	±	7.8	0.0*
Domestic	17.0	±	17.0	12.7	±	8.4	0.073
Total PA	40.5	±	17.8	40.1	±	9.4	0.59
Scoring Unit: Lifetime MET•Hours/Week							
Sports	14.1	±	18.4	20.9	±	24.7	0.0*
Occupation	48.9	±	55.6	53.6	±	27.7	0.0*
Domestic	50.6	±	49.8	38.2	±	26.3	0.04*
Total PA	113.6	±	82.3	112.7	±	42.0	0.15

Overall sensitivity analyses with *retest* L-PAQ data showed similar results with main analyses conducted for *baseline* L-PAQ data when the questionnaire was compared with the LT-PAQ.

5.5 Validity Study Part II: Comparative Studies of the Lifetime Physical Activity Questionnaire (L-PAQ) and the Chasan-Taber Physical Activity Questionnaire (CT-PAQ)

Following each LT-PAQ interview, subjects were asked to complete the paper-based CT-PAQ. This questionnaire was administered by subjects themselves following detailed instructions by the thesis authors. All interviewed validation study participants indicated their interest in completing the questionnaire. Of the 71 subjects interviewed at ARC, 51 completed the CT-PAQ immediately following the LT-PAQ interviews and 20 opted to fill out the CT-PAQ at home. Completion of the CT-PAQ by 51 participants at ARC gave the principal investigator the opportunity to record the length of time needed to complete the questionnaire. Subjects whose LT-PAQ interviews were conducted over the telephone were mailed a copy of the CT-PAQ. Reminder emails and telephone calls were sent to subjects complete the CT-PAQ at their homes.

Overall, the CT-PAQ was completed by 80 participants. For the 51 subjects who filled out the CT-PAQ at ARC, the average length of time for questionnaire completion was 18 minutes. The shortest duration for completion was 9 minutes while the longest duration for completion was 28 minutes.

5.5.1 L-PAQ Summary Scores

Baseline L-PAQ data for the 80 subjects who completed the CT-PAQ were scored for analysis for this study. Since the CT-PAQ measured physical activity across two domains - leisure and household activity - summary L-PAQ scores were calculated for similar domains of sports/recreation and domestic activity. Summary L-PAQ scores for participants who completed the CT-PAQ are summarized in Table 5.13. All scores were calculated for sports and domestic activity and for total activity. Please refer to Appendix II for corresponding frequency histograms with normal curves illustrating the distribution of the scores.

Table 5.13 Baseline L-PAQ Summary Scores for Comparison with CT-PAQ (n=80)

	Sports	Domestic	Total PA	Frequency Histogram Figure #
L-PAQ Scoring Unit: Total Lifetime Hours of Activity				
Mean	8299	70785	79084	Appendix II Figure A2.14 (page 178)
Std Dev	10679	96113	95671	
Median	4509	31551	42929	
Minimum	0	0	234	
Maximum	62208	562640	567041	
L-PAQ Scoring Unit: Lifetime Average Hours/Week				
Mean	3.4	20.5	23.9	Appendix II Figure A2.15 (page 179)
Std Dev	4.2	27.3	27.5	
Median	1.9	10.2	12.4	
Minimum	0	0	.1	
Maximum	21.4	155.6	158.6	
L-PAQ Scoring Unit: Lifetime MET•Hours/Week				
Mean	13.5	63.4	76.9	Appendix II Figure A2.16 (page 180)
Std Dev	16.5	85.1	85.3	
Median	8.5	31.4	49.4	
Minimum	.0	.0	.6	
Maximum	85.7	472.7	480.0	

5.5.2 CT-PAQ Summary Scores

CT-PAQ summary scores are expressed in two units of measurement: lifetime average hours/week and MET•hours/week. Scores for sports PA were calculated with and without the inclusion of respondent self-reports of walking as CT-PAQ authors reported on the lack of reliability of self-reports of walking. Summary scores for participants who completed the CT-PAQ are summarized in Table 5.14 for lifetime average hours/week and lifetime

MET•hours/week. Corresponding frequency histograms with normal curves illustrating the distribution of the scores are summarized in this chapter for lifetime average hours/week in Figure 5.7 (as this is the most readily interpreted scoring unit). Please refer to Appendix II for corresponding frequency histograms for MET•hours/week.

Table 5.14 CT-PAQ Summary Scores for Comparison with L-PAQ (n=80)

	Sports (avg hr/wk)		**Household (avg hr/wk)**	**Total PA (avg hr/wk)**		**Frequency Histogram Figure #**
	Includes Walking	Excludes Walking		Includes Walking	Excludes Walking	
CT-PAQ Scoring Unit: Lifetime Average Hours/Week						
Mean	6.5	3.2	11.8	18.3	15.0	Fig 5.7a and Fig 5.7b (page 92, 93)
Std Dev	7.3	3.6	15.0	18.7	15.7	
Median	4.1	2.1	5.9	10.5	9.3	
Minimum	0.14	0.0	0.0	0.14	0.14	
Maximum	35.3	17.3	75.6	108.6	80.9	
CT-PAQ Scoring Unit: Lifetime MET•Hours/Week						
Mean	30.3	18.8	34.9	65.3	53.7	Appendix II Figure A2.17a and FigureA2.17b (page 181, 182)
Std Dev	31.9	21.5	44.2	61.2	50.1	
Median	20.0	11.9	18.3	44.1	38.7	
Minimum	0.26	0.0	0.0	0.26	0.26	
Maximum	155.4	103.9	242.7	366.33	269.41	

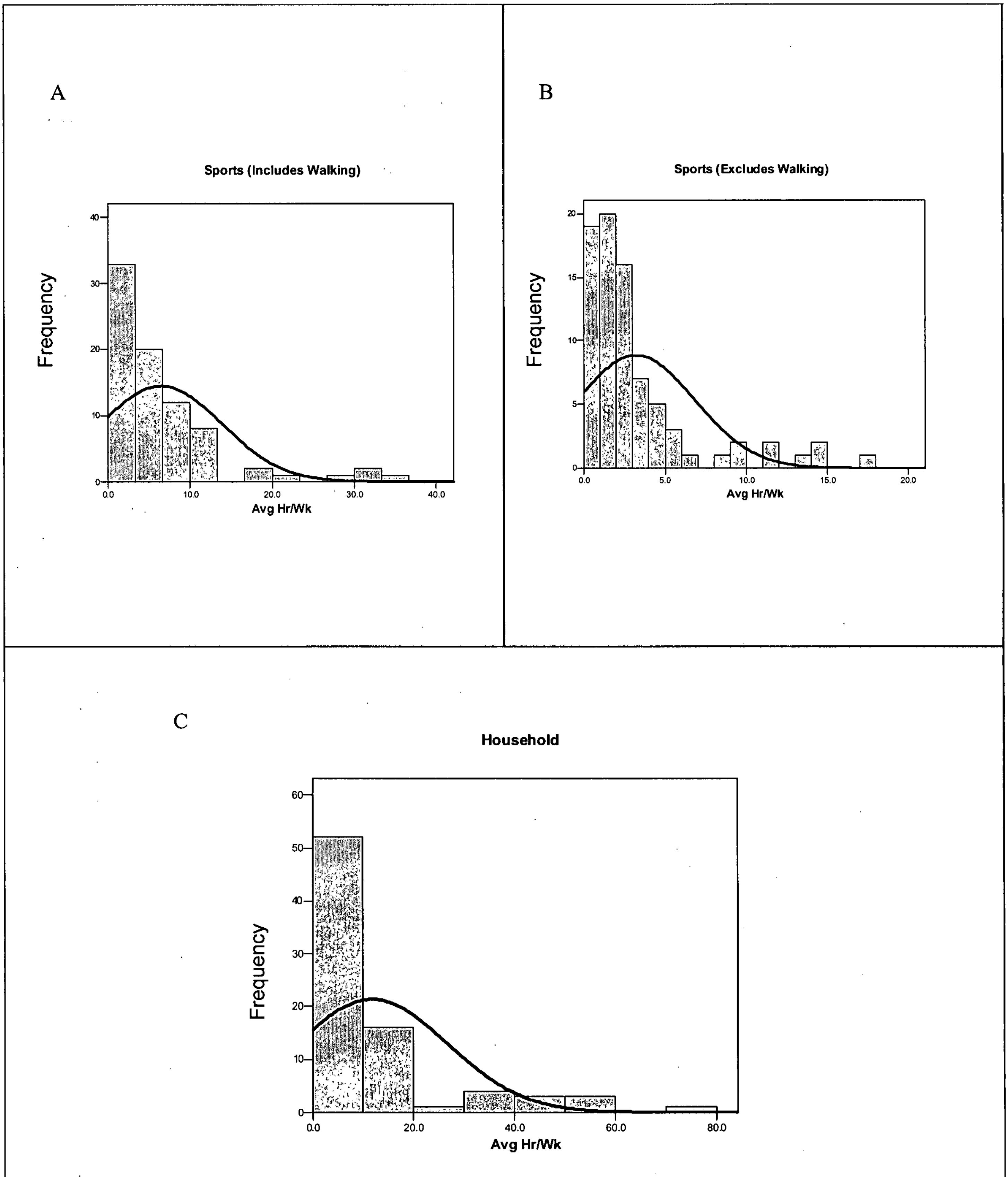
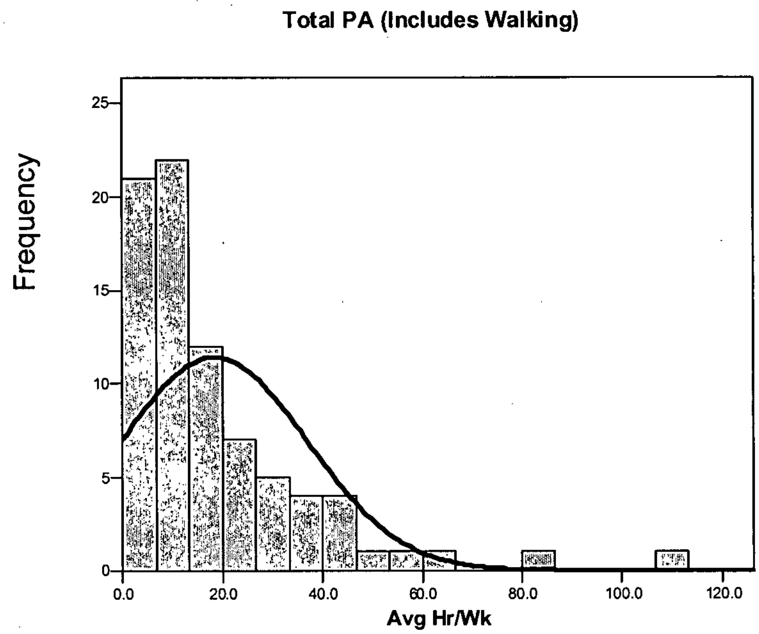


Figure 5.7(a). CT-PAQ Lifetime Average Hr/Wk Frequency Histograms for (A) Sports (Includes Walking), (B) Sports (Excludes Walking), and (C) Household PA

A



B

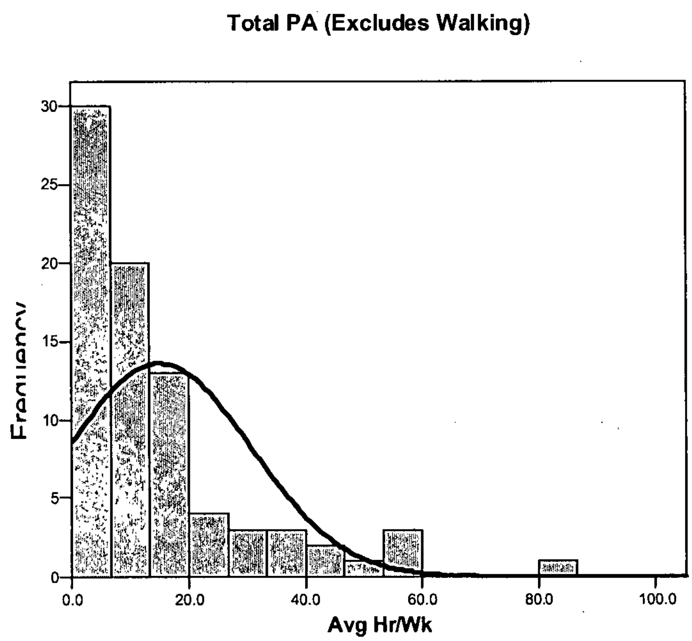


Figure 5.7(b). CT-PAQ Lifetime Average Hr/Wk Frequency Histograms for (A) Total PA (Includes Walking) and (B) Total PA (Excludes Walking)

5.5.3 Comparison of the L-PAQ and the CT-PAQ

Correlation coefficients were calculated between similar L-PAQ and CT-PAQ scoring units of lifetime average hours/week and lifetime MET•hours/week. Separate coefficients were obtained for CT-PAQ scores with both inclusion and exclusion of self-reports of walking in the sports domain.

Results of the correlation analysis show that the highest correlation coefficients were obtained for sports PA when self-reports of walking were excluded. This trend was consistent for both lifetime average hours/week and lifetime MET•hours/week scoring units (0.58 and 0.60, respectively). Looking at total PA, inclusion of walking resulted in higher correlations than when walking was excluded. Intermediate correlations were seen for domestic PA. Table 5.15 summarizes Spearman correlation coefficients calculated between similar scoring units of the L-PAQ and CT-PAQ. Scatter plots corresponding to the correlations are presented in this chapter for lifetime average hours/week. Please refer to Appendix II for scatter plots corresponding to correlations for MET•hours/week.

Table 5.15 Spearman Correlation Coefficients Between Baseline L-PAQ and CT-PAQ Summary Scores for Similar Scoring Units

Correlation #	L-PAQ	CT-PAQ	Spearman Correlation Coefficient	Scatter Plot Figure #
	Lifetime Avg Hr/Wk	Lifetime Avg Hr/Wk *(Walk)*		
1	Sports PA	Sports (Walk)	0.39	Fig 5.8 (page 96)
2	Household PA	Household PA	0.56	
3	Total PA	Total PA (Walk)	0.54	
	Lifetime Avg Hr/Wk	Lifetime Avg Hr/Wk *(No Walk)*		
4	Sports PA	Sports (Walk)	0.58	Fig 5.9 (page 97)
5	Household PA	Household PA	0.56	
6	Total PA	Total PA (Walk)	0.50	
	Lifetime MET•Hr/Wk	Lifetime MET•Hr/Wk *(Walk)*		
7	Sports PA	Sports (Walk)	0.38	Appendix II Figure A2.18 (page 183)
8	Household PA	Household PA	0.56	
9	Total PA	Total PA (Walk)	0.54	
	Lifetime MET•Hr/Wk	Lifetime MET•Hr/Wk *(No Walk)*		
10	Sports PA	Sports (Walk)	0.60	Appendix II Figure A2.19 (page 184)
11	Household PA	Household PA	0.56	
12	Total PA	Total PA (Walk)	0.49	

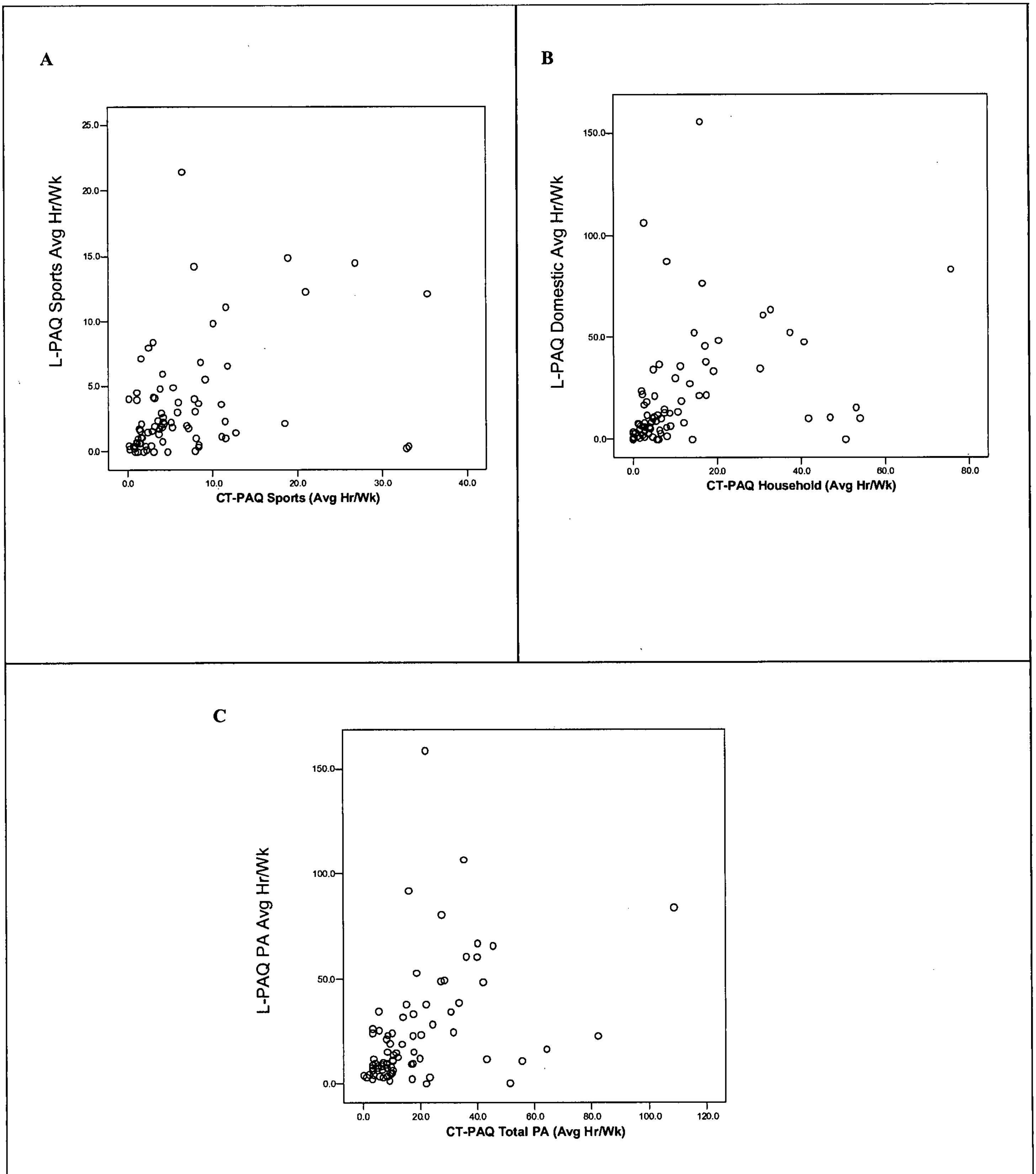
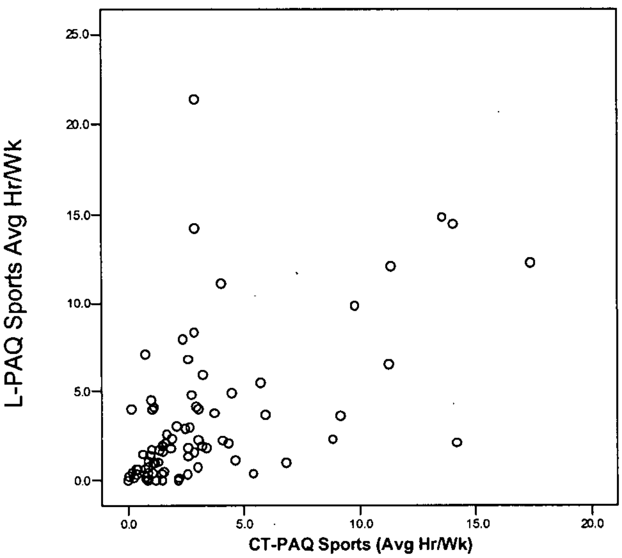
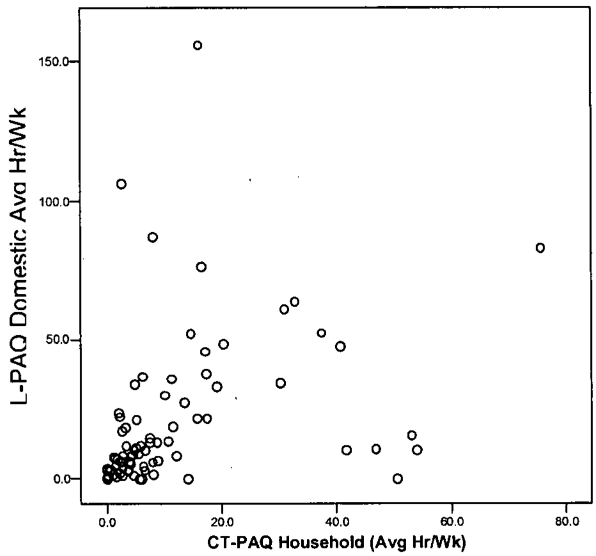


Figure 5.8 Scatter plots showing relationship between L-PAQ Avg Hr/Wk (y-axis) Versus CT-PAQ Avg Hr/Wk (Walking Included) (x-axis) for (A) Sports; (B) Domestic; and (C) Total PA

A



B



C

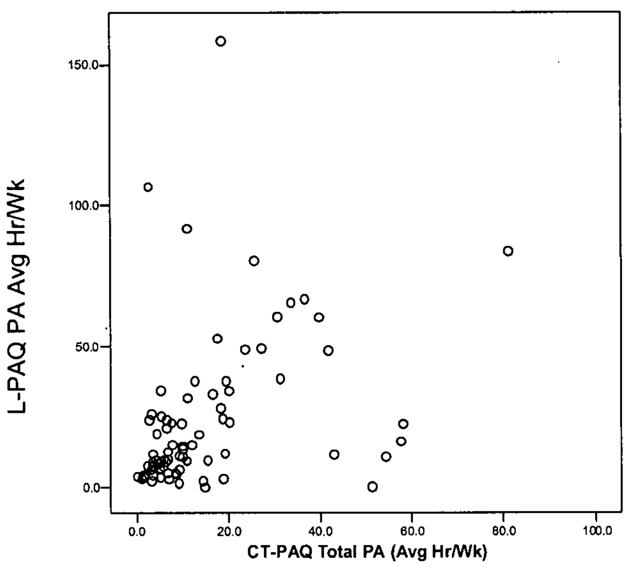


Figure 5.9 Scatter plots showing relationship between L-PAQ Avg Hr/Wk (y-axis) Versus CT-PAQ Avg Hr/Wk (Walking Excluded) (x-axis) for (A) Sports; (B) Domestic; and (C) Total PA

Physical activity scores for the L-PAQ and CT-PAQ were compared using Wilcoxon signed-rank tests for differences among pairs. Analyses were conducted for sports, domestic, and total physical activity for similar scoring units between the two questionnaires, lifetime average hours/week and MET•hours/week. For lifetime average hours/week, significant differences were seen in the measurement of domestic activity. Significant differences were also seen in the measurement of sports PA when CT-PAQ data included self-reports of walking were used. However, when CT-PAQ data excluded walking reports, there were no significant differences between the questionnaires. An opposite trend was seen when looking at total PA. Specifically, when CT-PAQ data included self-reports of walking, there were no significant differences between the two questionnaires in the measurement of total PA. With exclusion of walking from CT-PAQ data, there were significant differences in the measurement of total PA. Similar results were seen when looking at results of analyses with MET•hours/week. Table 5.16 summarizes results Wilcoxon signed-rank tests conducted to compare physical activity measured by the L-PAQ and the CT-PAQ for similar scoring units.

Table 5.16 Comparison of Physical Activity Scores Between Baseline L-PAQ and CT-PAQ

PA Domain	L-PAQ (mean)	CT-PAQ (mean)	p-value
Scoring Unit: Lifetime Average Hours/Week			
Sports (CT-PAQ includes walking)	3.4 ± 4.2	6.5 ± 7.3	0.0*
Sports (CT-PAQ excludes walking)	3.4 ± 4.2	3.2 ± 3.6	0.86
Domestic	20.5 ± 27.3	11.8 ± 15.0	0.0*
Total PA (CT-PAQ includes walking)	23.9 ± 27.5	18.3 ± 18.7	0.09
Total PA (CT-PAQ excludes walking)	23.9 ± 27.5	15.0 ± 15.7	0.0*
Scoring Unit: Lifetime MET•Hours/Week			
Sports (CT-PAQ includes walking)	13.5 ± 16.5	30.4 ± 31.9	0.0*
Sports (CT-PAQ excludes walking)	13.5 ± 16.5	18.8 ± 21.5	0.08*
Domestic	63.4 ± 85.1	34.9 ± 44.2	0.0*
Total PA (CT-PAQ includes walking)	76.9 ± 85.3	65.3 ± 61.1	0.1
Total PA (CT-PAQ excludes walking)	76.9 ± 85.3	53.7 ± 50.1	0.0*

Sensitivity analyses were conducted by repeating methods applied to *baseline* L-PAQ data to *retest* L-PAQ data and CT-PAQ data. Correlation coefficients were calculated between similar L-PAQ and CT-PAQ scoring units of lifetime average hours/week and lifetime MET•hours/week. Separate coefficients were obtained for CT-PAQ scores with both inclusion and exclusion of self-reports of walking in the sports domain. Results of the correlation analysis show that the highest correlation coefficients were obtained for domestic PA for both scoring units. These findings are inconsistent with the results of correlations between *baseline* L-PAQ and LT-PAQ. However, it is important to note that similar trends were seen between both baseline and retest L-PAQ data that higher correlations for sports were obtained when subject self-reports of walking were excluded from the CT-PAQ data. Intermediate correlations were seen for domestic activity. Table 5.17 summarizes results of the correlation analyses between similar *retest* L-PAQ and CT-PAQ scoring units. Since these analyses were done as supplementary to analyses for baseline L-PAQ, scatter plots corresponding to the correlations will not be presented.

Table 5.17 Spearman Correlation Coefficients Between Retest L-PAQ and CT-PAQ Summary Scores for Similar Scoring Units

Correlation #	L-PAQ (Scoring Units)	CT-PAQ (Scoring Units)	Spearman Correlation Coefficient
	Lifetime Avg Hr/Wk	**Lifetime Avg Hr/Wk *(Walk)***	
1	Sports PA	Sports (Walk)	0.45
2	Household PA	Household PA	0.65
3	Total PA	Total PA (Walk)	0.69
	Lifetime Avg Hr/Wk	**Lifetime Avg Hr/Wk *(No Walk)***	
4	Sports PA	Sports (Walk)	0.57
5	Household PA	Household PA	0.65
6	Total PA	Total PA (Walk)	0.63
	Lifetime MET•Hr/Wk	**Lifetime MET•Hr/Wk *(Walk)***	
7	Sports PA	Sports (Walk)	0.49
8	Household PA	Household PA	0.67
9	Total PA	Total PA (Walk)	0.71
	Lifetime MET•Hr/Wk	**Lifetime MET•Hr/Wk *(No Walk)***	
10	Sports PA	Sports (Walk)	0.61
11	Household PA	Household PA	0.67
12	Total PA	Total PA (Walk)	0.65

Comparison physical activity scores were also repeated for *retest* L-PAQ data and CT-PAQ data. Analyses were conducted for sports, domestic, and total physical activity for similar scoring units between the two questionnaires, lifetime average hours/week and MET•hours/week. For lifetime average hours/week, significant differences were seen in the measurement of domestic PA. Significant differences were also seen in the measurement of sports PA when CT-PAQ data included self-reports of walking were used. However, when CT-PAQ data excluded walking reports, there were no significant differences between the questionnaires. Similar results were seen when looking at results of analyses with

MET•hours/week. Table 5.18 summarizes results of Wilcoxon signed-rank tests conducted to compare questionnaire scores between *retest* L-PAQ and the CT-PAQ for similar scoring units.

Table 5.18 Comparison of Physical Activity Scores Between Retest L-PAQ and CT-PAQ

PA Domain	L-PAQ (mean)			CT-PAQ (mean)			p-value
Scoring Unit: Lifetime Average Hours/Week							
Sports (CT-PAQ includes walking)	2.2	±	3.0	6.4	±	7.5	0.0*
Sports (CT-PAQ excludes walking)	2.2	±	3.0	3.0	±	3.4	0.1
Domestic	16.5	±	16.6	11.3	±	14.9	0.01*
Total PA (CT-PAQ includes walking)	18.8	±	16.7	17.7	±	19.1	0.31
Total PA (CT-PAQ excludes walking)	18.8	±	16.7	14.3	±	15.8	0.005*
Scoring Unit: Lifetime MET•Hours/Week							
Sports (CT-PAQ includes walking)	13.1	±	17.6	29.8	±	32.0	0.0*
Sports (CT-PAQ excludes walking)	13.1	±	17.6	17.9	±	20.2	0.08
Domestic	49.1	±	48.8	33.2	±	44.0	0.0*
Total PA (CT-PAQ includes walking)	62.2	±	51.5	63.0	±	62.6	0.45
Total PA (CT-PAQ excludes walking)	62.2	±	51.5	51.1	±	50.7	0.011*

A final supplementary analysis was performed by comparing the two questionnaires which the L-PAQ was validated against. Using similar analyses techniques to the comparison with the L-PAQ, the LT-PAQ and the CT-PAQ were compared. The highest correlations were obtained for household PA for both lifetime average hours/week and MET•hours/week scoring units (0.82 and 0.78, respectively). The lowest correlations were obtained for sports when CT-PAQ data included subject reports of walking (0.54 for lifetime average hours/week and 0.48 for MET•hours/week). However, values for correlation coefficients increased somewhat when CT-PAQ data did not include walking (0.57 for both scoring units respectively). Table 5.19 summarizes results of correlation analyses between similar scoring units for the two questionnaires. Since these analyses were done as supplementary analyses, scatter plots corresponding to the correlations will not be presented.

Table 5.19 Spearman Correlation Coefficients Between the LT-PAQ and CT-PAQ Summary Scores for Similar Scoring Units (n=76)

Correlation #	L-PAQ (Scoring Units)	CT-PAQ (Scoring Units)	Spearman Correlation Coefficient
	Lifetime Avg Hr/Wk	**Lifetime Avg Hr/Wk *(Walk)***	
1	Sports PA	Sports (Walk)	0.54
2	Household PA	Household PA	0.82
3	Total PA	Total PA (Walk)	0.71
	Lifetime Avg Hr/Wk	**Lifetime Avg Hr/Wk *(No Walk)***	
4	Sports PA	Sports (Walk)	0.57
5	Household PA	Household PA	0.82
6	Total PA	Total PA (Walk)	0.73
	Lifetime MET•Hr/Wk	**Lifetime MET•Hr/Wk *(Walk)***	
7	Sports PA	Sports (Walk)	0.48
8	Household PA	Household PA	0.78
9	Total PA	Total PA (Walk)	0.63
	Lifetime MET•Hr/Wk	**Lifetime MET•Hr/Wk *(No Walk)***	
10	Sports PA	Sports (Walk)	0.57
11	Household PA	Household PA	0.78
12	Total PA	Total PA (Walk)	0.65

Questionnaire scores for the LT-PAQ and CT-PAQ were also compared using Wilcoxon signed rank tests of differences among pairs. Analyses were conducted for sports, domestic, and total physical activity for similar scoring units between the two questionnaires, lifetime average hours/week and MET•hours/week. For lifetime average hours/week, significant differences were seen in the measurement of sports (with inclusion and exclusion of walking), domestic, and total physical activity with exclusion of walking. Similar trends were seen for MET•hours/week scoring unit. Table 5.20 summarizes results of the comparison of LT-PAQ and CT-PAQ scores.

Table 5.20 Comparison of Physical Activity Scores Between LT-PAQ and CT-PAQ

PA Domain	LT-PAQ (mean)	CT-PAQ (mean)	p-value
Scoring Unit: Lifetime Average Hours/Week			
Sports (CT-PAQ includes walking)	4.38 ± 3.86	6.64 ± 7.51	0.02*
Sports (CT-PAQ excludes walking)	4.38 ± 3.86	3.25 ± 3.69	0.0*
Domestic	13.05 ± 9.08	11.41 ± 14.67	0.0*
Total PA (CT-PAQ includes walking)	17.43 ± 9.97	18.05 ± 18.71	0.07
Total PA (CT-PAQ excludes walking)	17.43 ± 9.97	14.66 ± 15.48	0.0*
Scoring Unit: Lifetime MET•Hours/Week			
Sports (CT-PAQ includes walking)	21.94 ± 25.21	30.85 ± 32.54	0.01*
Sports (CT-PAQ excludes walking)	21.94 ± 25.21	18.98 ± 21.89	0.03*
Domestic	39.12 ± 27.95	33.91 ± 43.41	0.0*
Total PA (CT-PAQ includes walking)	61.06 ± 36.85	64.76 ± 61.90	0.3
Total PA (CT-PAQ excludes walking)	61.06 ± 36.85	52.89 ± 50.14	0.0*

5.6 Validity Study Part III: Hypotheses Testing

Results for the hypotheses tests are presented in separate sections, corresponding to the relationships tested. An exploratory approach was used in these hypotheses tests and where relevant, interpretations of the results are also summarized. Results are tabulated and the final column of each table indicates the corresponding hypothesis tested (as previously described in Chapter 4.5.4).

5.6.1 Hypotheses Tests Based on Gender

Six hypotheses based on gender differences in PA were tested. Males were shown to have greater sports PA than females across all three L-PAQ scoring units. This confirms hypotheses H1 and H2. Males also have greater occupational PA scores for all three L-PAQ scoring units, supporting the hypotheses that they had higher participation in occupational activity and they performed occupational activities of higher intensities (H3 and H4). Opposite trends were seen for domestic PA with female having greater PA scores than males, confirming hypotheses H4 and H5. Results of hypotheses tests based on gender are summarized in Table 5.21.

Table 5.21. Results of Hypotheses Tests Based on Gender

L-PAQ Domain	Males *n=1571*	Females *n=2673*	t-test p-value *(mean)*	Hypothesis
	L-PAQ Scoring Unit: Total Lifetime Hours			
Sports	15045 ± 21250	7048 ± 12391	0.0*	**H1**
Occupational	84063 ± 34282	58768 ± 37299	0.0*	**H3**
Domestic	32083 ± 40653	92654 ± 82167	0.0*	**H5**
	L-PAQ Scoring Unit: Lifetime Average Hours/Week			
Sports	4.62 ± 6.51	2.26 ± 4.02	0.0*	**H1**
Occupational	25.61 ± 9.82	18.75 ± 11.49	0.0*	**H3**
Domestic	9.76 ± 12.31	29.14 ± 24.87	0.0*	**H5**
	L-PAQ Scoring Unit: MET• Hours/Week			
Sports	26.14 ± 38.33	13.64 ± 26.29	0.0*	**H2**
Occupational	76.58 ± 87.28	50.43 ± 71.07	0.0*	**H4**
Domestic	31.22 ± 38.09	87.60 ± 74.65	0.0*	**H5**

5.6.2 Hypotheses Tests Based on Education Level

The difference in intensity of occupational activity between different levels of education in males was tested. ANOVA results show significant differences between groups tested. Males who reported their elementary school as their highest level of education had the highest scores for intensity of occupational activity (130.4 METHW), followed by males with trade or technical school training (98.5 METHW). Males with college/university and post graduate education had the lowest scores with 60.1 METHW and 53.0 METHW respectively. Males with high school education had intermediate scores (81.7 METHW). Results, which support hypothesis H7 are summarized in Table 5.22.

Table 5.22. Results of Hypothesis Test of Education Level and Intensity of Occupational Activity in Males

L-PAQ Domain	Elementary *n=85*	High School *n=403*	Trade or Technical School *n=325*	College / University *n=454*	Post Grad Studies *n=291*	ANOVA p-value	Hypothesis
	L-PAQ Scoring Unit: MET•hours/week						
Occupational	130.4 ± 105.2	81.7 ± 94.0	98.5 ± 100.6	60.1 ± 73.1	53.0 ± 55.3	0.0*	**H7**

Another hypothesis tested was the relationship between education level and sports activity, specifically people with lower levels of education will have lower sports activity participation. Significant differences were seen for education levels compared. Results were consistent for both males and females that subjects with the highest level of education (college/university and post-graduate training) had higher sports PA scores than subjects with the lowest level of education (elementary school) across all three L-PAQ scoring units. Results of analyses confirm hypotheses H8 and H10 for males and H9 and H11 for females as summarized in Table 5.23.

Table 5.23. Results of Hypotheses Tests of Education Level and Sports Activity in Males and Females

Males							
L-PAQ Domain	**Elementary** *n=85*	**High School** *n=403*	**Trade or Technical School** *n=325*	**College / University** *n=454*	**Post Graduate Studies** *n=291*	**ANOVA p-value**	**Hypothesis**
	L-PAQ Scoring Unit: Total Lifetime Hours						
Sports	10,307 ± 13,336	15,308 ± 25,895	13,037 ± 23,630	14,351 ± 16,841	17,326 ± 19,458	0.08	**H8**
	L-PAQ Scoring Unit: Lifetime Average Hours/Week						
Sports	3.05 ± 3.9	4.67 ± 8.1	4.50 ± 6.7	4.48 ± 5.1	5.43 ± 6.5	0.04*	**H8**
	L-PAQ Scoring Unit: MET•Hours/Week						
Sports	17.60 ± 24.5	25.23 ± 45.3	25.87 ± 42.6	25.45 ± 28.5	31.57 ± 39.8	0.03*	**H10**
Females							
L-PAQ Domain	**Elementary** *n=112*	**High School** *n=827*	**Trade or Technical School** *n=417*	**College / University** *n=1014*	**Post Graduate Studies** *n=268*	**ANOVA p-value**	**Hypothesis**
	L-PAQ Scoring Unit: Total Lifetime Hours						
Sports	4,957 ± 10,722	5,818 ± 11,258	6,768 ± 11,016	7,575 ± 12,017	9,915 ± 17,545	0.0*	**H9**
	L-PAQ Scoring Unit: Lifetime Average Hour/Week						
Sports	1.48 ± 2.9	1.83 ± 3.5	2.14 ± 3.5	2.48 ± 4.2	3.2 ± 5.5	0.0*	**H9**
	L-PAQ Scoring Unit: MET•Hour/Week						
Sports	9.32 ± 19.9	10.96 ± 24.4	12.96 ± 24.0	14.92 ± 25.9	19.57 ± 34.8	0.0*	**H11**

5.6.3 Hypotheses Tests Based on General Health Status

Hypotheses based on three groupings of self-report of general health – excellent/very good, good, and poor/fair – were also evaluated. Of interest was the relationship between reported general health status and sports activity, specifically, we expected respondents who reported excellent/very good health to have higher sports PA scores than participants who reported poor/fair health. Analyses were done separately for males and females. Male subjects who reported excellent/very good health had the highest sports PA scores in both lifetime average hours/week (amount of activity) and MET•hours/week (intensity of activity), supporting hypotheses H11 and H13. Similar findings were seen in females, confirming hypotheses H12 and H14. Table 5.24 summarizes hypotheses tests based on subject reports of their general health status.

Table 5.24. Results of Hypotheses Tests Based on General Health Status and Sports Activity in Males and Females

Males					
L-PAQ Domain	**Group 1** Excellent/ Very Good Health *n = 851*	**Group 2** Good Health *n = 517*	**Group 3** Fair/Poor Health *n = 179*	**ANOVA p-value**	**Hypothesis**
	L-PAQ Scoring Unit: Lifetime Average Hours/Week				
Sports	5.38 ± 6.93	3.77 ± 5.44	3.46 ± 6.81	0.0*	**H11**
	L-PAQ Scoring Unit: MET•Hours/Week				
Sports	30.37 ± 37.46	21.04 ± 33.86	20.80 ± 50.47	0.0*	**H13**
Females					
L-PAQ Domain	**Group 1** Excellent/ Very Good Health *n =1315*	**Group 2** Good Health *n = 921*	**Group 3** Fair/Poor Health *n = 376*	**ANOVA p-value**	**Hypothesis**
	L-PAQ Scoring Unit: Lifetime Average Hours/Week				
Sports	2.48 ± 4.14	2.11 ± 2.11	1.82 ± 1.82	0.007*	**H12**
	L-PAQ Scoring Unit: MET•Hours/Week				
Sports	15.06 ± 27.71	12.55 ± 25.45	11.33 ± 22.80	0.016*	**H14**

5.6.4 Hypotheses Based on BMI

Hypotheses based on three groupings of current BMI – normal, overweight, and obese – were also evaluated. Of interest was the relationship between current BMI and sports activity where we expected respondents with lower BMI to have higher sports PA scores than participants with higher BMI. Analyses were done separately for males and females. There were no significant differences in sports PA when BMI in males were compared. Subjects who were categorized as overweight had the highest sports PA scores and subjects who were categorized as obese had the lowest sports PA scores. Results correspond to hypotheses H15 and H16. Findings were more conclusive with females with subjects with low BMI having the highest sports PA scores and subjects with high BMI, the lowest sports PA scores. Table 5.25 summarizes hypotheses tests based on subject BMI.

Table 5.25. Results of Hypotheses Tests Based on BMI and Sports Activity in Males and Females

Males					
L-PAQ Domain	**Normal** BMI 18.5-24.9 *n=481*	**Overweight** BMI 25.0-29.9 *n=694*	**Obese** BMI >30.0 *n=347*	**ANOVA p-value**	**Hypothesis**
	L-PAQ Scoring Unit: Lifetime Average Hours/Week				
Sports	4.69 ± 6.36	4.79 ± 6.90	4.18 ± 5.90	0.36	**H15**
	L-PAQ Scoring Unit: MET•Hours/Week				
Sports	26.91 ± 38.60	27.00± 37.25	23.65 ± 40.66	0.37	**H16**
Females					
L-PAQ Domain	**Normal** BMI 18.5-24.9 *n=934*	**Overweight** BMI 25.0-29.9 *n=857*	**Obese** BMI >30.0 *n =714*	**ANOVA p-value**	**Hypothesis**
	L-PAQ Scoring Unit: Lifetime Average Hour/Week				
Sports	2.61 ± 4.52	2.39 ± 4.20	1.71 ± 3.16	0.0	**H17**
	L-PAQ Scoring Unit: MET•Hour/Week				
Sports	15.88 ± 29.47	14.20 ± 27.17	10.45 ± 21.06	0.0	**H18**

Further analyses were conducted by obtaining Pearson correlation coefficients between subject BMI and L-PAQ sports PA scores. Negative correlations between BMI and sports were obtained for both males and females. Results are summarized in Table 5.26.

Table 5.26. Pearson Correlation Coefficients Between BMI and L-PAQ Sports PA Scores for Males and Females

Males	
Sports Lifetime Average Hours/Week	ρ = -0.027
Sports Lifetime MET• Hours/Week	ρ = -0.034
Females	
Sports Lifetime Average Hours/Week	ρ = -0.059
Sports Lifetime MET• Hours/Week	ρ = -0.098

VI. DISCUSSION

The psychometric or measurement properties of an instrument are important in epidemiologic studies. This study provided a comprehensive evaluation of the Lifetime Physical Activity Questionnaire (L-PAQ), an instrument measuring lifetime exposure to physical activity across three domains: sports/recreation, occupation, and domestic. There is no objective "gold standard" available for measuring total lifetime physical activity. In theory, development of this gold standard would involve prospectively following a cohort of subjects and measuring their physical activity levels using objective methods. The lack of such a gold standard lends to the complexity of validation of an instrument measuring lifetime PA with necessary steps including: 1) application of sound methods of instrument development to ensure face and content validity; 2) thorough pilot testing; 3) evaluation of the instrument's reliability or ability to produce consistent results; and most importantly, 4) establishing construct validity, the ability of the instrument to measure what it has been designed to measure. The first two steps represent instrument development while the last two steps involve the evaluation of the measurement properties of the instrument, the objective of this thesis.

Validation of the Lifetime Physical Activity Questionnaire involved a test-retest study to evaluate instrument reliability and a series of studies, including comparison to two questionnaires measuring similar constructs and hypotheses testing of L-PAQ constructs, to establish construct validity. In this chapter, the validation studies of the L-PAQ will be discussed by relating the results of each individual study to methodological aspects of its design. The strengths and limitations of each study will also be presented. Finally, a discussion of methodological issues associated with the measurement of lifetime physical activity will be discussed along with suggestions for improvement of the measurement of this construct.

6.1 Study Design

Validation studies of the L-PAQ were done within the Physical Activity and Joint Health cohort study. Since instruments are usually developed for use in epidemiologic studies, it seems intuitive to conduct validation studies within the context of the larger study to ensure the applicability and relevance of the measure to the target study population. Of the seven questionnaires measuring lifetime physical activity reviewed, two were validated within the study populations that were targeted for their use. The Retrospective Physical Activity Survey (RPAS) was validated within a clinical trial of historical physical activity (4) and its relation to adult bone parameters and the Historical Physical Activity Questionnaire (HPAQ) was validated within a clinical trial of a walking intervention (23). In these validation studies and in the validation studies conducted on the L-PAQ, a sample of the greater study population was recruited.

Since the PAJH cohort study involved Internet recruitment across Canada and participation in the validation studies involved face-to-face interviews at the Arthritis Research Centre in Vancouver, it was important that participants for the validation studies were representative of the participants in the cohort study. Comparisons between cohort study and validation study participants were performed across physical characteristics including age, height, and weight and demographic variables including marital status, ethnic origin, highest level of education obtained, and total household income were performed. The mean age of validation study participants was higher than the mean age of cohort study participants ($p=0.02$). However, comparison across other physical characteristics and demographic variables showed that validation study participants were representative of the overall cohort study participants.

6.2 Reliability

The reliability of the L-PAQ was evaluated in a test-retest study involving administration of the questionnaire on two separate occasions, with the intraclass correlation coefficient (ICC4) used as the estimate of instrument reliability. Correlation coefficients for the three physical activity domains and for total physical activity were calculated for the three scoring units of the L-PAQ, total lifetime hours, lifetime average hours/week, and lifetime MET•hours/week (Chapter 5.3, Table 5.12). For total lifetime hours, ICC4 ranged from 0.58 to 0.82; for lifetime average hours/week, ICC4 ranged from 0.57 to 0.83; and for lifetime MET•hours/week, ICC4 ranged from 0.59 to 0.80.

Across all three scoring units, the highest reliability coefficients for the L-PAQ were seen in the sports domain. These results are consistent with reliability studies of the CT-PAQ in which the authors reported the highest reliability coefficients for recreational activity (3). Both the L-PAQ and CT-PAQ incorporate lists of activities to query respondents on their sports participation. Use of such lists likely facilitated participant recall in the test-retest studies for both questionnaires, resulting in the highest correlation for the sports domain. Interestingly, findings of the reliability studies contrast to reliability studies of the LT-PAQ in which authors reported the lowest coefficients for sports activity (12). LT-PAQ authors attributed the lowest correlations for sports to the irregular nature of leisure activities and the large potential for subject variation. However, design of the LT-PAQ reliability study combined with questionnaire administration may have contributed to their findings. Specifically, four interviewers were used in the test-retest study of the LT-PAQ and the variability between interviewers' techniques was not accounted for in a separate inter-rater study.

Intermediate reliability coefficients were seen for occupational and domestic activity domains. For most people, their work life is constant and comprises a significant span of time, thus facilitating recall of occupational activities. A similar rationale could be applied to domestic activities, which for the most part, are performed on a routine basis. It has been shown that generic memory (of usual or common patterns) is more readily recalled than episodic memory (19).

There are many important factors to consider in the design and analysis of a reliability study of an instrument. One of these considerations is determining the length of time between administrations of the instrument. Too short of an interval may have respondents answering questions based on responses that they may remember from the previous administration while too long of an interval may result in changes in the construct being measured (16). There are differing opinions on the appropriate length of time between instrument administrations with some authors suggesting 2 days to 2 weeks as appropriate (Streiner 1995) to others suggesting a minimum of 1 month (20). Review of previous reliability studies of lifetime physical activity instruments show that most test-retest studies used intervals that were within recommended ranges. With the exception of the CT-PAQ for which authors used a one-year interval between questionnaire administrations, all of the reliability studies had intervals ranging from 2 weeks (LLPA and QUANTAP) to 8 weeks (RPAS and LT-PAQ). Since the test-retest study was one-part of a series of validation studies of the L-PAQ, which themselves were conducted within the PAJH cohort study, it was not feasible to recruit subjects for a repeat administration at the shorter recommended intervals. The average span between administrations of the L-PAQ was 8.01 months (median 8.23 months). The shortest interval between administrations was 2.54 months and the longest interval was 10.68 months. This interval is considerably longer compared to recommendations and to previous retest studies. Thus, one of the assumptions applied in the test-retest study was that no true changes in physical activity would have occurred between the two administrations of the L-PAQ (as we were interested in lifetime physical activity). However even with this assumption, it is important to recognize that long time interval between administrations of the L-PAQ would have virtually eliminated any memory effects which are probably common in test-retest studies and this may have reduced the correlations between administrations of the L-PAQ.

One method of adjusting for the effects of time in the test-retest study was the application of appropriate statistical techniques to assess reliability. The intraclass correlation coefficient (ICC4) was selected as the most appropriate estimator of the reliability of the L-PAQ since it excludes systematic differences in the measurements (administrations of the questionnaire). Since the L-PAQ was a self-administered questionnaire, there were no observers (interviewers) involved in its application, thus eliminating the observer error in the estimates of reliability. However, there were two important features of the L-PAQ reliability study to consider: a long time interval between administrations of the questionnaire and an intervention in the form of administration of the two other physical activity questionnaires (LT-PAQ and CT-PAQ). Use of ICC4 allowed for the adjustment for such variations due to systematic changes introduced by the study design.

Another measure of reliability commonly used by authors, but not used in the analysis of L-PAQ reliability, is the Pearson product-moment correlation (16). Based on regression techniques, the Pearson correlation is a measure of the linear relationship between two variables, or in measurement, scores between two instruments. The Pearson correlation coefficient is a more liberal measure and yields a higher estimate of reliability – which may artificially create confidence in an otherwise unreliable instrument. Another advantage of the intraclass correlation is that one could calculate a single coefficient in a situation where an instrument is administered by three observers. Using Pearson correlation, one would have to obtain three coefficients between the pairs of raters (Rater 1 vs. Rater 2; Rater 1 vs. Rater 3; and Rater 2 vs. Rater 3). However, since the L-PAQ was a self-administered instrument, this limitation of the Pearson correlation coefficient was not relevant to the reliability study.

Bland-Altman plots were constructed as a further investigation of the agreement between the baseline and retest administration of the L-PAQ. Since it was the difference of scores against the average of the scores between the two administrations that were plotted, interpretations are applied to actual scoring units (which contrast to the ICC which is a unitless value). The size of the interval of the limits of agreement (LOA) allow for the estimation of the precision of the instrument. The LOA is calculated as the mean difference of the scores plus 2 standard deviations for the upper limit and the mean difference of the scores minus 2 standard deviations for the lower limit. If the differences of the scores follow a normal distribution, 95% of them will lie between these two limits (39). This follows because the variation between subjects is not accounted for and the only thing considered is measurement error. Inspection of the Bland-Altman plots for the three physical activity domains and total physical activity showed that the differences of the scores followed a normal distribution (95% within LOA). Interval size of the LOA for sports was 7.7 hours/week, for occupation 32.8 hours/week, for domestic 71.2 hours/week, and for total physical activity 81.4 hours/week. Wide intervals for occupation, domestic, and total physical activity indicate a high degree of variation in the differences of the scores between baseline and repeat administrations of the L-PAQ. Possible explanations for these results could be the conditioning effect imposed by the overall validation study design or entry errors made by subjects. Although the limits of agreement were large across occupation, domestic, and total physical activity, true variation in these domains was also quite large, thus the reliability of the L-PAQ is still supported by these analyses.

Strengths and Limitations

Results of the test-retest showed that the L-PAQ is a reliable instrument. Test-retest studies seem straightforward -- an instrument is administered twice to the same group of subjects to determine if the instrument can yield consistent results. However, their complexity lies in the consideration of specific aspects of the study design, analysis, and interpretation of results. The strengths of the reliability study of the L-PAQ were the application of appropriate assumptions in the design of the study and in reporting of results.

The test-retest study of the L-PAQ was one part of a series of validation studies of the questionnaire. Prior to completion of the second administration of the L-PAQ, subjects were interviewed by the principal investigator for the construct validity studies of the LT-PAQ and completed the self-administered CT-PAQ. Thus, at the retest, subjects had multiple occasions of recalling previous physical activities. This was not the case at the baseline administration of the L-PAQ when subjects would have responded to recruitment over the Internet and began completing the questionnaire without prior conditioning about their past physical activity participation. This was perhaps one of the limitations of the reliability study. The recruitment process for the thesis validation studies was designed to enhance participation and encourage completion of all three questionnaires. Since the LT-PAQ interviews comprised the greatest potential for respondent burden (length of interviews plus subjects' commute to ARC) the studies were designed so that interviews for the LT-PAQ were conducted first and the self-administered questionnaires were given afterwards (with the option for participants to decline, if they chose). Thus, there was an intervention between the test and retest administrations of the L-PAQ. The effect of having the conditioning about physical activity would be different for each person, increasing variation and potentially reducing the correlation coefficients.

Suggestions for further reliability studies of the L-PAQ would be to conduct a test-retest study on a different sub-sample of PAJH cohort study participants. Since administration of the retest version of the L-PAQ would be done over the Internet, the sample need not be limited to Lower Mainland residents only. Two features of this suggested study would provide further information on the reliability of the L-PAQ: 1) potential large sample size (since recruitment could be done via mass electronic mail on a Canada-wide sample) and 2) removal of the intervening effect of prior exposure to other physical activity questionnaires (conditioning effect introduced in the L-PAQ validation studies). If such a study were to be conducted, similar assumptions that no true changes in physical activity would have occurred between the two administrations of the L-PAQ (as we were interested in lifetime physical activity) would have to be applied. A second suggested reliability study would involve a test-retest study in a new group of people (not the PAJH cohort study subjects) so that the effects of a shorter time interval between questionnaire administrations could be determined.

6.3 L-PAQ Validity

Since there is no gold standard for the measurement of lifetime physical activity, validity of the L-PAQ was established by applying the principles of construct validation. Establishing construct validity of an instrument is a complex process usually involving a series of studies which strengthen the 'nomological network' of interlocking beliefs or hypotheses about the construct measured by the instrument (21). Here, the construct measured by the L-PAQ is physical activity that a person has performed over his or her lifetime across three domains of sports/recreation, occupation, and domestic.

The validity of the L-PAQ was established in two-part studies. First the convergent validity of the L-PAQ was tested against existing questionnaires measuring similar constructs of lifetime or historical physical activity. The second part of the construct validation involved testing of a priori hypotheses of constructs measured by the L-PAQ.

The convergent validity of the L-PAQ was tested against the Lifetime Total Physical Activity Questionnaire, an interviewer-administered questionnaire (12). Correlation coefficients were calculated between similar scoring units for both questionnaires (lifetime average hours/week and lifetime MET•hours/week) for similar PA domains and total PA. Spearman correlation coefficients were reported due to the skewed distribution of PA scores for both questionnaires. For lifetime average hours/week, coefficients ranged from 0.41 (total PA) to 0.71 (domestic PA) and for lifetime MET•hours/week, coefficients ranged from 0.34 (total PA) to 0.71 (domestic PA). Correlations among measures of the same construct should fall in the midrange of 0.40 to 0.80 (16), thus, there is good correlation between the L-PAQ and the LT-PAQ.

For lifetime average hours per week, the strongest correlation between the L-PAQ and the LT-PAQ was seen in the comparison of the L-PAQ domestic domain and the LT-PAQ household domain (0.71) and shown in the scatter plot in Figure 5.8. Correlations for sports and occupational PA were mid-range (0.52 for sports and 0.55 for occupation) and the lowest correlations were obtained for total PA (0.41). Obtaining the lowest correlation for the total PA score was expected since calculation of total PA involved summation of the scores across the three PA domains, therefore propagating any errors associated in the measurement of each domain and resulting in the lower correlation.

The highest correlations were seen for the domestic physical activity. This is likely due to the nature of household activities which, for the most part, are done regularly over long periods of time. Such activities include meal preparation, laundry, and housekeeping - things required to run a household. Even seasonal activities such as gardening and lawn mowing in the summer months are often done routinely and thus, recall of these activities are more readily facilitated than sports activities. Further contributing to the strong correlation for domestic physical activity are the question formats in both questionnaires which were asked the question in general fashion. Specifically, the L-PAQ defined domestic activity across the four general areas of activity of childcare, elderly care, gardening, and housework and asked participants to report activity patterns in a typical week. Similarly, in administration of the LT-PAQ, interviewers probe subjects about domestic activity by asking them to think about general areas of housework, home repair and maintenance, gardening and yard work, and childcare and asking respondents to report activity patterns in a typical week.

Intermediate correlations were seen for the occupational activity domain. This again was likely due to the difference in the administration of the questionnaires as well as the question formats. In the L-PAQ, respondents were asked if they have ever had a job and a positive response prompted the system to questions on Job 1 including job title and items described in Chapter 4.2 to quantify participation and the frequency and duration of bodily movements involving the knee. This was followed with a question about whether they had another job and again, a positive response prompted the question on Job 2. Respondents had the opportunity to report on a total of 10 jobs. During administration of the LT-PAQ, principal investigator asked respondents about jobs that they have had over their lifetime by asking them to report job titles and main duties in an attempt to quantify physical activity according to the number of hours spent doing work-related tasks. The open format of the interview combined with the opportunity for interviewers to probe respondents potentially allowed for a greater number of jobs to be reported by individual and this might include part-time jobs held during adolescence and/or any seasonal jobs. This is in contrast to the L-PAQ where participants may have only reported on main occupations they have had. L-PAQ and LT-PAQ raw questionnaire data were examined to compare the average number of jobs reported by respondents. With the LT-PAQ interviews, respondents reported an average of 7.8 jobs with a range of 2 to 26 jobs reported while with the L-PAQ respondents reported an average of 2.8 jobs with a range of 1 to 7. Closer inspection of this raw data showed that the difference is in the reporting of jobs that were held during adolescence (for example newspaper routes and other after-school type work) and seasonal jobs held during vacations from post-secondary education.

The lowest correlations for sports PA was expected due to the nature of participation in sports and recreational activities. Whereas occupational activity are salient to most people and thus readily recalled and domestic activity are regularly performed, sports and recreational activities, whether they are done for exercise and fitness purposes are more occasional or irregular in nature. It has been shown that generic memory of usual or common patterns is more readily recalled than episodic memory (19). Despite arguments that many people may participate in scheduled activities (for example, sports team practices or weekly aerobics classes), participation may not necessarily be continuous. For example, a woman who may have participated in weekly aerobics classes prior to having children, stopped for a number of years, and resumed participation would be classified as having discontinuous participation.

The low correlation for sports activities between the L-PAQ and LT-PAQ could also be explained by the differing question formats and structure across the two questionnaires. While both questionnaires involved a "show-sheet" of possible sports/recreational activities, the L-PAQ list was shorter with 64 items and the LT-PAQ sheet shown to participants during interviews was longer with 104 items. It is intuitive that combining a longer list of items with the open-format of an interview would allow subjects greater opportunities to report sports activities in the LT-PAQ. This is confirmed by inspecting descriptive statistics for the lifetime average hours per week showed that the LT-PAQ mean and median (4.6 ± 4.6; median 3.5) were higher than L-PAQ mean and median (2.7 ± 3.5; median 1.5). Lower correlations in the sports domain could have also been due to differences in quantifying activity in the two questionnaires. As described in Chapter 4.2, the L-PAQ quantifies duration of sports participation by asking respondents to enter the age at which they started participation and the age at which they stopped participation. Calculating the difference between the ages entered would then yield duration of participation in years. The design of the L-PAQ did not allow for respondents to report intermittent sports participation. A typical scenario is a person who may have participated in a sport during one period of his or her life, stopped for a number of years, and resumed participation, perhaps at a different frequency from the previous participation. Such information was captured by the LT-PAQ but not by the L-PAQ. Several validation study participants who were completing the retest administration of the L-PAQ (after LT-PAQ interviews) had contacted the principal investigator with concerns and questions about reporting discontinuous sports participation.

Another possible explanation for the low correlation between the L-PAQ and the LT-PAQ for sports activity is the differences in the administrations of the questionnaires. Vuillemin et al. compared responses to the Modifiable Activity Questionnaire by the same subjects when the questionnaire was administered by an interviewer and when it was adapted for self-administration (50). Their results showed that higher levels of leisure activity were reported during interviews compared with self-administration of the questionnaire. The authors discussed that the presence of an interviewer may provide a subject with a more structured framework for his or her response thus allowing the subject to provide more information (48). Combining findings from this study with probing techniques incorporated with the LT-PAQ interviews may have indeed prompted respondents to provide greater information about participation in sports activity. Comparison of mean values for sports activity measured by the L-PAQ and LT-PAQ showed a significant difference between the two questionnaires with higher scores calculated with the LT-PAQ (mean lifetime average hours/week were 2.4 and 4.1 for the L-PAQ and LT-PAQ, respectively and p-value for independent samples t-test was 0.003).

The convergent validity of the L-PAQ was also evaluated against a questionnaire measuring lifetime physical activity developed by Chasan-Taber et al. (CT-PAQ) (3). Similar analyses were performed for this comparison (as done for the comparison with the LT-PAQ) with correlation coefficients calculated between similar scoring units for both questionnaires (lifetime average hours/week and lifetime MET•hours/week). For the CT-PAQ leisure domain, two scores were obtained: 1) with walking included; and 2) with walking excluded. Since the leisure activity score is used in the calculation of the total PA score, the total PA score was also calculated with and without walking. Such calculations were performed because authors of the CT-PAQ as well as the questionnaire from which it was adapted have commented that subject estimates of leisure activity were much less reproducible when self-reports of walking were included in the estimates (3, 4, 5).

Spearman correlation coefficients were reported due to the skewed distribution of PA scores for both questionnaires. With walking included, coefficients for lifetime average hours/week ranged from 0.39 (sports) to 0.56 (domestic) and for lifetime MET•hours/week, coefficients ranged from 0.38 (sports) to 0.56 (domestic). Exclusion of walking from calculation of PA scores yielded higher correlations for sports (0.58 for lifetime average hours per week; 0.60 MET•hours per week). Applying the same standard by Streiner that correlations among measures of the same construct should fall in the midrange of 0.40 to 0.80 (16), the L-PAQ is moderately correlated to the CT-PAQ.

The correlations for sports were lower when subject reports of walking were included in the calculations than when they were excluded for both scoring units. For lifetime average hours/week, correlations for sports with walking included was 0.39 and 0.58 with walking excluded and for MET•hours/week; correlations for sports with walking included was 0.38 and 0.60 with walking excluded. In their reports of the reliability of the CT-PAQ, authors excluded walking from their calculation of physical activity and subsequent estimates of reliability. While walking was listed as a leisure activity in the CT-PAQ, it is very likely that respondents reported on walking done outside of leisure pursuits (for example walking as a form of commuting). Thus, it is more appropriate to report correlations based on physical activity scores that exclude subject self-reports of walking.

The second part of the construct validity studies of the L-PAQ involved looking at the questionnaire itself to evaluate the robustness of measurement. Fourteen a priori hypotheses were developed based on rationalized relationships between constructs measured by the L-PAQ and variables collected as part of the PAJH survey.

The first set of hypotheses involved testing of differences between males and females. Specific hypotheses tested were higher sports and occupational activity L-PAQ scores in males than females. An opposite trend was hypothesized for L-PAQ domestic activity with females having greater scores than males. Results of independent t-tests comparing mean scores between males and females support these hypotheses. We found males to have higher scores for sports and occupational activity than females across all three L-PAQ scoring units. For domestic activity, we found females to have higher scores than males across all three L-PAQ scoring units.

These gender-related differences in physical activity measured by the L-PAQ are supported in the literature, including previous validation studies of the QUANTAP. Authors reported that mean time in hours spent in physical activity was significantly higher for males than females across the three dimensions of sports at school, leisure sport, and occupation (15). An opposite trend was shown for the dimension of daily activities with females having higher scores than males. The relationship between domestic activity and gender is further supported by previous researchers who have demonstrated that for women, household activity is a major contributor to weekly energy expenditure (40).

Hypotheses based on highest level of education completed by respondents were also tested. However, unlike the gender-based hypotheses for which previous authors have related to lifetime physical activity (15), there are no studies to date that have described the relationship of socioeconomic factors to *lifetime* physical activity. Thus the rationale applied to the hypotheses tests of education level and lifetime physical activity were based on social science literature relating participation in *current* physical activities to socioeconomic factors.

Baseline PAJH cohort study participants were grouped according to the highest level of education they have completed: 1) elementary school, 2) high school, 3) trades or technical school, 4) college/university, and 5) post-graduate studies. We tested the relationship between education level and intensity of occupational activity over lifetime in males. Specifically, we expected males reporting lower education levels (elementary school) to have the highest intensity of occupational activity than males reporting higher education levels (university/college and post-graduate study). We also expected males who attended trades and technical to have intermediate intensity of occupational activity. This hypothesis was tested using the MET•hours/week scoring unit as this provided an indicator of activity intensity. Results of one-way ANOVA showed significant differences between the five levels of education.

More importantly, expected trends amongst education levels were seen: males with elementary education had the highest scores (130.4 METHW), followed by males with trade or technical school training (98.5 METHW). Males with college/university and post graduate education had the lowest scores with 60.1 METHW and 53.0 METHW respectively. Males with high school education had intermediate scores (81.7 METHW). Such trends could be explained by the fact that men with elementary as their highest level of education were more likely to have high intensity, labour-based jobs while those with higher levels of education were more likely to have sedentary, office-type jobs. Interestingly, men who reported trade or technical school training had the second highest occupational activity intensity scores – which were likely reflective of the nature of trades-related work. Raw PAJHS data were inspected for subject reports of job titles to support the rationale for the findings. Some occupational descriptors for subjects with elementary education included "construction," "truck driver," "ship yard worker," "miner," and "commercial fisherman." While similar descriptors emerged in respondents reporting high school education, inspection of job titles showed that there were a great number of subjects who were also involved in less active occupations including clerical positions ("bank clerk," "mail clerk") and sales positions. Occupational titles reported by subjects with trade or technical school training included "mechanic," "electrician" and "pulp mill worker." Subjects who completed college or university reported such job titles as "teacher," "accountant" and "social worker." Finally subjects with post-graduate training reported the following titles: "engineer," "professional musician," "university lecturer" and "physician."

Findings are further supported by extending research that evaluated relationships between socioeconomic factors and *current* physical activity to *lifetime* physical activity. Fogelman et al. studied the relationship between socioeconomic and health factors and physical activity using the Baeke Physical Activity Questionnaire, an instrument measuring past month activity and showed that males with lower education levels had greater physical activity at work (41). This relationship was also shown by He et al. (42).

Another hypothesis tested was the relationship between education level and sports activity. The hypothesis that people with lower levels of education will have lower sports activity participation was tested in both males and females. Results were consistent for both genders that subjects with the highest level of education (college/university and post-graduate training) had higher sports PA scores than subjects with the lowest level of education (elementary school) across all three L-PAQ scoring units.

Droomers et al. showed this relationship in their study of educational differences in leisure-time physical *in*activity (43). Similar findings of high level of education and greater engagement in sports activities were shown by other authors (41, 44). However as with the rationale for the previous hypothesis on intensity of occupational activity and level of education in males, descriptors using *current* measures of physical activity were extended to our measure of *lifetime* physical activity.

Hypotheses based on three groupings of self-report of general health – excellent/very good, good, and poor/fair – were also evaluated. Of interest was the relationship between reported general health status and sports activity, specifically, we expected respondents who reported excellent/very good health to have higher sports PA scores than participants who reported poor/fair health. Analyses were done separately for males and females. Male subjects who reported excellent/very good health had the highest sports PA scores in both lifetime average hours/week (amount of activity) and MET•hours/week (intensity of activity). Similar findings were seen in females.

The relationship between general health status and sports activity was tested on two L-PAQ scoring units, lifetime average hours/week and MET•hours/week. The exact wording of the question was "In general would you say your health is" and subjects had the option of reporting one of "Excellent, Very Good, Good, Fair, or Poor." Such wording would likely prompt respondents to think about present health and thus, evaluating the relationship between general health and L-PAQ scores on units of total lifetime hours of activity would not be relevant. Lifetime average hours/week was the most relevant scoring unit to test the hypothesis because its averaged nature take into account subjects' current activity and it was the most readily interpreted unit. MET•hours/week is also an averaged unit and provides an indicator of the intensity of the activity.

While results of the analyses on reported general health status and lifetime sports activity supported our rationalized relationships between these two constructs, these were perhaps the weakest set of hypotheses in the series. One reason for this, as previously described, was that the exact wording of the question on general health was "In general would you say your health is." Such wording would likely prompt respondents to think about present health and thus, it might not be relevant to relate a question that possibly prompted responses on present health with the construct of physical activity done in the past. However, by testing the hypothesis on averaged L-PAQ units which factored in subject age, we were able to account for some degree of present time. There is some support in the literature for our findings. Kaplan et al. studied variables predictive of physical activity in older adults (>65 years) and reported that absence of chronic conditions, injuries, and functional limitations and lower levels of psychological distress were associated with higher frequencies of leisure activity (51). It seems intuitive that these variables would contribute to respondents' perception of the construct of general health. It is important to note that Kaplan's study looked at relationships of demographic and psychosocial factors with *current* leisure physical activity. Thus similar to previous discussions of rationale for hypotheses tests conducted, descriptors using *current* measures of physical activity were extended to our measure of *lifetime* physical activity.

Finally, hypotheses based on three groupings of current BMI – normal, overweight, and obese – and sports activity were also evaluated. There were no significant differences in sports PA when BMI in males were compared. Subjects who were categorized as overweight had the highest sports PA scores and subjects who were categorized as obese had the lowest sports PA scores. Findings were more conclusive with females with subjects with low BMI having the highest sports PA scores and subjects with high BMI, the lowest sports PA scores. Further analyses were conducted by obtaining Pearson correlation coefficients between subject BMI and L-PAQ sports PA scores. Negative correlations between BMI and sports were obtained for both males and females. Findings were similar to findings in the validation study of the QUANTAP where authors also reported negative correlations with subjects who had lower percent body fat having higher scores for leisure time activity (15).

Strengths and Limitations

Establishing the validity of questionnaires measuring lifetime PA has been recognized as a challenge by many researchers (3, 8, 12). One of the strengths of the validation study of the L-PAQ was the application of sound principles of construct validation. Validity studies of the L-PAQ involved the use of supportive studies – an improvement over previous validity studies of lifetime physical activity instruments which have not applied such methodology. In addition, the validity studies have focused on the actual construct measured by the L-PAQ. Review of the validity studies of previous lifetime physical activity questionnaires showed that validity studies for three questionnaires were focused on questionnaire versions that measured physical activity over shorter time frames.

Perhaps one of the most important parts of the construct validity studies was the comparison of the L-PAQ with the LT-PAQ. The Lifetime *Total* Physical Activity Questionnaire is a well-developed instrument which measure similar domains of activity. Important features of the LT-PAQ include the incorporation of cognitive techniques both prior to the interview process (through use of Life Events Calendars) and during the interview process to facilitate respondent recall. Further supporting utility of the questionnaire are its demonstrated use in two separate epidemiologic studies of cancer. Appropriate steps were taken to ensure the quality of LT-PAQ interviews including thorough pilot testing of interview techniques, refinement of the interview script applied in the studies, recording and playback of a random sample of interviews, and on-going record keeping of experiences throughout the interviews. Finally, mean values for activity measured by the LT-PAQ interviews in this study were comparable to mean values for activity reported in the two epidemiologic studies applying the LT-PAQ, further ensuring the quality of interview methods applied.

Another strength of the validity studies was the use of a series of hypotheses tests to examine relationships between physical activity constructs measured by the L-PAQ and characteristics of subjects completing the questionnaire. Relationships between lifetime physical activity and gender and lifetime physical activity and BMI with L-PAQ measurement agree with previous findings by Vuillemin et al. (15). In addition, this is the first study to show relationships between lifetime physical activity and education level. The rationale for these hypotheses tests were based on relationships between current physical activity and education level. Previous authors have applied this methodology of extending what is known about current activity and an outcome of interest to lifetime activity and an outcome of interest as seen in Kriska's study of historical physical activity and adult bone parameters (4).

One of the limitations of the L-PAQ validity studies, which could be extended as a general limitation of validity studies of lifetime physical activity questionnaires is the inability to compare questionnaire responses to objective measures of physical activity. Such a study would be cumbersome and expensive as it would involve outfitting subjects with an objective measure of choice such as an activity monitor and prospectively following them over an extended timeframe of study. In addition to using an objective measure, also suggested would be the supplementary use of subjective measures or a current physical activity questionnaire. Given the potential study costs and tremendous subject burden, such a study would not be very feasible.

The assessment of the construct validity of an instrument is an on-going process, with each study strengthening the nomological network' of interlocking beliefs or theories about the construct being measured and the validity of the instrument (21). Thus, suggestions for future studies L-PAQ validity studies are not necessarily to improve methods but to continue building this nomological network. Further hypotheses testing of relationships between lifetime physical activity measured by the L-PAQ and characteristics of PAJH respondents could be conducted as new relationships between *current* activity and sociodemographic factors are reported.

6.4 Issues Associated with the Measurement of Lifetime Physical Activity

There are many issues associated with the measurement of lifetime physical activity using self-report instruments. In this section, these issues will be highlighted in context with the validation studies of the L-PAQ. In addition, some suggestions for improvement of measurement of lifetime physical activity, both specific to the L-PAQ and to instruments in general, will be discussed.

Responder bias, which could be considered as an umbrella term to include the factors that influence a person's response which make it deviate from the true answer, is an important consideration in administration of self-report instruments (16). Social desirability is a type of responder bias that is relevant in the measurement of physical activity. Adams defined social desirability as "the defensive tendency of individuals to portray themselves in keeping with perceived cultural norms" (52). It is closely related to social approval which the same authors defined as the "need to obtain a positive response in a testing situation" (52). Social desirability may have come to effect in the domestic activity domain with subjects overestimating time spent in household activities. Mean values for domestic activity measured by the L-PAQ (19.3 hours/week) were higher than mean values measured by the LT-PAQ (12. hours/week) and the CT-PAQ (11.8 hours/week) for lifetime average hours/week. It is possible that in an attempt to portray themselves as responsible individuals, weary of household duties, subjects overestimated their reports of the hours spent performing domestic activity. This was especially problematic in reporting of hours spent in childcare. Inspection of raw PAJHS data showed that women tended to report higher weekly hours spent caring for their children (greater than the applied ceiling of 126 hours/week). It is also possible that for these women, they tended to overestimate their household activities to compensate for lack of or lesser amounts of activities in other domains, especially sports and recreation.

These rationale are supported by a study of the effects of both social desirability and social approval biases in self-reports of physical activity in females (52). Using a validated social desirability score (Marlow-Crowe Social Desirability Scale) to describe their subjects, authors compared subject responses to two self-administered, 7-day physical activity recall surveys against objective reference measures and showed that social desirability was associated with over-reporting of both light and moderate activities (52). Interestingly, authors commented that for their subject population, household activity was a major contributor to time spent in light activity. Authors provided similar explanations for the effect of social desirability in their study: first, societal norms for women to be "good caretakers" in the home, and second, compensating in reporting of time spent in household activity when there were no activities to report for leisure and sports questions (52). The routine nature of activities performed in the home might also facilitate the effects of social desirability bias. Authors further proposed that people may be more prone to over-report on more highly prevalent, routine activities than activities that are performed less frequently or not performed at all (52).

Another important bias recognized by researchers measuring lifetime physical activity is recall bias. When epidemiologists refer to this type of bias, they usually refer to the differential bias seen in case-control studies where subjects with the specific disease under study (cases) tend to report exposures differently than those subjects without the disease (controls) (54). According to Durante, recall bias can also affect the measurement of physical activity if subjects consistently over-report activities such that relations between physical activity and outcomes of interest may artificially emerge (53).

A related problem to recall bias which also has important effects on the measurement of lifetime physical activity is limitations in recall (54). Gordis describes limitations in recall as both a problem in recollecting information and also a problem with subjects simply not having the information requested (54). This problem has perhaps received greater attention by researchers with activities focusing on both characteristics of the instrument and characteristics of respondents (53) to address issues associated with limitations in recall.

There is a small body of research on the subject determinants of physical activity recall. There are two papers – both by the same author – that are relevant to this thesis topic of measurement of lifetime physical activity. In 1999, Falkner studied the quality of recalled information on physical activity in the distant past by comparing subject self-reports completed in 1960 to recalled information obtained in between 1992-1996 (55). Authors showed that recall was best for weekday light and weekday moderate activities. The same authors further reported on characteristics of study participants in an attempt to relate these to their ability to recall information (56). Findings were inconclusive in that results did not show any associations between subject characteristics of gender, age, education level, marital status, BMI and quality of recall (54).

There is a growing emphasis on studying characteristics of the instrument and the underlying mental processes involved in the question-answering process or a more task-oriented approach. Researchers recognize that these may have greater implications for obtaining accurate histories of physical activity (53). In an editorial by Friedenreich, an important undertaking in epidemiology is to improve participant recall in retrospective research by understanding the process of information retrieval and delineating factors that predict accurate and reliable recall (19). Applying theories of cognitive psychology to the process of questionnaire completion identified four stages involved in responding to a question: 1) question comprehension, 2) information retrieval, 3) estimation and judgment, and 4) response formulation (19). During the first stage, a subject perceives, interprets, and then stores the question in short-term memory. In the second stage of retrieval, the subject uses the information in the question to generate retrieval cues for use to search their memory. The third stage involves the integration and evaluation of the retrieved information and if this information satisfies the retrieval objectives, the subject will then move on to the fourth stage and formulate a response (53). This is not a unidirectional process since a subject may also deem the information from their formulated response as inappropriate, thus prompting them to go through the decision making steps again.

While the application of cognitive methods are more applicable to interviewer-administered instruments, an understanding of this question-answering process has important implications in the design and development of lifetime physical activity instruments, discussion of results of validation studies, and formulation of suggestions for current physical activity questionnaires. The first stage, question comprehension, is perhaps the most relevant for application in self-administered questionnaires and pertains to the optimal phrasing of the question to generate the appropriate, relevant response from a subject. Perhaps one challenge with self-administered physical activity questionnaires is that subjects may fail to interpret the meaning of the question (intended by the developer). As was seen when a number of subjects overestimated their reported weekly hours spent in childcare in the L-PAQ (greater than the applied ceiling of 126 hours/week). Equally likely to social desirability bias was subjects' misinterpretation of the question. The actual wording of the question was, "How many hours per week did you care for children, on average?" It was possible that without specifying that of interest was the physical aspect of this activity, subjects may have interpreted the literal meaning of the question which would explain why a number of people reported 168 hours/week (which literally translates into a response that they care for their children 24 hours/day, 7 days/week). A suggested improvement to physical activity questionnaires is to ensure that questions are worded to facilitate subject interpretation of the intended meaning of the question. It is important throughout a questionnaire to specify the interest in the physical component of the activity. This is especially important for the L-PAQ domestic activity domain (which was the final section) due to the number of potential questions participants were required to answer before this section. A reminder about the theme of physical activity throughout the questionnaire would be important given the potential for respondent burden (mental fatigue) with this instrument.

A related suggestion would be determining if different ordering of questionnaire sections would affect reliability and validity of the L-PAQ (by facilitating subject response or minimizing subject mental fatigue). The L-PAQ was organized such that respondents answered questions on sports activities first, followed by occupation, and then domestic activity. In contrast, the LT-PAQ interviews were structured such that respondents were asked about occupation, household, and then sports activities. Drawing from the thesis author's experiences with the LT-PAQ interviews, section order was important in helping respondents with their recall. Validation study subjects anecdotally commented that they found recall of occupational events and household activities more difficult and appreciated having them covered earlier in the interviews. Many found recall of sports activities easier – this may have been due to the nature of sports activities or the fact that they would have been well primed for this section having already covered two domains of activity.

One of the novel features of the L-PAQ was its Internet-based administration which allowed for the incorporation of computer adaptive technology to maximize efficiency of data collection while minimizing respondent burden. Another suggestion for questionnaire improvement is to use the technology to provide cues and set response limits to some questions such that calculated average hours/week of activities do not exceed 168 hours/week (as was seen when baseline L-PAQ was cleaned prior to analyses for the validation studies). Another suggestion would be to divide respondents' lifetimes into time-periods similar to ones used in the CT-PAQ, HLAQ, and RPAS. These questionnaires divided subjects' lifetime into four age periods: 1) 14-21 years; 2) 22-34 years; 3) 35-50 years; and 4) 50+ years. Given the large span covered by the fourth age period, a further breakdown into two periods, 50-65, and 65+ may provide greater insights into activities performed during these age periods. As previously discussed in Chapter 6.4 a number of validation study participants commented that the design of the L-PAQ did not allow for them to report intermittent sports participation. Dividing lifetime into age periods would allow for such information to be captured since an individual's participation in certain activities may vary across different points in their lives. Placing questions within time frames may also facilitate participant recall of physical activities that they have performed over their lifetime.

Finally, interpretation and response behaviours towards certain questions should also be considered in the development of self-administered questionnaires and literature outside of measurement and epidemiology would provide tremendous insight into this concept. According to experts assembled together by the Women's Health Initiative by the National Institutes of Health/Centre for Disease Control and Prevention, in women, the physical intensity of a task is often confused with the emotional demand of the activity (48). This has important implications for questions on domestic activity and may have been a factor in responses to the L-PAQ since a number of women over-reported on the time spent caring for children. Perhaps, their responses were based on their recollected emotional demand of caring for children and not necessarily the actual activity component of the task. Ergonometric literature on self-reports of physical demands of jobs could also lend insight into responses to the occupational activity domain of the L-PAQ (Burdorf 1997). Of interest is subjects' reporting on the duration (number of hours) of 9 bodily movements (sitting, standing, standing while holding objects >50 lbs, walking, walking while carrying objects >50 lbs, pushing objects over 75 lbs, using heavy tools, squatting continuously, and kneeling continuously) which were used in L-PAQ calculations to derive MET•hours/week for occupational activity. These data will be further used in estimating loading on the knee and hip and thus, it is important to investigate subjects' response patterns and behaviours to these questions. During scoring and cleaning of the L-PAQ, data for some subjects were coded as missing since the weekly sum of the 9 bodily movements for these subjects exceeded 168 hours/week, indicating an overlap in the number of hours reported for individual movements. Occupational tasks are often comprised of multiple activities (and body movements) and often, people do not separate these tasks when reporting. Other factors that would be of interest are explored gender differences in the reporting of physical demands of occupation (Hooftman, 2005). Overall, this body of literature could be applied not only in the development of appropriate questions to ask about physical activities at work but also in the interpretation of studies applying the instrument of interest.

Suggestions for improvement measurement of lifetime physical activity are not only applicable to the L-PAQ but for other instruments measuring similar constructs. These suggestions were informed practically by the review of existing quantitative history physical activity questionnaires and experiences with the validation studies. More importantly, these suggestions were also informed by findings from the validation studies.

6.5 Conclusions

The psychometric properties of the Lifetime Physical Activity Questionnaire (L-PAQ) were evaluated in comprehensive validation studies. Information across epidemiology, exercise science, social science, and cognitive psychology were integrated and applied in the validation studies of this instrument. The L-PAQ demonstrated very good reliability for sports activity, good reliability for occupation, and moderate reliability for domestic and total physical activity. Based on the results of comparison with other instruments measuring similar constructs, the L-PAQ demonstrated very good validity for domestic activity and moderate validity for sports and occupational activity. The robustness of the measurement of lifetime physical activity by the L-PAQ was supported by the results of hypotheses tests between constructs measured by the questionnaire and descriptors of survey respondents. The relationship between lifetime physical activity measured by the L-PAQ and gender agree with previous findings. In addition, this was the first study to show relationships between lifetime physical activity and education level.

Overall, validation studies of the L-PAQ contribute to the existing knowledge of measurement of lifetime physical activity and provide evidence for the utility of this instrument. Previous validity studies of instruments measuring lifetime physical activity have not looked at the actual construct measured or used supportive studies as the ones used to validate the L-PAQ. Psychometric properties define the quality of an instrument and this has a limiting influence on the ability of epidemiologic studies to detect significant associations between physical activity and health outcomes. Results of the L-PAQ validation studies will be useful in the interpretation of findings of the PAJH cohort study through the estimation of measurement error and its effects on risk estimates for lifetime physical activity and joint health.

Bibliography

1. Montoye H. Introduction: Evaluation of some measurements of physical activity and energy expenditure. Medicine & Science in Sports & Exercise 2000.

2. Shephard R. Limits to the measurement of habitual physical activity by questionnaires. British Journal of Sports Medicine 2003; 37:197-206.

3. Chasan-Taber L, Erickson J, McBride J, Nasca P, Chasan-Taber S, Freedson P. Reproducibility of a self-administered lifetime physical actvitiy questionnaire among female college alumnae. American Journal of Epidemiology 2002; 155(3):282-289.

4. Kriska A, Sandler R, Cauley J, Laporte R, Hom D, Pambianco G. The assessment of historical physical activity and its relation to adult bone parameters. American Journal of Epidemiology 1988; 127(5):1053-1063.

5. Kriska A, Knowler W, Laporte R, Drash A, Wing R, Blair S et al. Development of questionnaire to examine relationship of physical activity and diabetes in Pima Indians. Diabetes Care 1990; 13:401-411.

6. Caspersen C, Powell K, Christenson G. Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. Public Health Reports 1985; 100:126-131.

7. Caspersen C. Physical activity epidemiology: Concepts, methods, and applications to exercise science. Exercise & Sport Sciences Reviews 1989; 17:423-462.

8. Kriska A, Caspersen C. Introduction to a collection of physical activity questionnaires 41. Medicine & Science in Sports & Exercise 1997; 29(Suppl6):5-9.

9. Laporte R, Montoye H, Caspersen C. Assessment of physical activity in epidemiologic research: Problems and prospects. Public Health Reports 1985; 100:131-146.

10. Ainsworth B, Jacobs D, Leon A. Validity and reliability of self-reported physical activity status: The Lipid Research Clinics questionnaire. Medicine & Science in Sports & Exercise 1993; 25:92-98.

11. Godin G, Shephard R. A simple method to assess exercise behavior in the community. Canadian Journal of Applied Sport Sciences 1985; 10:141-146.

12. Friedenreich C, Courneya K, Bryant H. The lifetime total physical activity questionnaire: Development and reliability. Medicine & Science in Sports & Exercise 1998; 30(2):266-274.

13. Kahn H. The relationship of reported coronary heart disease mortality to physical activity of work. American Journal of Public Health 1963; 53:1058-1067.

14. Menotti A, Puddu V. Ten-year mortality from coronary heart disease among 172,000 men classified by occupational physical activity. Scandinavian Journal of Work, Environment & Health 1979; 5:100-108.

15. Vuillemin A, Guillemin F, Denis G, Huot J, Jeandel C. A computer-assisted assessment of lifetime physical activity: Reliability and validity of the QUANTAP software. Revue d'epidemologie et de sante publique 2000; 48:157-167.

16. Streiner D, Norman G. Health Measurement Scales: A practical guide to their development and use. 2nd ed. New York: Oxford University Press, 1995.

17. Ainsworth B, Haskell W, Leon A, Montoye H, Sallis J, Paffenbarger R. Compendium of physical activities: Classification of energy costs of human physical activities. Medicine & Science in Sports & Exercise 1993; 25:71-80.

18. Ainsworth B, Haskell W, Whitt M, Irwin M, Swartz A, Strath S et al. Compendium of physical activities: An update of activity codes and MET intensities. Medicine & Science in Sports & Exercise 2001.

19. Friedenreich C. Editorial: Improving long-term recall in epidemologic studies. Epidemiology 1994; 5(1):1-4.

20. Armstrong B, White E, Saracci R. Principles of exposure measurement in epidemiology. New York: Oxford University Press, 1992.

21. Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychological Bulletin 1955; 52:281-302.

22. Dawson J, Juszczak E, Thorogood M, Foster C, Marks S, Dodd C et al. Distant past exercise in women: Measures may be reliable, but are they valid? Medicine & Science in Sports & Exercise 2003; 35(5):862-866.

23. Winters-Hart C, Brach J, Storti K, Trauth J, Kriska A. Validity of a questionnaire to assess physical activity in older women. Medicine & Science in Sports & Exercise 2004; 36(12):2082-2087.

24. Dan A, Wilbur J, Hendricks C, O'Connor E, Holm K. Lifelong physical activity in midlife and older women. Psychology of Women Quarterly 1990; 14:531-542.

25. Ainsworth B, Jacobs D, Leon A, Richardson M, Montoye H. Assessment of the accuracy of physical activity questionnaire occupational data. Journal of Occupational Medicine 1993; 35:1017-1027.

26. Paffenbarger R, Wing A, Hyde R. Physical activity as an index of heart attack risk in college alumni. American Journal of Epidemiology 1978; 108:161-175.

27. Kriska A, Laporte R, Pettitt D, Chalres M, Nelson R, Kuller L et al. The association of physical activity with obesity, fat distribution and glucose intolerance in Pima Indians. Diabetologia 1993; 36:863-869.

28. Vuillemin A, Guillemin F, Jouanny P, Denis G, Jeandel C. Differential influence of physical activity on lumbar spine and femoral neck bone mineral density in the elderly population. Journal of Gerentology 2001; 56A:B248-B253.

29. Friedenreich C, Bryant H, Courneya K. Case-control study of lifetime physical activity and breast cancer risk. American Journal of Epidemiology 2001; 154(4):336-347.

30. Friedenreich C, McGregor S, Courneya K, Angyalfi S, Elliot F. Case-control study of lifetime total physical activity and prostate cancer risk. American Journal of Epidemiology 2004; 159(8):740-749.

31. Washburn R, Montoye H. The assessment of physical activity by questionnaire. American Journal of Epidemiology 1986; 123:563-576.

32. Lamb K, Brodie D. The assessment of physical activity by Leisure-Time Physical Activity Questionnaires. Sports Medicine 1990; 10:159-180.

33. Vuillemin A. Revue des questionnaires d'evaluation de l'active physique. Revue d'epidemologie et de sante publique 1998; 46:49-55.

34. Felson D, Lawrence R, Dieppe P, Hirsch P, Helmick C, Jordan J et al. Osteoarthritis new insights. Part 1: The disease and its risk factors. Annals of Internal Medicine 2000; 133:635-646.

35. Felson D, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis & Rheumatism 1998; 41(8):1343-1355.

36. McAlindon T, Wilson P, Aliabadi P, Weissman B, Felson D. Level of physical activity and the risk of radiographic and symptomaic knee osteoarthritis in the elderly: The Framingham Study. American Journal of Medicine 1999; 106:151-157.

37. Hannan M, Felson D, Anderson J, Naimark A. Habitual physical activity is not associated with knee osteoarthritis: The Framingham Study. Journal of Rheumatology 1993; 20(4):704-709.

38. Newton P, Mow V, Gardner T, Buckwalter J, Albright J. The effect of lifelong exercise on canine articular cartilage. American Journal of Sports Medicine 1997; 25(3):282-286.

39. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;307-310.

40. Wilbur J, Holm K, Dan A. A quantitative survey to measure energy expenditure in midlife women. Journal of Nursing Management 1993; 1(1):29-40.

41. Fogelman Y, Bloch B, Kahan E. Assessment of participation in physical activities and relationship to socioeconomic and health factors. The controversial value of self-perception. Patient Education & Counseling 2004; 53:95-99.

42. He X, Baker D. Differences in leisure-time, household, and work-related physical activity by race, ethnicity, and education. Journal of General Internal Medicine 2005; 20:259-266.

43. Droomers M, Schrijvres H, Van De Mheen H, Mackenbach J. Educational differences in leisure-time physical inactivity: A descriptive and explanatory study. Social Science and Medicine 1998; 47:1665-1676.

44. Burton N, Turrell G. Occupation, hours worked, and leisure-time physical activity. Preventive Medicine 2000; 31:673-681.

45. Mei Z, Grummer-Strawn L, Pietrobelli A, Goulding A, Goran M, Dietz W. Validity of body mass index compared with other body-composition screening indexes for the assessment of body fatness in children and adolescents. American Journal of Clinical Nutrition 2002; 75:978-985.

46. Garrow J, Webster J. Quetelet's index (W/H2) as a measure of fatness. Internatonal Journal of Obesity 1985; 9:147-153.

47. BMI — Body Mass Index: About BMI for Adults [homepage on the Internet]. Atlanta, Georgia: Center for Disease Control and Prevention. Available from http://www.cdc.gov/doc.do/id/0900f3ec80093d70/

48. Masse L, Ainsworth B, Tortolero S, Levin S, Fulton J, Henderson K et al. Measuring physical activity in midlife, older, and minority women: Issues from an expert panel. Journal of Women's Health 1998; 7:57-67.

49. Artazcoz L, Borrell C, Benach J, Cortes I, Rohlfs I. Women, family demands and health: The importance of employment status and socio-economic position. Social Science and Medicine 2004; 59:263-274.

50. Vuillemin A, Opper J, Guillemin F, Essermeant L, Fontvieille A, Galan P et al. Self-administered questionnaire compared with interview to assess past-year physical activity. Medicine & Science in Sports & Exercise 2000; 32(6):1119-1124.

51. Kaplan M, Newsom J, McFarland B, Lu L. Demographic and psychosocial correlates of physical activity in late life. American Journal of Preventive Medicine 2001; 21:306-312.

52. Adams S, Matthews C, Ebbeling C, Moore C, Cunningham J, Fulton J et al. The effect of social desirability and social approval on self-reports of physical activity. American Journal of Epidemiology 2005; 161:389-398.

53. Durante R, Ainsworth B. The recall of physical activity: Using a cognitive model of the question-answering process. Medicine & Science in Sports & Exercise 1996; 28:1282-1291.

54. Gordis L. Epidemiology. 2nd ed. Philadelphia: W.B. Saunders Company, 2000.

55. Falkner K, Trevisan M, McCann S. Reliability of recall of physical activity in the distant past. American Journal of Epidemiology 1999; 150(2):195-205.

56. Falkner K, McCann S, Trevisan M. Participant characteristics and quality of recall of physical activity in the distant past. American Journal of Epidemiology 2001; 154:865-872.

57. Burdorf A, Rossignol M, Fathallah F. Challenges in assessing risk factors in epidemiologic studies on back disorders. American Journal of Industrial Medicine 1997; 32:142-152.

59. Hooftman W, van der Beek A, Bongers P. Gender differences in self-reported physical and psychosocial exposures in jobs with both female and male workers. Journal of Occupational and Environmental Medicine 2005; 47:244-252.

APPENDICES

Appendix I. Study Items

A1.1 Certificate of Ethical Approval

A1.2 L-PAQ List of Sports Activities and Assignment of MET Values

Table A1.1 List of Sports/Recreational Physical Activity and Corresponding Metabolic Equivalent (MET) Intensity Level

	Sports Activity	MET		Sports Activity	MET
1	Aerobics/Calistenics	6.5	**35**	Sailing	3.0
2	Backpacking/Hiking	6.0	**36**	Scuba Diving	7.0
3	Badminton	4.5	**37**	Skateboarding	5.0
4	Baseball/Softball	5.0	**38**	Skating	7.0
5	Basketball	6.0	**39**	Ski Machine	7.0
6	Bicycling	8.0	**40**	Cross Country Skiing	8.0
7	Bowling	3.0	**41**	Downhill Skiing	6.0
8	Boxing	12.0	**42**	Sledding/Luge/Bobsled	7.0
9	Broomball	7.0	**43**	Snorkelling	5.0
10	Canoeing	3.5	**44**	Snow Shoeing	8.0
11	Circuit Training	8.0	**45**	Snowboarding	8.0
12	Cricket	5.0	**46**	Snowmobiling	3.5
13	Curling	4.0	**47**	Soccer	7.0
14	Fencing	6.0	**48**	Squash	12.0
15	Field Hockey	8.0	**49**	Stair-climber	9.0
16	Floor Hockey	8.0	**50**	Surfing	3.0
17	Football	8.0	**51**	Swimming	**6.5**
18	Frisbee/Ultimate	3.0	**52**	Table Tennis	4.0
19	Golf	4.5	**53**	Tai Chi	4.0
20	Gym Classes	5.5	**54**	Tennis	7.0
21	Gymnastics	4.0	**55**	Track & Field	6.0
22	Health Club	5.5	**56**	Trampoline	3.5
23	Ice Hockey	8.0	**57**	Treadmill	5.5
24	Hunting	5.0	**58**	Volleyball	4.0
25	Jogging/Running	7.0	**59**	Walking	3.5
26	Jumping/Jump Rope	5.0	**60**	Water Polo	10.0
27	Martial Arts	10.0	**61**	Water Volleyball	3.0
28	Equestrian/Polo	4.0	**62**	Weight Lifting	6.0
29	Racquetball	7.0	**63**	White Water Rafting	5.0
30	Rock/Mountain Climbing	8.0	**64**	Wrestling	6.0
31	Rollerblading	7.0	**65**	Activity 1	
32	Rowing	3.5	**66**	Activity 2	
33	Rowing (Stationary)	7.0	**67**	Activity 3	
34	Rugby	10.0	**68**	Activity 4	

**MET Codes obtained using Compendium of Physical Activities: An update of activity codes and MET Intensities (Ainsworth,30)*

A1.3 L-PAQ Assignment of MET Values to Bodily Movements for Calculations for Occupation Section

Table A1.2 List of Bodily Movements and Corresponding Metabolic Equivalent (MET) Intensity Level

	Bodily Movement	MET
1	Sitting	1.5
2	Standing	2.3
3	Standing and holding or moving objects over 50 lbs	4.0
4	Walking	3.3
5	Walking and carrying objects over 50 lbs	6.5
6	Moving or pushing objects over 75 lbs	7.5
7	Using heavy tools	8.0
8	Squatting continuously	4.0
9	Kneeling continuously	1.0

A1.4 L-PAQ List of Domestic Activities and Assignment of MET Values

Table A1.3 List of Domestic Activities and Corresponding Metabolic Equivalent (MET) Intensity Level

	Domestic Activity	MET
1	Caring for children	2.5
2	Caring for elderly	4.0
3	Gardening	4.0
4	Housework	2.8

A1.5 Lifetime Total Activity Questionnaire (LT-PAQ) Items

A1.5.1 LT-PAQ Interview Script

Introduction

Thank you for taking part in this phase of our study. I will now be conducting an interview in which I will be asking you questions about your lifetime physical activity patterns. I am interested in the **level** and **amount** of physical activity that you have done over your lifetime. There are three parts to this interview: employment and volunteer activities, household activities, and exercise/sports activity. In each part, I will ask you about your lifetime activity starting with your childhood and ending with your reference year. **The reference year is the year before the interview in this study.**

For each part, I will ask you to tell me:

1. the age when you started and ended an activity
2. the **number of months** per year, **days per week** and **hours per day** that you did each activity
3. whether the level of activity was **light, moderate** or **heavy**

I will provide you with a set of cards which indicate levels or intensity of activity for you to use to describe your activities during our interview. I will also provide you with an activity minimums page to remind you throughout the interview of the minimum number of hours that you must have done an activity in order for it to be included. At any time during the interview, please feel free to ask any questions you may have. We will now begin with your **occupational activities.**

Occupational Activities

I am interested in employment and volunteer work that you did during your lifetime and the physical demands of that work. We will start with your first job and end with the job that you had in your reference year.

The minimum number of hours of employment or volunteer activity that needs to be done for it to be included is:

128 hours total per year, or
2.5 hours per week per year, or
If seasonal, **8** hours per week for 4 months

For each job, I will ask you:

1. the age when you started and ended
2. the **number of months** per year, **days per week** and **hours per day** that you worked
3. whether the level of activity for that job was **sedentary, light, moderate** or **heavy**
4. whether you walked, cycled or went by rollerblades, or ran to work

I will now give you a set of intensity cards. I will read these intensity categories to you. Category 1 is **sedentary** work and applies to jobs that require only sitting with minimal walking. An example would be an office job. Category 2 is **light** and applies to jobs that require a minimal amount of physical effort such as standing and slow walking. There is no increase in heart rate and no perspiration. An example would be a cashier in a store. Category 3 is **moderate** and applies to jobs that require carrying light loads (about 5 to 10 lbs), continuous walking, and are mainly performed indoors. These activities would increase the heart rate slightly and may cause some light perspiration. An example would be a stockperson in a store. Category 4 is **hard** and applies to jobs that require carrying heavy loads (greater than 10 lbs), brisk walking, climbing. These jobs require mainly outdoor activity and would increase the heart rate substantially and cause heavy sweating. Examples would include heavy construction, farming, and trade jobs that involve a high level of activity.

Job 1

We will now begin. What was the very first job you had?

a. What was your job title?
b. What is the description of this occupational activity? Or
What did you mostly do on the job, physically speaking? Did you mainly sit, stand, walk, etc? I can code up to 3 main activities. What would you say you did for 5-6 hours out of your 8 hour day?
c. At what age did you start this job?
d. At what age did you end this job?
e. How many months per year did you work at this job?
f. How many days per week did you work at this job?
g. How many hours per day and minutes her day did you work at this job?
h. What was the physical intensity of the job?
i. Did you walk, cycle or went by rollerblades, or ran to work
j. How much time per day you spent walking, cycling or rollerblading, or running to work?

Sample Probe: We do count babysitting, delivering newspapers, part-time jobs as occupations. Did you do any of these?

Continue with subsequent occupations until the last one reported

Household Activities

Now I am going to ask you about household activities. Household activities include 4 things: **housework** (cooking, cleaning, laundry, ironing, vacuuming); **yardwork** (grass, snow, vehicle maintenance); **repairs and maintenance on your house**; and **childcare.** It basically includes anything to do with your home and property. Also included are hobbies and working in the shop.

This is different from your occupational activities as rather than telling me about each job that you did over your lifetime, I would like you to tell me about patterns of activity that you did during different periods in your life. It may help if you consider what a typical day or week was for you. Then please think about how many **hours** of household, gardening, yard work or do-it-yourself jobs around your home that you did in that typical day or week. The minimum number of hours of household activity that needs to be done for it to be included is:

112 hours total per year, or

2.15 hours per week per year, or

If seasonal, **7** hours per week for **4 months**

I will know give you a set of intensity cards. I will read these intensity categories to you as well as the description of each. Category 2 is **light household activity** and these are activities that require minimal effort such as those activities that are done sitting, standing, or with slow walking. Generally, they do not require much physical effort. Examples include: ironing, washing dishes, cooking, vacuuming, dusting, light hobbies, and do-it-yourself jobs. Category 3 is **moderate household activity** and these are activities that are not exhausting, that increase the heart rate slightly, and may cause some light perspiration. Examples would include scrubbing and polishing floors, lawn mowing with an electric lawn mower, shoveling snow, and playing and caring for young children. Category 4 is **heavy household activity** and these are activities that increase the heart rate and cause heavy sweating. These activities include those that require lifting, moving heavy objects, rubbing vigorously for fairly long periods, activities that cause sweating or faster heartbeat. Examples include moving furniture, mowing with a push lawn mower, digging up a garden, or renovating a house.

Household Activity Pattern 1

TO MEN: "Did you have any household chores as a child? What about as an adult"

TO WOMEN and MEN who said YES: "When did you begin doing household chores for just over 2 hours each week? That would be about 20 minutes a day."

-What age where you at the beginning of this time period of household activity pattern?
-What would you say your pattern was? Was it a daily activity, a once-a-week activity, or a combination?
-What age where at the end of this time period of household activity pattern? Or How long did this pattern go on?
-How many months per year did you perform this pattern of household activity during this period?
-How many days per week did you perform this pattern of household activity during this period?
-How many total hours and minutes per typical day did you perform this pattern of household activity during this period
-Of this total time per day, how many hours per day did you spend doing light activity?
-Of this total time per day, how many hours per day did you spend doing moderate activity?
-Of this total time per day, how many hours per day did you spend doing heavy activity?

Household Activity Pattern 2

It's important to think about my questions on household activity in terms of time periods in your life. If there has been a change in the pattern that you report, then this will be classified as a different pattern. Moving on to a different period in your life, please tell me what your pattern of household activity was. Again I am interested in patterns of activity that you have done **for at least 7 hours per week for 4 months per year.** Again, it may help to consider what a typical day or week is for you. Then think about how many **hours** of household, gardening, yard work or do-it-yourself jobs around your home that you did in a typical day or week. For seasonal activities such as gardening, you can report those separately from all other household activities that are done all year.

-What age where you at the beginning of this time period of household activity pattern
-What age where at the end of this time period of household activity pattern
-How many months per year did you perform this pattern of household activity during this period
-How many days per week did you perform this pattern of household activity during this period
-How many total hours and minutes per typical day did you perform this pattern of household activity during this period
-Of this total time per day, how many hours per day did you spend doing light activity?
-Of this total time per day, how many hours per day did you spend doing moderate activity?
-Of this total time per day, how many hours per day did you spend doing heavy activity?

Sample Probes for Life Changes:

Men: Did you move to a bigger/smaller house, which might change your pattern of household activities? If so, how old were you when you moved, and what did your pattern become?

Women who are mothers: When did you get married / have your first child? Did that change your pattern of time spent doing household activities?

Continue with subsequent household activity patterns (if there are pattern changes) until the last pattern reported (that of the reference year)

Exercise/Sports History

In this last section, I am going to ask you about exercise/sports activity that you did during your lifetime. This section is similar to the occupational activity section and rather than thinking in patterns of activity during different time periods in your life, I am going to ask you about exercise or sports activities that you did during your lifetime starting with your childhood and continuing to your reference year. The minimum number of hours of household activity that needs to be done for it to be included is:

32 hours total per year, or
40 minutes per week per year, or
If seasonal, **2** hours per week for **4 months**

Besides sports and exercise, we are also interested in knowing whether you walked or biked to work. If you have done this, please report all the information as for the other sports activities. I will know give you an Exercise Show Sheet which is a list of exercise and sports activities, you might find this useful in remembering activities that you have done. I will also give you a set of intensity cards. Please note that on the front of the card, you will find an intensity level from 2 to 4 as well as the category of that intensity. Category 2 is **light exercise/sports activity.** These are activities that require minimal effort such as those activities that are done sitting, standing, or with slow walking. They do not require much physical effort. Category 3 is **moderate exercise/sports activity.** These are activities that are not exhausting, that increase the heart rate slightly, and may cause some light perspiration. Category 4 is **heavy exercise/sports activity.** These are activities that increase the heart rate and cause heavy sweating. These activities include those that require lifting and cause heavier sweating or faster heartbeat.

Interview:

Exercise/Sports 1

-What is the description of this activity?
-At what age did you start doing this activity?
-At what age did you end doing this activity?
-What is your usual pattern of doing this activity? **(Frequency)**
- -# times per day or
- -# times per week or
- -# times per month or
- -# times per

-How much time did you spend doing this activity in hours and minutes? **(Duration)**
-Using the intensity cards that I have given you, what intensity category would you classify this activity as?

A1.5.2 LT-PAQ Data Collection Form

Lifetime Record of Occupational Activities

Minimums:	128 hours per year 8 hours per week for 4 months 2.5 hours per week, year round

No	Age Start	Age End	Job Title	Description of Occupational Activity	CODE	Mo / Yr	Days / Wk	Time/Day		I	Walk, Bike or Car	Mo / Yr	Days / Wk	Time/Day	
								Hr	Min					Hr	Min
1															
2															

Lifetime Record of Household Activities

Minimums:	112 hours per year 7 hours per week for 4 months 2.15 hours per week, year round
Categories:	Housework; Home repair and maintenance; Gardening; Childcare

No.	Age Started	Age Ended	Number of Months/Year	Number of Days/Week	Time per Day		Hours per day spent in activities that were in category:		
					Hrs	Min	2	3	4
1									
2									

Lifetime Record of Exercise/Sports Activities

Minimum:	32 hours per year 2 hours per week for 4 months 40 minutes per week, year round
Activity done at least 10X in lifetime	

No.	Sports Activity	Code	Age Started	Age Ended	Frequency of Activity (# times)				Time/Activity		Intensity
					Day	Week	Month	Year	Hours	Min	
1											
2											

A1.5.3 LT-PAQ Formulae for Calculations
(Friedenreich 1998)

1. Average number of hours per week spent in occupational activity over lifetime:

$$= \Sigma \left[\frac{\text{(Age end – Age start) * (mo/year) * (4.33 wks/mo) * (no. days/wk) * (hr/day)/]52}}{\text{Number of years}} \right.$$

2. Average number of hours per week spent in household activity over lifetime:

$$= \Sigma \ \frac{\text{[(Age end – Age start) * (mo/yr) * (4.33 wks/mo) * (no. days/wk) * (hr/day)]/52}}{\text{Number of years}}$$

3. Average number of hours per week spent in exercise/sports activities over lifetime:

(If respondent reported per day)

$$= \Sigma \ \frac{\text{[(Age end – Age start) * 365 d/yr * (No. times/day) * (Hr/exercise session)] / 52}}{\text{Number of years}}$$

(If respondent reported per week)

$$= \Sigma \ \frac{\text{[(Age end – Age start) * 52 wk/yr * (No. times/wk) * (Hr/exercise session)] / 52}}{\text{Number of years}}$$

(If respondent reported per month)

$$= \Sigma \ \frac{\text{[(Age end – Age start) * 12 m/yr * (No. times/mo) * (Hr/exercise session)] / 52}}{\text{Number of years}}$$

(If respondent reported per year)

$$= \Sigma \ \frac{\text{[(Age end – Age start) * (No. times/yr) * (Hr/exercise session)] / 52}}{\text{Number of years}}$$

A1.6 Chasan-Taber Physical Activity Questionnaire

Lifetime Physical Activity Questionnaire (Female)

Step 1: Please place a **check mark** in the first column next to each activity that you have **ever participated in more than 10 times** during your lifetime.

Step 2: For those activities you have checked, **proceed to the right** answering the questions in the columns above.

Have you ever participated in any of the following?		**During the Past Year**		**During ages 51-65 yrs (a 15-year span)** *If younger than 51, skip to next column*			**During ages 35-50 yrs (a 15-year span)**			**During ages 22-34 yrs (a 12-year span)**			**Between age of onset of menstruation to 21 yrs**		
		No. of Months	**Total no. of hours per week**	**No. of years (15 max)**	**Typical no. of months per year**	**Typical no. of hours per week**	**No. of years (15 max)**	**Typical no. of months per year**	**Typical no. of hours per week**	**No. of years (12 max)**	**Typical no. of months per year**	**Typical no. of hours per week**	**No. of years**	**Typical no. of months per year**	**Typical no. hours per week**
Leisure Time Activities															
Walking for exercise (outdoor, indoor at mall, treadmill)															
Hiking															
Stair-climbing machine															
Jogging (outdoor/treadmill)															
Bicycling (stationery/outdoor)															
Horseback riding															
Dancing (social/ballet/tap)															
Gymnastics															
Calisthenics/toning exercises															
Yoga															
Aerobics/Jazzercise															
Lifting weights															
Swimming for exercise (ie. laps)															

Have you ever participated in any of the following?		During the Past Year		During ages 51-65 yrs (a 15-year span) *If younger than 51, skip to next column*			During ages 35-50 yrs (a 15-year span)			During ages 22-34 yrs (a 12-year span)			Between age of onset of menstruation to 21 yrs		
	Check if Yes ↓	**No. of Months**	**Total no. of hours per week**	**No. of years (15 max)**	**Typical no. of months per year**	**Typical no. of hours per week**	**No. of years (15 max)**	**Typical no. of months per year**	**Typical no. of hours per week**	**No. of years (12 max)**	**Typical no. of months per year**	**Typical no. of hours per week**	**No. of years**	**Typical no. of months per year**	**Typical no. hours per week**
Rowing/canoeing/ kayaking/rowing machine															
Water skiing															
Skiing/downhill															
X-country skiing/ski machine															
Skating (ice, roller, in-line)															
Tennis															
Other racquet sports															
Softball/baseball															
Golf (use golf cart)															
Golf (walking)															
Volleyball															
Basketball															
Bowling															
Other															
Household Activities															
Gardening/yard work															
Grooming and feeding children															
Playing with children (walk/run them)															
Light housecleaning (sweep/dust)															
Heavy housecleaning (scrub floors/vacuum)															

A1.6.1 CT-PAQ Formulae for Calculations
(Chasan-Taber 2002)

1. Calculate the average participation for each activity in a given time period to arrive at an estimate in units of **hours of week**

$$\frac{\text{(yr participated in activity) x (mo/yr) x (4 wk/mo) x (hr/week)}}{\text{(total yr in period) x (52 wk/year)}}$$

APPENDICES

Appendix II. Figures Corresponding to Study Results

Appendix 2.1 Reliability Study Figures

The following figures represent frequency histograms corresponding to scoring of the baseline L-PAQ and repeat administration L-PAQ for the test-retest reliability study. Each figure is a group of frequency histograms (with normal curves) corresponding to scores for each of the three physical activity domains and for total PA for each of the three L-PAQ scoring units: 1) total lifetime hours, 2) lifetime average hours/week, and 3) lifetime MET•hours/week. Figures A2.1 to A2.3 correspond to baseline L-PAQ scoring and Figures A2.4 to A2.6 corresponding to repeat L-PAQ scoring.

Bland-Altman plots were constructed as a graphical description of the reliability of the L-PAQ. Figures for lifetime average hours/week were presented in the main body of the thesis as they correspond to the most readily interpreted L-PAQ scoring unit. Figure A2.7 is the Bland-Altman plot for total lifetime hours for sports, occupation, domestic, and total PA. Figure A2.8 is the Bland-Altman plot for lifetime MET•hours/week for sports, occupation, domestic, and total PA.

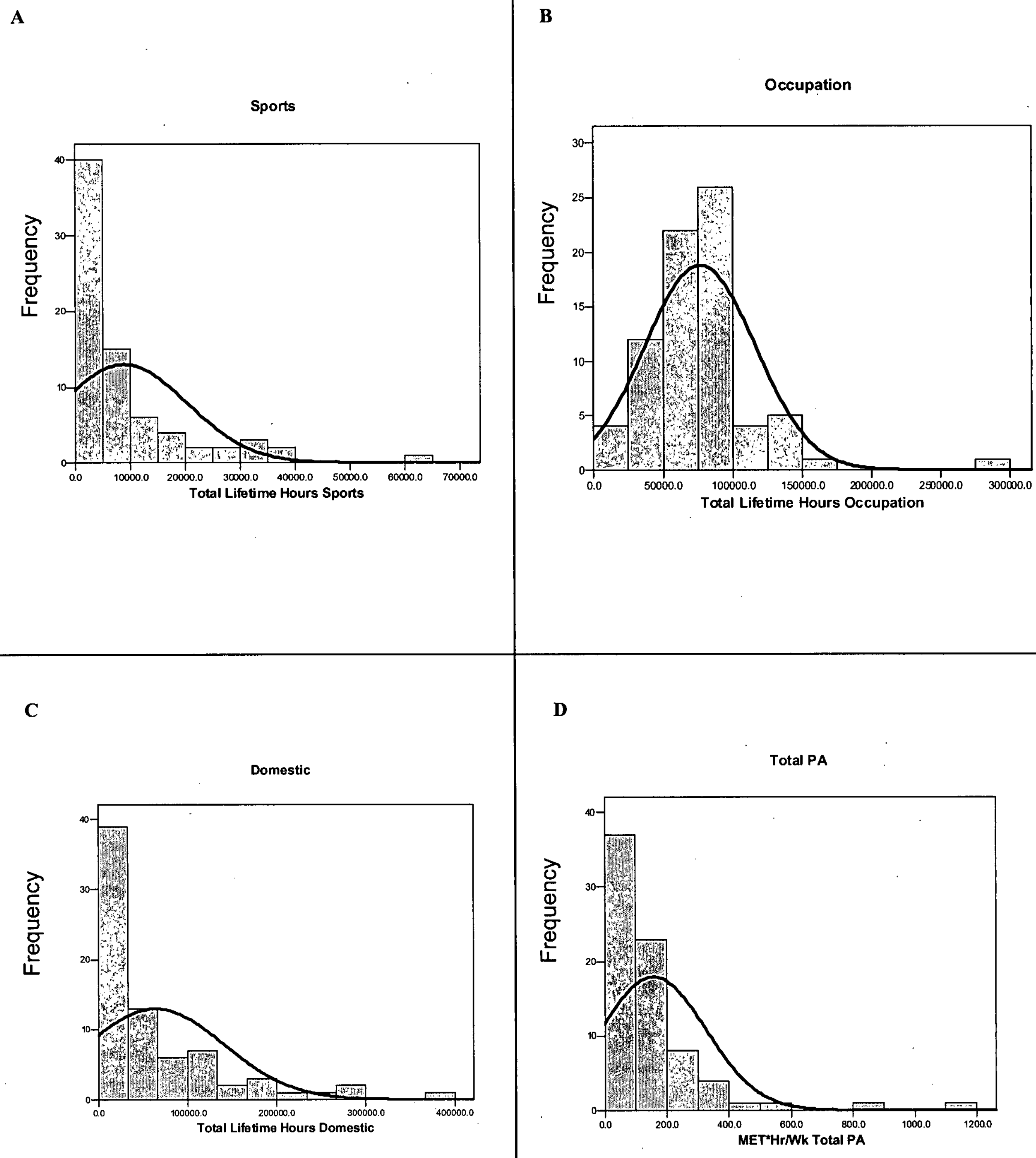


Figure A2.1. Baseline L-PAQ Total Lifetime Hours Frequency Histograms for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

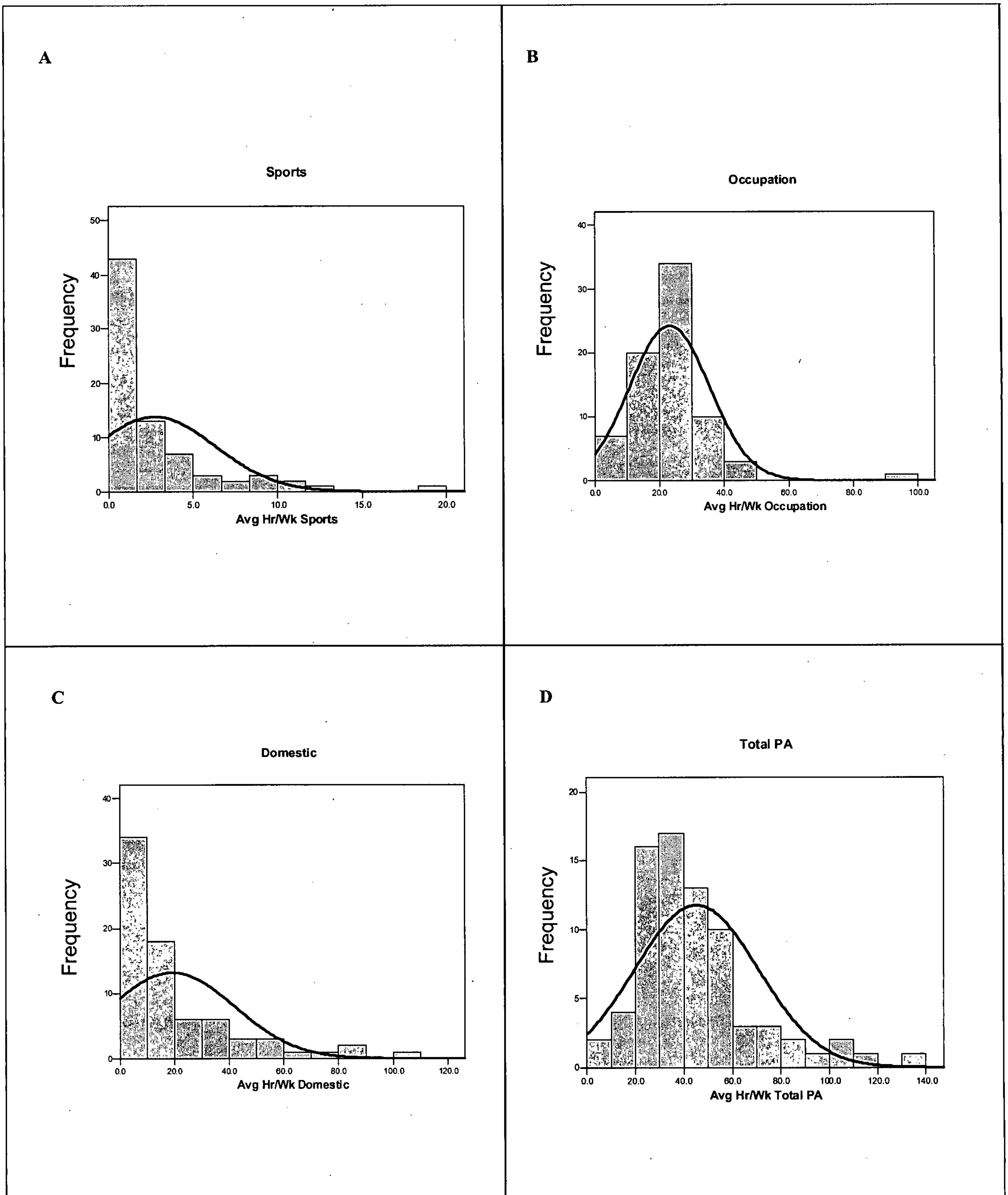


Figure A2.2. Baseline L-PAQ Lifetime Average Hr/Wk Frequency Histograms for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

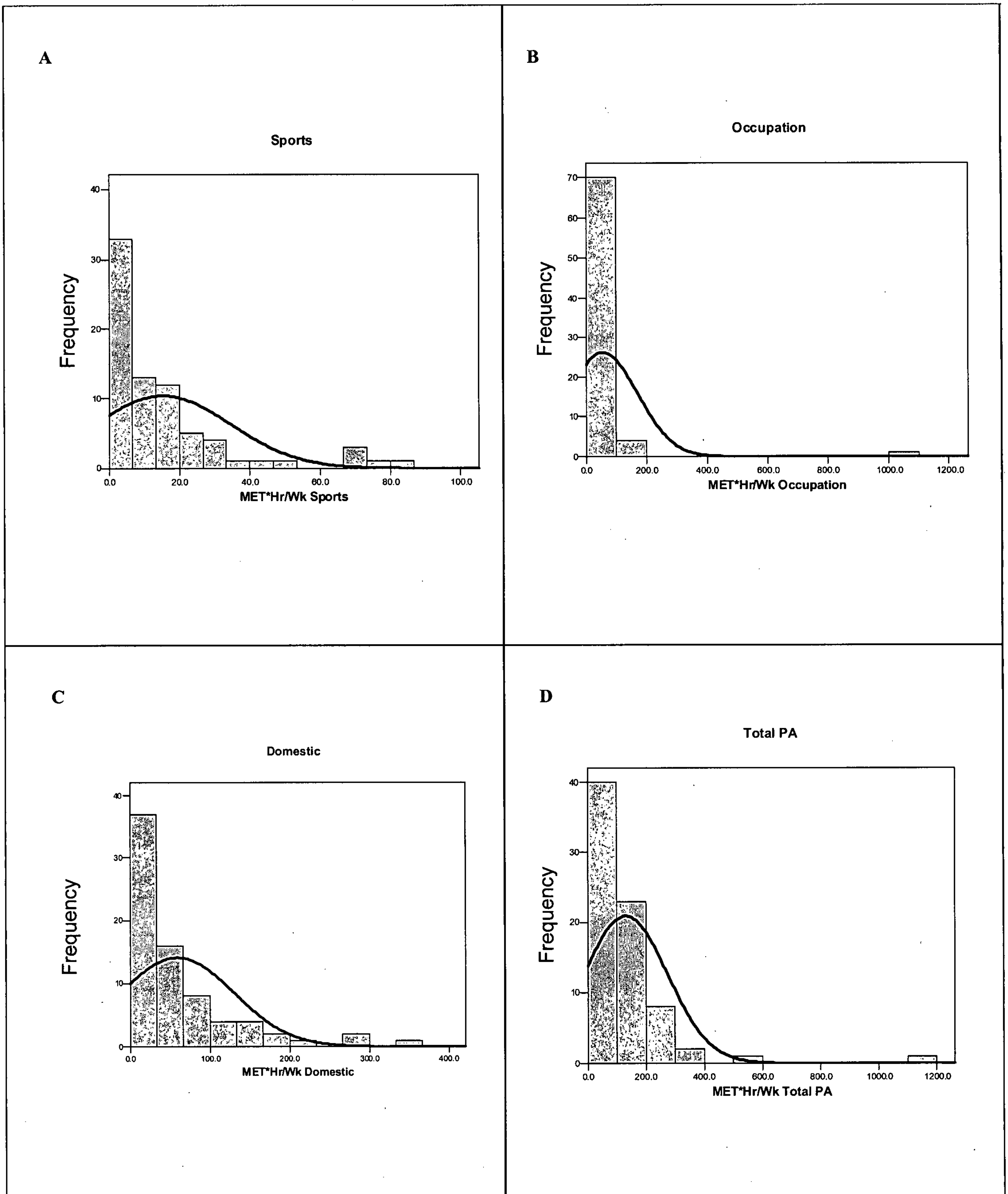


Figure A2.3. Baseline L-PAQ Lifetime MET•Hr/Wk Frequency Histograms for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

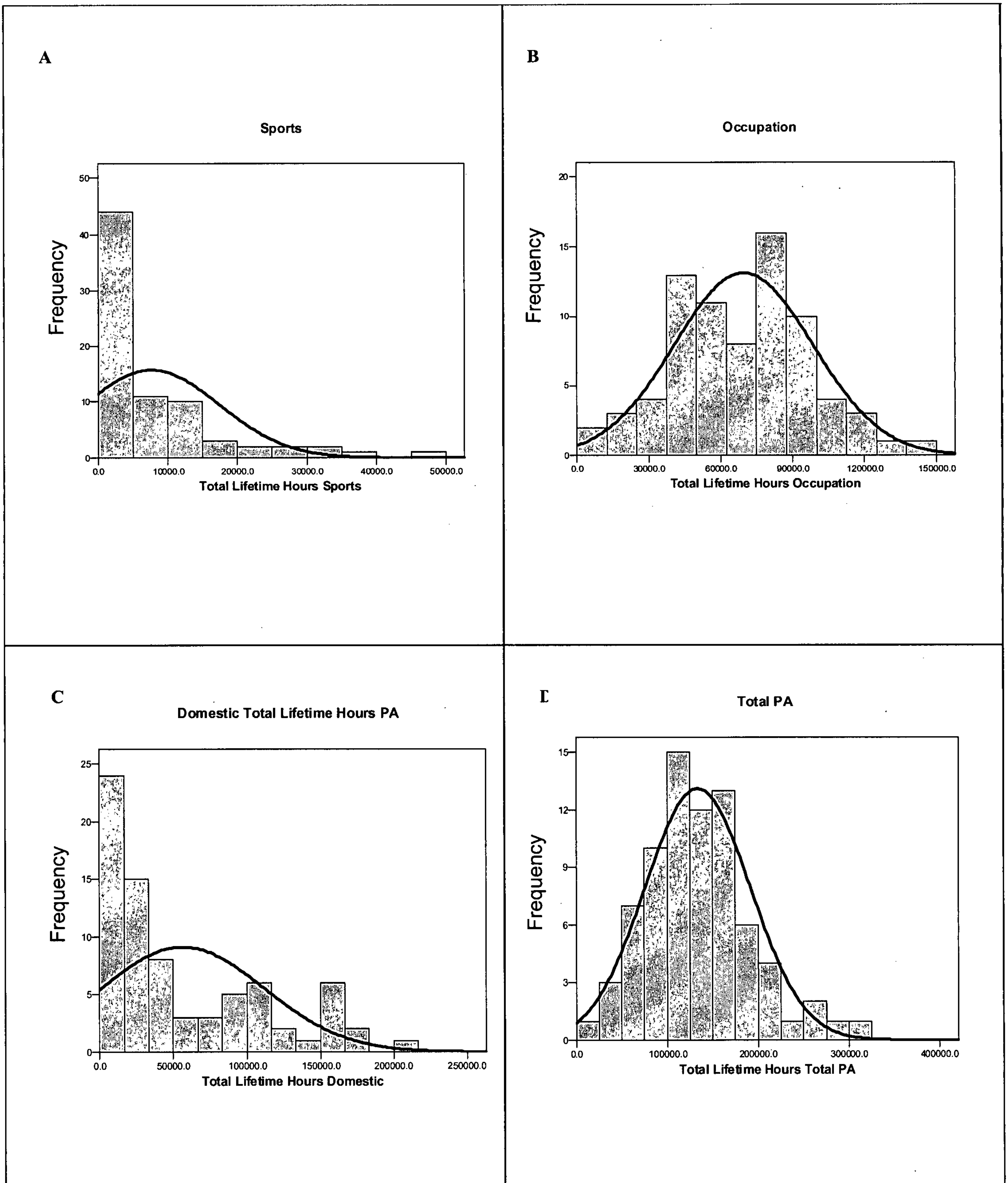


Figure A2.4. Retest L-PAQ Total Lifetime Hours Frequency Histograms for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

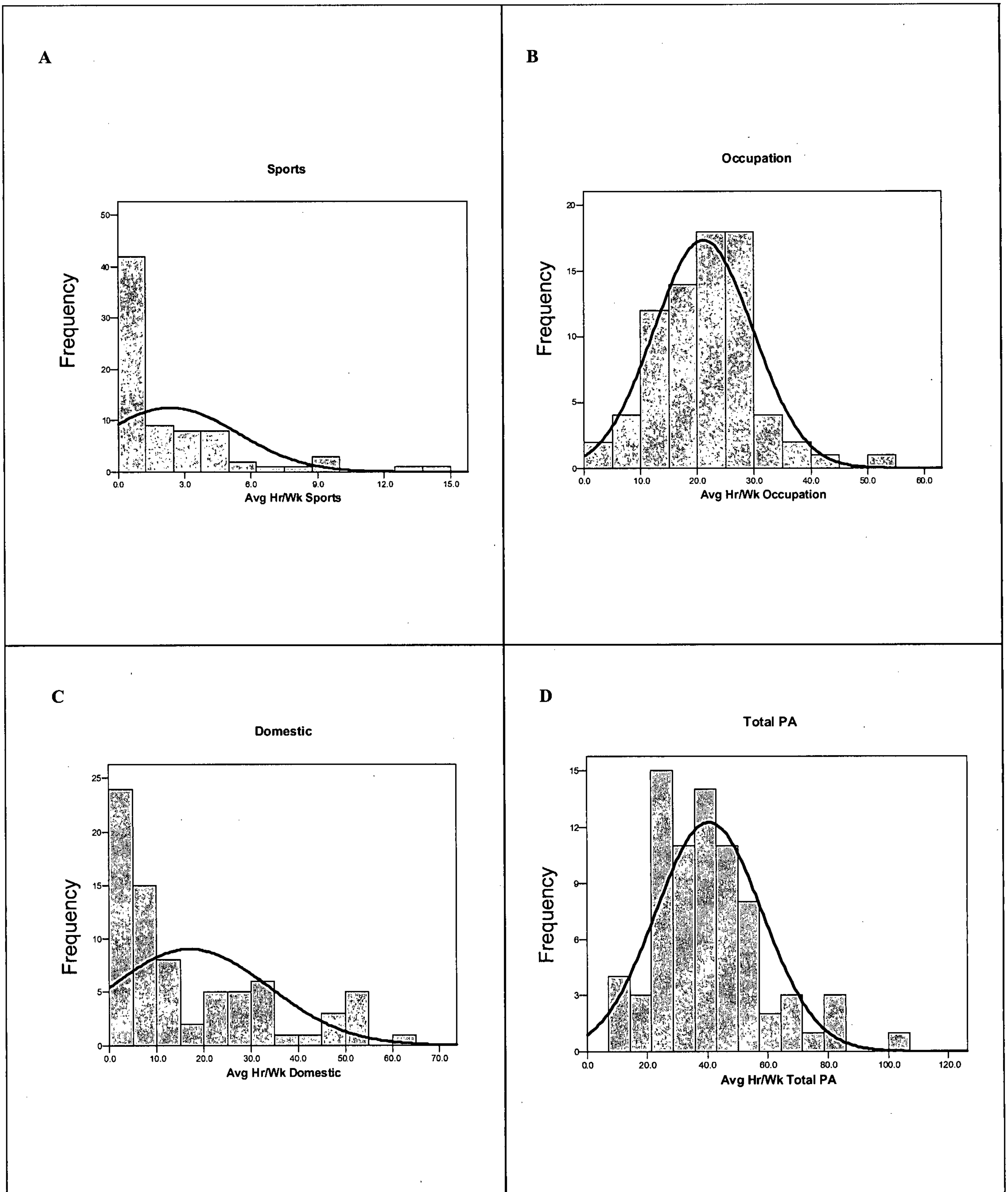


Figure A2.5. Retest L-PAQ Lifetime Average Hr/Wk Frequency Histograms for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

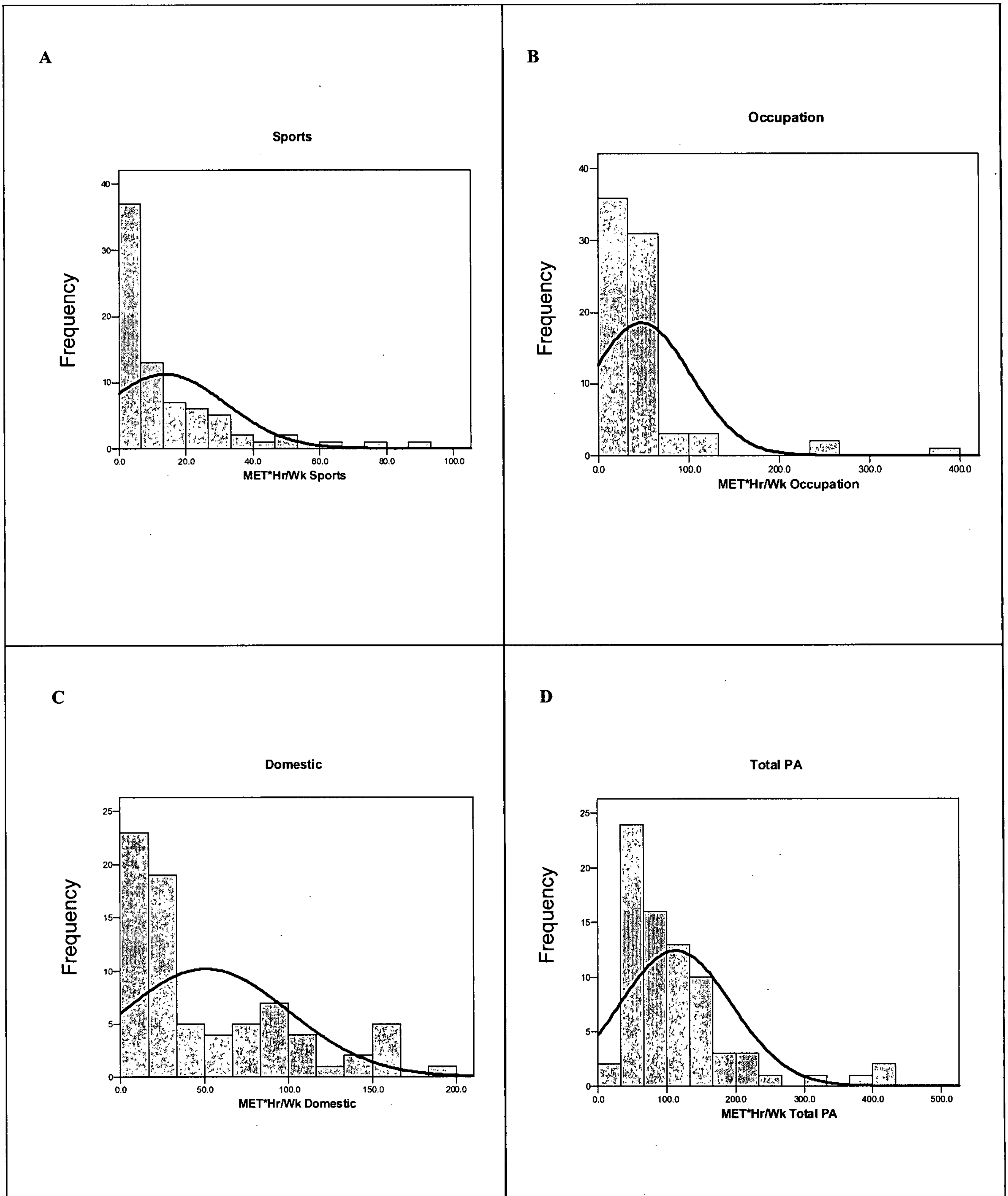
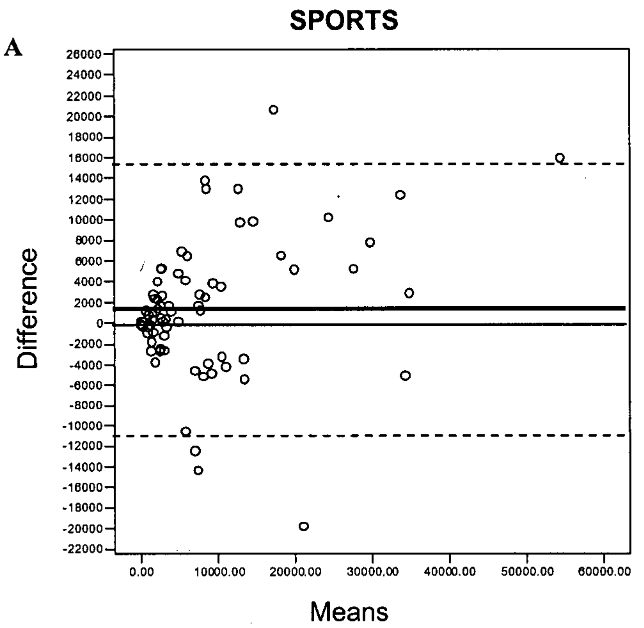
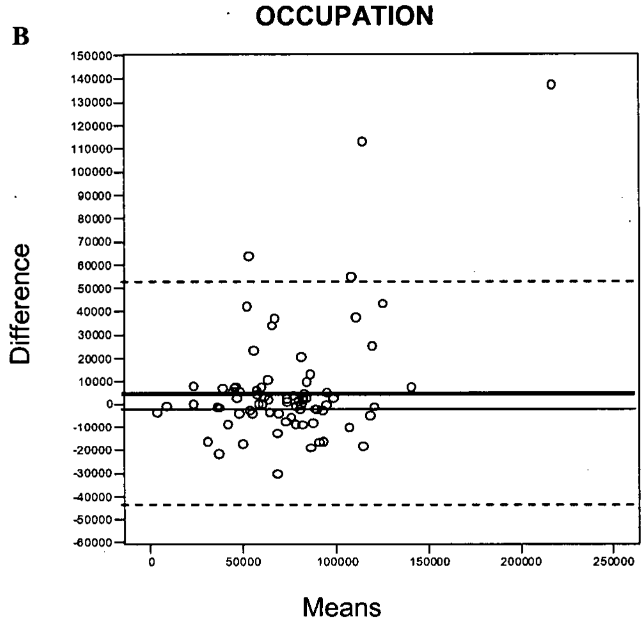


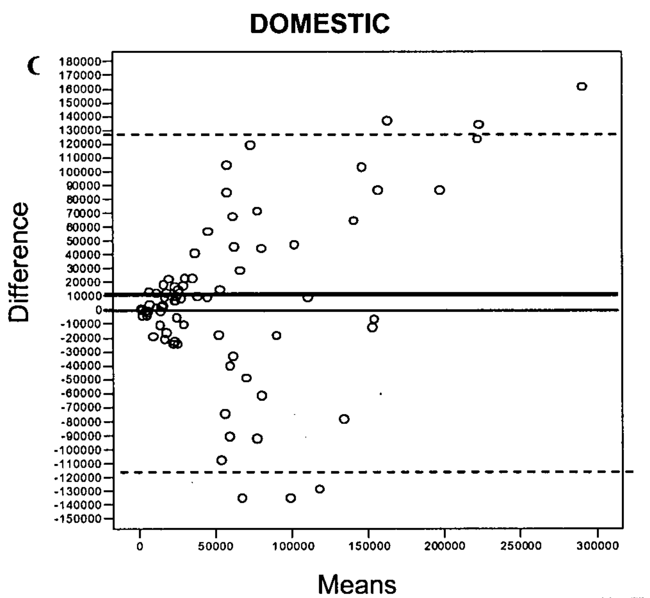
Figure A2.6. **<u>Retest L-PAQ</u> Lifetime MET•Hr/Wk Frequency Histograms for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity**



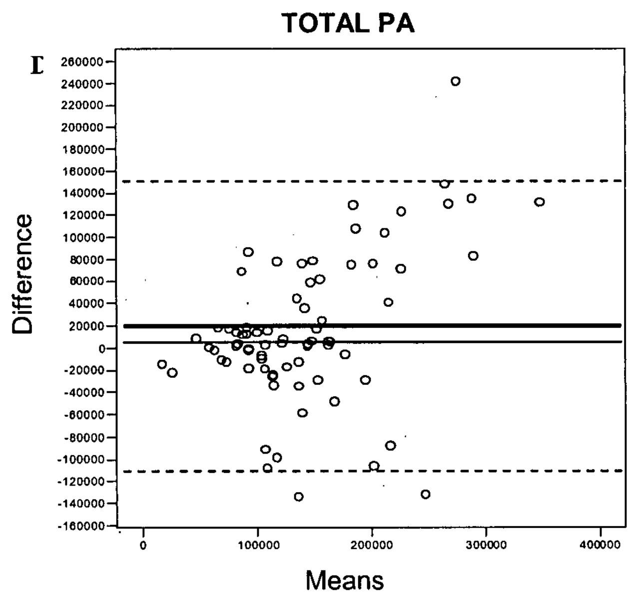
Mean of differences	1404.6
SD	6,424.9
2 x SD	12,849.8
Limits of Agreement	-11,445.2 to 14,254.4
SEM	4,543.0



Mean of differences	6,693.3
SD	25,926.9
2 x SD	51,853.8
Limits of Agreement	-45,160.5 to 58,547.1
SEM	18,333

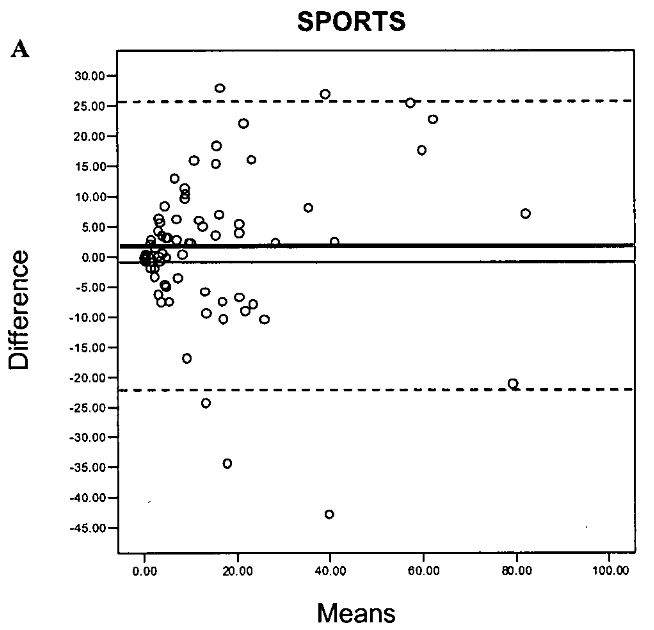


Mean of differences	8,476.3
SD	60,014.1
2 x SD	120,028.2
Limits of Agreement	-111,551.9 to 128,504.5
SEM	42,436.4

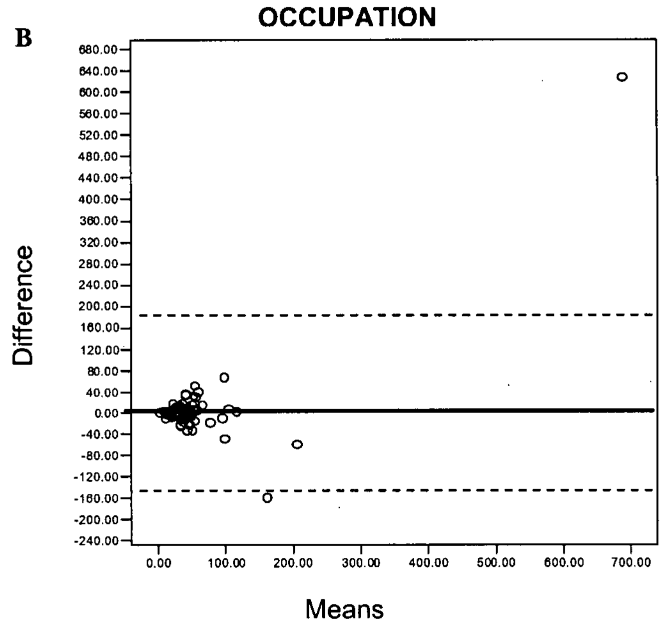


Mean of differences	16,574.2
SD	67,145.8
2 x SD	134,291.6
Limits of Agreement	-117,717.4 to 150,865.8
SEM	47,479.2

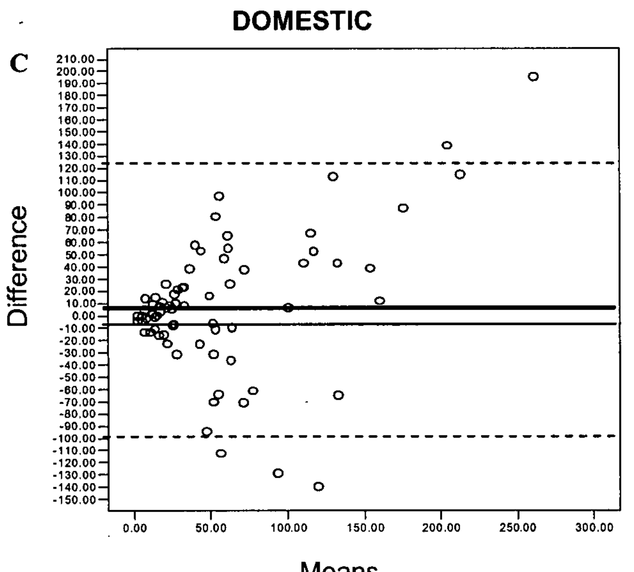
Figure A2.7. Bland-Altman Plots for Total Lifetime Hours (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity



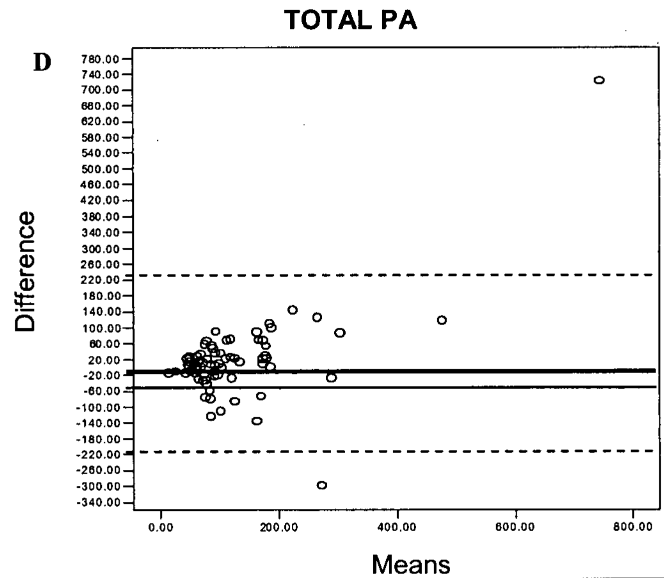
Mean of differences	1.58
SD	11.97
2 x SD	23.94
Limits of Agreement	-22.36 to 25.52
SEM	8.46



Mean of differences	7.74
SD	77.37
2 x SD	154.74
Limits of Agreement	-147 to 162.48
SEM	54.71



Mean of differences	8.74
SD	54.90
2 x SD	109.80
Limits of Agreement	-101.60 to 118.54
SEM	38.82



Mean of differences	18.07
SD	105.24
2 x SD	210.48
Limits of Agreement	-192.41 to 228.55
SEM	74.4

Figure A2.8. Bland-Altman Plots for Lifetime MET • Hours/Week (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity

Appendix 2.2 Validity Study Part I Figures

The following figures represent frequency histograms corresponding to results of analysis for the comparison between the L-PAQ and the LT-PAQ. Figures A2.9 to A2.11 correspond to scoring of the L-PAQ dataset used for this analysis. Each figure is a group of frequency histograms (with normal curves) corresponding to scores for each of the three physical activity domains and for total PA for each of the three L-PAQ scoring units: 1) total lifetime hours, 2) lifetime average hours/week, and 3) lifetime MET•hours/week. Figures A2.12 corresponds to frequency histograms for the MET•hours/week scoring unit for the LT-PAQ.

Scatter plots corresponding to the correlations between L-PAQ and LT-PAQ for lifetime average hours/week were presented in the main body of the thesis. Figure A2.13 represents correlations between L-PAQ and LT-PAQ for lifetime MET•hours/week.

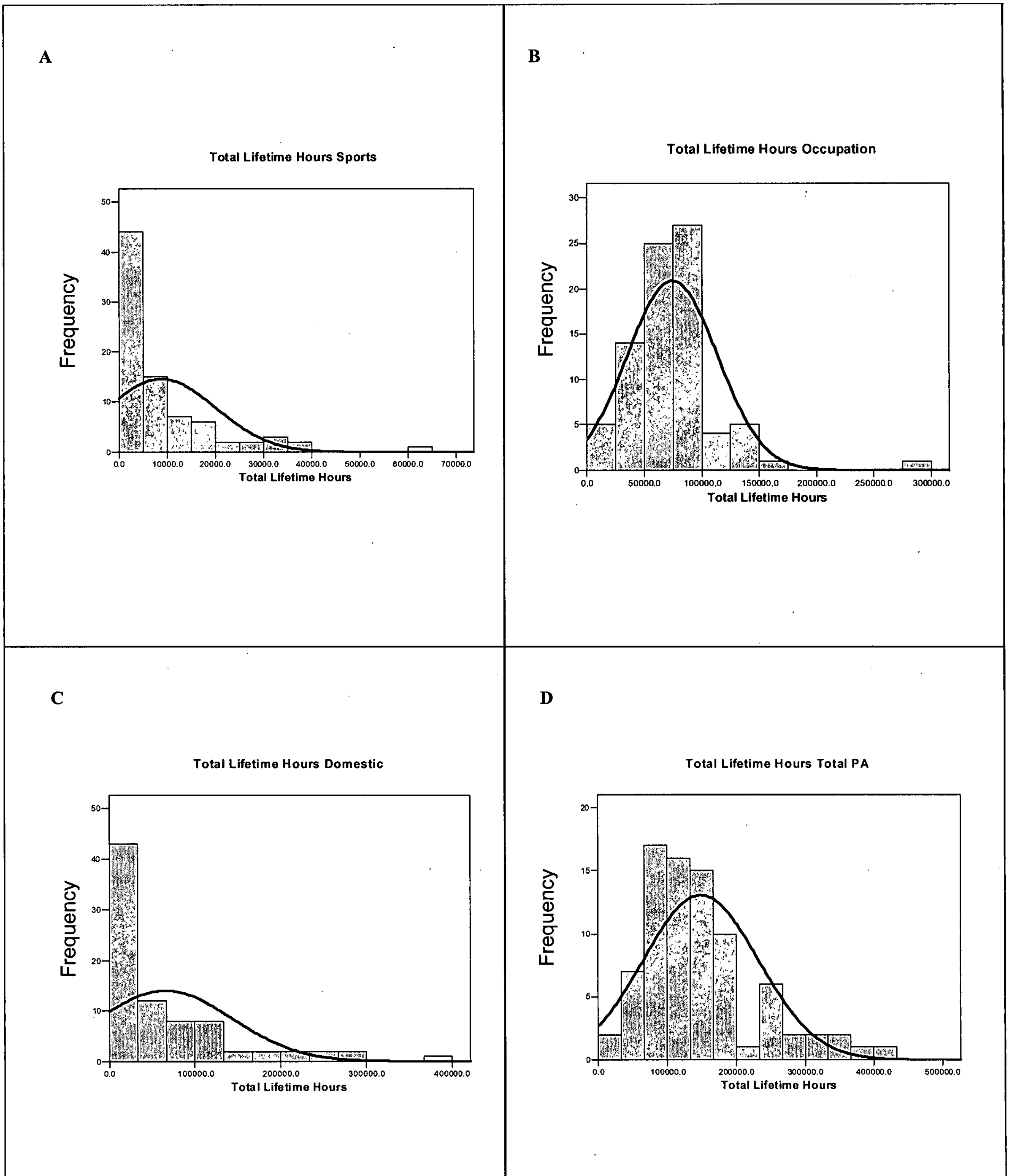


Figure A2.9. <u>Baseline L-PAQ</u> Total Lifetime Hours Frequency Histograms for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity (Data used for comparison with LT-PAQ)

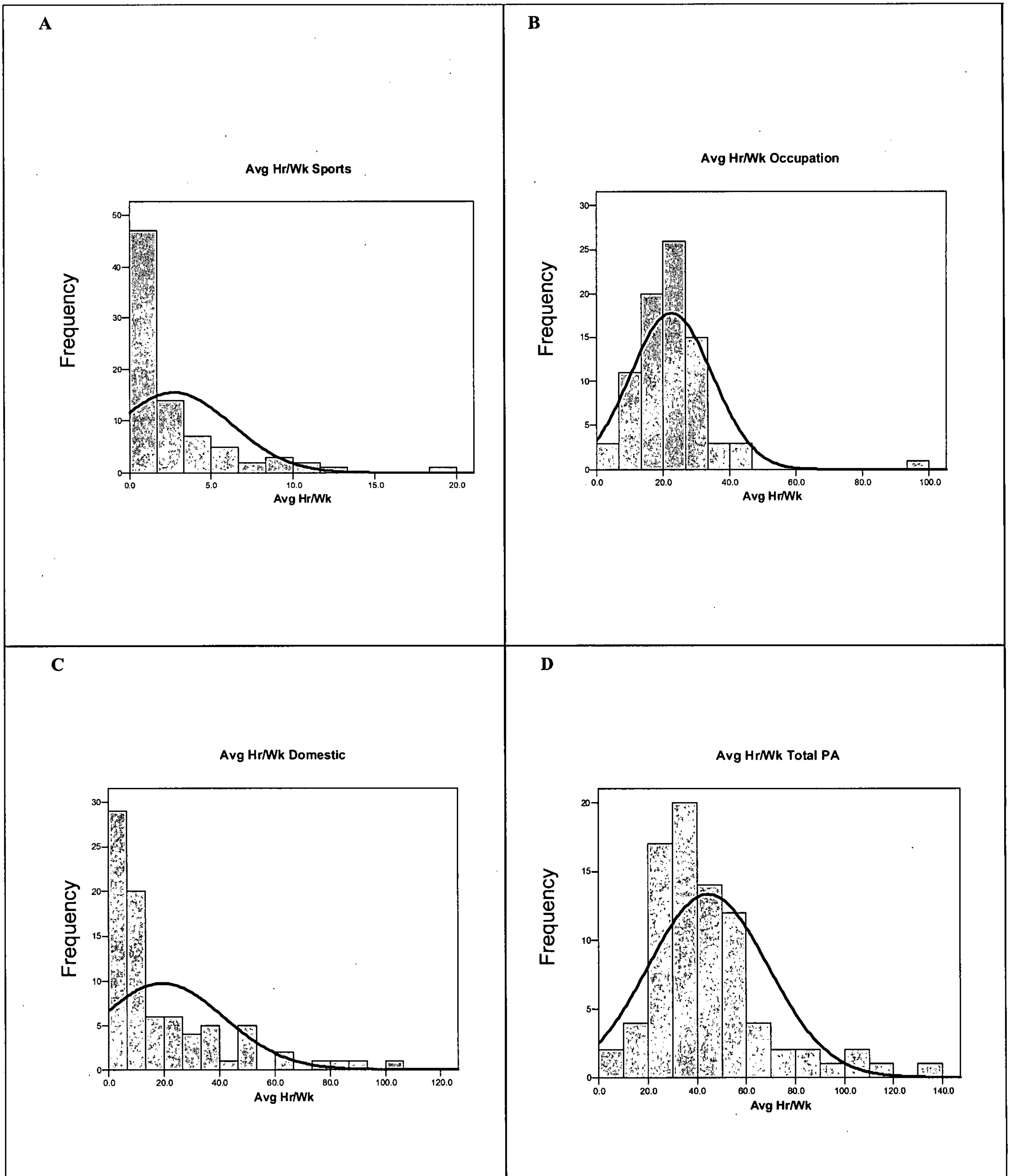


Figure A2.10. Baseline L-PAQ Lifetime Average Hr/Wk Frequency Histograms for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity (Data used for Comparison with LT-PAQ)

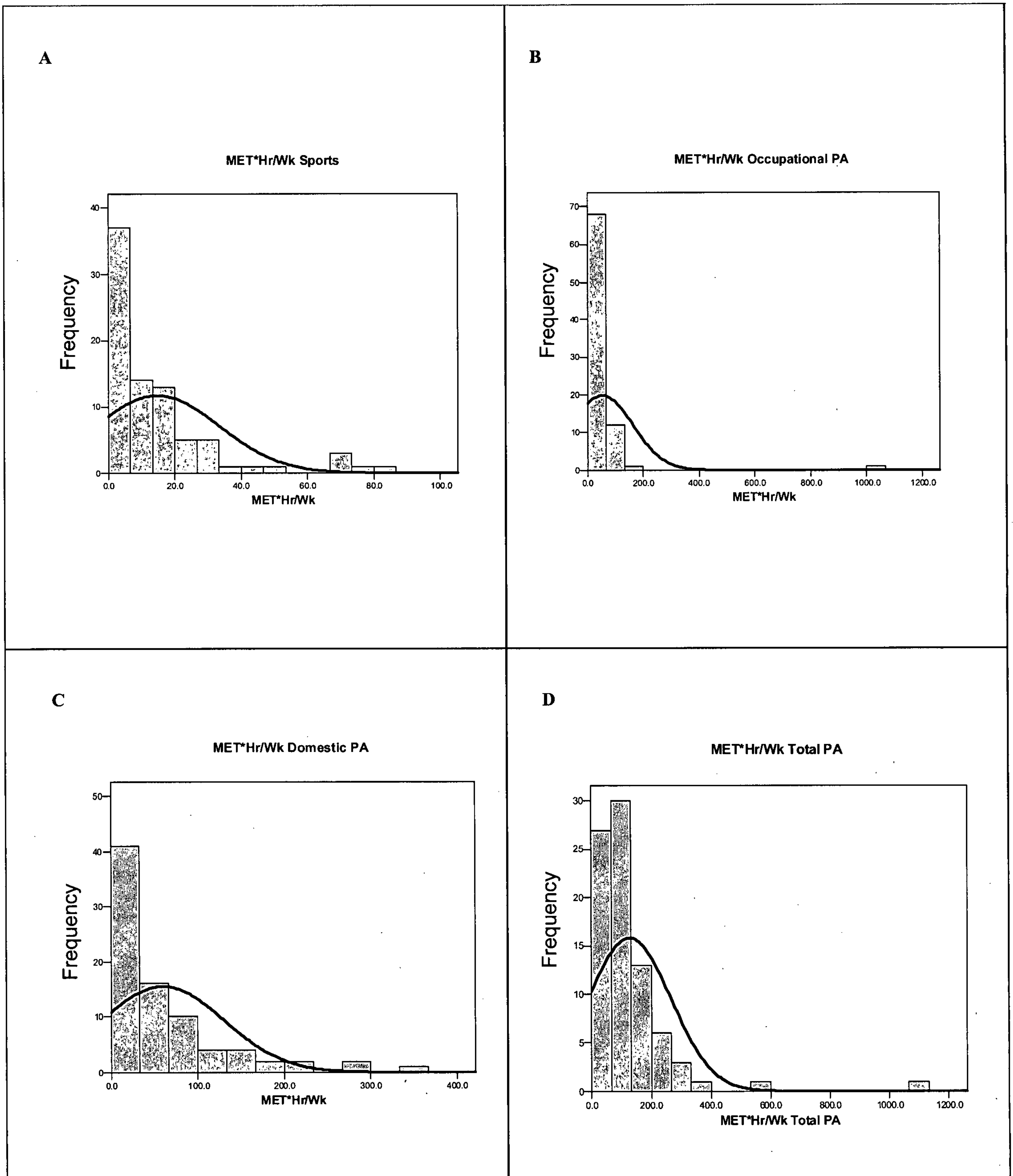


Figure A2.11. Baseline L-PAQ MET•Hr/Wk Frequency Histograms for (A) Sports, (B) Occupation, (C) Domestic, and (D) Total Physical Activity (Data Used for Comparison with L-PAQ)

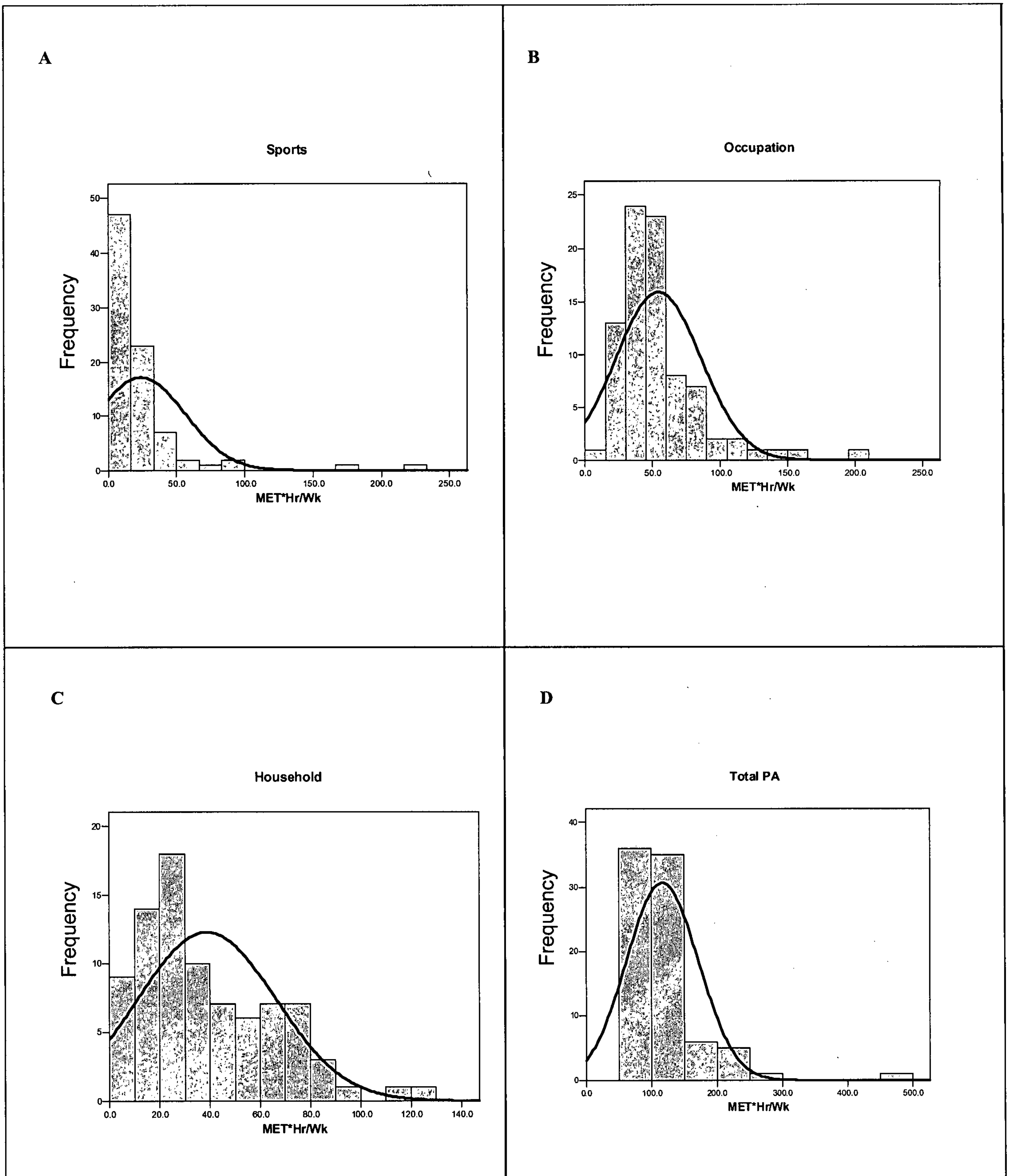


Figure A2.12. LT-PAQ MET•Hr/Wk Frequency Histograms for (A) Sports, (B) Occupation, (C) Household, and (D) Total Physical Activity (Data Used for Comparison with L-PAQ)

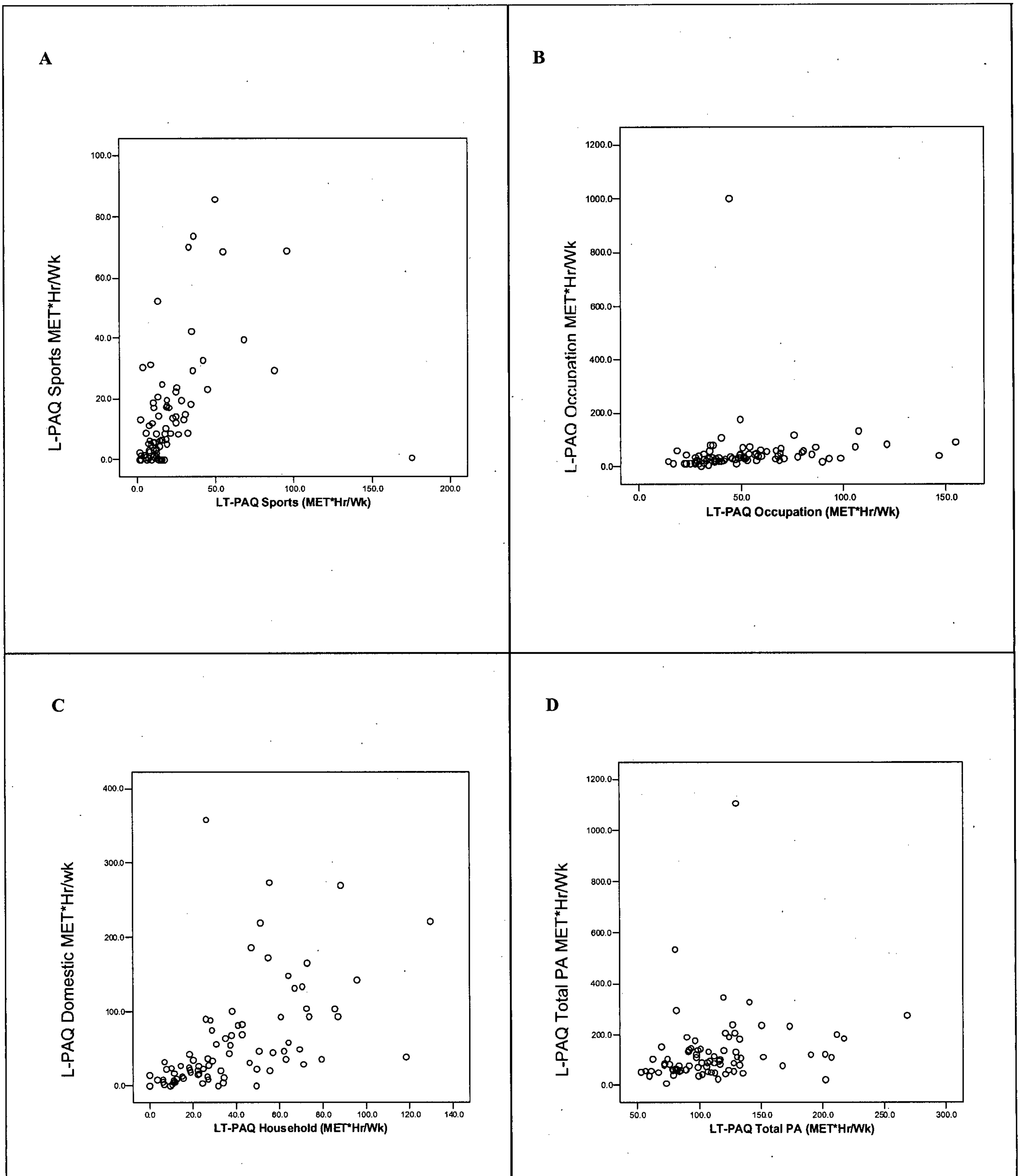


Figure A2.13. Scatterplots showing relationship between L-PAQ MET•Hr/Wk (y-axis) Versus LT-PAQ MET•Hr/Wk (x-axis) for (A) Sports; (B) Occupation; (C) Domestic; and (D) Total PA

Appendix 2.3 Validity Study Part II Figures

The following figures represent frequency histograms corresponding to results of analysis for the comparison between the L-PAQ and the CT-PAQ. Figures A2.14 to A2.16 correspond to scoring of the L-PAQ dataset used for this analysis. Each figure is a group of frequency histograms (with normal curves) corresponding to scores for each of the three physical activity domains and for total PA for each of the three L-PAQ scoring units: 1) total lifetime hours, 2) lifetime average hours/week, and 3) lifetime MET•hours/week. Figures A2.17 corresponds to frequency histograms for the MET•hours/week scoring unit for the CT-PAQ.

Scatter plots corresponding to the correlations between L-PAQ and LT-PAQ for lifetime average hours/week were presented in the main body of the thesis. Figure A2.18 represents correlations between L-PAQ and CT-PAQ for lifetime MET•hours/week.

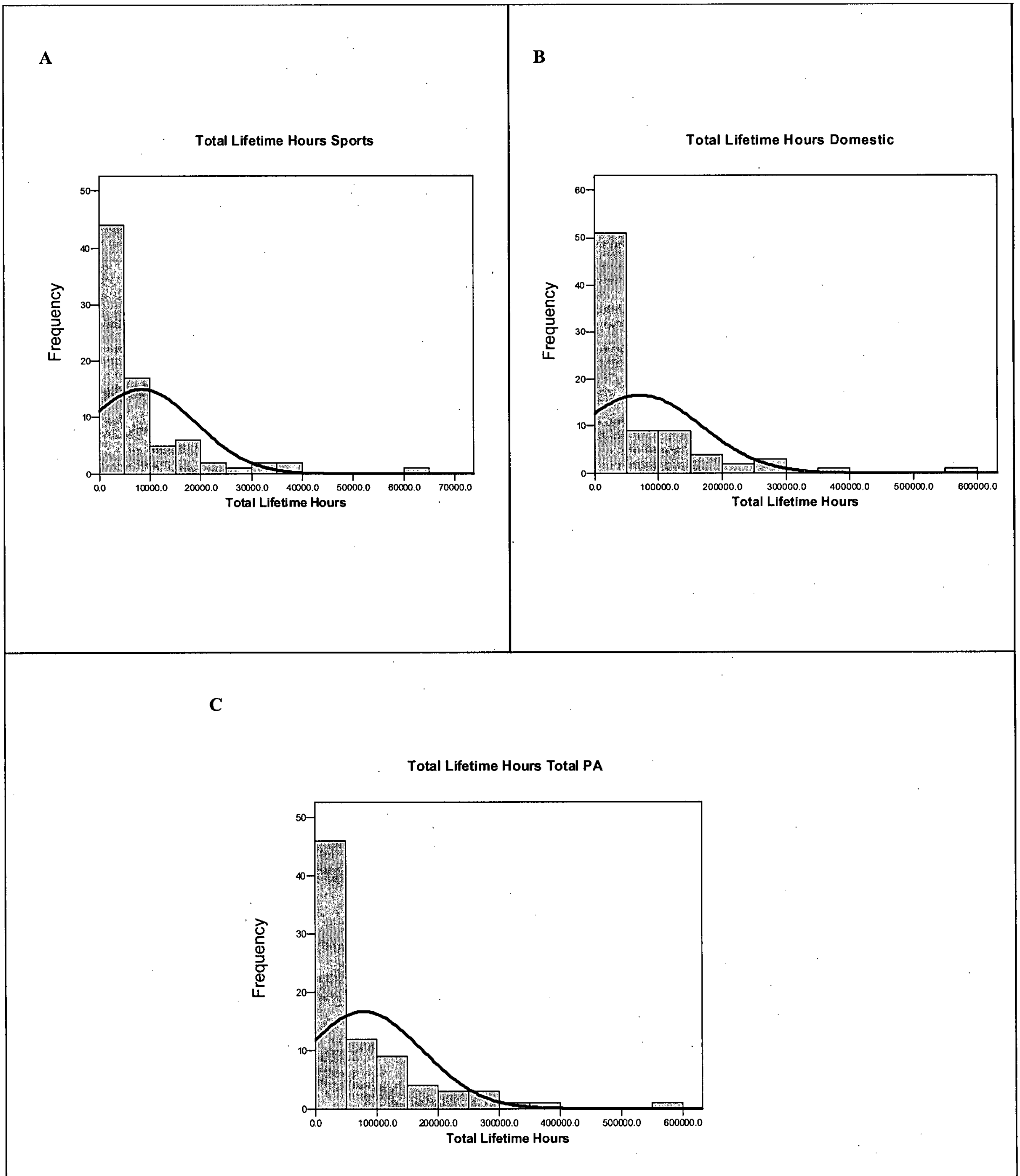


Figure A2.14. <u>Baseline L-PAQ</u> Total Lifetime Hours Frequency Histograms for (A) Sports; (B) Domestic; and (C) Total PA (Data Used for Comparison with CT-PAQ)

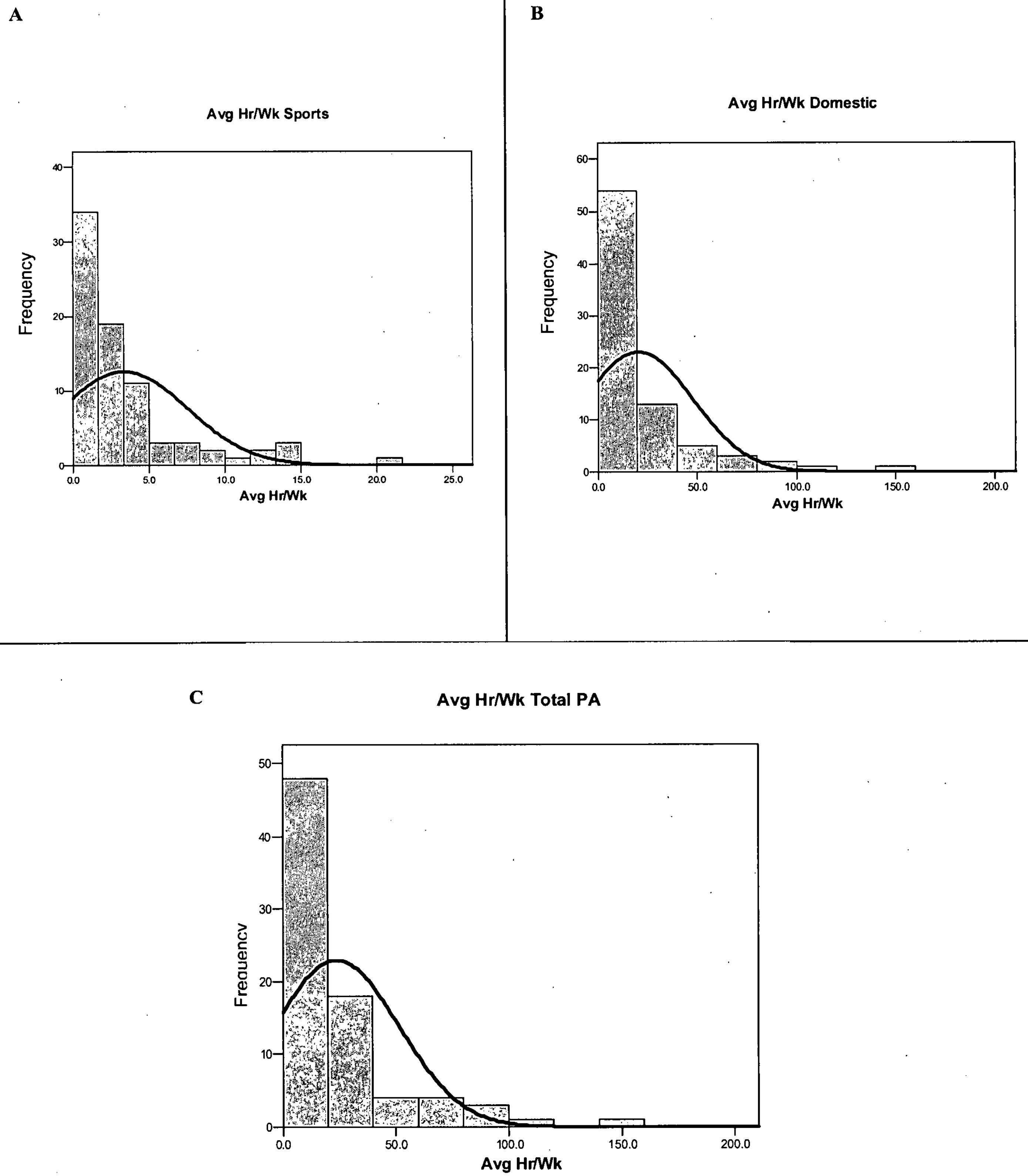


Figure A2.15 Baseline L-PAQ **Average Hours/Week Frequency Histograms for (A) Sports; (B) Domestic; and (C) Total PA (Data Used for Comparison with CT-PAQ)**

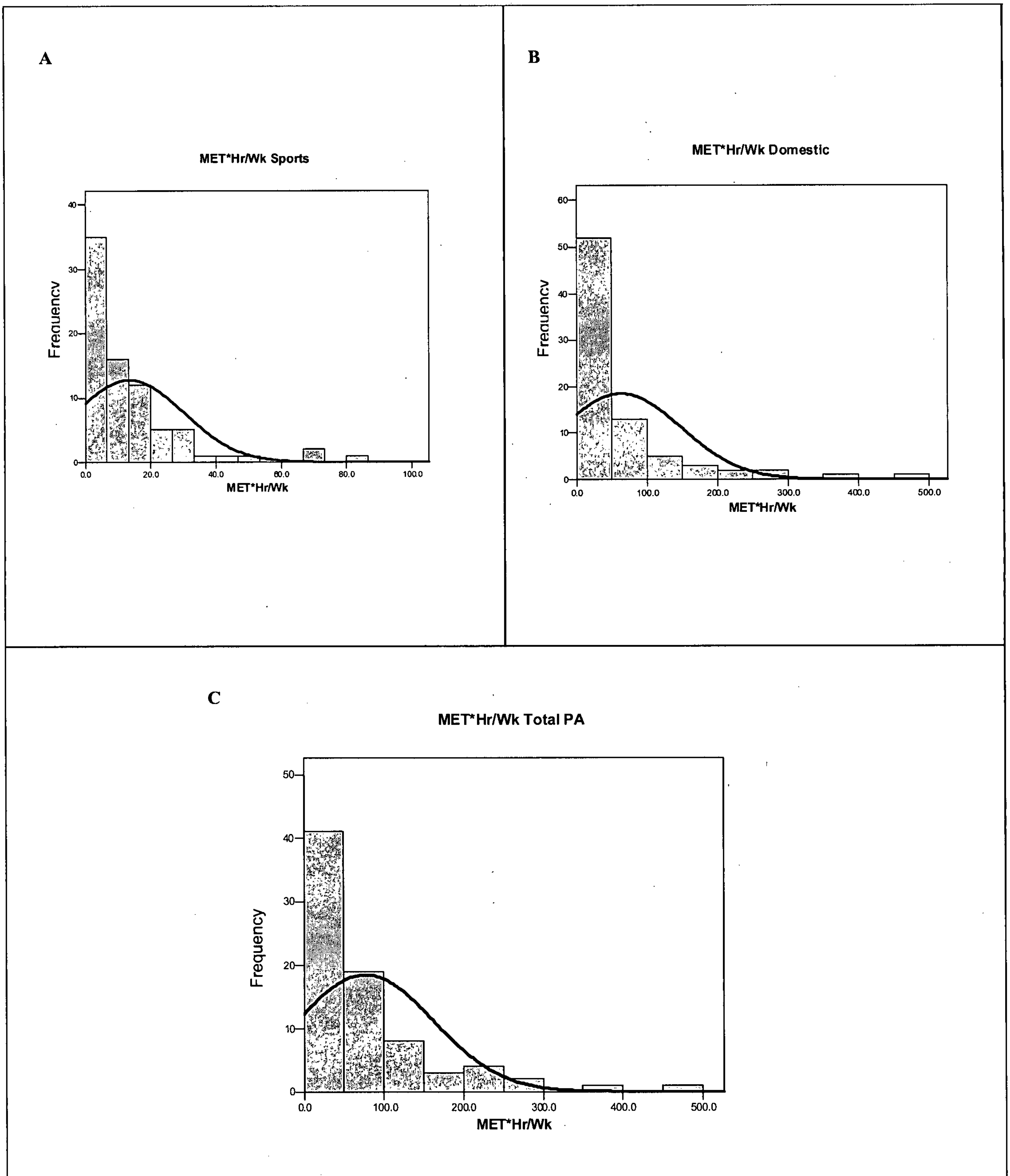


Figure A2.16. <u>Baseline L-PAQ</u> MET•Hours/Week Frequency Histograms for (A) Sports; (B) Domestic; and (C) Total PA (Data Used for Comparison with CT-PAQ)

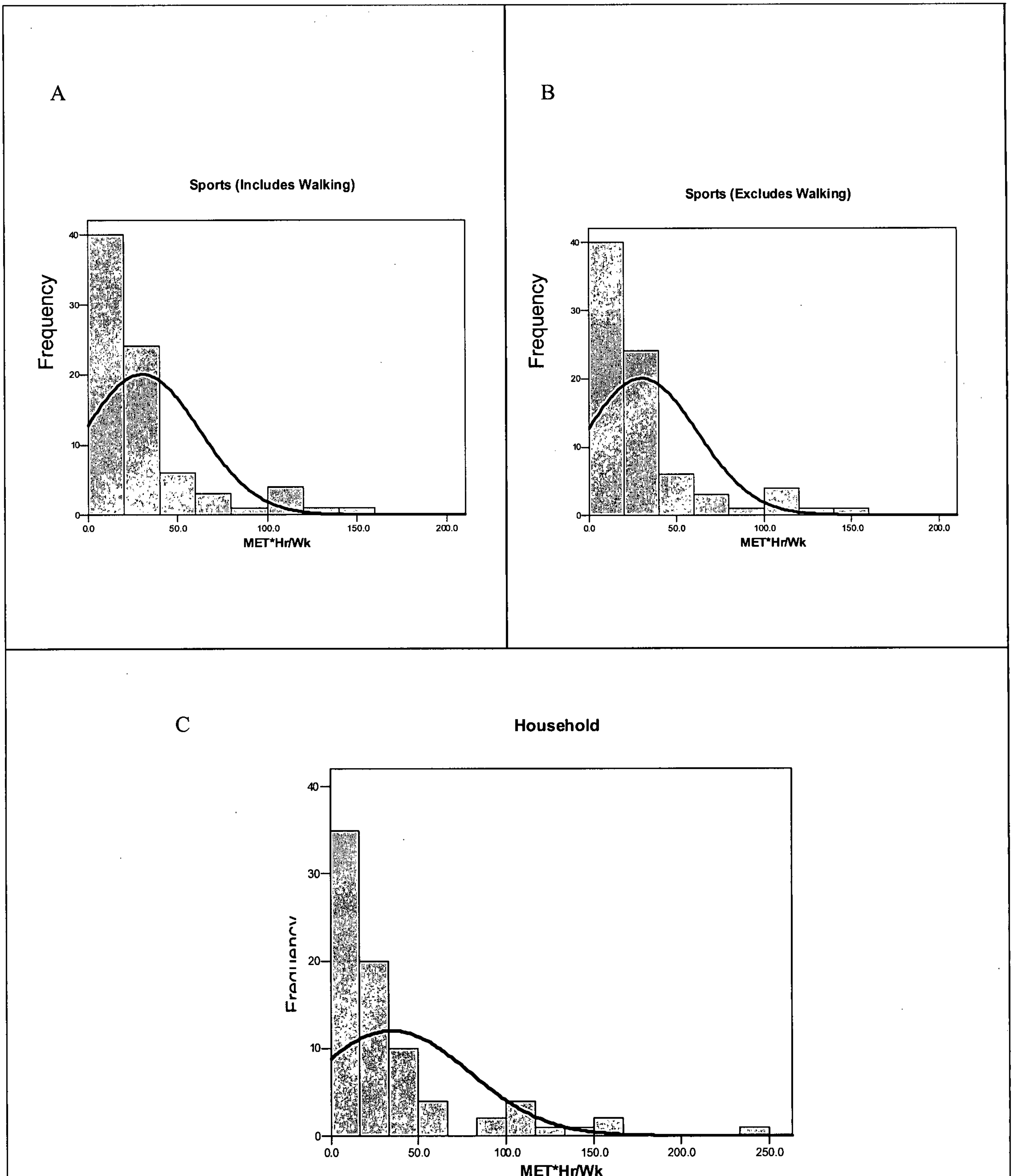
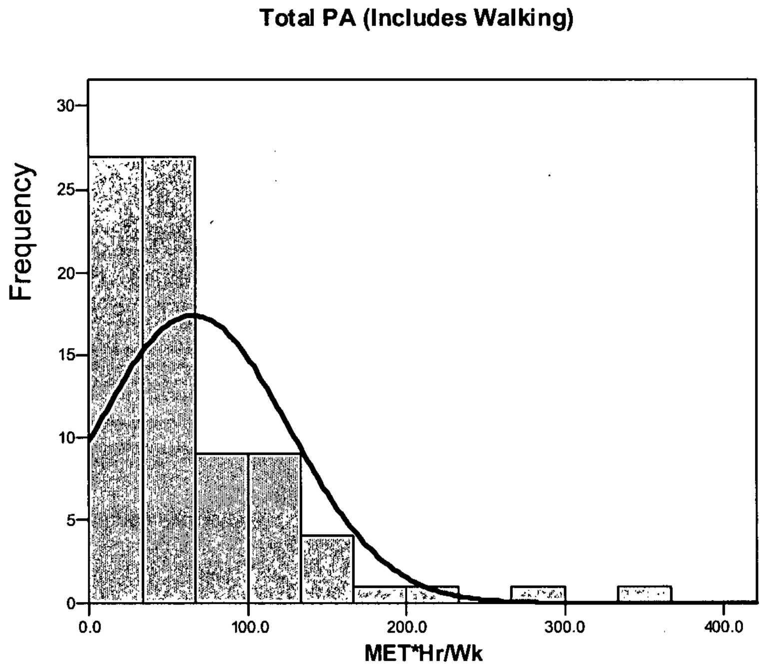


Figure A2.17(a). CT-PAQ Lifetime MET•Hr/Wk Frequency Histograms for (A) Sports (Includes Walking), (B) Sports (Excludes Walking), and (C) Household PA (Data Used for Comparison with L-PAQ)

A



B

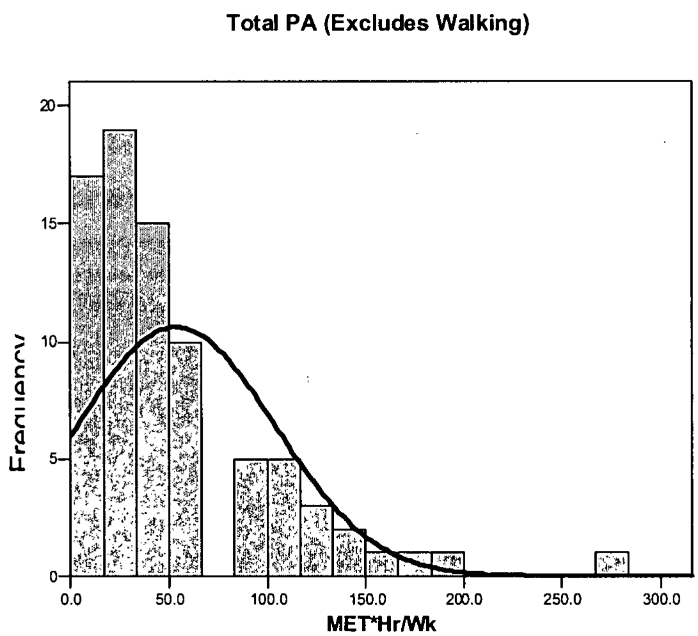


Figure A2.17(b). CT-PAQ Lifetime MET•Hr/Wk Frequency Histograms for (A) Total PA (Includes Walking) and (B) Total PA (Excludes Walking) (Data Used for Comparison with L-PAQ)

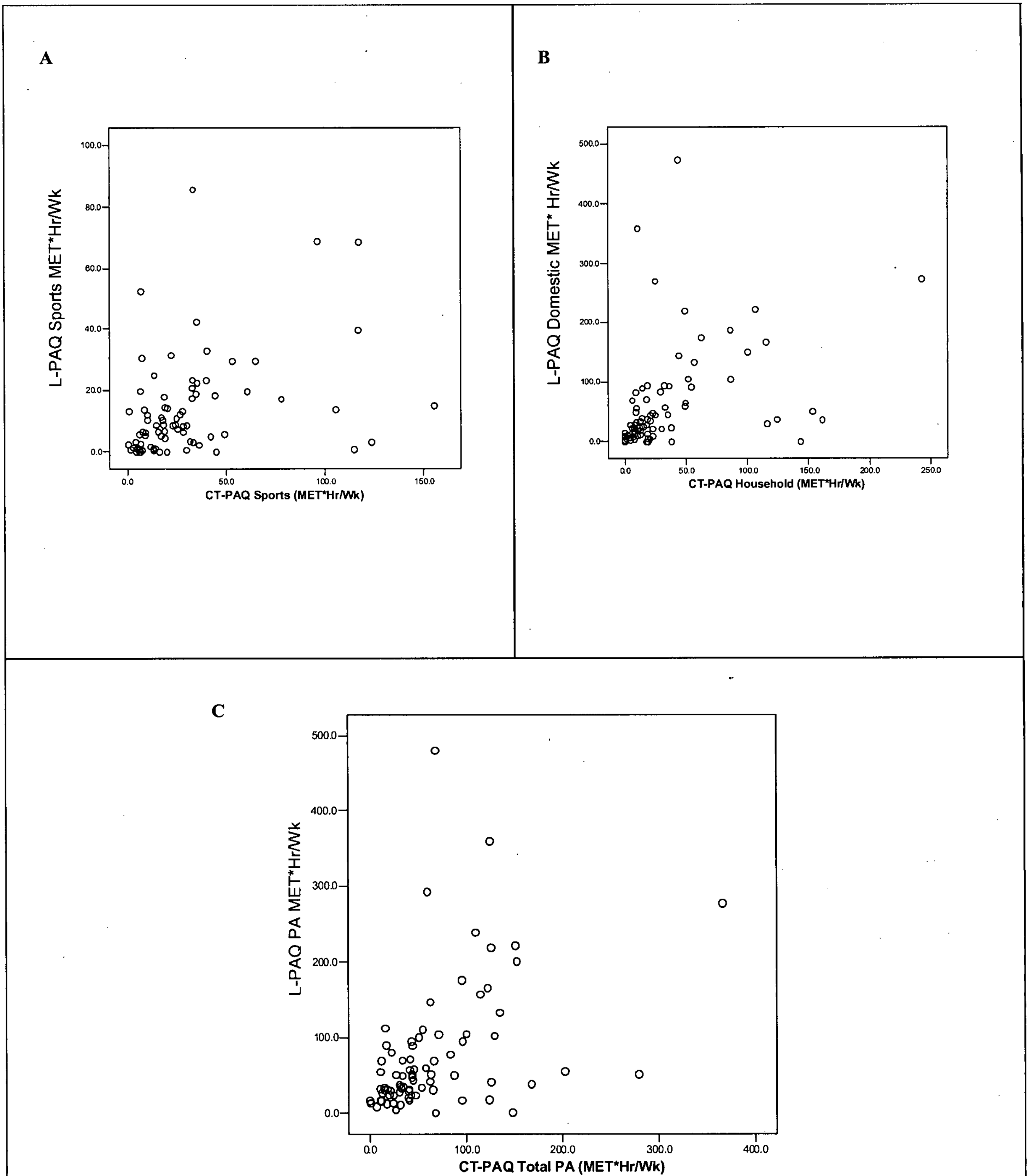


Figure A2.18 Scatter plots showing relationship between L-PAQ MET•Hr/Wk (y-axis) Versus CT-PAQ MET•Hr/Wk (Walking Included) (x-axis) for (A) Sports; (B) Domestic; and (C) Total PA

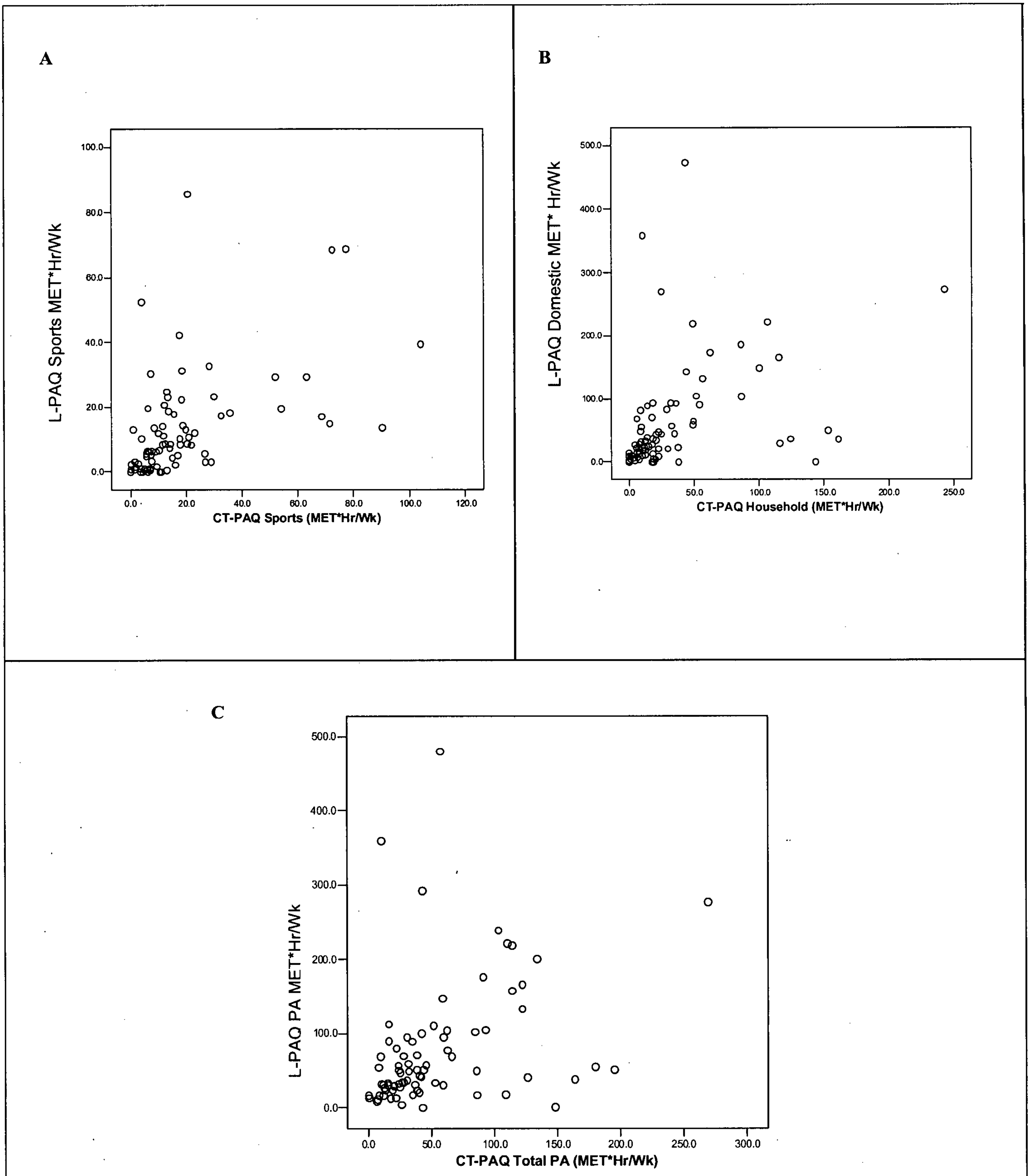


Figure A2.19 Scatter plots showing relationship between L-PAQ MET•Hr/Wk (y-axis) Versus CT-PAQ MET•Hr/Wk (Walking Excluded) (x-axis) for (A) Sports; (B) Domestic; and (C) Total PA